nanoString TECHNOLOGIES®



nSolver™ 3.0

Analysis Software User Manual

NanoString Technologies®, Inc.

530 Fairview Ave N
Seattle, Washington 98109

www.nanostring.com

Telephone: 206.378.6266
888.358.6266

E-mail: info@nanostring.com

Molecules That Count®

Translational Research • Gene Expression • miRNA Expression • Epigenomics • Copy Number Variation

MAN-C0019-06, April 2016

FOR RESEARCH USE ONLY. Not for use in diagnostic procedures.

Intellectual Property Rights

This nSolver™ Analysis Software user manual and its contents are the property of NanoString Technologies, Inc. ("NanoString"), and is intended for the use of NanoString customers solely in connection with their operation of the nCounter® Analysis System. The nCounter Analysis System (including both its software and hardware components) and this User Manual and any other documentation provided to you by NanoString in connection therewith are subject to patents, copyright, trade secret rights, and other intellectual property rights owned by or licensed to NanoString. No part of the software or hardware may be reproduced, transmitted, transcribed, stored in a retrieval system, or translated into other languages without the prior written consent of NanoString. For a list of patents, see Patents | NanoString Technologies.

Limited License

Subject to the terms and conditions of sale of the nCounter Analysis System, NanoString grants you a limited, non-exclusive, non-transferable, non-sublicensable, research use only license to use this proprietary nSolver software with the nCounter Analysis System only in accordance with this manual, the manual for the nCounter Analysis System, and other written instructions provided by NanoString. Except as expressly set forth in the terms and conditions, no right or license, whether express, implied, or statutory, is granted by NanoString under any intellectual property right owned by or licensed to NanoString by virtue of the supply of this software or the proprietary nCounter Analysis System. Without limiting the foregoing, no right or license, whether express, implied, or statutory, is granted by NanoString to use the nSolver Analysis Software or nCounter Analysis System with any third party product not supplied or licensed to you by NanoString, or recommended for use by NanoString in a manual or other written instruction provided by NanoString.

Trademarks

NanoString Technologies, NanoString, nCounter, nSolver, nDesign Gateway, and Molecules That Count are registered trademarks or trademarks of NanoString Technologies, Inc., in the United States and/or other countries. All other trademarks and/or service marks not owned by NanoString that appear in this manual are the property of their respective owners.

Copyright

©2016 NanoString Technologies, Inc. All rights reserved.

Contents

1. Introduction 7

nSolver Analysis Software Overview 7

nSolver 3.0 Release Details 8

System Requirements 8

Support Options and Contacts 8

nSolver Basics 8

Quick Start 9

Workflow Wizards 10

Home Screen 10

Raw Data Main View Screen 11

Experiment Main View Screen 12

2. Importing Data 13

Importing RLF Files 13

Importing RCC Files 16

3. Quality Control Metrics and Flags 19

GX and GX Plex^2 Quality Control 21

miRNA and miRGE Quality Control 22

CNV Quality Control 23

Protein Quality Control 24

GeneProtein Quality Control 25

'Other' Quality Control 26

QC Parameters and Guidelines 27

Common QC Metrics—Imaging 27

Common QC Metrics—Control Probes 27

CNV-Specific Metrics 27

Reviewing QC Results 28

Re-Running Quality Control Checks 32

4. Adding Lane/Sample Descriptions and Gender to Raw Data 35

Gender 35

Descriptions 36

5. Adding Custom Annotations to Probes 39

Probe Annotations 39

6. Viewing Raw Data 43

7. Studies 47

8. Experiments 51

Using the Experiment Design Wizard 51

Creating a New Experiment 52

Creating Experiments from Multiple RLFs 53

Selecting Files to Use in the Experiment 53

Creating a CrossRLF Experiment 54

Overlapping Probes 55

Annotating Lanes/Samples for Grouping and Ratio Building 55

Background Subtraction 60

Background Subtraction Using Negative Controls 60

Background Subtraction for GeneProtein 60

Background Subtractions Using a Specified Value 61

Background Subtraction Using Blank Lane 62

Normalization 63

Gene Expression Normalization Dialog Box 65

GeneProtein Normalization Dialog Box 66

Protein Normalization Dialog Box 67

miRNA Normalization Dialog Box 68

CNV Normalization Dialog Box 69

miRGE Normalization Dialog Box 70

Other Normalization Dialog Box 71

Inter-CodeSet Calibration (Cross RLF and Plex^2 Only) 72

Estimating Fold Changes 73

Understanding Baselines 74

Choosing a Baseline 74

Creating a MultiRLF Experiment 77

Selecting Experiments to Merge in MultiRLF Experiments 77

Overlapping Probes 78

Merging Normalized Data for MultiRLF Experiments 78

Selecting Columns for the Merged Data Table 79

Estimating Fold Changes 80

Editing an Experiment 84

9. Experiment Output—Data Tables and Reports 88

Viewing Analyzed Data 88

Normalized Data Table 90

Grouped Data Table 91

Ratio Data Table 92

Plex^2 Assay Data Tables 94

Significance Testing 96

Error Model—(DE Call) 98

T-test ... 98
False Discovery Rate ... 98
Experiment Reports ... 99

10. Exporting Data ... 102

Exporting All Data ... 103
Exporting a Subset of a Data Table ... 103
Custom Export ... 105
RCC Collector Format Export ... 109
Partek® Genomics Suite™ Export ... 111
BioDiscovery Nexus Copy Number™ Export ... 112
nDesign™ Gateway Integration ... 114

11. Analyses and Visualizations ... 116

Analysis Wizard ... 116
Analysis Properties ... 117
Histogram Filter ... 118
Samples to Be Included in Analysis ... 120
Genes/Probes to Be Included in Analysis ... 121
Agglomerative Cluster (Heat Map)-Specific Parameters ... 122
Scatter Plot Parameters ... 123
Analyses/Visualizations ... 123
Common Settings (Excludes Heat Maps) ... 123
Agglomerative Cluster (Heat Map) ... 126
Violin Plots ... 127
Box Plots ... 130
Scatter Plot ... 134
Histogram Plot ... 137
Managing Saved Analyses/Visualizations ... 138
Viewing Analyses ... 138
Editing Analyses ... 138
Analysis Report ... 138
Advanced Analyses ... 139
Third-party Software Requirements ... 139
Installing New Advanced Analyses ... 139
Changing the R Home Path ... 140
Advanced Analysis Wizard ... 141
Advanced Analysis Properties ... 142
Sample Annotations ... 142
Import Annotations ... 143
Gene Annotations ... 145
Custom Parameters for Advanced Analyses ... 147
Managing Advanced Analyses ... 148
Viewing Analyses ... 148
Editing Analyses ... 150

Analysis Report 150

12. Data Management 152

Views/Tabs 152
Raw Data Main View/Tab 152
Experiment Main View/Tab 152
List and Properties Minor Views/Tabs 153

Filter/Search 154
Search Based on Fold Change 154
Search for a List (Such as Genes) 155

Table Functionality 156
Sort 156
Customizing Columns Displayed 159
Row Counts 161
Copying Data in Tables 161
Export / Save 162
Close 162

13. Glossary 164

14. Java TreeView Copyright Notice 166

15. R.app Copyright Notice 168

1 Introduction

nSolver Analysis Software Overview

NanoString Technologies nCounter assays are designed to provide a single-tube, ultra-sensitive, reproducible, and highly-multiplexed method for detecting nucleic acid targets across all levels of biological expression. These assays provide a method for direct detection of targets using molecular barcodes without the use of reverse transcription or amplification. nCounter assays are processed on the fully-automated Prep Station and data are collected and tabulated by the Digital Analyzer for each cartridge. This user manual describes the nSolver Analysis Software that you can use to analyze the nCounter data.

This product allows customers using Research Use Only (RUO) products to quickly and easily perform Quality Control (QC), normalization, and basic analyses of their nCounter Analysis System data without the need to purchase an additional software package. In addition, this product integrates with, or exports data in formats compatible with, commercially-available software packages designed for more sophisticated analyses and visualizations (see *Chapter 14, Glossary* for a definition of terms). The nSolver product allows customers to do the following:

- Import Data
- Annotate Genes & Add Sample Descriptions
- Define Experimental Analysis (Annotate Samples)
- Perform Quality Control
- Normalize
- Build Ratios
- View Tables
- Integrate with Pathway Applications and nDesign Gateway
- Export Data
- Create Analyses and Visualizations
- Search

The nSolver product provides tailored support of the following assay types:

- Copy Number Estimate (CNV), including Plex2 Assay
- Gene Expression (GX), including Plex2 Assays
- GeneProtein (combination mRNA and Protein)
- microRNA (miRNA)
- miRGE (combination miRNA and mRNA)
- Protein
- Other customer developed applications

NOTE: If you are upgrading from Version 1 of the nSolver Analysis Software and want to migrate data from that version, contact the nSolver support team at **support@nanostring.com**

nSolver 3.0 Release Details

The nSolver 3.0 release includes expansion of supported nCounter assay types to include Protein and GeneProtein assay types, support for data generated on *SPRINT* systems and multiRLF data merging.

If you are upgrading from nSolver 2.6, your data will be copied into the nSolver 3.0 database automatically the first time you launch nSolver 3.0.

System Requirements

The nSolver Analysis Software is compatible with both Macintosh (10.10–10.11) and Windows (8.1 and 10) operating systems. Java 1.7 or higher must be installed separately on computers using Macintosh, but for Windows, the nSolver installer automatically installs the correct version of Java.

Advanced Analyses are available in nSolver versions 2.5 and higher, and require the use of R, a statistics software package freely available to the public. There is an option to install R automatically during installation of nSolver on Windows operating systems. If using a Macintosh operating system, then R must be installed separately before using the Advanced Analyses feature. Customers using Macintosh 10.9 and later must also install XQuartz. Links are provided to download R and XQuartz on the NanoString website and in this user manual.

Separate installers are provided for Mac and Windows. It is essential that you install nSolver using the appropriate installer. If you have any questions or concerns about the installer please email the nSolver support team at **support@nanostring.com**.

Support Options and Contacts

Please contact **support@nanostring.com** for any issues or requests. In addition to this manual, please visit the customer website for other documentation, including the ***nCounter® Expression Data Analysis Guide*** and the ***nCounter® Data Analysis Guidelines for Copy Number Variation (CNV).***

nSolver Basics

The nSolver System enables users to analyze and manage data sets in a variety of ways. The seven steps of a basic analysis (Import, QC, Normalization, Ratio Building, Export, Analysis/Visualizations, and Save) are assisted by four wizards. Perform data management using two Main View screens (Raw Data and Experiment). The buttons contained within the main view screens and taskbars provide a wide variety of options. There is also a home screen that highlights the first steps of analysis to assist new users. Complete descriptions of the various wizards and screens are presented in later chapters.

Quick Start

1. Unzip the folder(s) containing your RCC data.
2. Launch the nSolver Analysis Software application.
3. Click on Import RCC files. Navigate to the folder containing decompressed RCC files and select the files to be imported. If necessary, select the type of assay (GX, miRNA, CNV, miRGE, Protein, GeneProtein or Other).
4. Choose QC parameters as desired. Click Import.
5. Select the experiment tab.
6. Choose New Study. Fill in any information desired to track with your study and then click Save.
7. Choose New Experiment. Fill in any information desired to track with experiment and then click Next.
8. Select CodeSets to be used on the left side of the window. Filter out any undesired lanes and then click Next. All lanes displayed on the right side of the window will be used to create experiment.
9. Add any extra sample annotations if desired. Click Next.
10. Adjust normalization parameters. Click Next.
11. Define any desired sample expression ratios by selecting control or baseline sample(s), all pairwise comparisons, etc. If no ratio calculations are required, uncheck the box at the top left. Click Next.
12. Assign any names for ratios built in Step 10. Click Finish.
13. View data by clicking the arrow next to the experiment name on the left hand side. Select the data set to be viewed, select the samples to be viewed, and then click Table.
14. Create visualizations such as heat maps, box plots, violin plots, and scatter plots using the analysis wizard. Select the experiment to be analyzed and launch the wizard using the button in the menu bar. Filter out unwanted samples or probes and customize your analysis using options provided in the visualization windows.

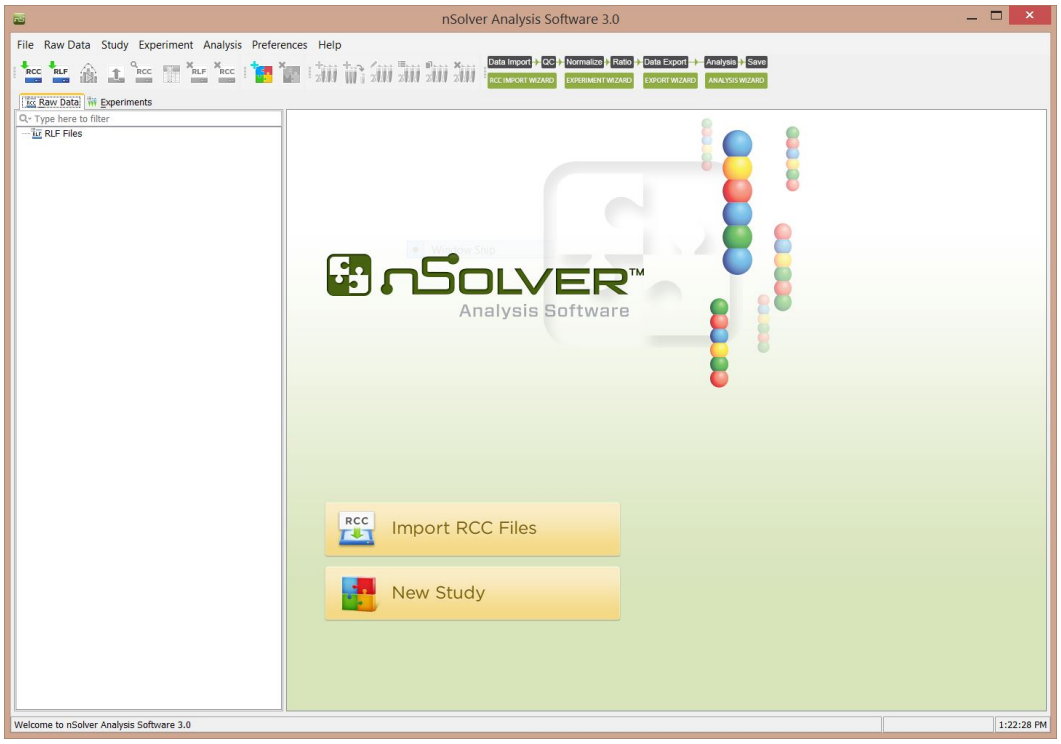
Workflow Wizards

The typical workflow of data analysis throughout the nSolver software is depicted in the figure below. Four wizards cover the six steps. The wizards are designed to streamline user actions and provide guidance for the appropriate analysis parameters based on assay type. Overall, the nSolver Analysis Software is very flexible and allows users to compare different outputs and make small modifications quickly and easily.



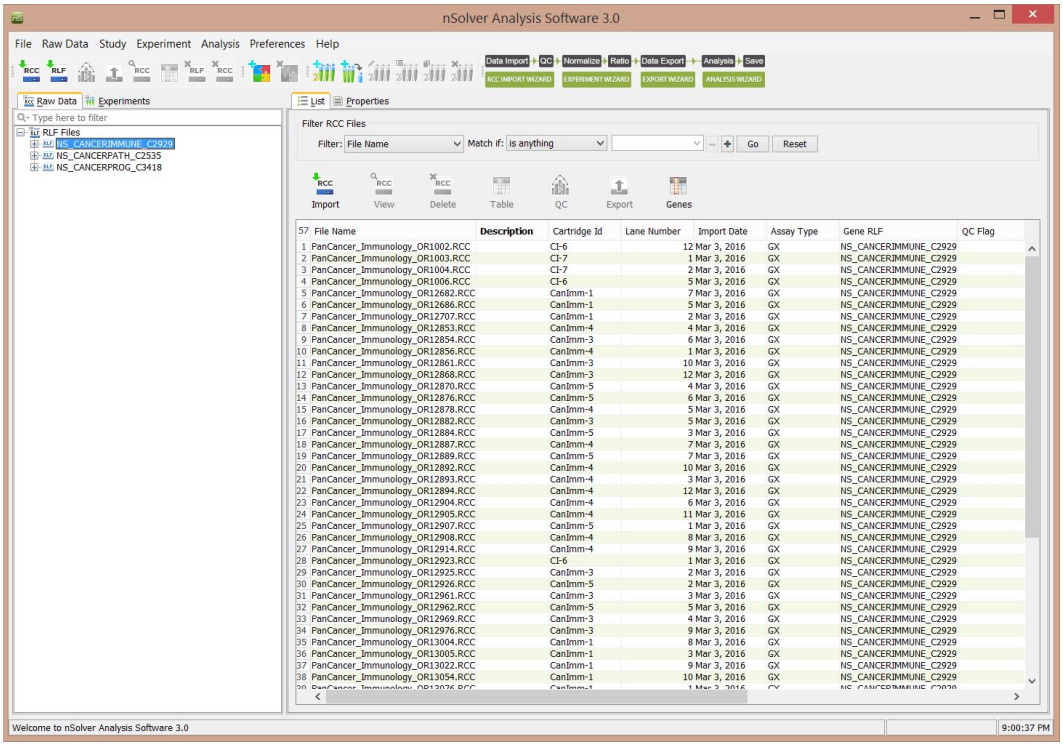
Home Screen

When first launching the nSolver Analysis Software, the **Home Screen** appears. From the home screen you can import RCC files or create a new study by selecting the **Main Action Buttons** at the middle of the screen. The **Active Icons** in the task bar at the top perform the same actions. You can search and select the RCC file data previously imported using the **Navigation Tree** on the left side of the screen.



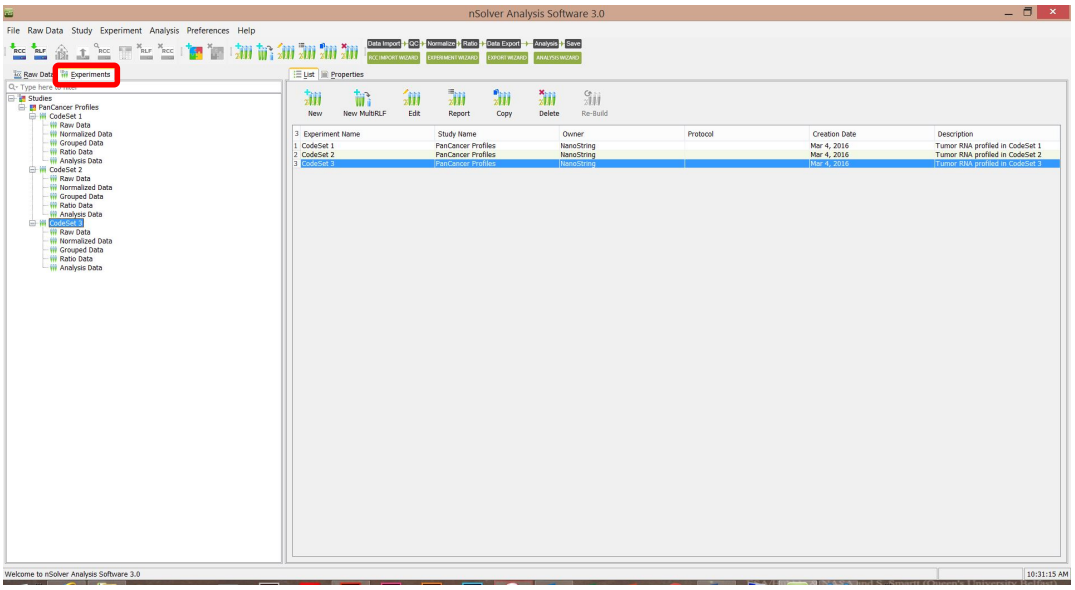
Raw Data Main View Screen

All imported **Raw Data** (RCC files) are organized by the RLF name. The **Raw Data Main View** screen contains a variety of searching and sorting options and displays the selected RCC data files. The screen contains icons that allow you to view raw data files individually or in tables, re-execute QC, and export. You can also use the raw data navigation tree and task bar icons in the Raw Data Main View.



Experiment Main View Screen

The **Experiment Main View** screen allows you to create and manage experiments. Use the icons within the view to create new experiments; edit, copy, delete, or re-build existing experiments; or create a report for an experiment. From the **Experiment** navigation tree, select the various tables output by experiment, file, and treatment group attributes.



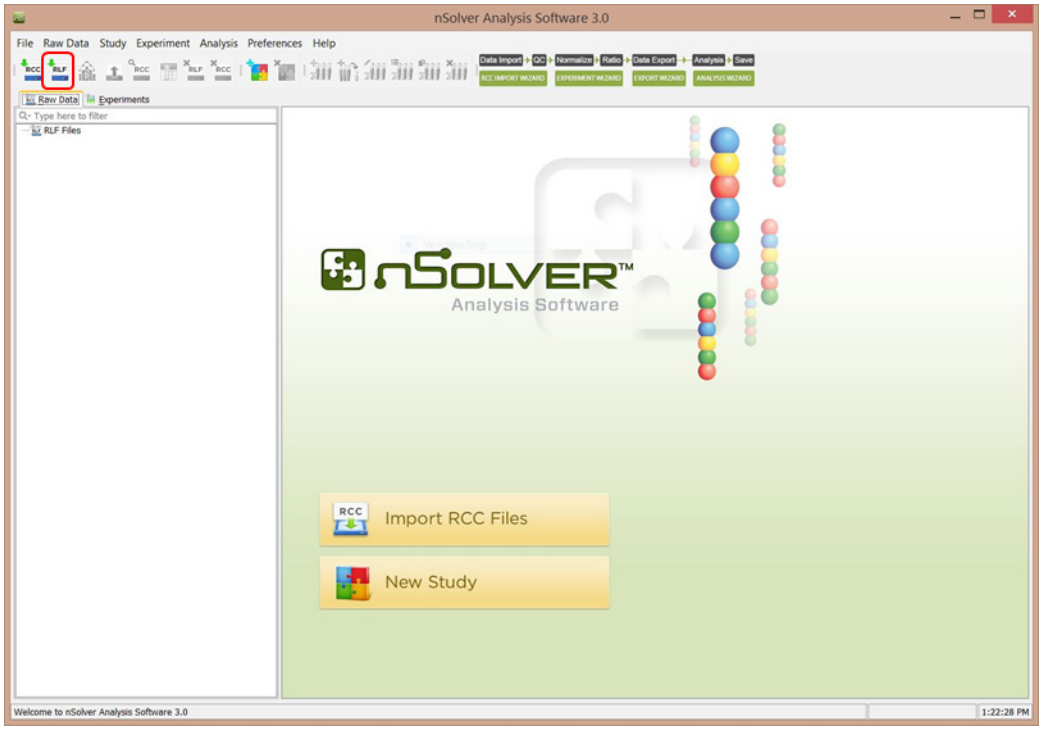
2 Importing Data

This chapter contains directions on how to import your nCounter RCC data and RLF files.

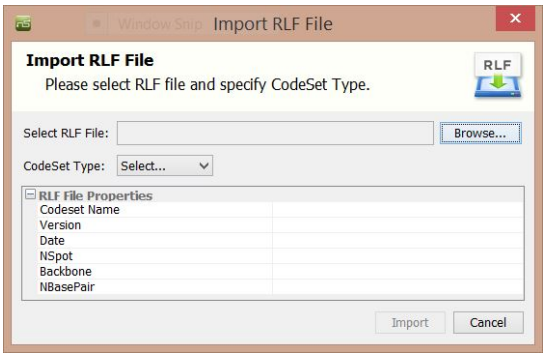
Importing RLF Files

Although importing in RLF files (Reporter Library Files) is optional, this is recommended and required for some features. RLF files contain additional information about your CodeSet, such as Probe ID and Target Sequence that can be stored, searched, and exported.

1. To import an RLF file, click the RLF import icon in the task bar at the top of the screen.

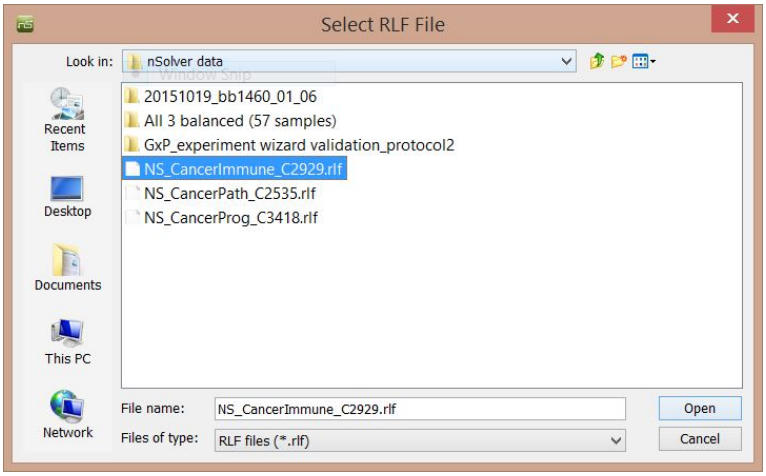


2. The RLF Import Screen will appear. Select the CodeSet type and click on the **Browse** button to import the RLF file.

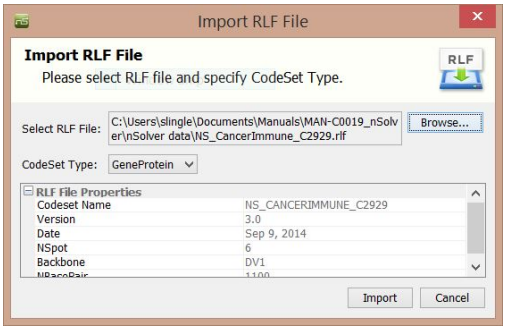


NOTE: Import Plex² Expression Assay RLFs as GX or CNV as appropriate. The nSolver application automatically detects that the files are derived from a Plex² Assay kit.

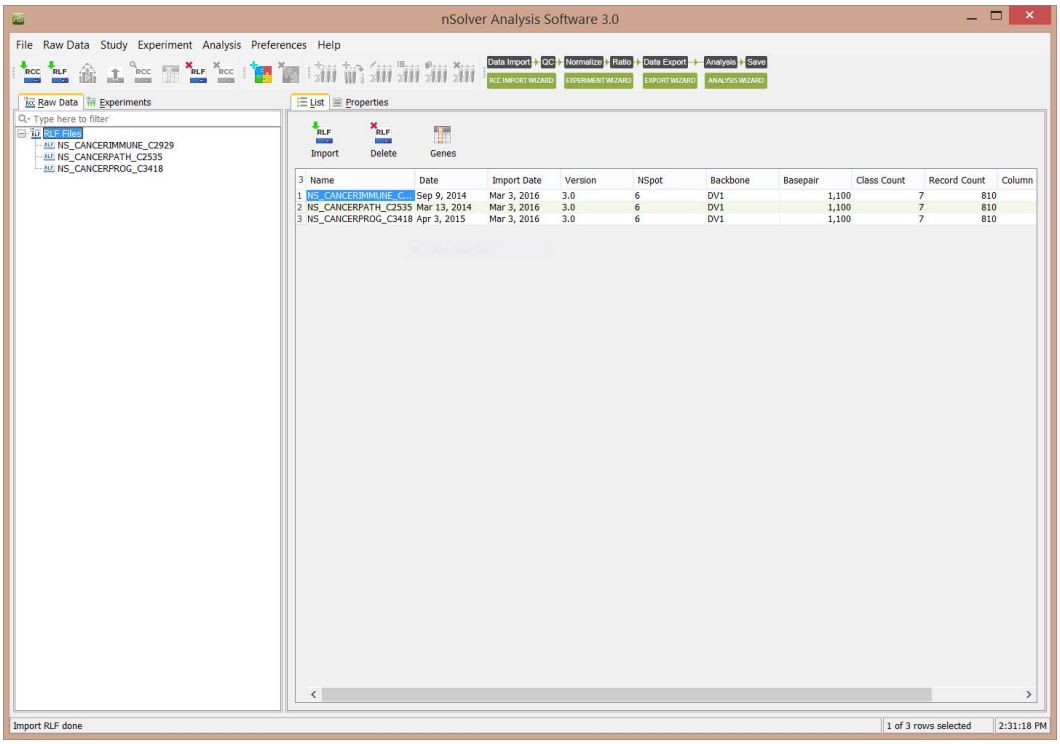
3. Navigate to the appropriate RLF file, select it, and click Open.



4. Click the Import Button to complete the import.



5. The RLF file will be imported into the nSolver application, and you can view it in the Raw Data tab from the navigation tree on the left.

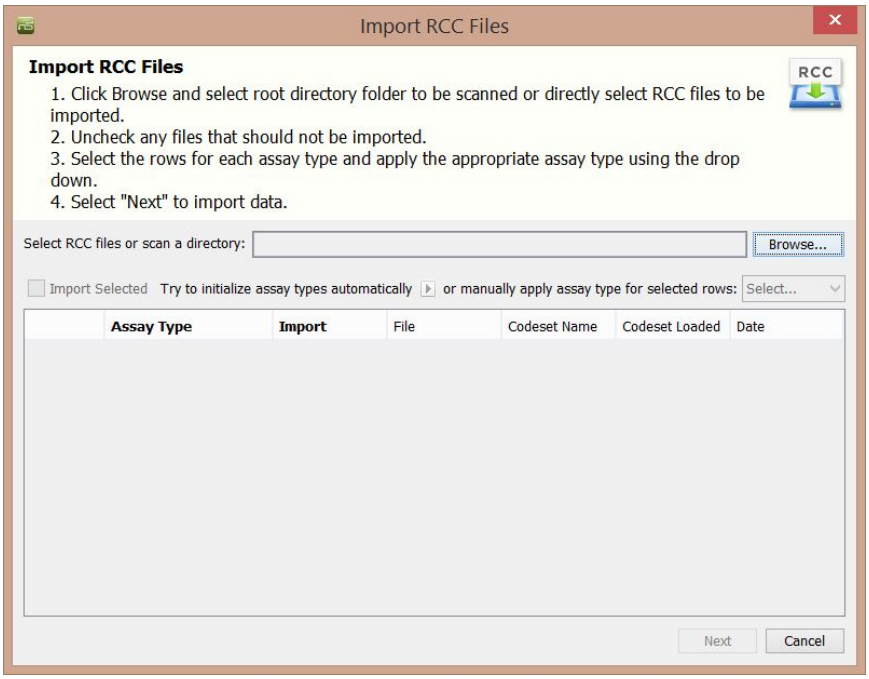


Importing RCC Files

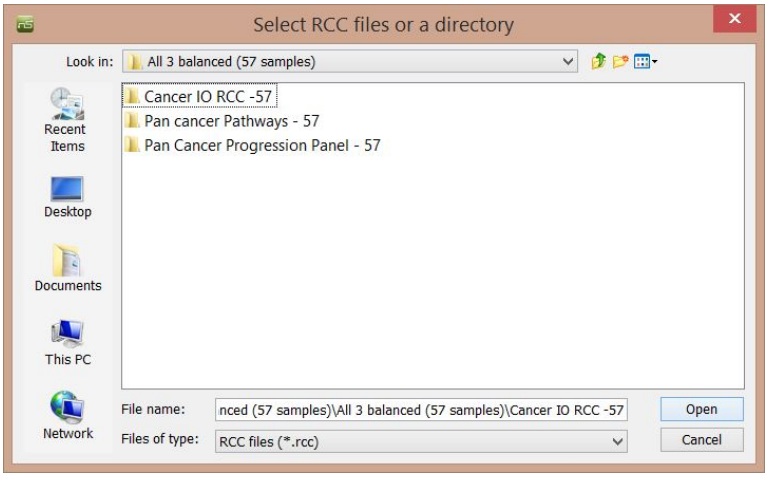
Raw data files are downloaded from the Digital Analyzer as compressed RCC files. Before importing them into nSolver, your data must be decompressed. nSolver does not automatically decompress folders or files and cannot recognize that RCC files are contained within a zipped file.



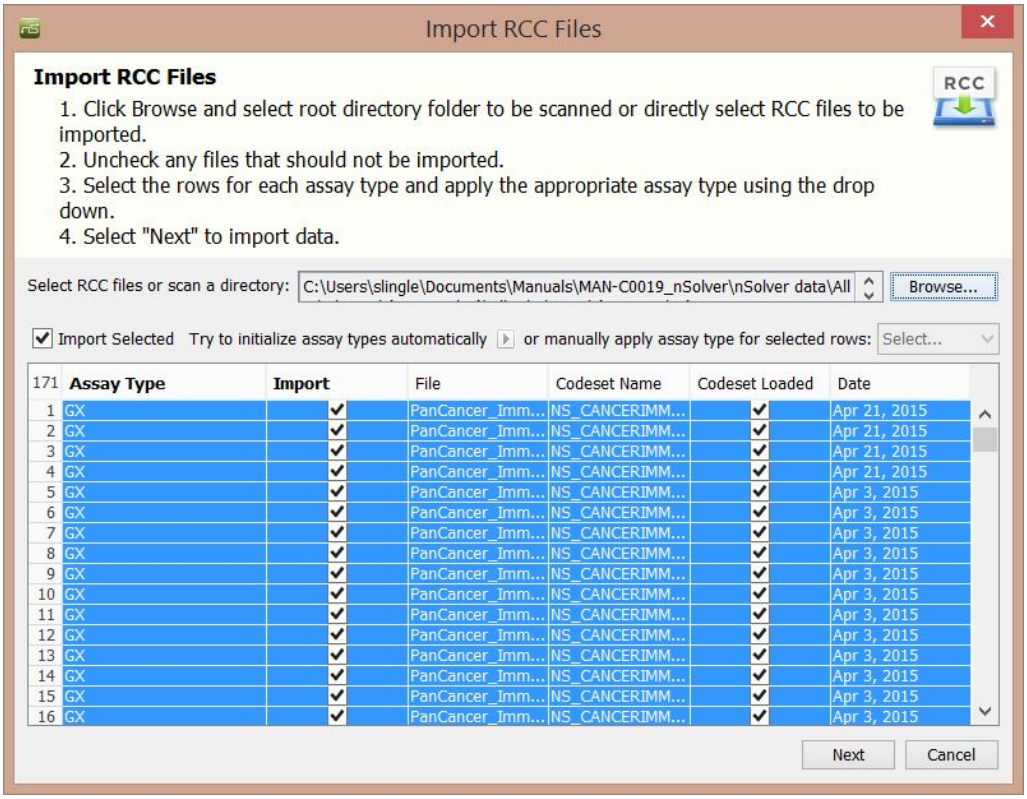
1. To import decompressed RCC data, click the **RCC** import icon in the task bar at the top of the screen or the **Import RCC Files** button on the home screen as shown below.
2. This will open the **RCC File Import Wizard**. Click the **Browse…** button to choose the RCC files to import.



3. Use multi-select (use the Shift, Control, or Command button while using the mouse) to import multiple RCC files. Alternatively, just click on the folder(s) that contain those files and the nSolver system will scan the sub-directories under the selected folder for all RCC files.



4. The RCC files will appear in the import wizard as shown below.

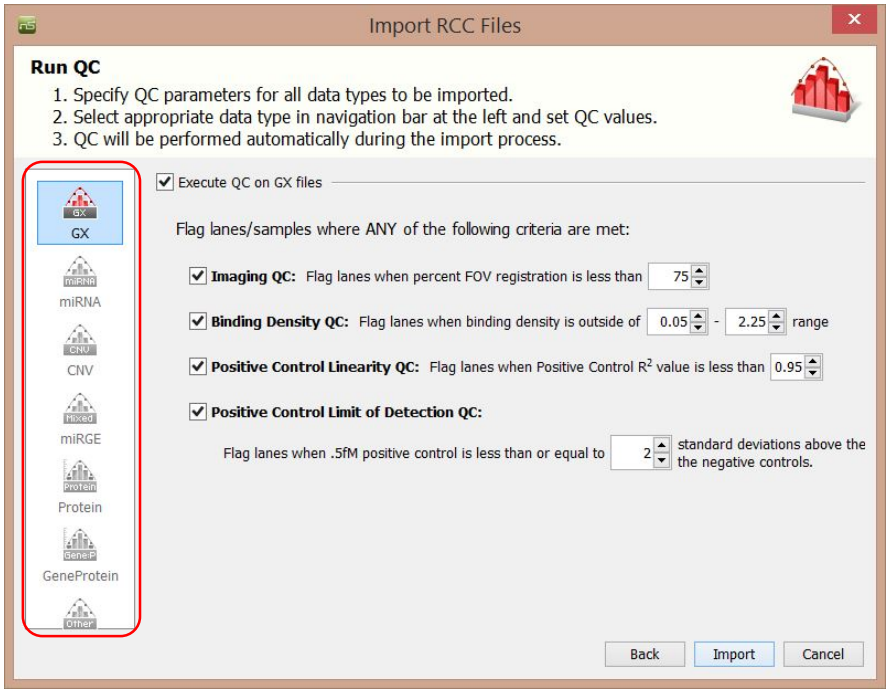


NOTE: The nSolver Quality Control parameters are tailored specifically for each assay type. In addition, for miRNA and miRGE assay types, the background subtraction denoted with '+++' in the RCC and RLF files will be automatically performed during import.

For all RCC files that are to be imported, **the assay type must be defined** (if no RLF or RCC files for the CodeSet have been imported). Within the table, there are two columns containing user selections:

- Required: Assay Type—For all RCC files marked to be imported, assign the assay type by first selecting the appropriate rows, and then selecting the assay type from the drop down on the right side. You can also assign each assay type individually from the drop down under the Assay Type column. If any RCC files have the assay type in the file name, then clicking on  automatically fills in those assay types. Available assay types are GX, miRNA, CNV, miRGE, Protein, GeneProtein, and Other. Plex2 Expression Assays should be imported as GX or CNV as appropriate. nSolver automatically recognizes that the files are derived from a Plex2 Expression assay.
- Optional: Import Selected—Deselect the check box(es) below the Import column for any RCC files that you do not want to import into the nSolver system at this time. If a file has been previously loaded, it is not necessary to uncheck the Import Checkbox as nSolver will recognize this and will not import the same file twice.

5. After completing your selections, click **Next** to proceed to the nSolver Quality Control Metrics screen.



NOTE: It is **critical** to choose the appropriate assay type when importing data into the nSolver application.

The Quality Control parameters and downstream normalization parameters are tailored specifically for each assay type. In addition, for miRNA and miRGE assay types, the background subtraction denoted with '+++' in the RCC and RLF files will be automatically performed during import.

3 Quality Control Metrics and Flags

This chapter contains directions on how to use nSolver to apply Quality Control parameters to your nCounter RCC data. It also provides guidelines and tips for troubleshooting data that have fallen below the Quality Control thresholds.

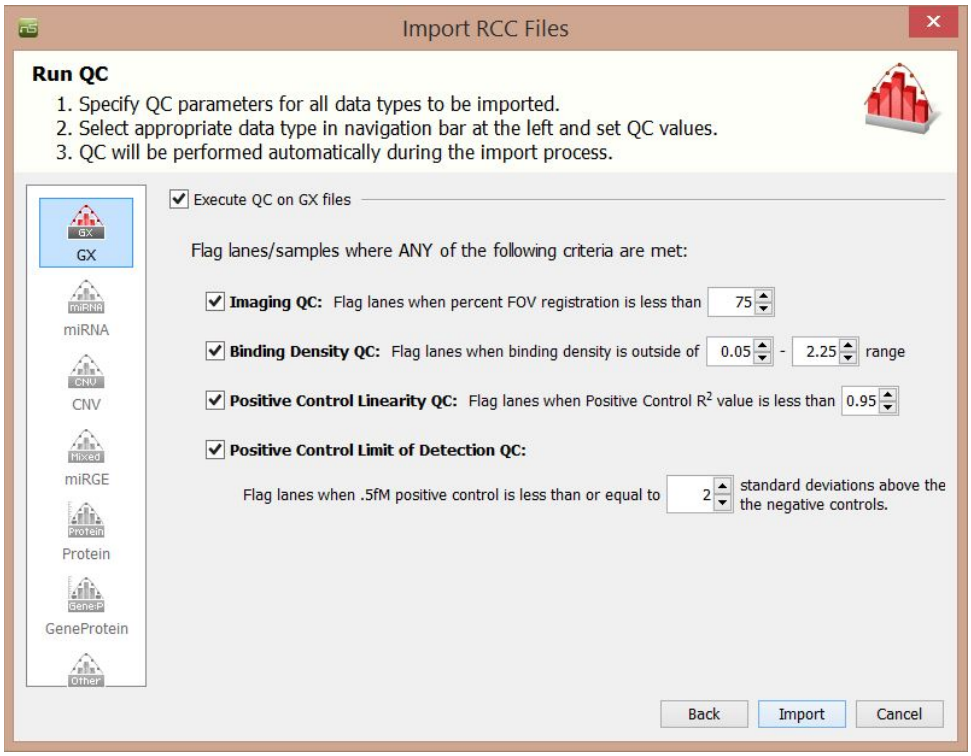
Every assay type has a suggested set of optional QC parameters. To change the value of a parameter, simply enter the desired value in the appropriate box, or deselect the check box to ignore entirely.



NOTE: It is important to understand that QC flags are cautionary and do not prevent the user from performing any downstream analysis. Data from flagged assays should be reviewed carefully before being included in further analysis.

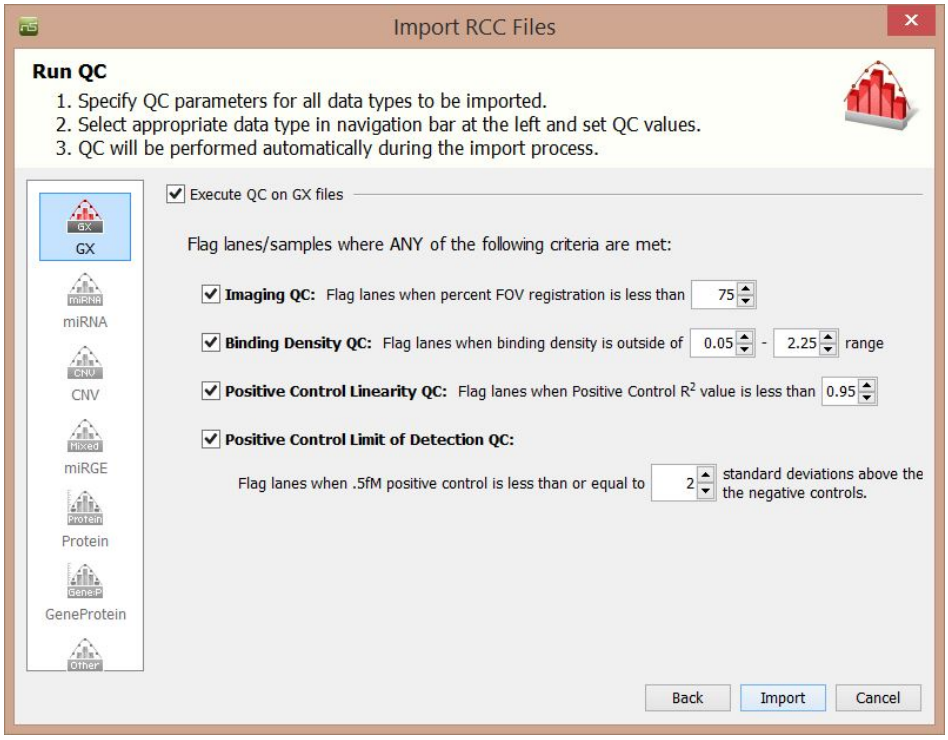
After selecting the **Next** button on the import screen, the QC screen will appear. For every assay type you are importing, you can choose to activate/deactivate different QC tests and edit the associated thresholds. To view and change QC options for multiple assay types within a single import action, choose the appropriate icon on the left side.

For all assay types, you can choose to omit a particular QC test by deselecting the appropriate check box, or you can choose to omit all of the QC tests by unselecting the **Execute QC** check box at the top. Performing QC test using the default setting is recommended for new users.



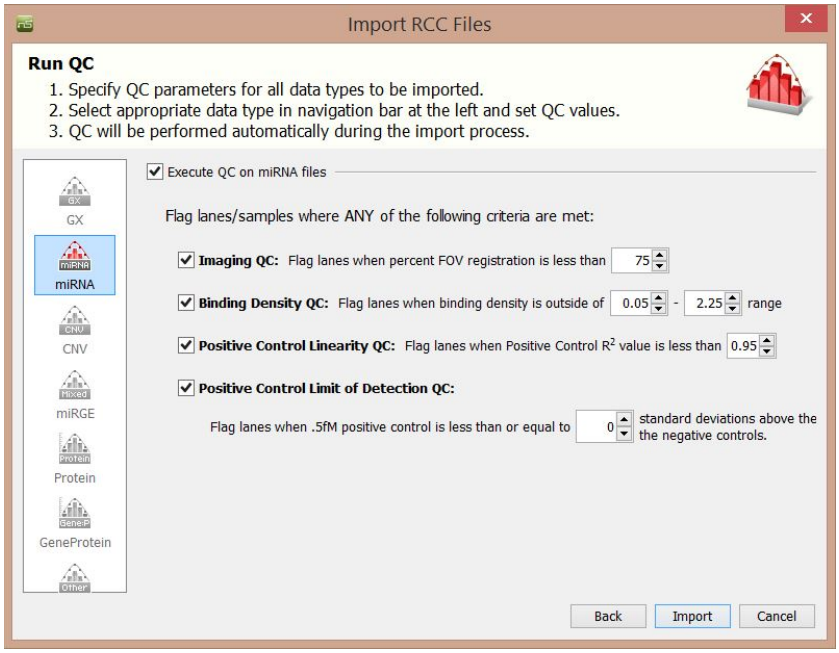
After choosing parameters for all data to be imported, click the **Import** button at the bottom. A dialogue box will appear indicating progress.

GX and GX Plex^2 Quality Control



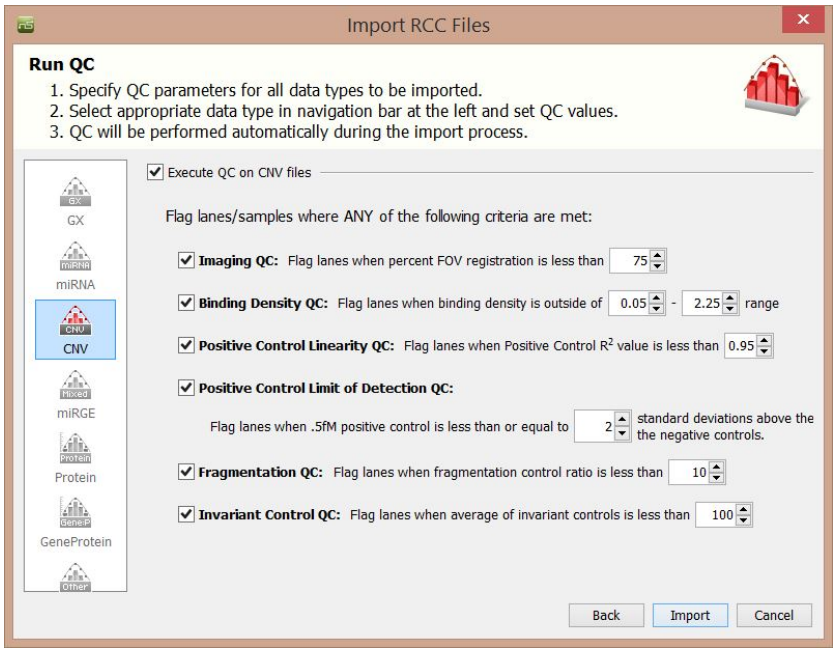
miRNA and miRGE Quality Control

NOTE: For miRNA and miRGE assays, the background subtraction denoted with '+++' in the RCC and RLF files is automatically performed during import.

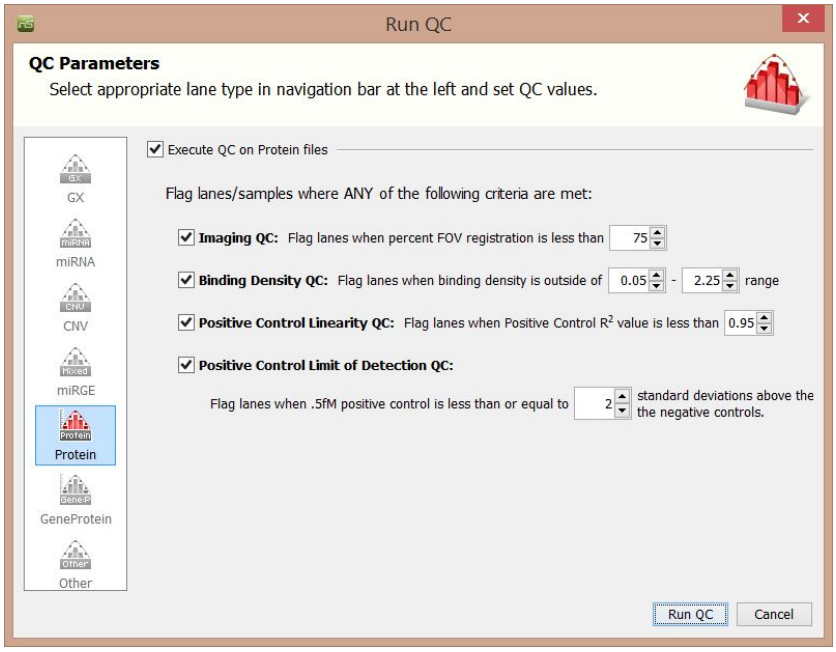


CNV Quality Control

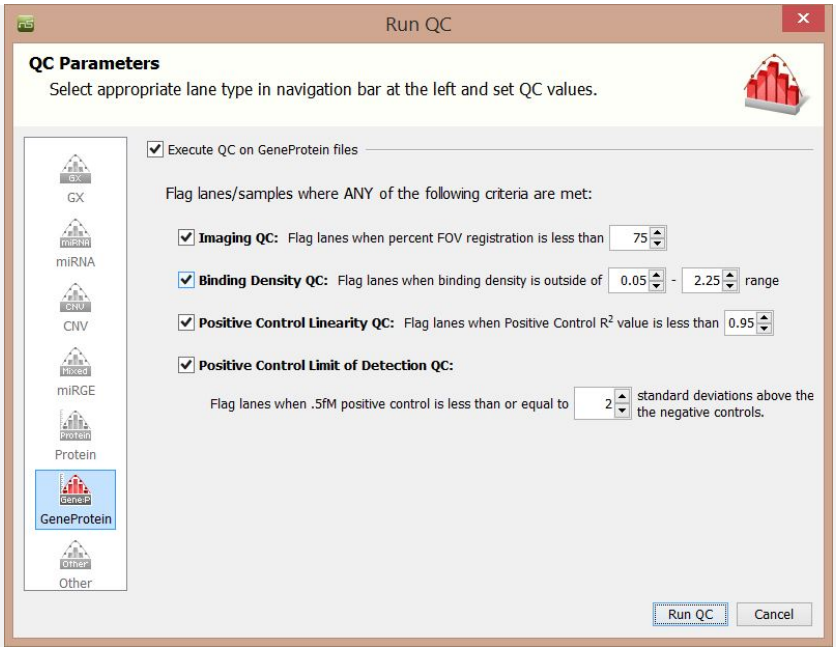
NOTE: The **Fragmentation QC** criterion is only relevant if Alu1 digestion was used as the method for fragmentation. If you did not use Alu1 fragmentation, please uncheck the "Fragmentation QC" boxes.



Protein Quality Control

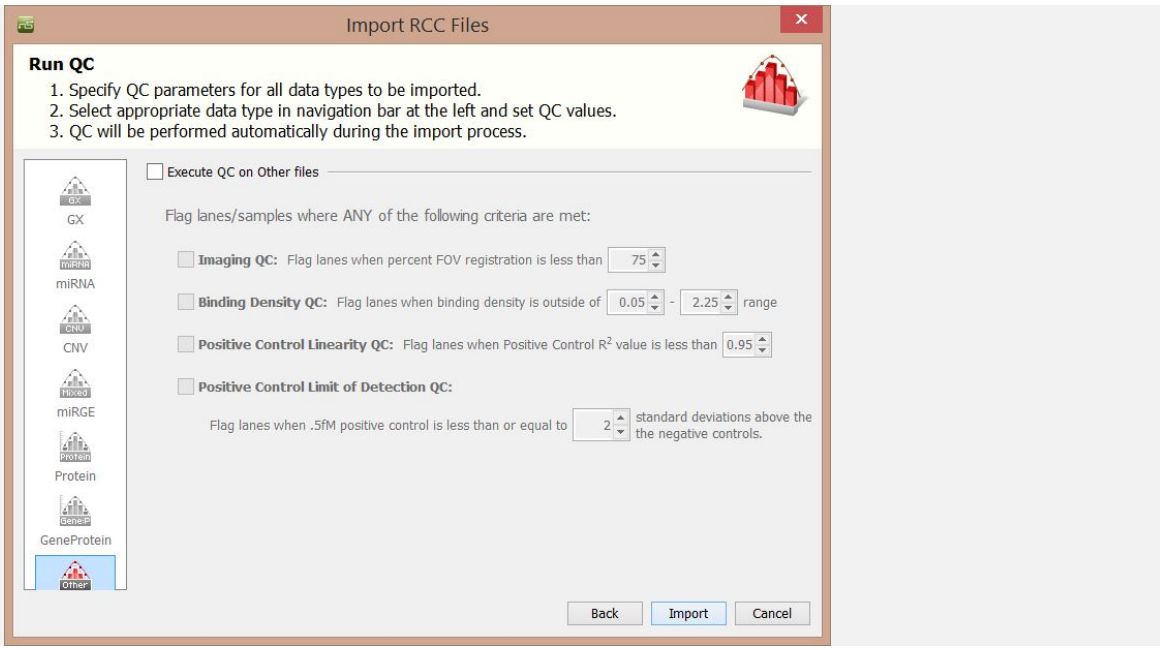


GeneProtein Quality Control



'Other' Quality Control

NOTE: For the assay type 'Other,' QC is turned off by default. To run QC for the other type, select the **Execute QC on Other files** check box and then adjust parameters as desired.



QC Parameters and Guidelines

There are four quality control parameters common to all nCounter assays and two that are specific to the CNV assay. The four common parameters use imaging thresholds and control probes for each lane. It is important to note that you can adjust all of the QC thresholds. Although thresholds represent the typical levels of performance, altering these parameters may be appropriate to prevent flagging of useful data due to the custom nature of many of the nCounter assays.

Common QC Metrics—Imaging

- **Imaging QC:** Imaging QC refers to the percentage of fields of view (FOVs) successfully counted by a Digital Analyzer scan of a lane. When a substantial percentage of FOVs are not successfully counted, there may be issues with the resulting data. Furthermore, consistently reduced percentages can be indicative of an issue associated with the instrumentation. Please refer to the ***nCounter® Expression Data Analysis Guide*** for additional information about counted FOVs and troubleshooting.
- **Binding Density QC:** Due to the nature of the nCounter technology, analysis of some samples may produce too many or too few probes to be accurately counted by the Digital Analyzer. When too many probes are present, the Digital Analyzer is not able to distinguish each and every probe present in the lane. When too few fluorescent species are present, the Digital Analyzer may have difficulty focusing on the lane surface. Therefore, a measurement of mean binding density (spots per square micron) is provided with each lane scanned. The linear range of counting extends from 0.05 to 2.25 spots per square micron for assays run on an nCounter *MAX* or *FLEX* system. The range is 0.05 to 1.18 spots per square micron for assays run on the nCounter *SPRINT* system. While it is possible to obtain useful data outside of this range, we recommend that you check the control and housekeeping probe data for consistency with other lanes within the experiment. Please refer to the ***nCounter® Expression Data Analysis Guide*** for additional information about binding density and troubleshooting.

Common QC Metrics—Control Probes

- **Positive Control Linearity QC**: Each nCounter assay contains a variety of positive control probes targeting molecules added by NanoString during production of the CodeSet. Positive control linearity QC performs a correlation analysis in $\log_2$ space between the concentrations of added targets and the resulting counts. Low correlation values may indicate an issue with the hybridization reaction and/or instrument performance. Please refer to the ***nCounter® Expression Data Analysis Guide*** for additional information about troubleshooting hybridization and/or instrument performance.
- **Positive Control Limit of Detection QC**: This QC metric uses the positive and negative control probes to confirm the limit of detection for each nCounter assay. It is expected that the 0.5 fM positive control probe (Pos_E) produces raw counts significantly higher than the mean of the negative control probes. The number of standard deviations used to estimate significance can be modified by the user. Due to the differences in control architecture for the miRNA assays, defaults values for the detection QC are the mean of negative control probes for miRNA assays and 2 standard deviations above the mean of negative control probes for all other assays. Please refer to the ***nCounter® Expression Data Analysis Guide*** for additional information about miRNA control architecture and troubleshooting hybridization and/or instrument performance.

NOTE: If the calculated standard deviation of all of the negatives is less than 4, then the standard deviation is set to 4 counts. Additionally, lanes are flagged when the 0.5 fM positive control is less than **or equal** to the specified standard deviations above the mean of the negative controls.

CNV-Specific Metrics

NOTE: Because the nCounter Custom CNV assays target DNA rather than RNA and provide copy number estimates, additional controls and QC parameters are included to ensure high quality results.

- **Fragmentation QC:** To accurately measure relative amounts of DNA, it is important to fragment the sample to create targets of similar sizes. Four fragmentation control probes are included with each Custom CNV assay (two positive and two negative). The probes target oligonucleotides included with the CNV Sample Preparation reagents. Two of the targeted oligonucleotides contain the restriction site used for fragmentation and two do not. After a successful sample preparation procedure, the positive fragmentation probes should exhibit reduced counts with respect to the negative fragmentation probes. This QC metric compares the ratio of the negative to the positive fragmentation probes.

Please refer to the ***nCounter® Data Analysis Guidelines for Copy Number Variation (CNV)*** for additional information about fragmentation control architecture and troubleshooting sample preparation reactions.

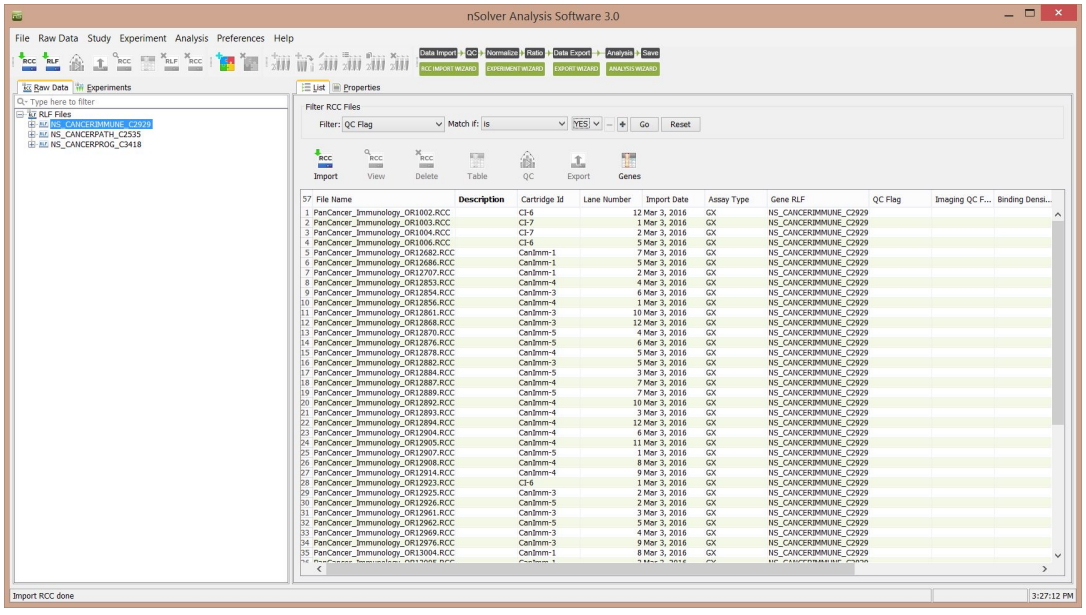
NOTE: If using a method other than NanoString's DNA Sample Preparation Kit (Alu1 digestion) to fragment the sample, it is important to disable Fragmentation QC or to ignore the results as the Fragmentation QC probes require Alu1 digestion to function properly.

- **Invariant Control QC:** Each custom CNV assay contains a set of invariant control probes that target regions of the genome that have been measured as two copies in a majority of individuals. Because the invariant probes are critical for normalization of differences in input amount between samples, the count levels for these probes should be well within the linear range of the system. A mean invariant count level of over one hundred counts typically ensures normalization will be performed within the linear range of the system. Mean invariant count levels of less than 100 counts may result in less accurate normalization and/or copy number estimates. Please refer to the ***nCounter® Data Analysis Guidelines for Copy Number Variation (CNV)*** for additional information about invariant control probes and troubleshooting.

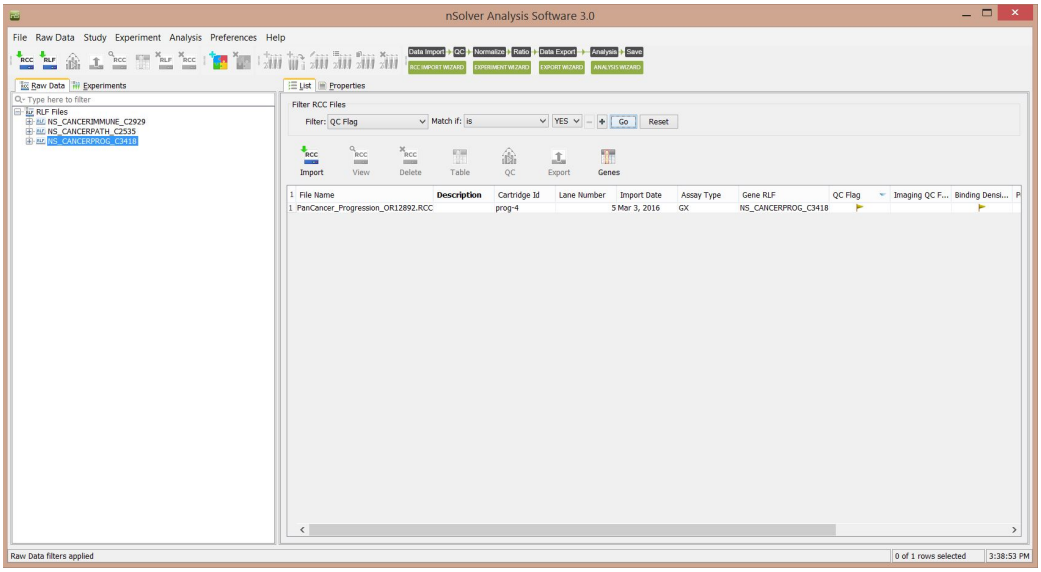
Reviewing QC Results

NOTE: As previously mentioned, it is important to note that the QC flags presented are simply an indication that closer examination of the lane is warranted prior to proceeding with analysis. Often times, flags are raised due to intended attributes of a particular CodeSet or sample. QC flags do NOT necessarily mean that data are unreliable. After importing and running QC on your data, the next logical step is to examine the data for QC results.

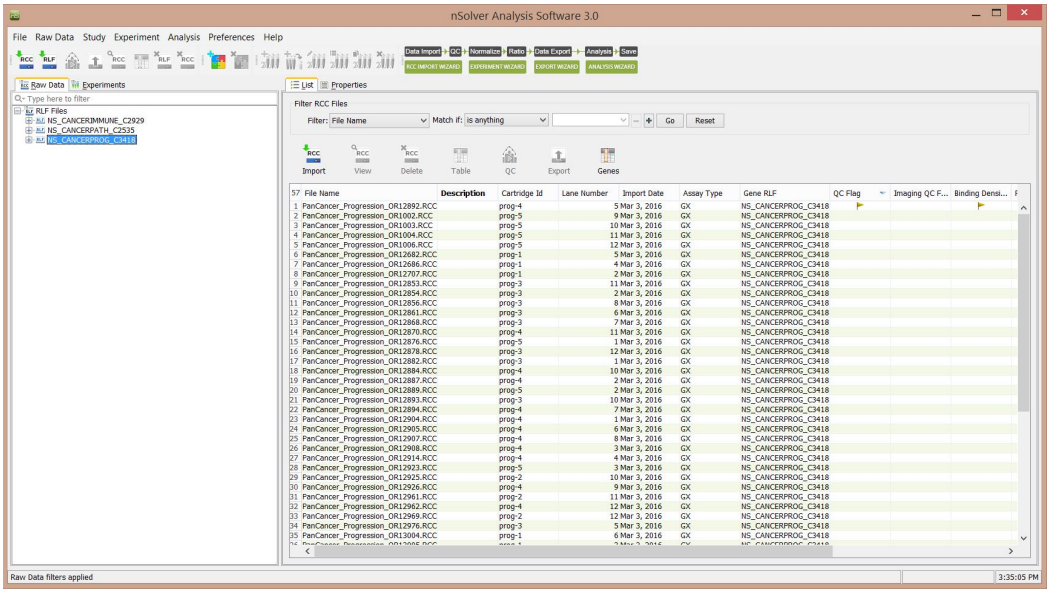
After importing RCC files, you can find them in the Raw Data Main View in the navigation tree on the left side. The RCC files are organized in the navigation tree by CodeSet RLF. There are several ways that you can examine the QC results of the data you have just imported. For example, after selecting the CodeSet (RLF) name from the Raw Data viewer in the navigation tree, search for only flagged data by setting the filter to find files where the QC Flag is **YES** as shown below:



Search results will look similar to the following:



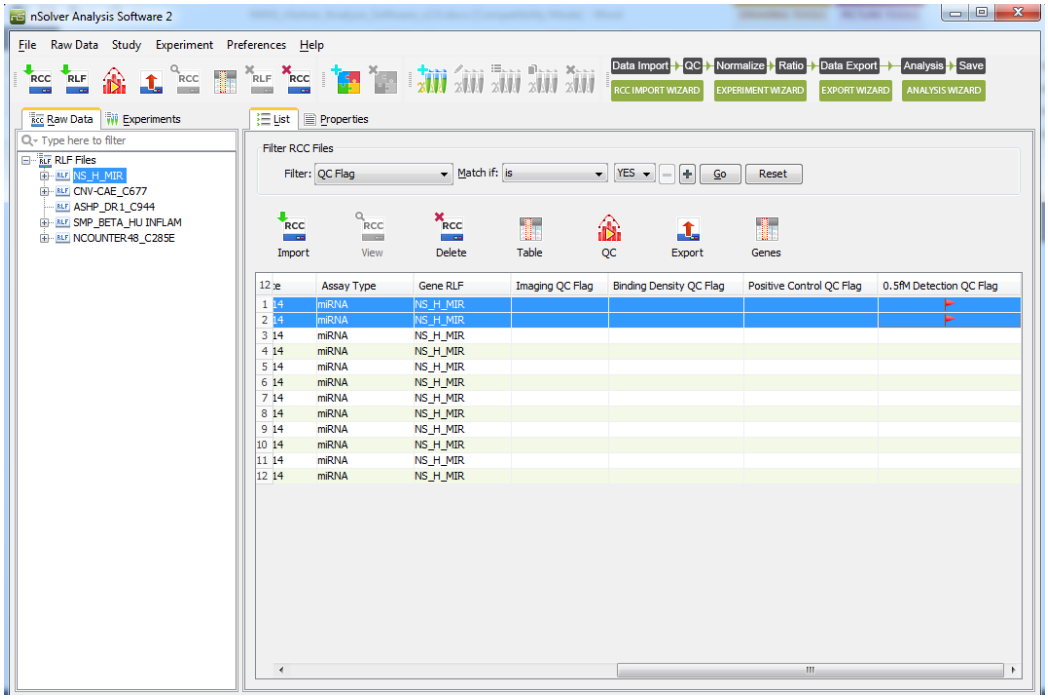
Alternatively, you can sort by clicking on the column header labeled **QC Flag** to simply view all the QC flagged files on top as shown below:



There are a number of QC Flags as described previously. If any of the individual QC tests fail, that particular item displays a flag. Also, if any QC test fails, the Overall QC Flag displays. When viewing data tables, column headers in **red** indicate a QC Flag is associated with that lane.

To further examine which QC tests failed and why, there are several columns that are useful.

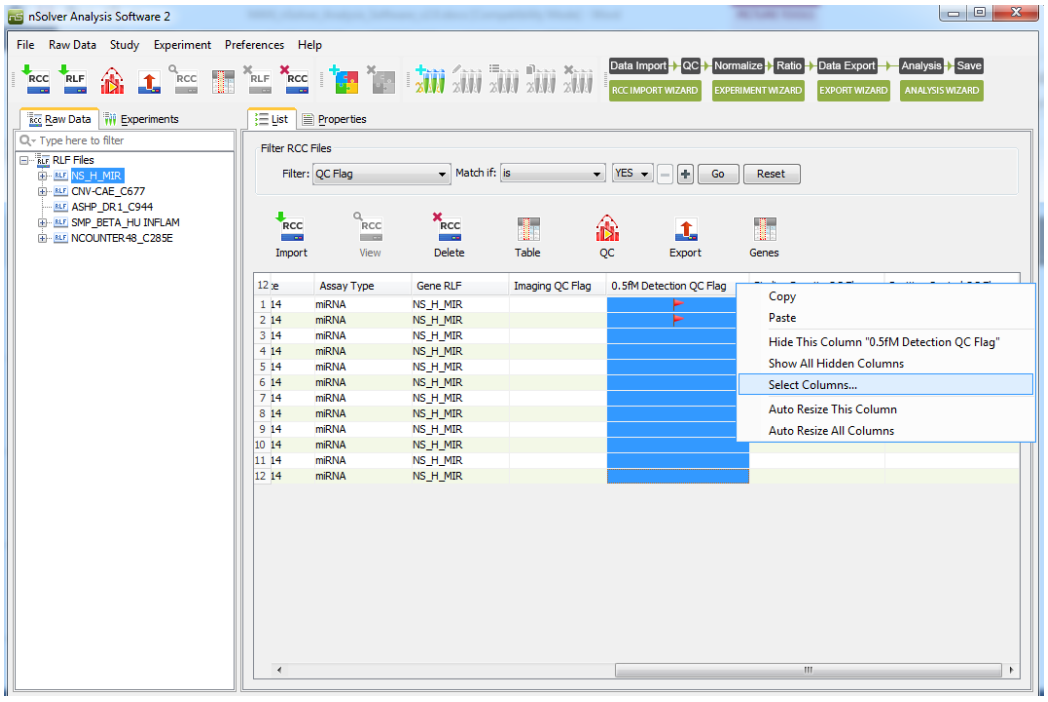
Scroll to the right side of the main raw data viewer and notice the different flags for your data that did not pass QC.



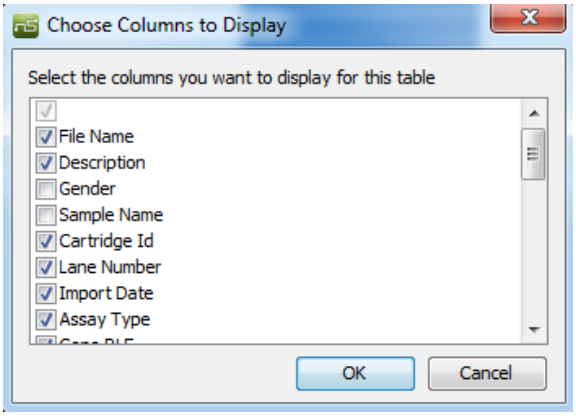
In the example above, note that for two samples/lanes, the 0.5 fM Detection QC Flag is displayed. For the other samples/lanes, no QC Flag is displayed.

There are several columns that are hidden by default in the application to keep the tables manageable. However, if QC flags are displayed, we advise you to examine the QC values that were calculated.

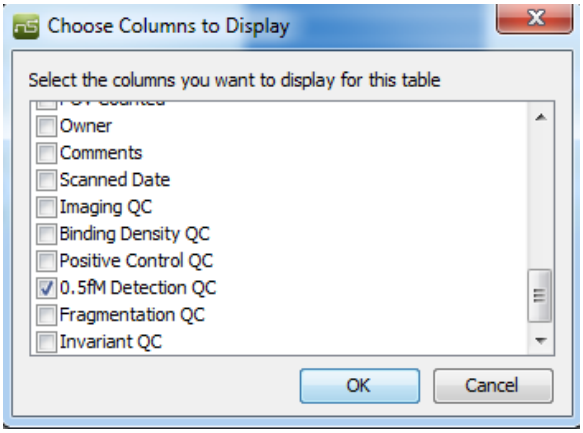
To show the hidden columns, right-click on the column header and choose **Select Columns…** as displayed below:



The Choose Columns display menu will appear. Select the checkbox for any of the fields that are not checked to add the column(s) to the table displayed.



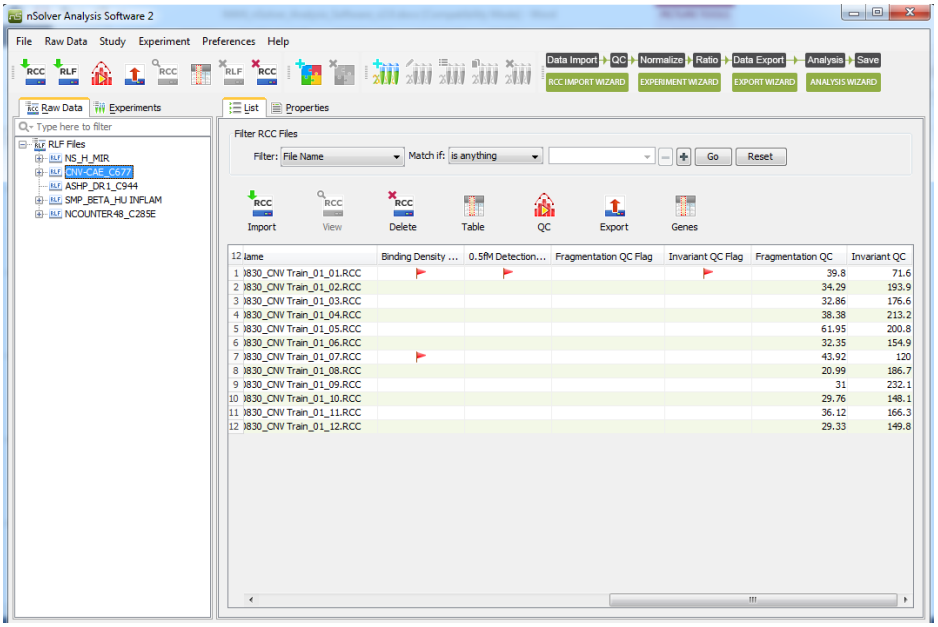
Scroll down to view the different QC values that were calculated for each sample/lane. Check the boxes to see the QC values that were calculated for the flagged samples/lanes. Click **OK**.



The new columns are added to the right side of the table viewer. The values contained in each of the columns are as follows:

- **Imaging QC**—displays the percentage of FOVs counted for the lane.
- **Binding Density QC**—displays the mean binding density for the lane in spots per square micron.
- **Positive Control QC**—displays the linear correlation coefficient for the positive control probes within the lane.
- **0.5 fM Detection QC**—displays the mean of the negative controls for the lane plus the user defined number of standard deviations.
- **Fragmentation QC (CNV assays only)**—displays the ratio of counts between the average of the two negative fragmentation control probes versus the average of the two positive fragmentation control probes.
- **Invariant QC (CNV assays only)**—displays the average counts for the invariant probes

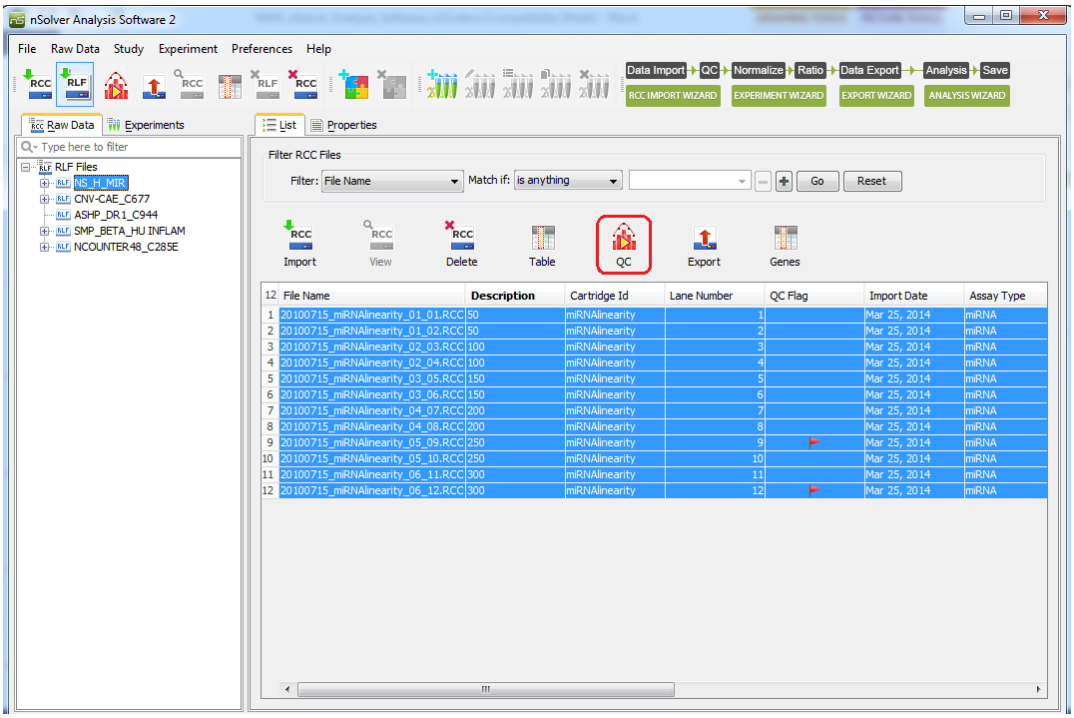
NOTE: 'Fragmentation QC' is only relevant if Alu1 digestion was used as the method for fragmentation. If you did not use Alu1 fragmentation, please unselect the 'Fragmentation QC' checkbox.



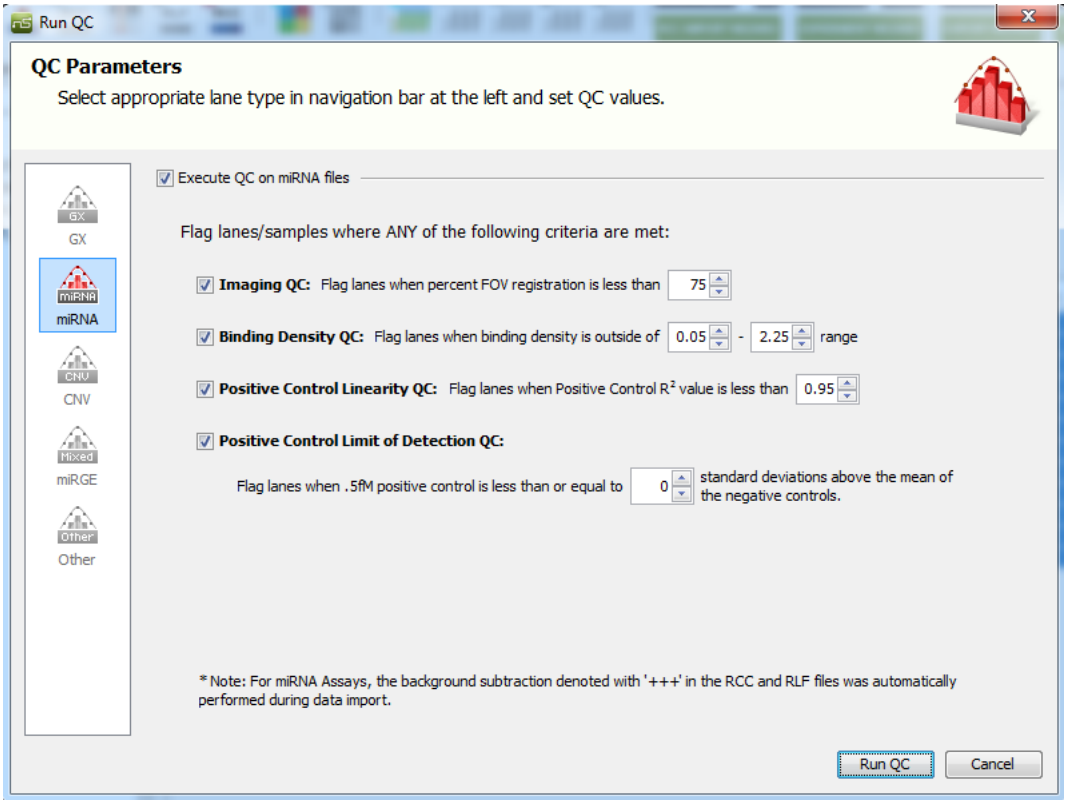
Re-Running Quality Control Checks

Because many nCounter CodeSets are custom designed, in some cases the parameters for QC should be adjusted to fit the needs of specific data sets or experimental designs. If you ran QC using the default parameters and observed flags that seem to be set unnecessarily, you can either ignore the flags or re-run QC.

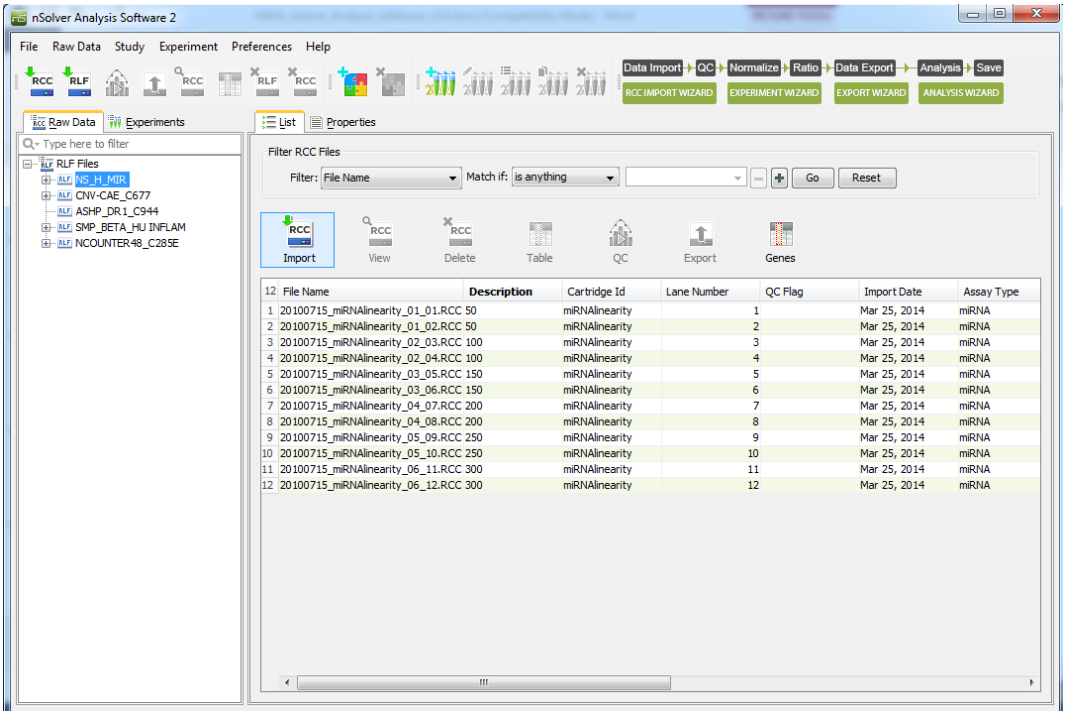
Re-run QC anytime you are viewing raw data files from the main viewer by selecting the appropriate lanes/samples and click the **QC Button**.



This opens the QC Wizard for the assay type(s) selected. Change the thresholds as appropriate.



Choose **Run QC**, and the QC calculations will be re-calculated and flags adjusted commensurately. Notice that there are no flags in the screen shown below.


nSolver Analysis Software 2
File Raw Data Study Experiment Preferences Help
Data Import QC Normalize Ratio Data Export Analysis Save
RCC IMPORT WIZARD EXPERIMENT WIZARD EXPORT WIZARD ANALYSIS WIZARD
Raw Data Experiments
List Properties
Type here to filter
RLF Files
NS_H_MIR
CNV-CAE_C677
ASHP_DR1_C944
SMP_BETA_HU INFLAM
NCOUNTER48_C285E
Filter RCC Files
Filter: File Name Match if: is anything Go Reset
Import View Delete Table QC Export Genes
12 File Name Description Cartridge Id Lane Number QC Flag Import Date Assay Type
1 20100715_miRNAlinearity_01_01.RCC 50 miRNAlinearity 1 Mar 25, 2014 miRNA
2 20100715_miRNAlinearity_01_02.RCC 50 miRNAlinearity 2 Mar 25, 2014 miRNA
3 20100715_miRNAlinearity_02_03.RCC 100 miRNAlinearity 3 Mar 25, 2014 miRNA
4 20100715_miRNAlinearity_02_04.RCC 100 miRNAlinearity 4 Mar 25, 2014 miRNA
5 20100715_miRNAlinearity_03_05.RCC 150 miRNAlinearity 5 Mar 25, 2014 miRNA
6 20100715_miRNAlinearity_03_06.RCC 150 miRNAlinearity 6 Mar 25, 2014 miRNA
7 20100715_miRNAlinearity_04_07.RCC 200 miRNAlinearity 7 Mar 25, 2014 miRNA
8 20100715_miRNAlinearity_04_08.RCC 200 miRNAlinearity 8 Mar 25, 2014 miRNA
9 20100715_miRNAlinearity_05_09.RCC 250 miRNAlinearity 9 Mar 25, 2014 miRNA
10 20100715_miRNAlinearity_05_10.RCC 250 miRNAlinearity 10 Mar 25, 2014 miRNA
11 20100715_miRNAlinearity_06_11.RCC 300 miRNAlinearity 11 Mar 25, 2014 miRNA
12 20100715_miRNAlinearity_06_12.RCC 300 miRNAlinearity 12 Mar 25, 2014 miRNA

4 Adding Lane/Sample Descriptions and Gender to Raw Data

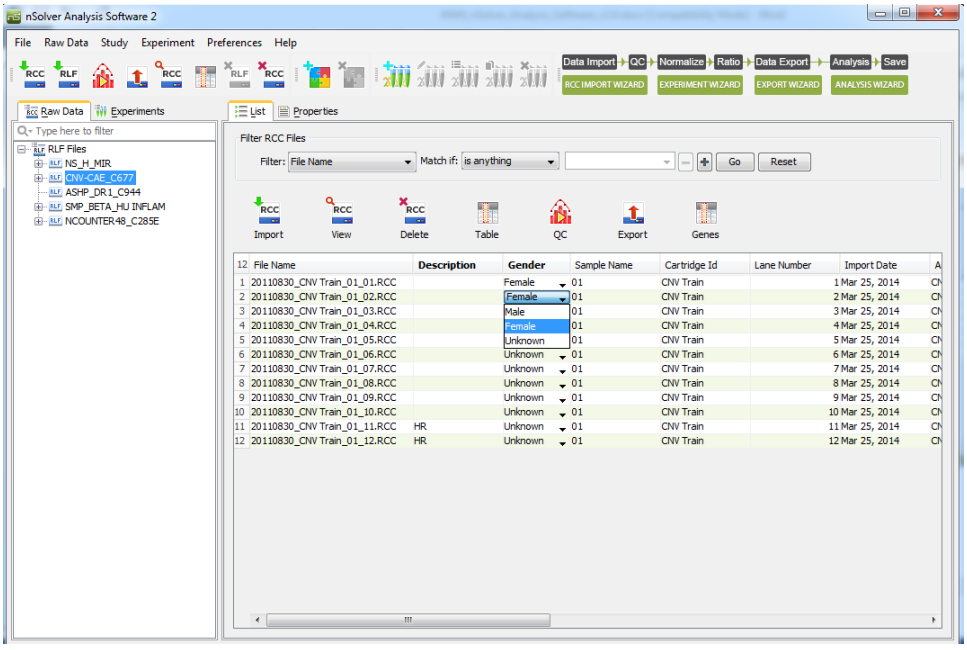
This chapter contains information on how to add lane/sample descriptions and gender to your raw data.

Gender

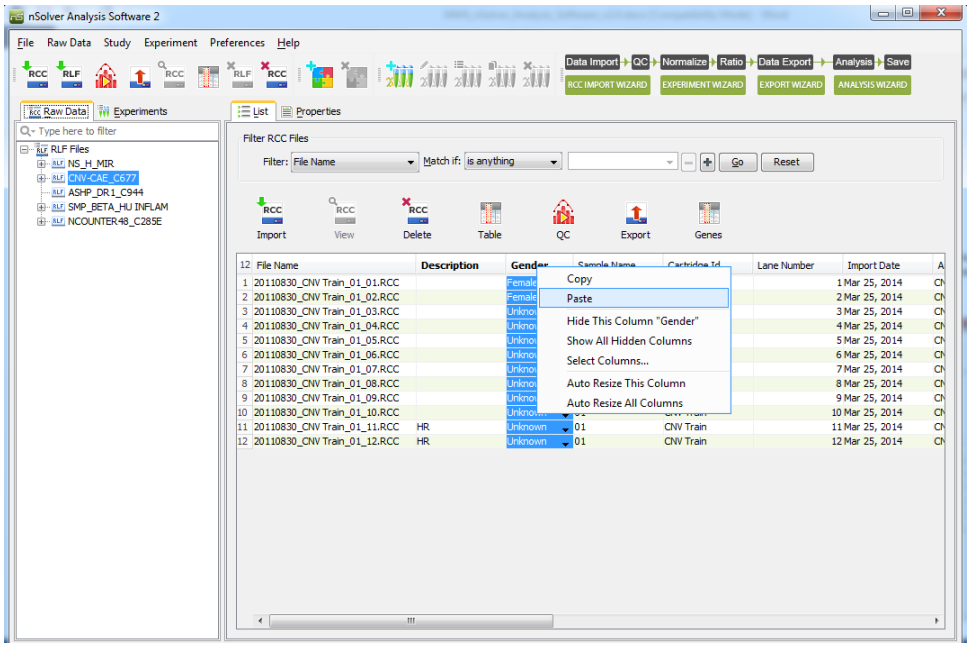
You have the option to assign gender to your lanes/samples for any assay type, but it is required for proper copy number estimate calculations for your CNV data as explained below.

Copy number estimates are calculated for autosomal probes by determining the ratio of experimental to control sample and then multiplying by 2. For the sex chromosome probes, the system considers the sex of the reference sample. Copy number estimates generated from male reference samples multiply the X and Y chromosome probe ratios by 1. For female reference samples, the copy number estimates multiply for X chromosome probes ratios by 2 and do not display copy number estimates for the Y chromosome. For unknown reference samples, copy number estimates are not displayed for the sex chromosome probes.

Gender is displayed by default on the raw data views for CNV data. Gender information is not utilized for assay types other than CNV. For those other assay types, the gender column is hidden by default. To learn how to display hidden columns in tables, see *Chapter 13 – Data Management*. To enter gender data for your samples/lanes, select either **Male** or **Female** from for the appropriate lane/sample the Raw Data view.



If you have the gender saved in a spreadsheet, you can also use the copy and paste function to enter the appropriate values in bulk. Copy the values from a text or Excel® file to your clipboard (do not copy the column heading, if any). Within nSolver, right click on the desired column header and paste in the values. For more information on the copy and paste features, see *Chapter 13 – Data Management*.

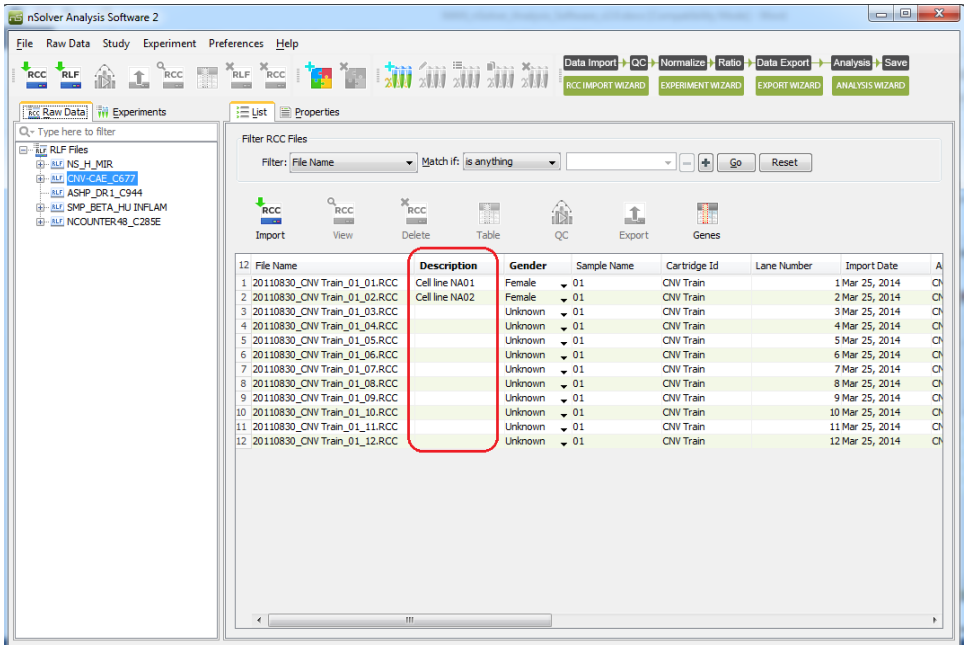


Descriptions

At times you may want to assign annotation or description to your lanes/samples that are informative for sample/lane tracking but are not appropriate annotations to be used for analysis.

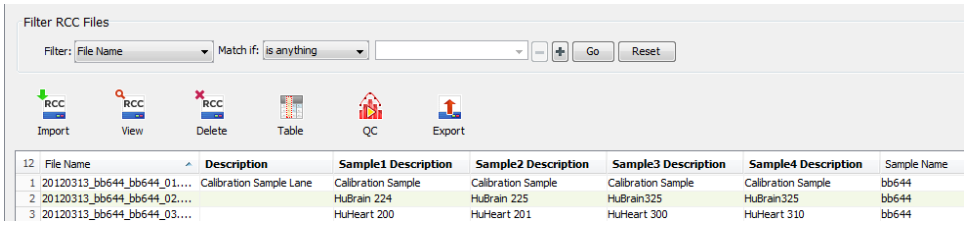
To enter annotations or "descriptions" to your lanes/samples that will not be used for downstream analysis, type the desired information for the appropriate lane/sample in the Raw Data view. Alternatively, use the copy and paste function to enter your values in bulk as described above and in *Chapter 13 – Data Management*.

NOTE: Description fields are used solely by the user to identify and filter samples. Annotations are used to create treatment groups in downstream analyses and are added as part of the Experiment Wizard.



For Plex^2 Assays, use the Sample Description columns to identify each of the samples in that lane. We recommend that you create a spreadsheet that you can use to associate sample names with each lane. You can paste the contents of the spreadsheet into the raw data main view columns for Sample Descriptions provided they are in the correct orientation (Lanes/RCC files as rows and sample names as columns in order from 1-4 based on the associated sub-CodeSet). The Sample ID column in the raw data main view will only indicate the text entered into the CDF for sample name.

After importing your Plex^2 Assay files, identify each sample using the Sample Description fields in the Raw Data main view. Sample1 refers to the sample hybridized against Sub-CodeSet1 (Endogenous1). Note that lane descriptions can also be added for each lane.



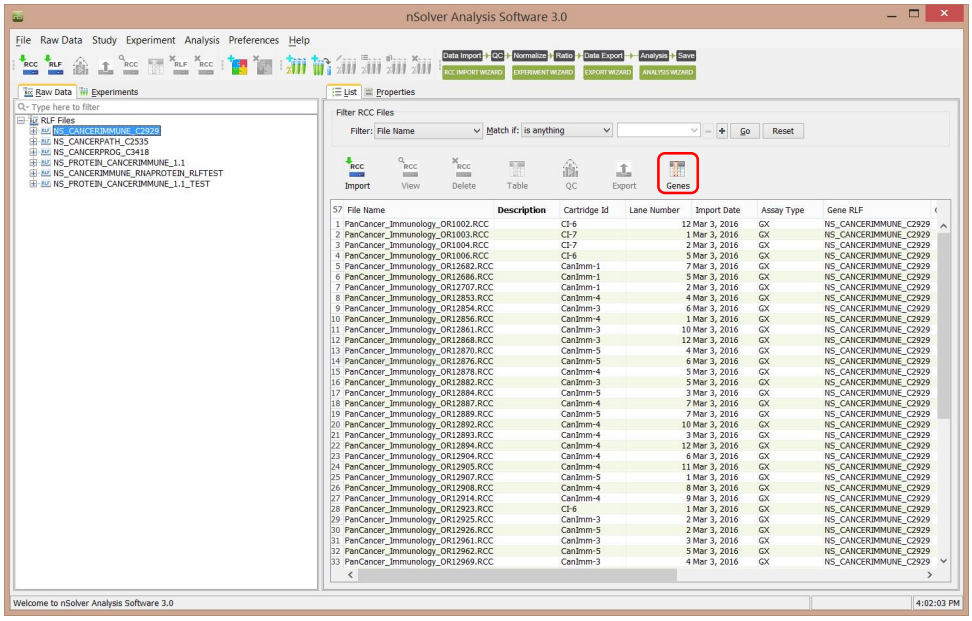
5 Adding Custom Annotations to Probes

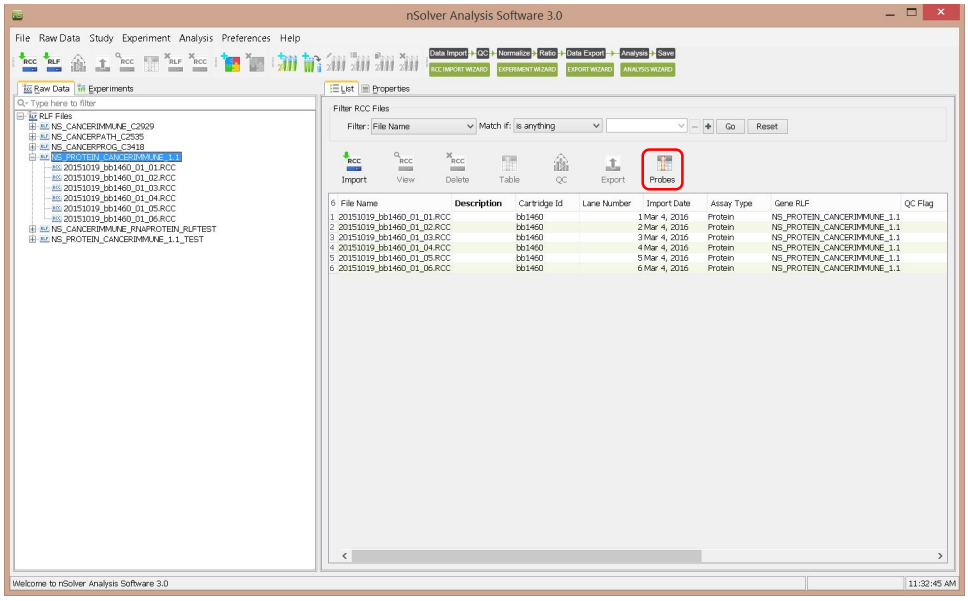
This chapter contains information on how to customize annotations for probes in your assay. nSolver screens and dialog boxes refer to the probes for genetic material as genes.

Probe Annotations

Assign custom annotations to the probes in your assay that are informative for exports, and particularly useful for filtering when creating visualizations.

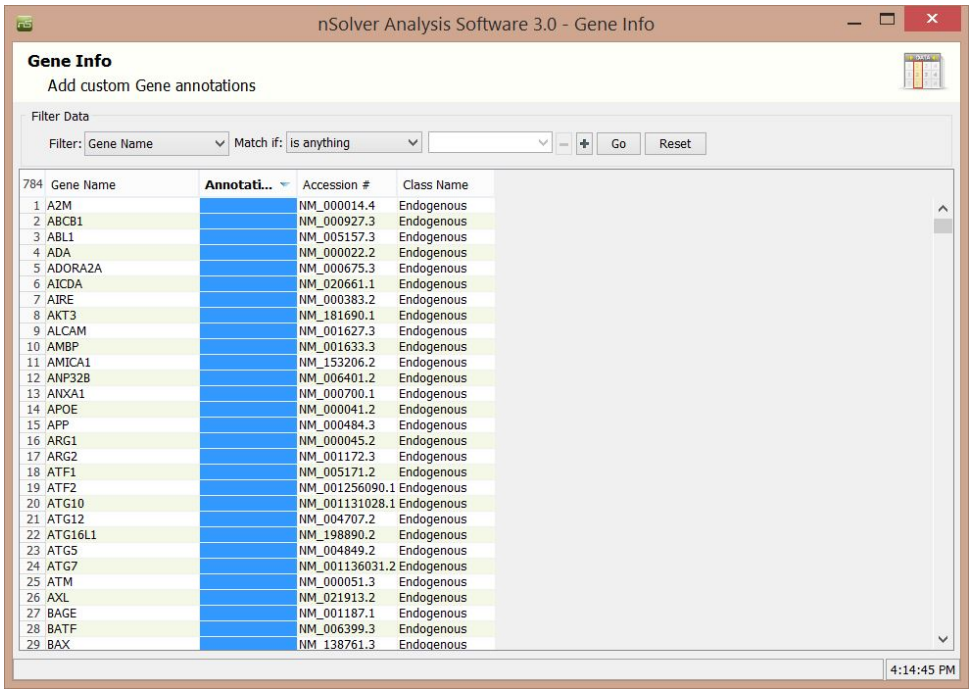
To enter annotations for your probes, click the Raw Data tab and select the RLF file containing the probes you want to annotate. Then select the **Genes** or **Probes** icon depending on your assay type. .



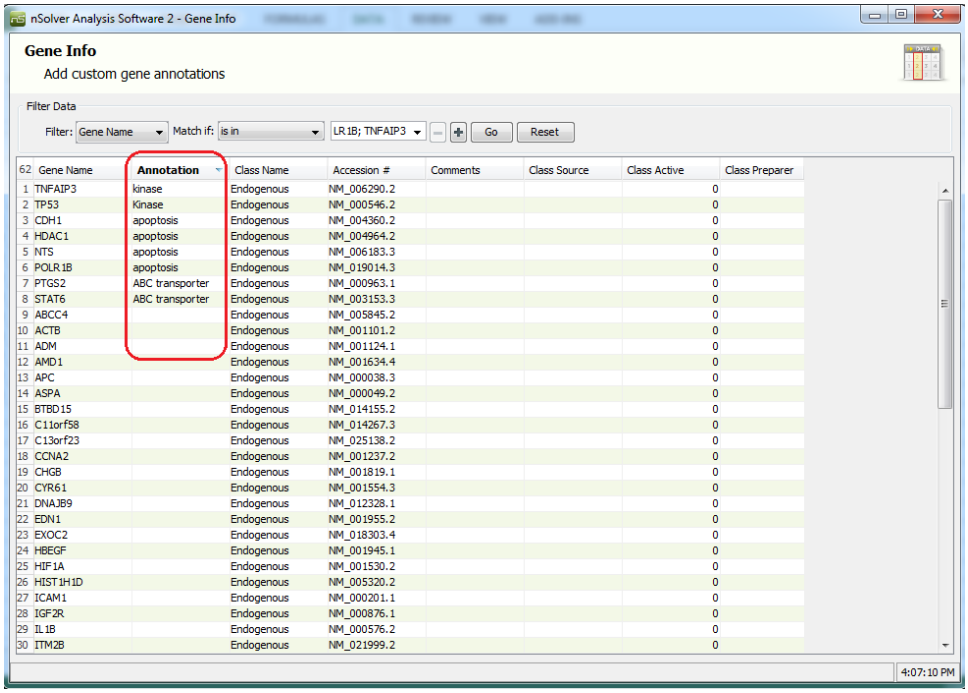


NOTE: It is important to remember that this field is a text field. If it is being used for ordering purposes within the analysis wizard and probe annotations, then it may be advisable to use letters A-Z in the beginning of this field so it sorts properly. Numbers will also be sorted as text fields and may only be evaluated by the first digit (for example, “20” will be come before “3”).

A table will appear containing all probes for the CodeSet. Use the filter at the top to display only the probes you want to annotate. Then type in the annotations as desired. Alternatively, use the copy and paste function to enter your values in bulk as described above and in *Chapter 13 – Data Management*.



The genes will appear and can be visualized and exported in all subsequent nSolver tables and visualizations for this CodeSet.



6 Viewing Raw Data

This chapter contains information on the different ways to view your raw data.

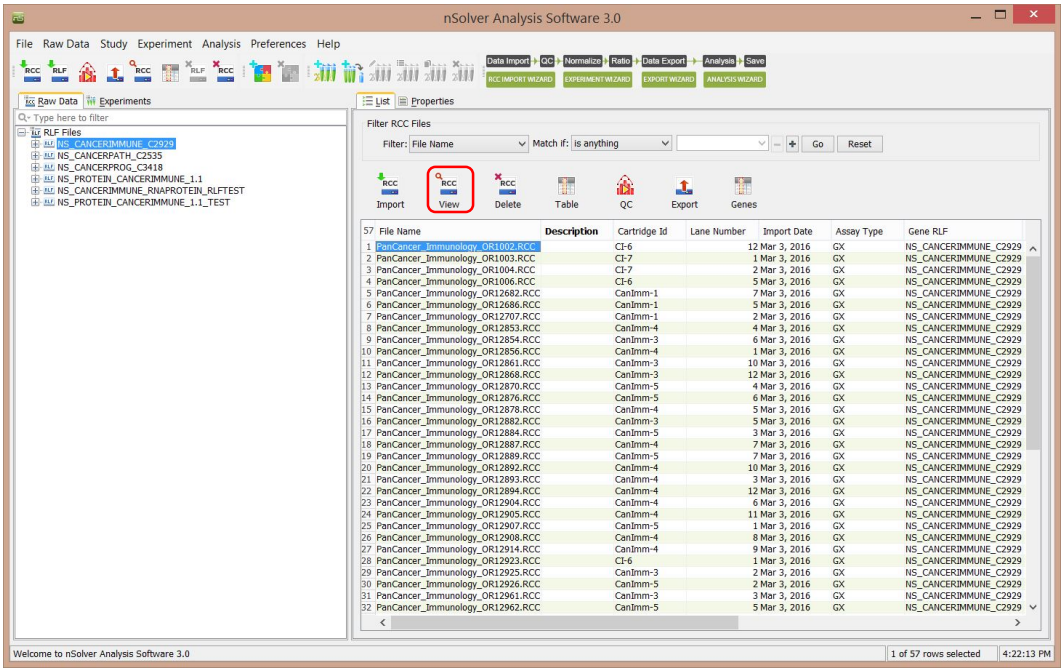
After data are imported and examined for QC, frequently you will want to view the data or perhaps even export it prior to normalizing and creating fold changes.

There are two main views for examining raw data. Access both from the Raw Data View in the left hand navigation window. You can search on your raw data names by typing into the filter, or simply click on the appropriate CodeSet (RLF name) to view all raw data loaded for the CodeSet.

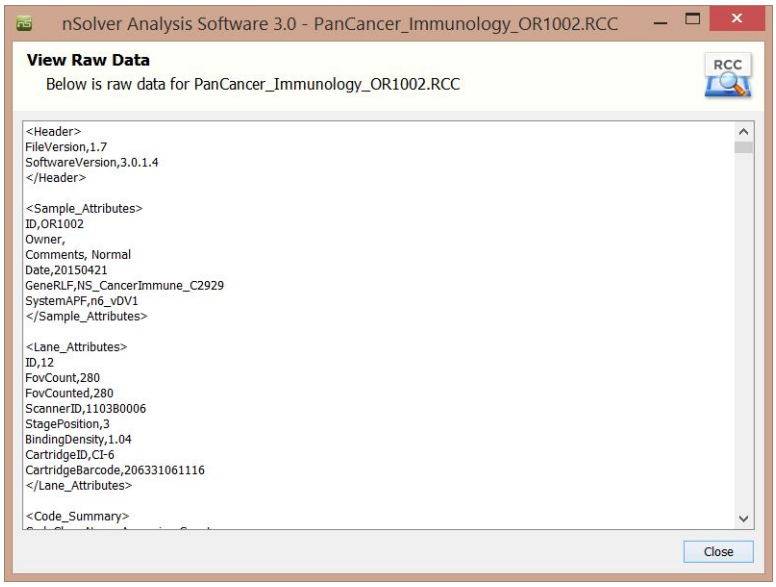
The actual RCC file that was imported can be viewed within the nSolver application. Only one file can be viewed at a time.

NOTE: This view can be particularly useful when checking miRNA and miRGE files to view the background subtraction values that are automatically calculated upon import. The View RCC file is the only way to view the actual raw, unadjusted counts and background subtraction values for miRNA and miRGE.

To view the individual RCC file, select the appropriate sample/lane and click on the **View** icon as shown below:

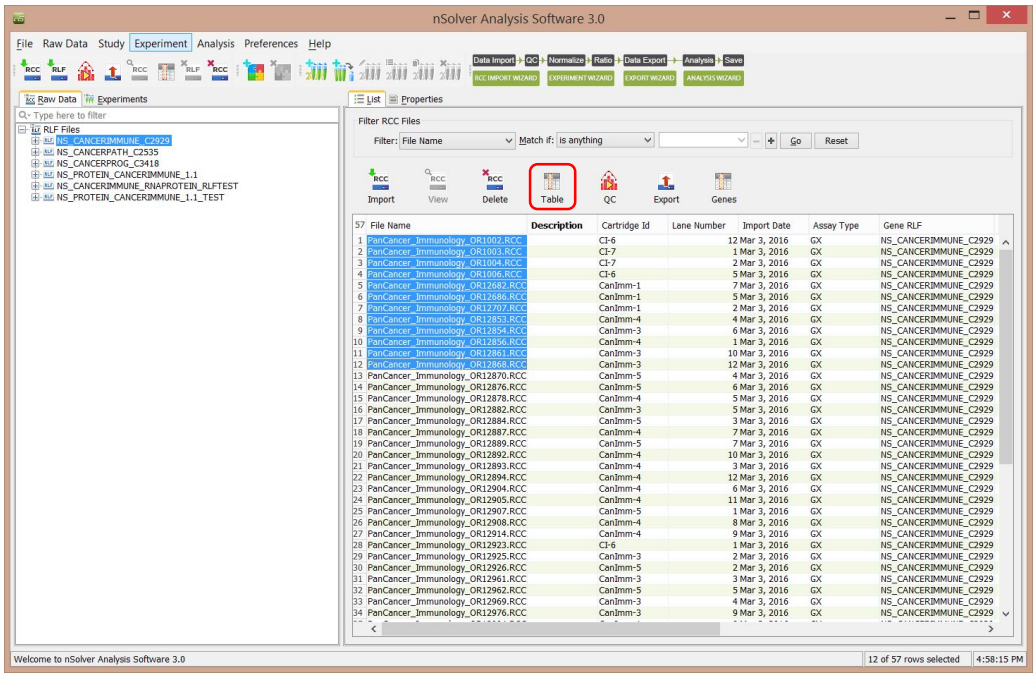


The RCC file will open in the View Raw Data dialog box.

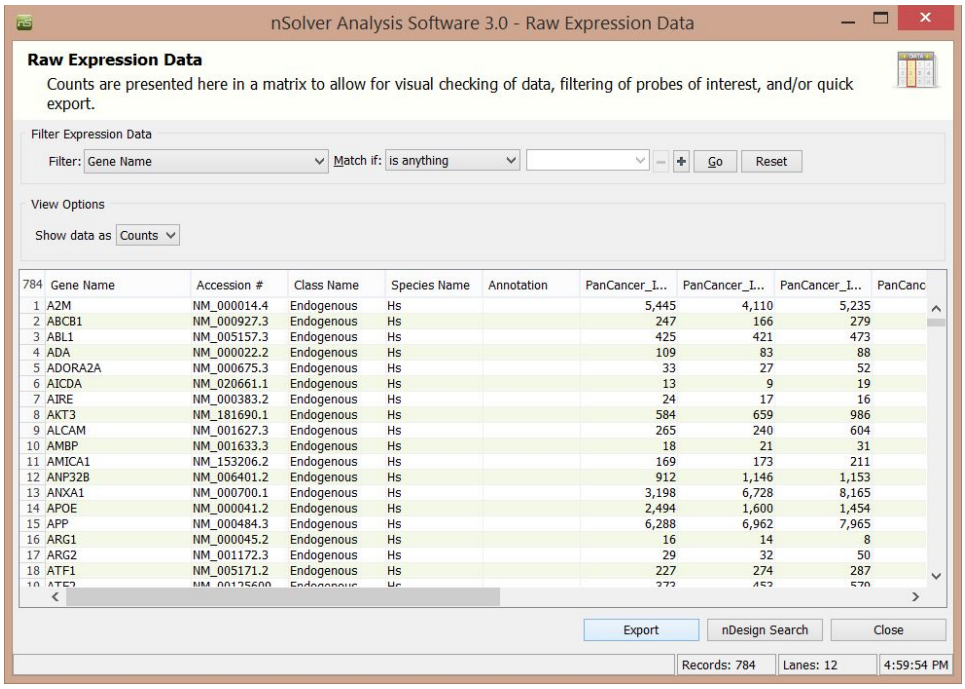


More frequently, you may want to examine raw data for multiple lanes or samples at the same time. The Raw Table view is much more flexible than the View RCC option. The Raw Table view allows searching, sorting, and exporting data in addition to the other features that are available in all tables in the system.

To open a table view of lanes/samples, select the appropriate rows from the main window and click on the **Table** icon as shown below:



The Raw Expression Data dialog box enables you to search the table for a particular list or item, sort, rearrange columns, show hidden columns, and export in a variety of formats. For more information on functions available in tables, see the Table Functionalities section in *Chapter 13 – Data Management*.

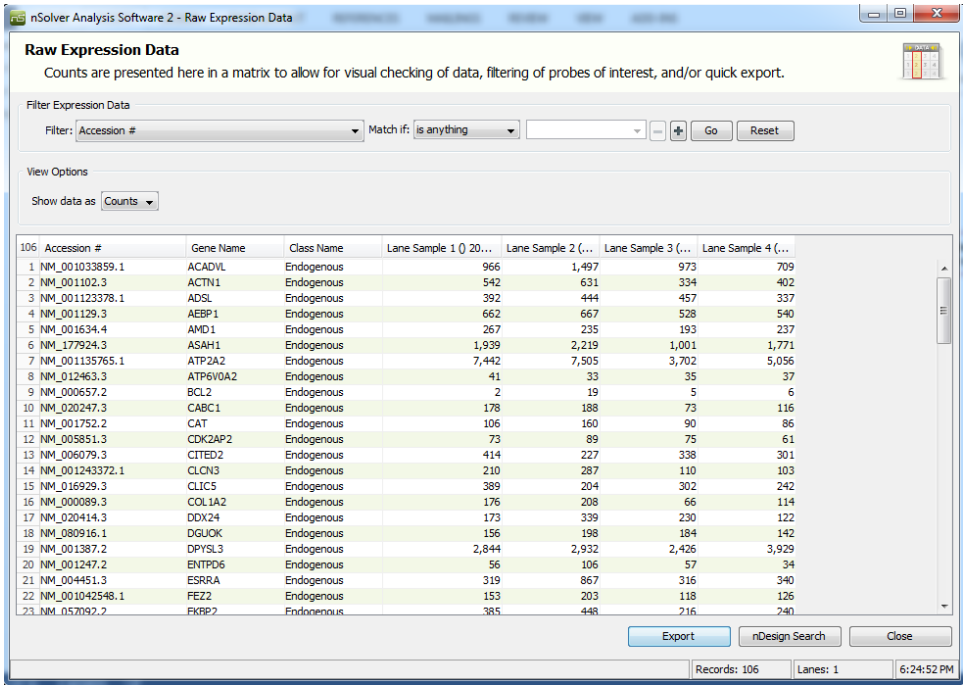


NOTE: Column headers in **red** indicate a QC Flag is associated with the lane.

Columns available in Raw Data table are as follows:

- **Accession #**
- **Gene Name**
- **Class Name**
- **Species Name** (hidden by default; RLF import required)
- **Probe Name** (hidden by default; RLF import required)
- **Target Sequence** (hidden by default; RLF import required)
- **Comments** (hidden by default; RLF import required)
- **Class Source** (hidden by default; RLF import required)
- **Positive Flag** (hidden by default; RLF import required)
- **Class Active** (hidden by default; RLF import required)
- **Class Preparer** (hidden by default; RLF import required)
- **Class Date** (hidden by default; RLF import required)
- **Avg Count** (hidden by default)—Can be used to obtain a sense for the count level for the gene across the samples (very low counts may indicate lower confidence in expression levels observed)
- **Min Count** (hidden by default)—Can be used in conjunction with Ave Count to obtain a sense for the count level for the gene across the samples
- **Max Count** (hidden by default)—Can be used in conjunction with Ave Count to obtain a sense for the count level for the gene across the samples
- **% CV** (hidden by default)—%Coefficient of Variation - Can be used to identify and verify housekeeping genes
- **File Name(s)**

NOTE: For Plex^2 Assays, genes associated with each sample/sub-CodeSet are separated by the **Class Name** column (*e.g.*, Class Name: Endogenous1=sub-CodeSet 1, Endogenous 2=sub-CodeSet 2). Housekeeping genes are also separated by Class Name (e.g., Class Name: Housekeeping1=Housekeeping genes for sub-CodeSet 1).



106	Accession #	Gene Name	Class Name	Lane Sample 1 () 20...	Lane Sample 2 (...	Lane Sample 3 (...	Lane Sample 4 (...
1	NM_001033859.1	ACADVL	Endogenous	966	1,497	973	709
2	NM_001102.3	ACTN1	Endogenous	542	631	334	402
3	NM_001123378.1	ADSL	Endogenous	392	444	457	337
4	NM_001129.3	AEBP1	Endogenous	662	667	528	540
5	NM_001634.4	AMD1	Endogenous	267	235	193	237
6	NM_177924.3	ASAH1	Endogenous	1,939	2,219	1,001	1,771
7	NM_001135765.1	ATP2A2	Endogenous	7,442	7,505	3,702	5,056
8	NM_012463.3	ATP6V0A2	Endogenous	41	33	35	37
9	NM_000657.2	BCL2	Endogenous	2	19	5	6
10	NM_020247.3	CABC1	Endogenous	178	188	73	116
11	NM_001752.2	CAT	Endogenous	106	160	90	86
12	NM_005851.3	CDK2AP2	Endogenous	73	89	75	61
13	NM_006079.3	CITED2	Endogenous	414	227	338	301
14	NM_001243372.1	CLCN3	Endogenous	210	287	110	103
15	NM_016929.3	CLIC5	Endogenous	389	204	302	242
16	NM_000089.3	COL1A2	Endogenous	176	208	66	114
17	NM_020414.3	DDX24	Endogenous	173	339	230	122
18	NM_080916.1	DGUOK	Endogenous	156	198	184	142
19	NM_001387.2	DPYSL3	Endogenous	2,844	2,932	2,426	3,929
20	NM_001247.2	ENTPD6	Endogenous	56	106	57	34
21	NM_004451.3	ESRRA	Endogenous	319	867	316	340
22	NM_001042548.1	FEZ2	Endogenous	153	203	118	126
23	NM_057092.2	FKBP2	Endogenous	385	448	216	240

7 Studies

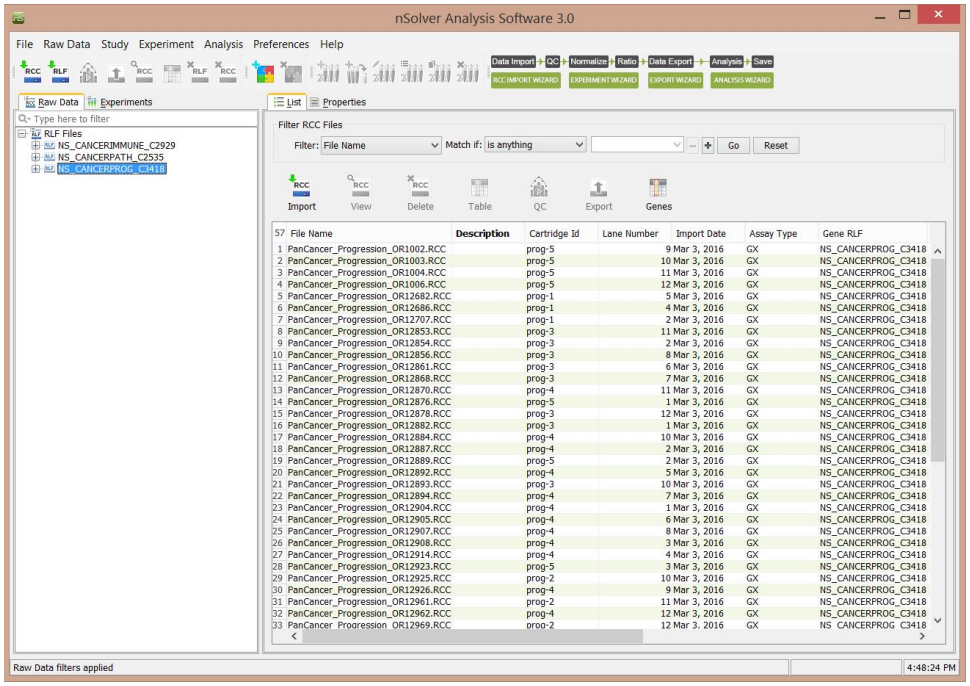
This chapter contains information on how to create, edit, and delete studies.

Studies allow you to logically organize your experiments, data, and analysis results. You can set up studies to organize your data in any way, for example, based on business (such as departments) or scientific areas (such as Oncology Biomarkers). By setting up studies and using them throughout the system, all downstream analyses and operations can be grouped, making all of your data easy to find. A study can contain multiple experiments and span multiple CodeSets and assay types. Similarly, the same RCC files can be used across multiple experiments (although gene annotations are preserved at the CodeSet/RLF level).

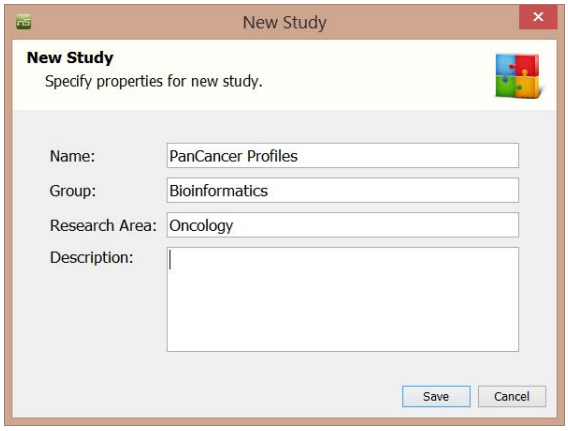
In contrast, an experiment is the particular data set for which you are using to answer the biological question. Experiments are explained in detail in *Chapter 9, Experiments*.

NOTE: In order to create an experiment, you must first create a study.

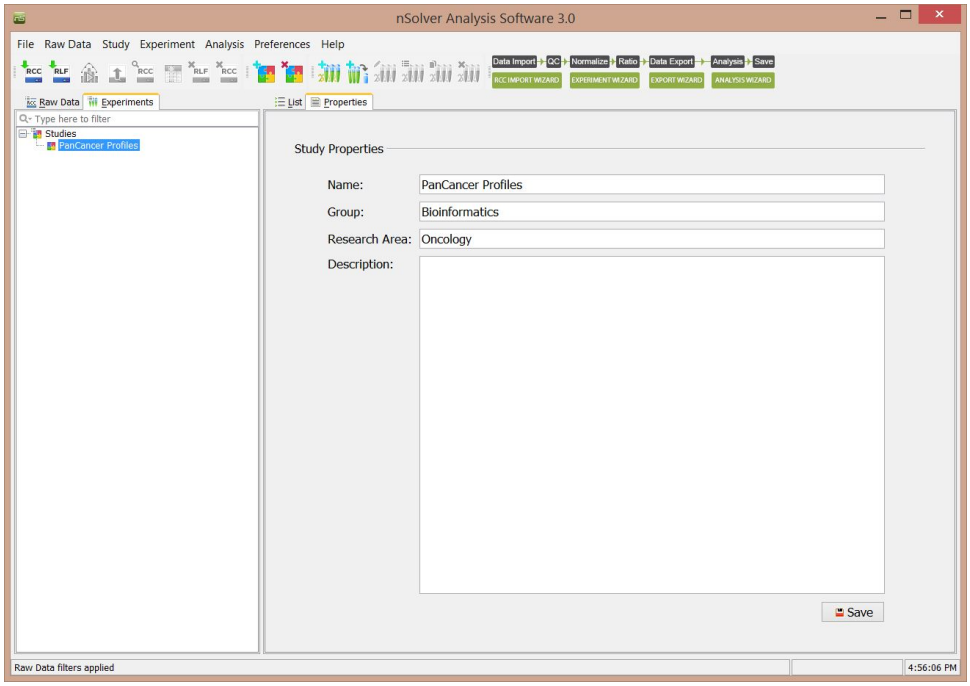
To create a new study, click on the **New Study** icon from the top menu or select it from the Study dropdown menu.



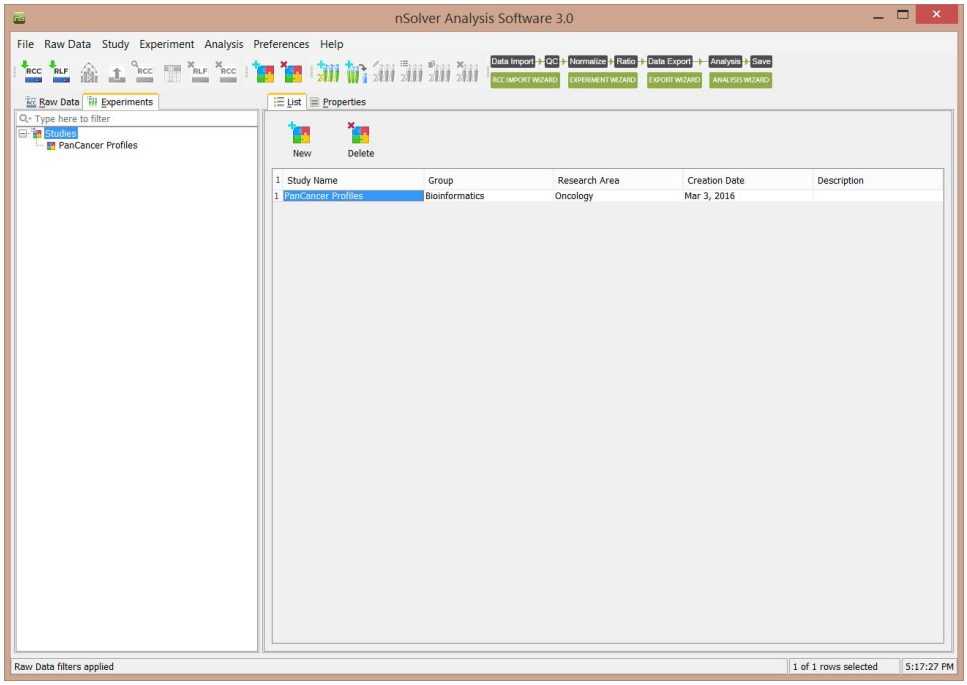
The New Study dialog box will appear. Enter a unique study name. The Group, Research Area, and Description fields are optional. Click **Save** to save the study in the database.



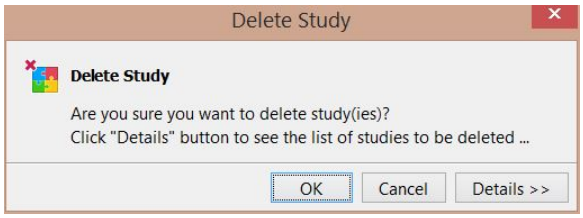
To edit a study, select the study from the navigation tree on the left side and click on the **Properties** tab. Edit your study and click **Save**.



You can delete studies that do not have experiments associated with them. If a study contains experiments, you must delete the experiments before you can delete the study. To delete a study, click on **Studies** at the top of the navigation tree, select from the list of studies, and click the **Delete** icon.



At the prompt, confirm that you want to delete the study by clicking **OK**.



Your study will be deleted.

8 Experiments

This chapter contains information about how to create, edit, copy, and delete experiments using the Experiment Design Wizard. This process is required to normalize, group, and create fold change estimates.

nSolver uses the term 'Experiment' to describe the annotations, background subtraction, normalization, and comparisons (ratios) required to answer the biological question of interest. Each experiment generates four different levels of data that correspond to each step in analysis.

- Raw data are the group of raw samples that were included in the experiment.
- Normalized data are the same samples after undergoing the background subtraction and normalization calculations based on the parameters chosen.
- Grouped data are the geometric mean of the normalized samples within each treatment group as defined by the user.
- Ratio data are the ratio of each sample to the baseline sample or ratio of the grouped data to the baseline chosen by the user in the Experiment Wizard if sample annotation columns were added.
- An analysis group is created but will remain empty until analyses are run.

The exception to this format is the MultiRLF Merge experiment, which does not have Raw data.

There are three types of experiments that you can run in nSolver 3.0:

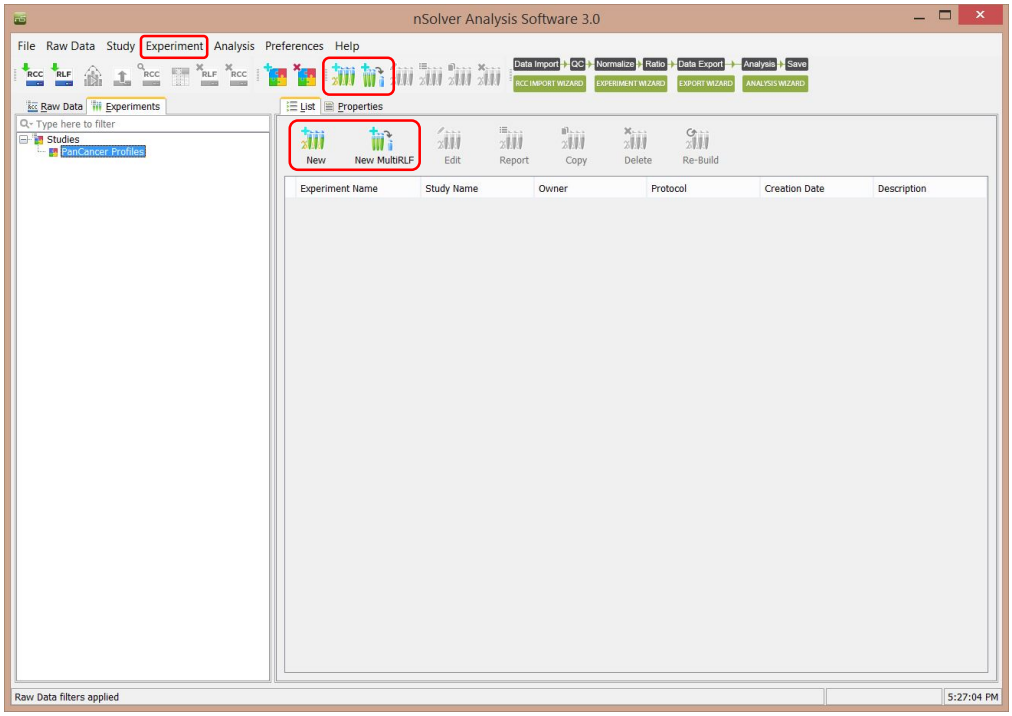
- **Standard experiment**—create an experiment using data for samples run with a single RLF or CodeSet.
- **CrossRLF experiment**—consolidate data for sets of non-identical samples run across multiple CodeSets or CodeSet RLFs. Select raw data from two or more distinct non-Plex2 CodeSets of the same assay type to create a CrossRLF experiment. CrossRLF experiments will only contain data for probes shared between all selected RLFs.
- **MultiRLF experiment**—consolidate data from a set of identical samples run across two or more CodeSets to be aggregated. Select previously merged experiments to merge normalized data. A Sample ID is required to be entered to identify matched samples that were run across different CodeSets. Data may be centered to counts observed for probes common between two or more CodeSets with which the samples were run.



Using the Experiment Design Wizard

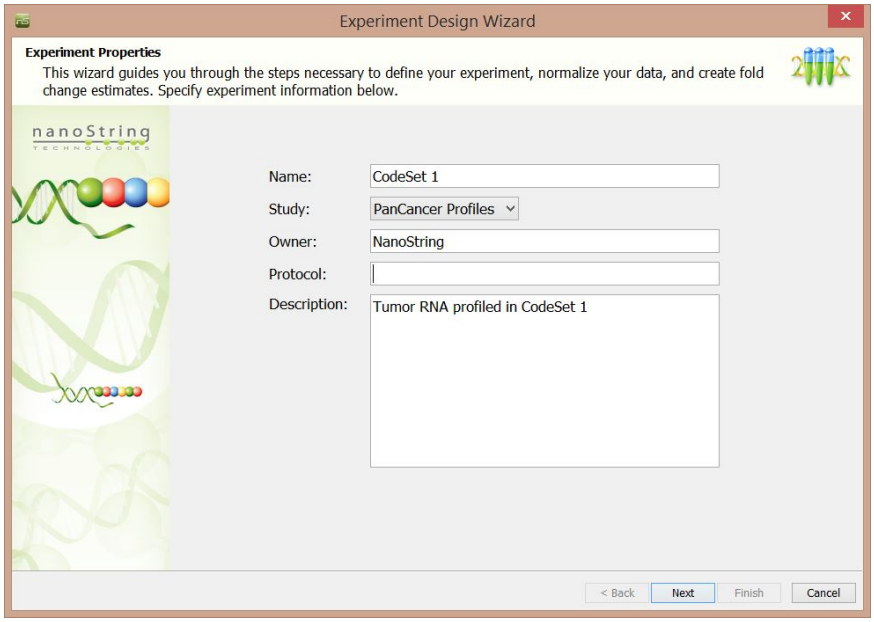
The Experiment Design Wizard is a comprehensive tool for creating experiments. This wizard allows you to create the analysis framework for your experiments using data imported into the nSolver application. Create a new experiment in one of three ways:

- select Experiment from the menu pulldown, or
- click on the New Experiment or New MultiRLF Experiment icons on the menu bar, or
- click the Experiment tab from the navigation tree on the right and select the New or New MultiRLF icon.



Creating a New Experiment

When you select New Experiment, you will see the Experiment Properties dialog box. Enter a unique experiment name and choose the study associated with your experiment. The Owner, Protocol, and Description fields are optional but highly recommended for future reference. Click **Next** to proceed to the next screen of the wizard.

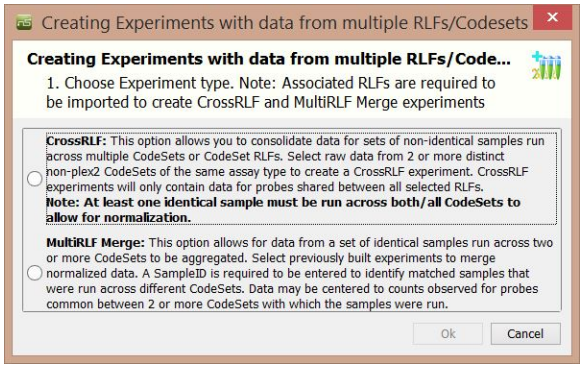


Creating Experiments from Multiple RLFs

When you select New MultiRLF Experiment, you can select a CrossRLF or a MultiRLF Merge experiment.

NOTE: A CrossRLF experiment allows you to compare samples that are different except for the calibrator sample and run on different CodeSets. At least one identical sample must be run across multiple CodeSet RLFs which is used to calibrate counts of overlapping probes between Codesets.

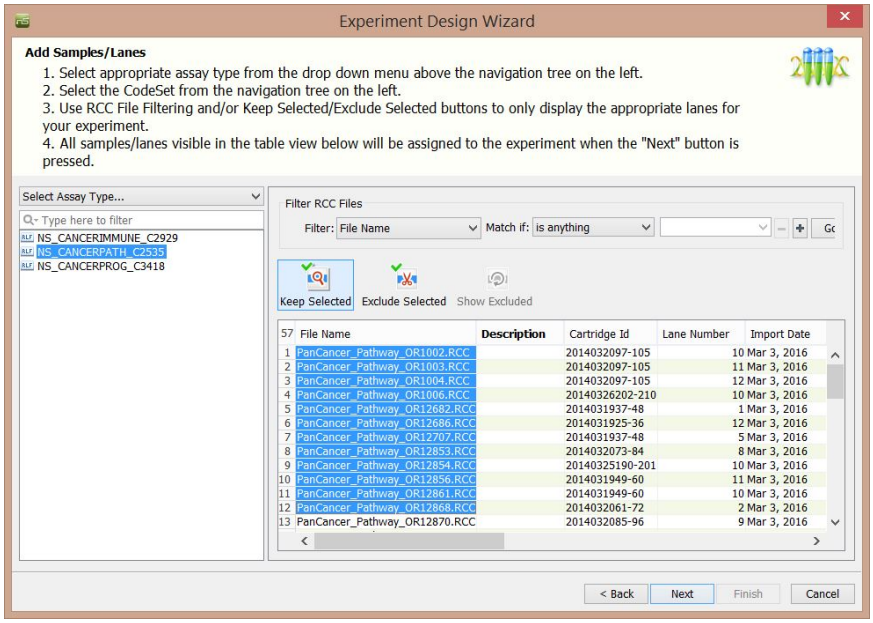
A MultiRLF Merge uses previously-built experiments and allows you to merge and compare data from a set of identical samples across multiple CodeSets. MultiRLF Merge allows normalized data for identical samples to be scaled based on the expression of overlapping probes between CodeSets and then combine the scaled data across all CodeSets for the samples.



Use the radio button to select the type of experiment you want to create and click **OK**. You will see the Experiment Properties dialog box. Enter a unique experiment name and choose the study associated with your experiment. Owner, protocol, and descriptions are optional but recommended. Click **Next** to proceed to the next screen of the wizard.

Selecting Files to Use in the Experiment

The Add Sample/Lanes screen will appear. On this page, choose the raw data that you want to add to your experiment. There are several ways to search for the data you are looking for and select it for use in your experiment.

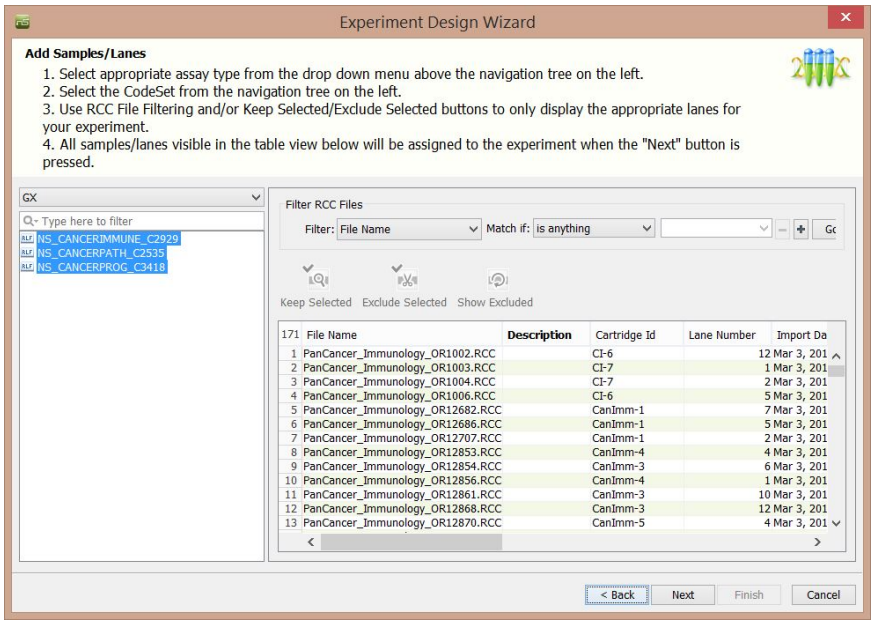


1. From the navigation tree on the left, choose the assay type for your experiment. Choose from GX, miRNA, CNV, miRGE, or Other. Only CodeSets of the chosen assay type will be displayed.
2. Select the CodeSet of interest from the navigation tree. All files imported for the selected CodeSet will display in the table view in the main window.
3. Use the filters at the top of the main window to display only files of interest.
4. Select rows and use the Keep Selected or Exclude Selected icons to filter out any unwanted samples/lanes. Show excluded displays all files once again.
5. After filtering to only display lanes/samples of interest for your experiment, click Next. All samples/lanes displayed in the table are included in the experiment.

Creating a CrossRLF Experiment

The Add Sample/Lanes screen will appear. On this page, choose the raw data that you want to add to your experiment. There are several ways to search for the data you are looking for and select it for use in your experiment.

NOTE: For all cross-CodeSet (CrossRLF) experiments, at least one calibration sample (the same sample) should be tested against each CodeSet.

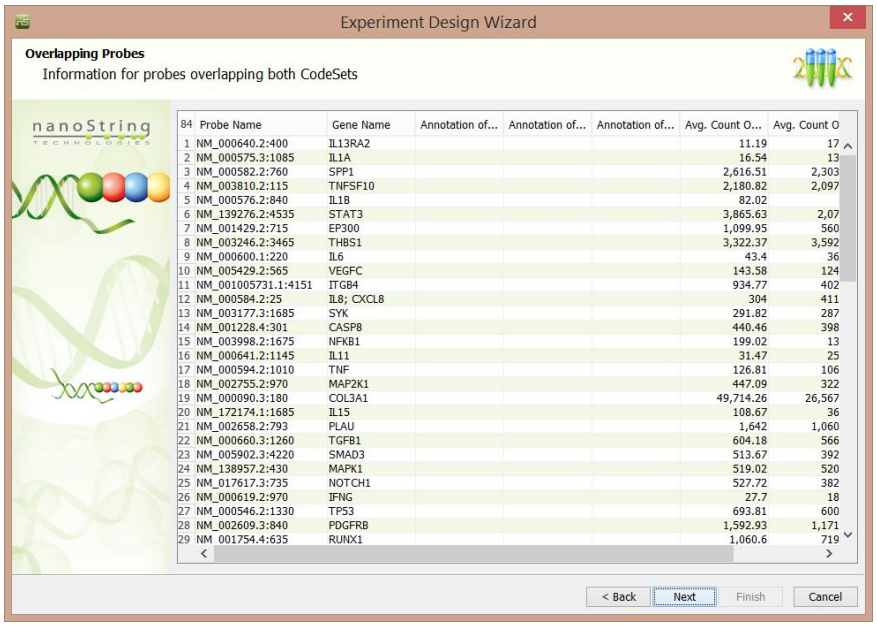


To create an experiment using samples from multiple CodeSets (different RLF files), all of the information from the previous section applies. Additionally, all desired RLF files from the navigation tree on the left need to be selected using Ctrl-click (or Command-click for Macintosh uses). When the additional RLFs are chosen, the table on the right will expand accordingly.

NOTE: All RLF files used for the experiment must be imported in order to proceed to the Cross RLF Overlapping Probes dialog box.

Overlapping Probes

For CrossRLF experiments, the Overlapping Probes dialog box displays information about the overlapping genes across selected RLFs and experiments. CrossRLF experiments can be created even if a single probe overlaps, but it is advisable that most of the probes in all of the CodeSets overlap. The resulting data tables will contain only the overlapping probes.

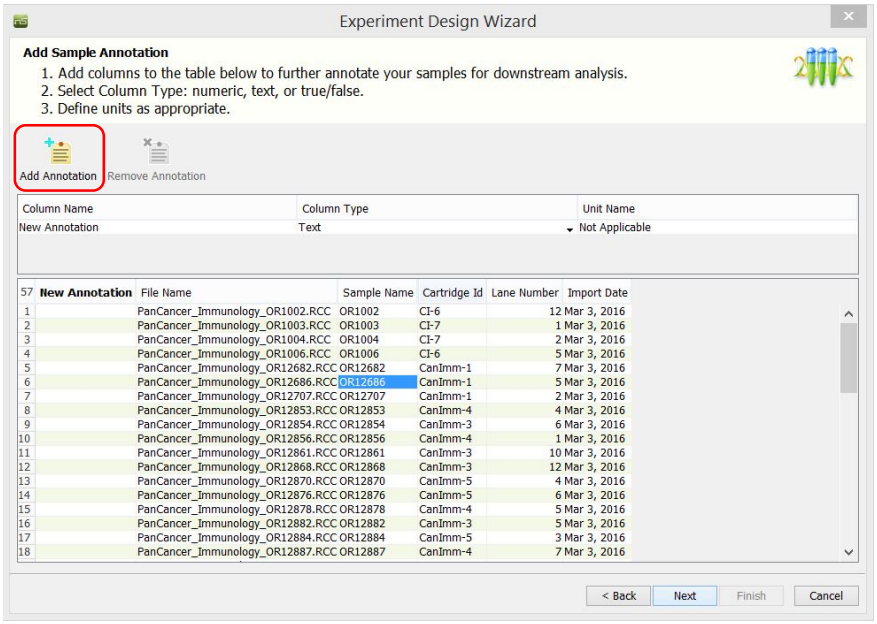


84	Probe Name	Gene Name	Annotation of...	Annotation of...	Annotation of...	Avg. Count O...	Avg. Count O
1	NM_000640.2:400	IL13RA2				11.19	17
2	NM_000575.3:1085	IL1A				16.54	13
3	NM_000582.2:760	SPP1				2,616.51	2,303
4	NM_003810.2:115	TNFSF10				2,180.82	2,097
5	NM_000576.2:840	IL1B				82.02	
6	NM_139276.2:4535	STAT3				3,865.63	2,07
7	NM_001429.2:715	EP300				1,099.95	560
8	NM_003246.2:3465	THBS1				3,322.37	3,592
9	NM_000600.1:220	IL6				43.4	36
10	NM_005429.2:565	VEGFC				143.58	124
11	NM_001005731.1:4151	ITGB4				934.77	402
12	NM_000584.2:25	IL8; CXCL8				304	411
13	NM_003177.3:1685	SYK				291.82	287
14	NM_001228.4:301	CASP8				440.46	398
15	NM_003998.2:1675	NFKB1				199.02	13
16	NM_000641.2:1145	IL11				31.47	25
17	NM_000594.2:1010	TNF				126.81	106
18	NM_002755.2:970	MAP2K1				447.09	322
19	NM_000090.3:180	COL3A1				49,714.26	26,567
20	NM_172174.1:1685	IL15				108.67	36
21	NM_002658.2:793	PLAU				1,642	1,060
22	NM_000660.3:1260	TGFB1				604.18	566
23	NM_005902.3:4220	SMAD3				513.67	392
24	NM_138957.2:430	MAPK1				519.02	520
25	NM_017617.3:735	NOTCH1				527.72	382
26	NM_000619.2:970	IFNG				27.7	18
27	NM_000546.2:1330	TP53				693.81	600
28	NM_002609.3:840	PDGFRB				1,592.93	1,171
29	NM_001754.4:635	RUNX1				1,060.6	719

NOTE: All RLF files used for the experiment must be imported to proceed to the next page of the Experiment Wizard for CrossRLF experiments. If the experiment is a CodeSet Plus experiment, it is important that the original RLF names do not change, as the CodeSet Plus is identified by file name and will not force a reference sample to be identified.

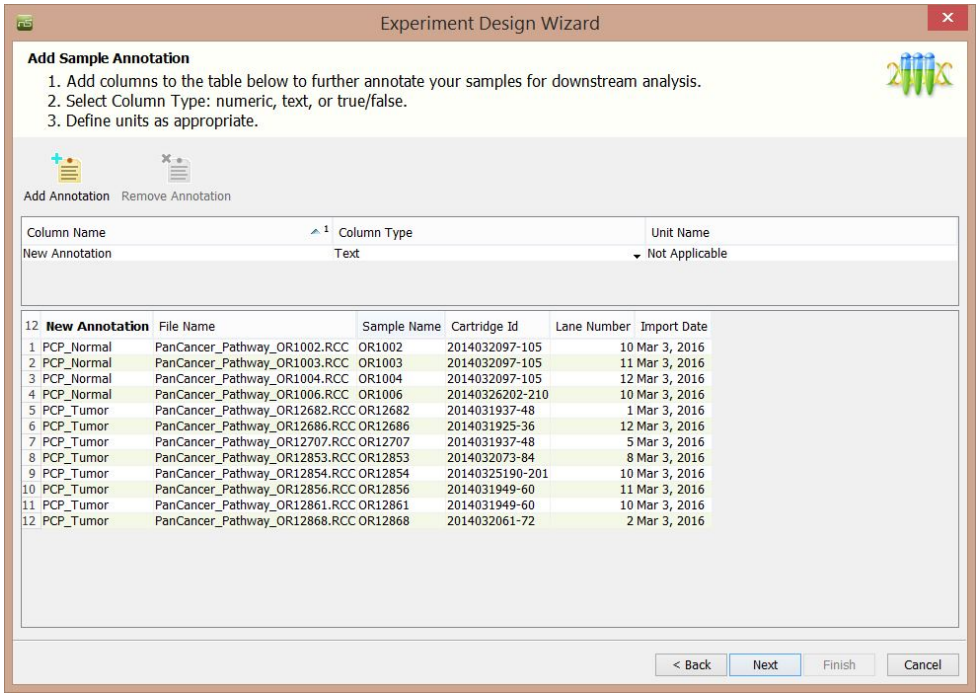
Annotating Lanes/Samples for Grouping and Ratio Building

The next step for all experiments is the Add Sample Annotation dialog box. Use this to add columns to annotate each file and group it for downstream analysis. Click on the **Add Annotation** icon to add a new column as shown below.



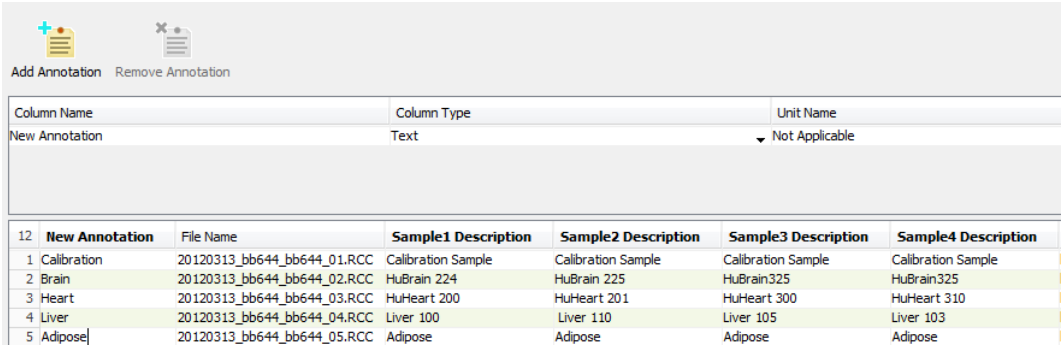
Type in the column name and choose the **Column Type**. Remember that this will be used for downstream analysis; all rows must have data entered.

NOTE: To add annotation that will not be used for downstream analysis see *Chapter 5 – Adding Lane/Sample Descriptions and Gender to Raw Data*.



After selecting the **Column Type**, add your annotation to the columns. You can type directly into each of the cells or copy and paste from an nSolver table, text file, or spreadsheet such as Excel®. To copy and paste from Excel®, select the rows from your spreadsheet as shown below (excluding the header) and use Ctrl-C (or Command-C) to copy to your clipboard.

As part of adding an annotation column, it is important to designate some type of experimental condition. Examples of common annotations are time, concentration, and compound. After adding an annotation column, type in the values associated for each lane/sample. For example, time may have three values entered—1 hour, 2 hours, and 3 hours.



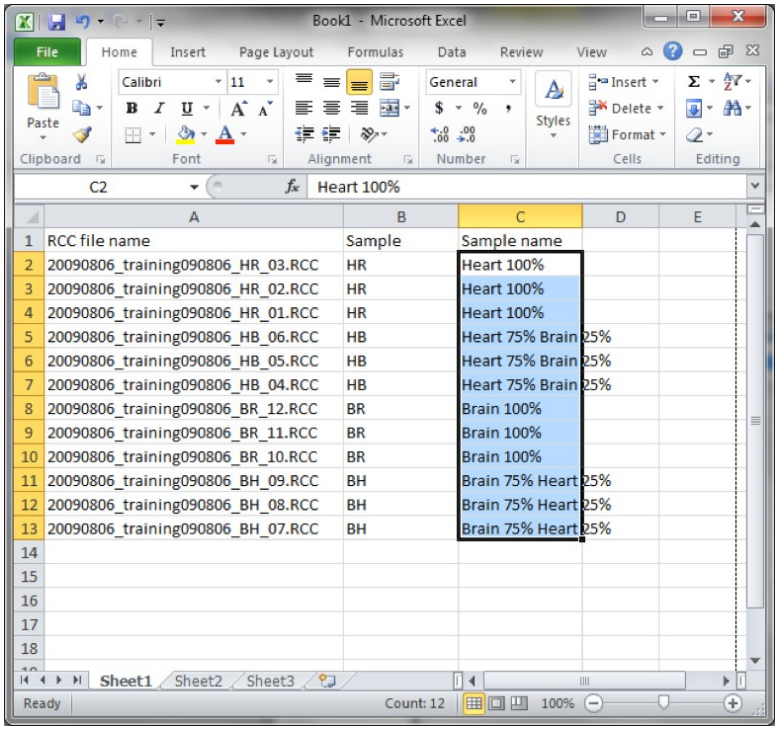
12	New Annotation	File Name	Sample1 Description	Sample2 Description	Sample3 Description	Sample4 Description
1	Calibration	20120313_bb644_bb644_01.RCC	Calibration Sample	Calibration Sample	Calibration Sample	Calibration Sample
2	Brain	20120313_bb644_bb644_02.RCC	HuBrain 224	HuBrain 225	HuBrain325	HuBrain325
3	Heart	20120313_bb644_bb644_03.RCC	HuHeart 200	HuHeart 201	HuHeart 300	HuHeart 310
4	Liver	20120313_bb644_bb644_04.RCC	Liver 100	Liver 110	Liver 105	Liver 103
5	Adipose	20120313_bb644_bb644_05.RCC	Adipose	Adipose	Adipose	Adipose

A more complex example:

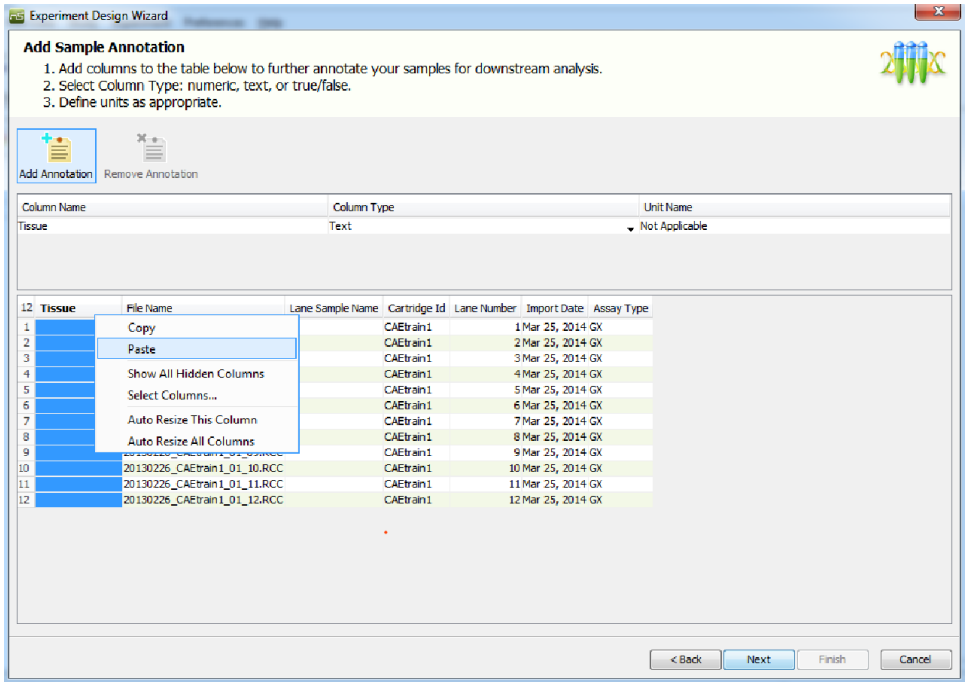
Three-by-Four Experimental Design					
		Time			
		1 Hour	2 Hours	3 Hours	4 Hours
Treatment	Control	Ctrl @ 1 hr	Ctrl @ 2 hrs	Ctrl @ 3 hrs	Ctrl @ 4 hrs
	Compound A	A @ 1 hr	A @ 2 hrs	A @ 3 hrs	A @ 4 hrs
	Compound B	B @ 1 hr	B @ 2 hrs	B @ 3 hrs	B @ 4 hrs

In this example, the two annotation columns to be added are **Treatment** and **Time**. The factor named compound has three different unique values or levels added, and the column named time has four levels. The intersection of a level within each of the factors is considered a treatment group. For example, compound A at 2 hours (A @ 2 hrs) is a treatment group. Also, each treatment group can have an unlimited number of replicates. For example, you might have three different measures of compound A at 1 hour, hence three replicates.

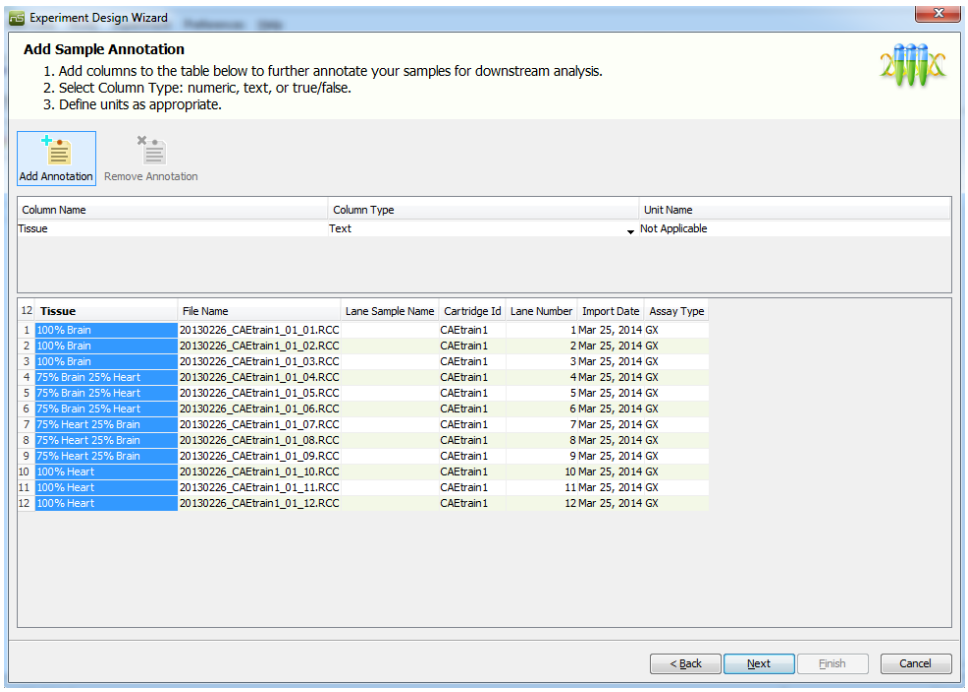
For a more simplified example, the spreadsheet below has 12 lanes/samples. Listed as the sample name are 4 different values. For nSolver, in this case, 1 annotation column should be added, called **Sample Name**. The type should be Text and then the 4 different values can be copied and pasted into the nSolver system. This will create 4 treatment groups, which will later be used to determine the baseline when building ratios.



Then right-click on the column header of the associated column and select **Paste**.



The values will be pasted in appropriately.

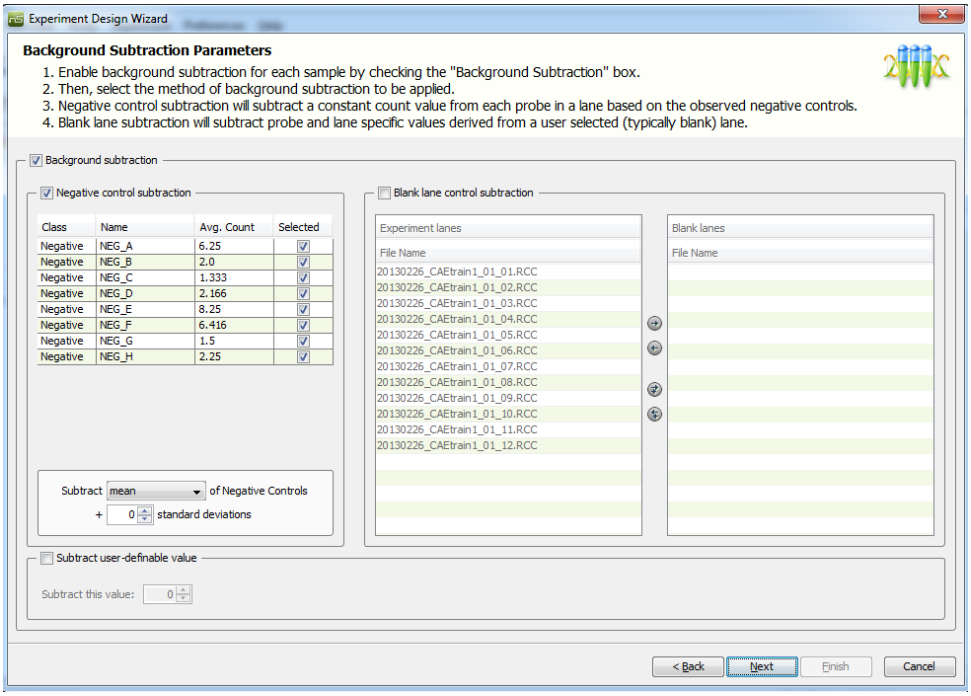


Background Subtraction

nSolver provides options for subtracting a background count level from each of your probes. You can apply background subtraction in three basic ways: using the negative controls, applying a user-specified value, or running and specifying a blank (usually water) lane. In all three cases, background subtraction is applied before normalization. To apply any of the methods, first select the checkbox to run background subtraction in the upper left hand corner, and then select the desired option.

Background Subtraction Using Negative Controls

To apply background subtraction using the negative controls, select the checkbox to run **Background subtraction** and choose the **Negative control subtraction** option. Unselect any of the negatives that you do not want used in the calculation. For the remainder, choose to use their mean, median, geometric mean, or maximum. If mean is desired, enter the number of standard deviations to add to the mean of the negatives that will be used.

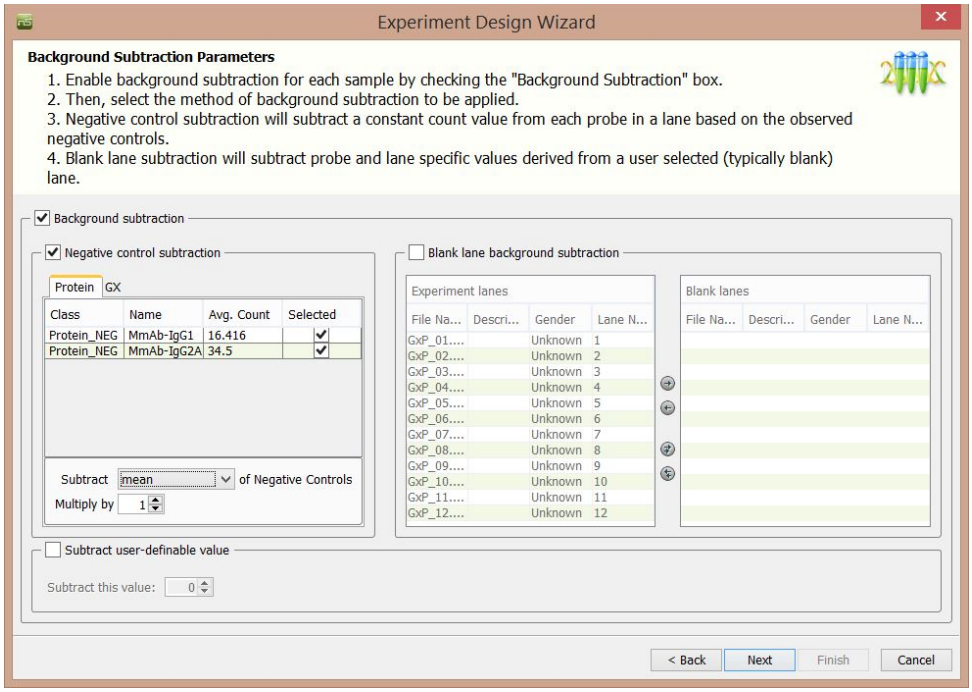


A constant count value calculation based on the mathematical operation chosen will be subtracted from each probe based on the selected negative controls for the associated lane.

Background Subtraction for GeneProtein

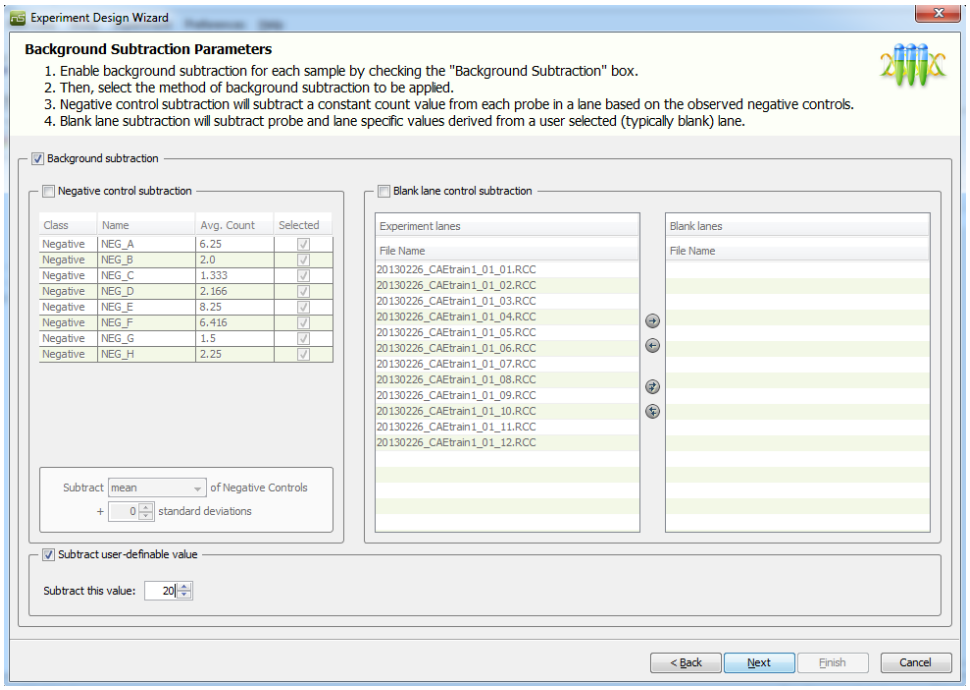
GeneProtein Background subtraction for GeneProtein and Protein assay types is similar to other assay types .To apply background subtraction using the Protein sample-specific negative control probes in a Protein assay , select the check boxes for Protein_NEG probes. Choose the mean, median, geometric mean, or maximum, and any multiple of this value to be subtracted from the raw protein counts.

To apply background subtraction using the Protein sample-specific negative control probes in a GeneProtein assay, click the Protein tab and select the check boxes for Protein_NEG probes. Choose the mean, median, geometric mean, or maximum, and any multiple of this value to be subtracted from the raw protein counts.



Background Subtractions Using a Specified Value

To apply background subtraction by specifying a number, select the checkbox to run **Background subtraction**, select the **Subtract user-definable value** option, and enter a count value.

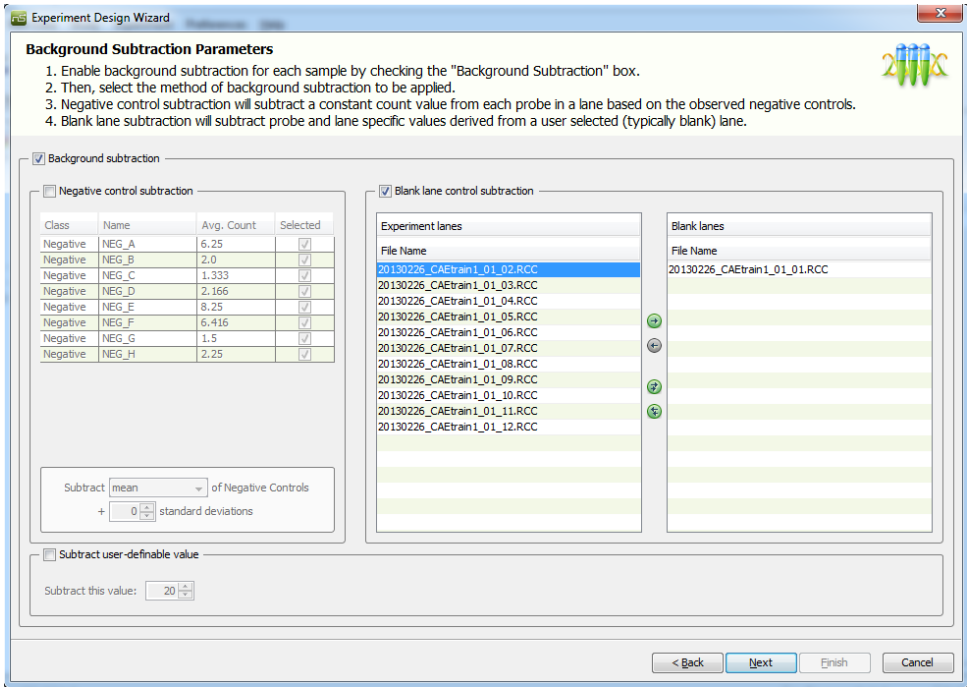


The specified count value will be subtracted from each probe.

Background Subtraction Using Blank Lane

To apply blank lane background subtraction, select the checkbox to run **Background subtraction** and select the **Blank lane control subtraction** option. Select one or more of the blank (water) lanes used to run the cartridge, then click on the right arrow to move that lane to the right side of the window.

Variability in background counts in NanoString data appears to arise primarily from probe-to-probe differences. nSolver uses the blank lane(s) to estimate each probe's unique expected background level and then subtract different levels of background from each of the probes in the experiment lanes.



NOTE: If more than one lane is used for blank lane control subtraction, then the mean of the raw count will be computed to determine the value that will be used for subtraction.

After you have added the appropriate annotations, click **Next** to proceed to the Normalization Parameters screen.

Normalization

Normalization refers to a type of data transformation that balances counts between lanes, allowing you to make meaningful biological comparisons. Due to the unique nature of nCounter data, specific normalization options are offered for each of the nCounter assay types. Please see NanoString's ***nCounter® Expression Data Analysis Guide*** and ***nCounter® Data Analysis Guidelines for Copy Number Variation (CNV)*** for detailed information on the recommended normalization options for each assay type.

Within the nSolver application, the recommended normalization method for each assay type is presented by default. However, you may choose to change the normalization method from the normalization screen.

A description of the defaults and various options for each type are described in the table below:



Experiment Wizard Normalization and Defaults Options Summary

Assay Type	Default	Options
GX *and* **Plex² GX**	• Geometric mean of the Positive Control Probes AND • Geometric mean of the Housekeeping Probes	• Mean or Geometric mean of any/all of the Positive Control Probes • Mean or Geometric mean of any/all of the Content Probes • Scale to X for all content
miRNA	• Geometric mean of Top 100 miRNAs (probes with highest 100 counts)	• Mean or Geometric mean of any/all of the Positive Control Probes • Mean or Geometric mean of any/all of the Content Probes • Scale to 'X' for all Content Probes
CNV *and* **Plex² CNV**	• Mean of Invariant Probes	• Mean & Geometric mean of any/all Positive Control Probes • Mean of any/all of the Content Probes
miRGE	• mRNA Geometric mean of the Positive Control Probes AND Average of Housekeeping Probes • miRNA Geometric mean of the Positive Control Probes	• mRNA: Mean & Geometric mean of any or ALL positive controls • miRNA: Mean & Geometric mean of any or all content probes • miRNA: Mean & Geometric mean of any or ALL positive controls • miRNA: Mean & Geometric mean of any or all content probes • miRNA: Scale to X for all content
Protein	• Geometric mean of Positive Control Probes AND • Geometric mean of Content Normalization Probes	• Mean & Geometric mean of any or ALL Positive Control Probes • Mean & Geometric mean of any or all content probes • Sample Input Calibration counts

Other	• Geometric mean of positive controls AND • Mean of Housekeeping gene	• Mean & Geometric mean of any or ALL positive controls • Mean & Geometric mean of any or all content probes • Scale to X for all content

NOTE: % Coefficient of Variation (%CV) is displayed in the CodeSet Content Normalization box to help identify genes that display lower variability or more stable expression across the samples.

While the normalization parameters screen is tailored for each assay type, there are some basic methods available to alter the parameters.

1. Use checkboxes to activate/deactivate normalization based on the type of control sequence(s) desired.

   Select the checkboxes next to Positive Control Normalization and CodeSet Content Normalization to activate normalization based on positive controls or CodeSet Content, respectively.

2. Use checkboxes to select Positive Control Probes to include in Positive Control Normalization.

   Select the checkboxes next to each positive control probe to include its usage in positive control normalization calculations. Note that although the normalization is applied to the gene counts, it is not applied to the positive controls; they will appear the same as before normalization.

3. Use arrow buttons to move selected probes from CodeSet Content to Normalization Codes.

   All CodeSet Content probes listed in the Normalization Codes box will be used for normalization. Use the single arrow button to move any combination of probes into and out of the CodeSet Content and Normalization Codes boxes. Use the double arrow buttons to move all probes back and forth.

   For GeneProtein Normalization, use radio buttons to choose either **Protein_Cell_Norm**, which allows you to automatically select pre-specified control probes (similar to Housekeeping genes for mRNA), or to choose **All** proteins for normalization.

   For GeneProtein assays, select the Protein tab for normalization of protein probe content. Select the GX tab for normalization of mRNA content.

4. Use radio buttons to select an entire group of CodeSet Content probes for normalization (miRNA/miRGE/Other only).

   Use the radio buttons at the top of the CodeSet Content box to select the **Top 100** count probes, all **Housekeeper** probes, or **All** content probes into the Normalization Probes box with a single click. The average counts used to determine the Top 100 count probes display next to each CodeSet Content probe.

5. Use the dropdown within the lower calculation box to select mean or geometric mean.

   Use the dropdown menus located below each of the Normalization Probe boxes to select the type of mean used to calculate the normalization factor. Use mean if normalization probes are of similar expression levels or use geometric mean if probes reflect widely varying levels of expression. Note that using geometric mean for Positive Control normalization weights the low-concentration positive controls equally to the high-concentration controls.

6. Use the Scale to 'X' radio button when using all Content Probes for Normalization (only miRNA and Other).

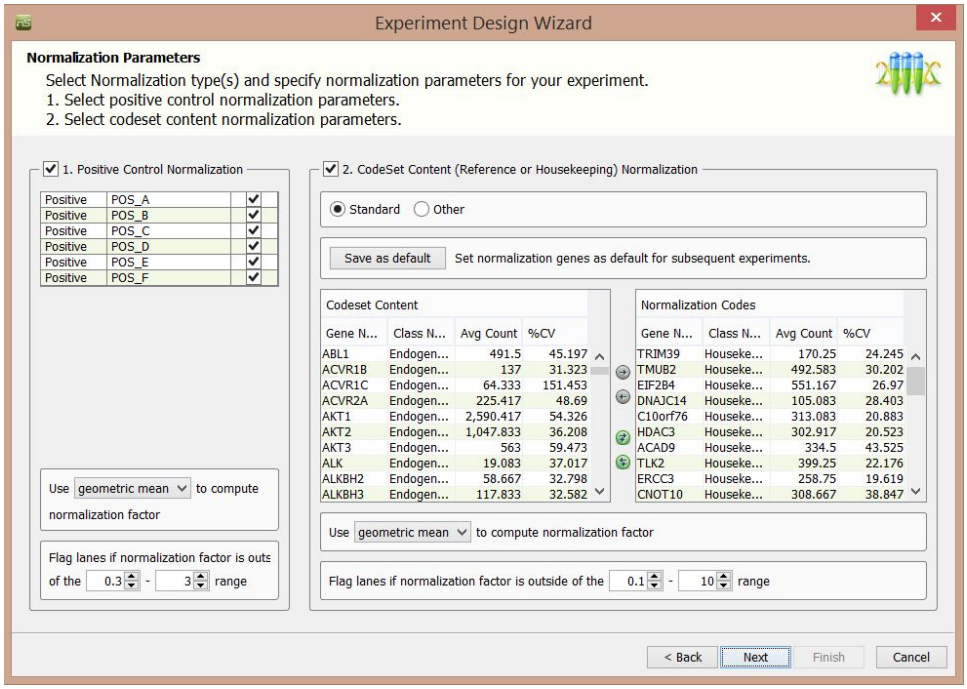
   Select the Scale to 'X' radio button to define a value to which the sum of all probes will be normalized.

7. Select the Set as Default button to set the genes used for content normalization to be used by default in future experiments.

Gene Expression Normalization Dialog Box

Normalization Flags display when the calculated normalization factors are unusually large or small. These ranges can be adjusted to suit your specific data set. Normalization flags are only flags and do not eliminate data from downstream analysis. To eliminate flagged files from downstream analysis, we recommend that you copy or edit the experiment, removing flagged files.

The normalization parameters vary depending on what type of data you are using. The example below is of the gene expression normalization screen.



NOTE: For Plex² Expression Assays, nSolver performs CodeSet Content Normalization within each sub-CodeSet. For example, sub-CodeSet 1 will be normalized with only sub-CodeSet 1 probes for all RCC files within the experiment. In the resulting normalized data output file, columns will contain normalized counts for all four sub-CodeSets for each RCC file. Positive Control Normalization may also be used in conjunction with CodeSet Content Normalization, but we do not recommend using only Positive Control Normalization. Please refer to the Plex² product literature for more information on the design and utility of the positive controls for Plex² Expression Assays.

GeneProtein Normalization Dialog Box

Normalization parameters can be selected separately for Protein and mRNA-specific probes in a GeneProtein assay by clicking the Protein or GX tab.

GeneProtein Normalization options for the Protein content in GeneProtein assays are identical to Protein assays. Use the radio buttons to select Standard or Input Calibration for CodeSet Content Normalization.

For Standard Protein Normalization, select either **Protein_Cell_Norm**, which automatically selects Protein_CELL_NORM probes, or **All**, which selects all protein probes as normalization Codes.

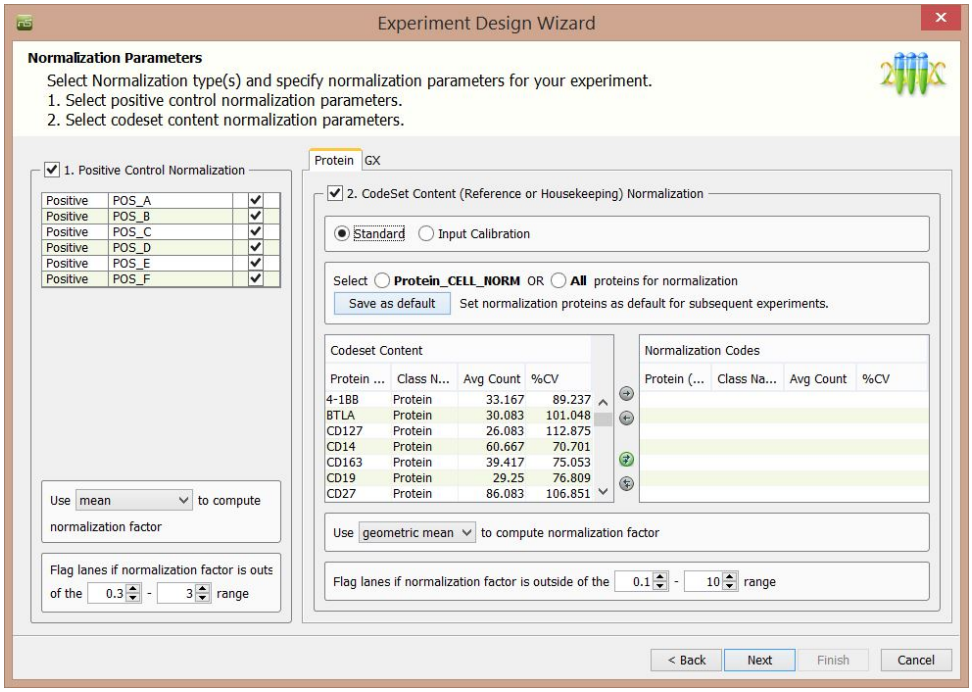
In the Standard method, select any set of one or more probes. Use the single arrows to move probes between the CodeSet Content table and the Normalization Codes table. Use the double arrows to move entire probe sets between the CodeSet Content table and the Normalization Codes table.

Input calibration allows protein probe counts to be normalized to the sample quantity used in the Protein assay, such as the number of cells or the unit quantity of protein.

Select Input Calibration radio button. The default value is 1 which assumes equal / comparable sample quantity across all lanes. Replace the default values with numeric values representing Sample input quantity. A sample input calibration factor is calculated where,

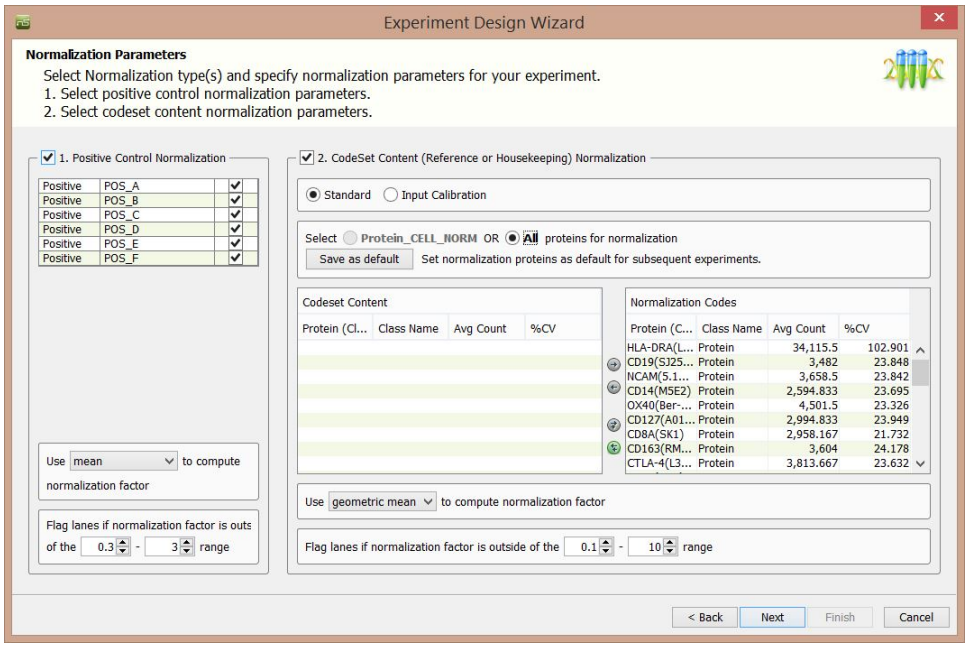
Sample input calibration factor = minimum input calibration factor entered on all samples ÷ value entered for that sample

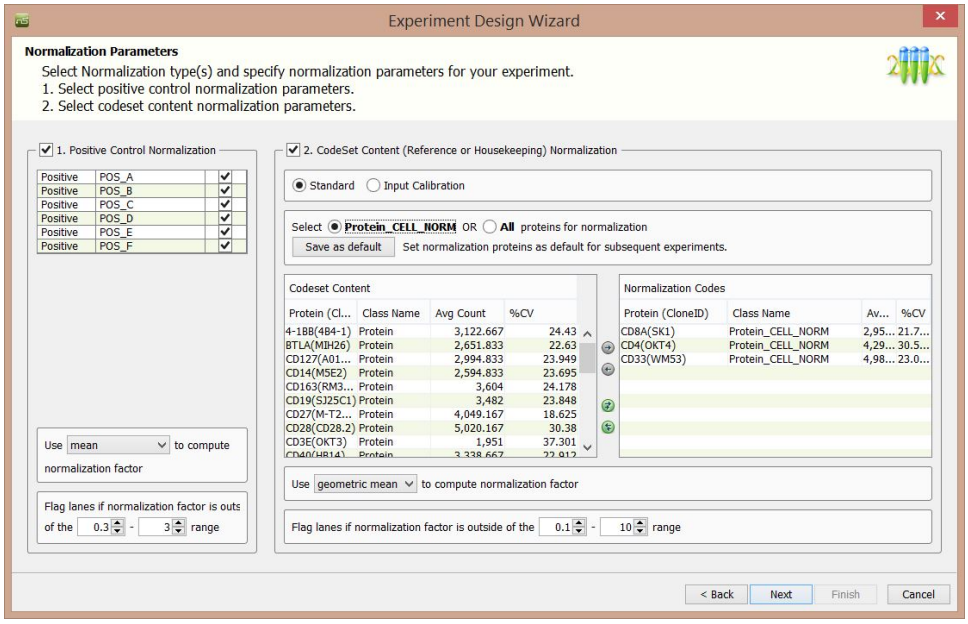
Sample input calibration factor will then be applied to all Protein assay specific probes.



Protein Normalization Dialog Box

For Protein Normalization, select either **Protein_Cell_Norm**, which automatically selects Protein_CELL_NORM probes or allows you to select a one or more probes to use as normalization codes, or **All**, which automatically selects all protein probes to use as normalization codes.

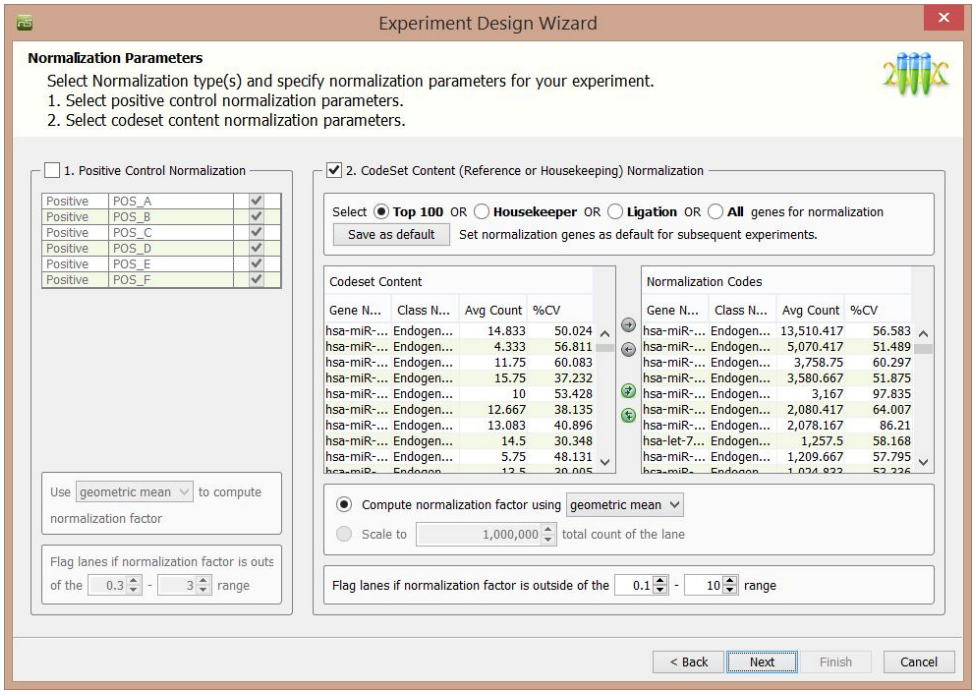




miRNA Normalization Dialog Box

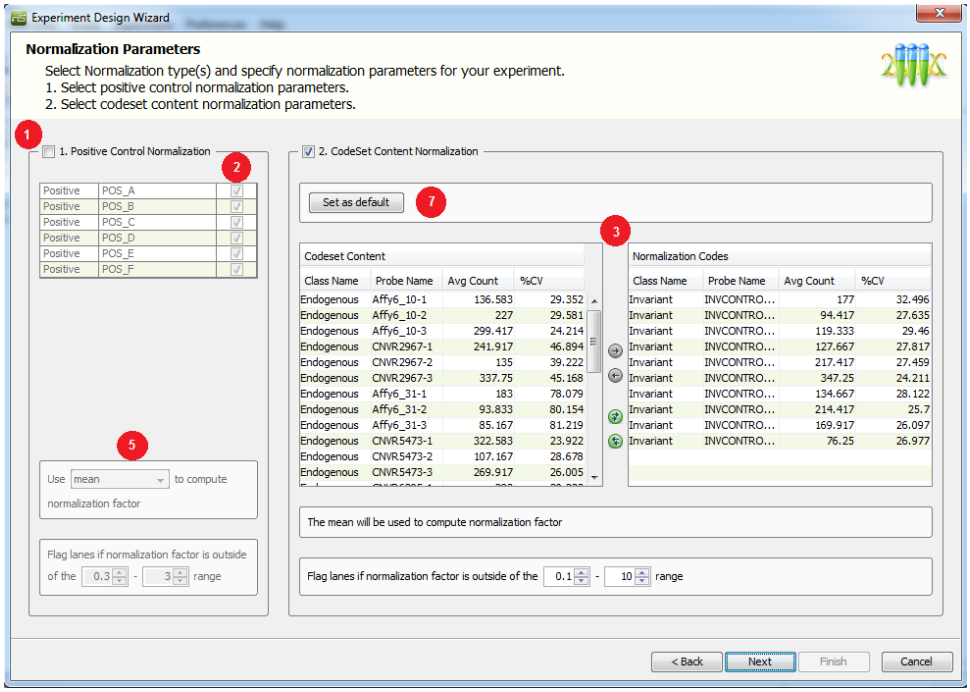
NOTE: For miRNA assays, the background subtraction denoted with '+++' in the RCC and RLF files is automatically performed during import.

For the CodeSet Content (Reference or Housekeeping) Normalization, select the **Top 100**, **Housekeeper**, **Ligation,** or **All** radio button to automatically select the respective set of probes as normalization codes to calculate normalization factor.



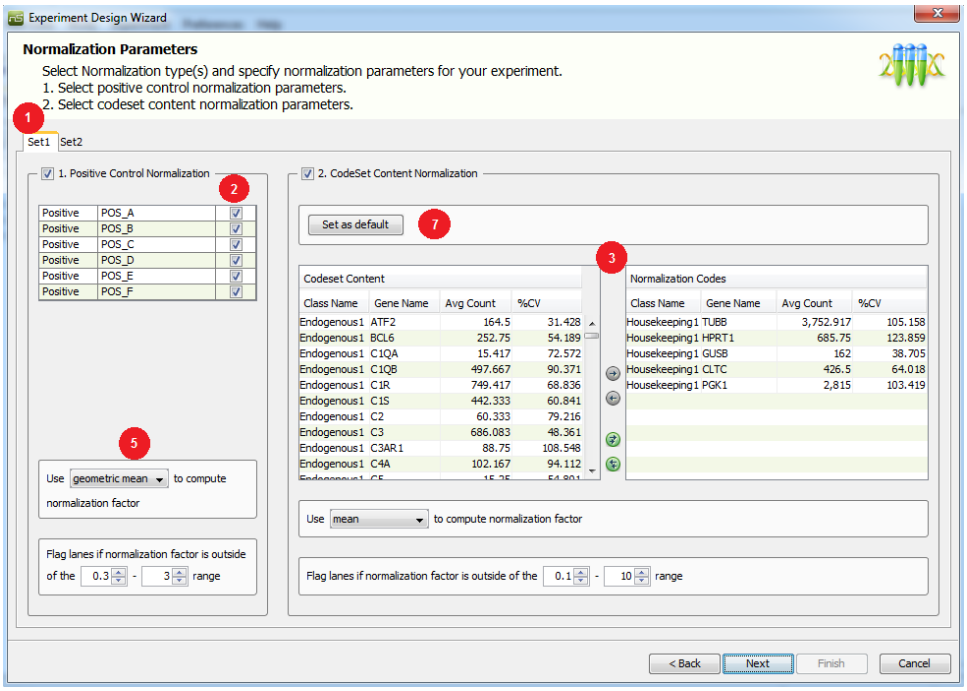
NOTE: If you prefer to use spike-ins to normalize your miRNA assays, select all current probes in the normalization codes window and move them back to the left hand side using single left pointing arrow or simply click the left pointing double arrows . Then select all probes with Class Name of 'SpikeIn' and move these to the Normalization codes window by clicking the arrows to move the probes to the right.

CNV Normalization Dialog Box

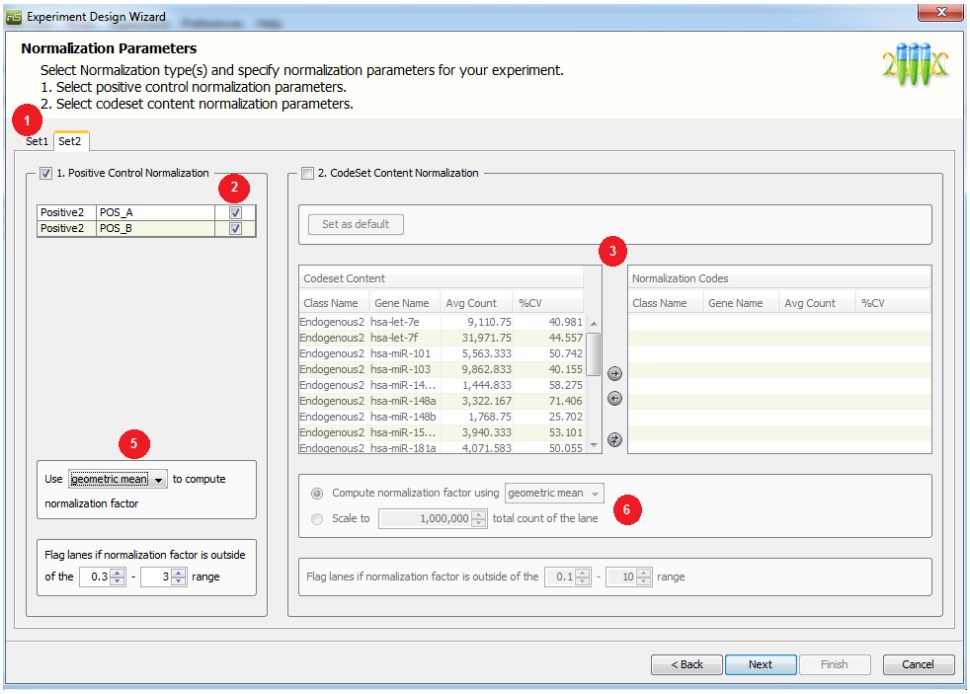


miRGE Normalization Dialog Box

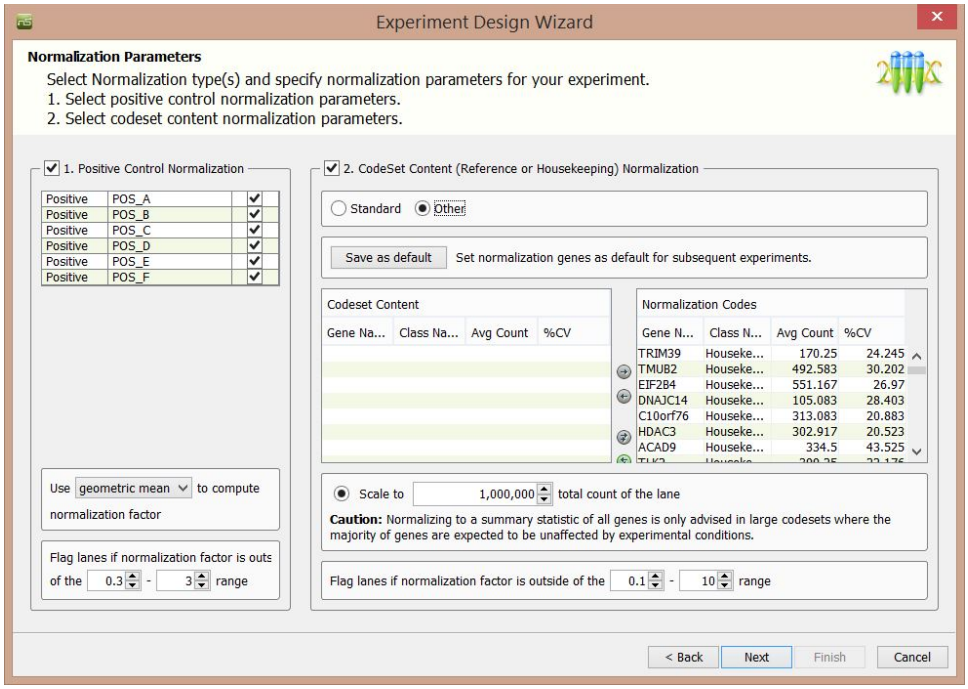
It is important to note that for miRGE data, there are two normalization options, one for mRNA and one for miRNA. Set 1 is mRNA. The normalization options and defaults are displayed below:



Click on the Set 2 tab in the upper left corner to specify the miRNA normalization options, as shown below.

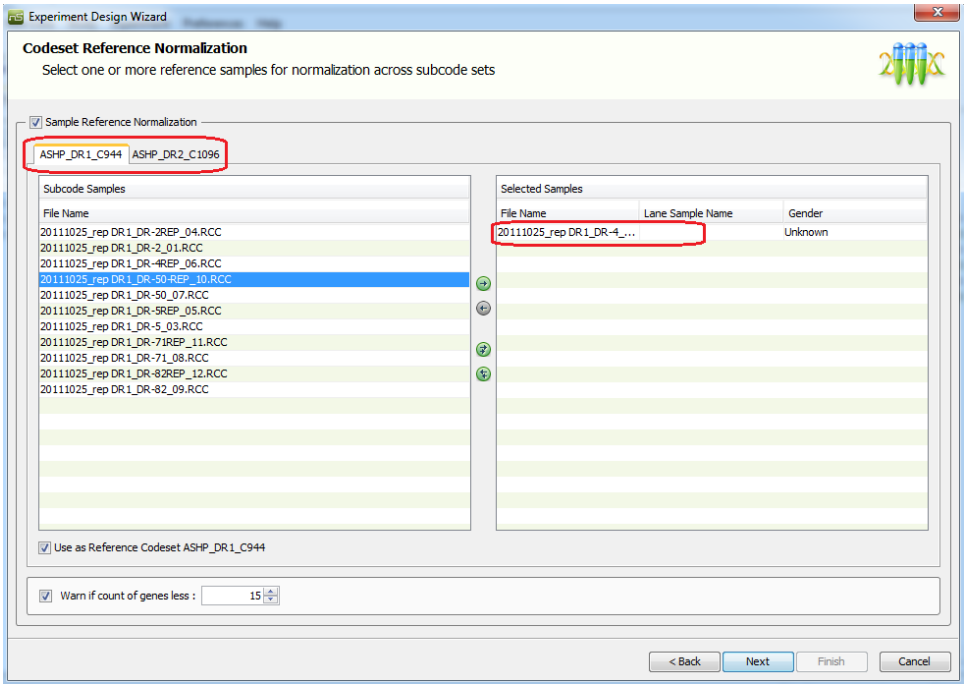


Other Normalization Dialog Box



Inter-CodeSet Calibration (Cross RLF and Plex^2 Only)

Plex^2 and Cross RLF experiments (with the exception of CodeSet Plus experiments) require repeat testing of at least one sample (calibration sample(s)) with all CodeSets. The calibration sample(s) allows for precise comparisons to be made at the normalized data level between samples tested with different CodeSets. To specify the reference sample, select the file from the left-hand side and move it to the right-hand side. Then select the next tab to specify the same sample for the other CodeSet(s).

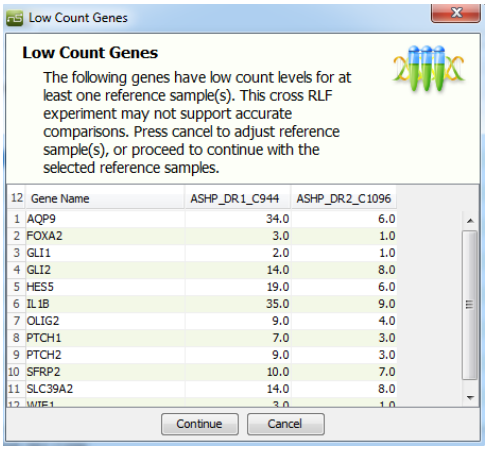


CodeSets are calibrated as follows:

1. Data are normalized separately for each CodeSet as is described in NanoString's nCounter® Expression Data Analysis Guide and nCounter® Data Analysis Guidelines for Copy Number Variation (CNV).
2. Ratios are built against the selected calibration samples in the window above. Calibration samples must be specified for each CodeSet in the screen above. If more than one calibration sample is selected, the counts for each gene will be averaged across all calibration samples.
3. The resulting ratios are multiplied by the normalized counts for the calibration samples in the reference CodeSet. The Reference CodeSet is selected via a checkbox in the bottom left corner of the screen above.
4. The normalized data level for this experiment will now reflect the calibrated counts.

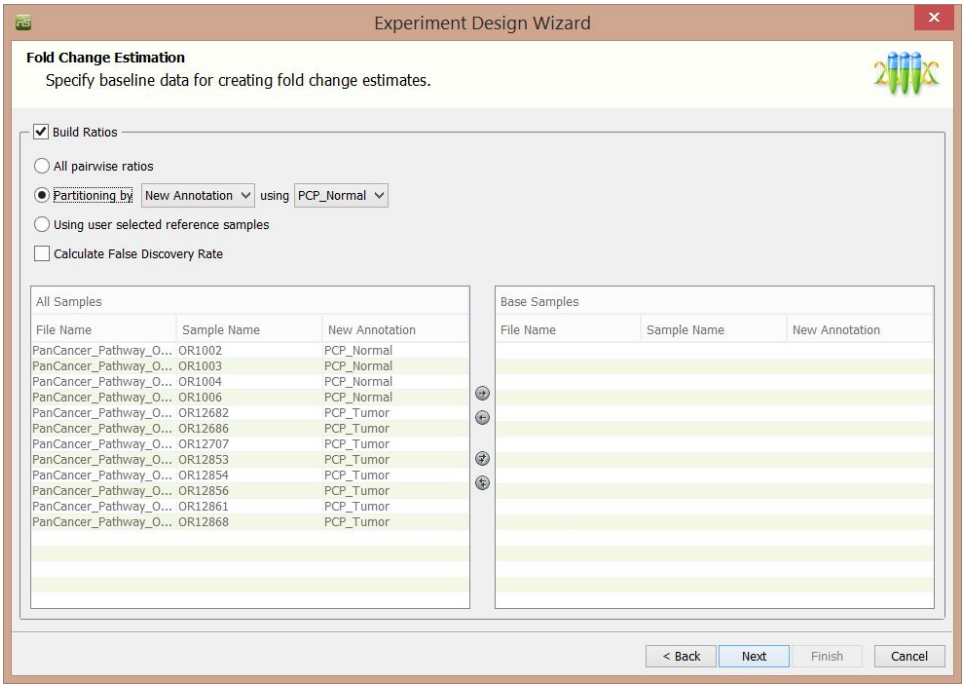
A warning will be displayed showing those genes that have counts below 15 in the reference sample(s) for any CodeSet. If many genes have low counts in the calibration samples for any CodeSet, the quality of the calibration for that probe may be diminished. This warning can be adjusted as desired on the screen above by changing the number at the bottom. Click the **Continue** button to proceed to the next page of the Experiment Wizard or **Cancel** to stay on the same page and adjust the calibration samples to be used for normalization.





Estimating Fold Changes

After selecting the normalization parameters click **Next** to proceed to the Fold Change Estimation Dialog Box.



Fold change is computed by taking the mean of the normalized lanes associated with the same treatment group from the Annotation Parameters screen and then dividing by the baseline specified on this screen. There are three ways to

choose the baseline for computing ratios and fold change estimates.

- **All pairwise ratios**—This method uses each treatment group as a baseline and creates ratios in all possible unique combinations. Note, if more than 200 ratios will be created, this option is disabled.
- **Partitioning by**—This is the most commonly-used method. When using 'Partitioning by,' one of the annotation values is the baseline for all ratios. See the example below for more information on the partitioning by option.
- **User-selected reference lanes**—This method allow you to select any sample lane including those already used in treatment groups, to be used as a baseline in building ratios. Select the samples/lanes from the table on the left hand side and then move the ones that you want to use in the baseline to the right hand side using the arrow button. You can move any lanes/samples into and out of the box using the single arrow buttons. Use the double arrow buttons to move all lanes/samples back and forth. The baseline is the mean of all lanes you select. Using this option results in a total number of experiments equal to the total number of treatment groups (one ratio experiment for each treatment group using the same baseline).

If sample annotations columns were added previously on the Add Sample Annotations page of the Experiment Wizard, and if more than one sample is added to the treatment groups, then a t-test will be automatically run as the experiment is built. If you would also like to calculate the False Discovery Rate, choose that option on this page by selecting the checkbox.

Understanding Baselines

Defining annotation columns earlier in the Experiment Wizard allows you to select baselines based on one of the annotation columns that is part of the design. An example of using this type of baseline is when you have a control as part of your design. For example, if you are comparing different compounds and have added a control condition, you may wish to create ratios using the control as the baseline.

		Time			
		1 Hour	2 Hours	3 Hours	4 Hours
Treatment	Control	Ctrl @ 1 hr	Ctrl @ 2 hrs	Ctrl @ 3 hrs	Ctrl @ 4 hrs
	Compound A	A @ 1 hr	A @ 2 hrs	A @ 3 hrs	A @ 4 hrs
	Compound B	B @ 1 hr	B @ 2 hrs	B @ 3 hrs	B @ 4 hrs

Selecting the level **Control** as the partitioning baseline builds the following ratios. Within each ratio, the treatment with one of two compounds is compared to the Control, but the time of treatment is the same:

- Compound A @ 1 hour / Control @ 1 hour
- Compound B @ 1 hour / Control @ 1 hour
- Compound A @ 2 hours / Control @ 2 hours
- Compound B @ 2 hours / Control @ 2 hours
- Compound A @ 3 hours / Control @ 3 hours
- Compound B @ 3 hours / Control @ 3 hours
- Compound A @ 4 hours / Control @ 4 hours
- Compound B @ 4 hours / Control @ 4 hours

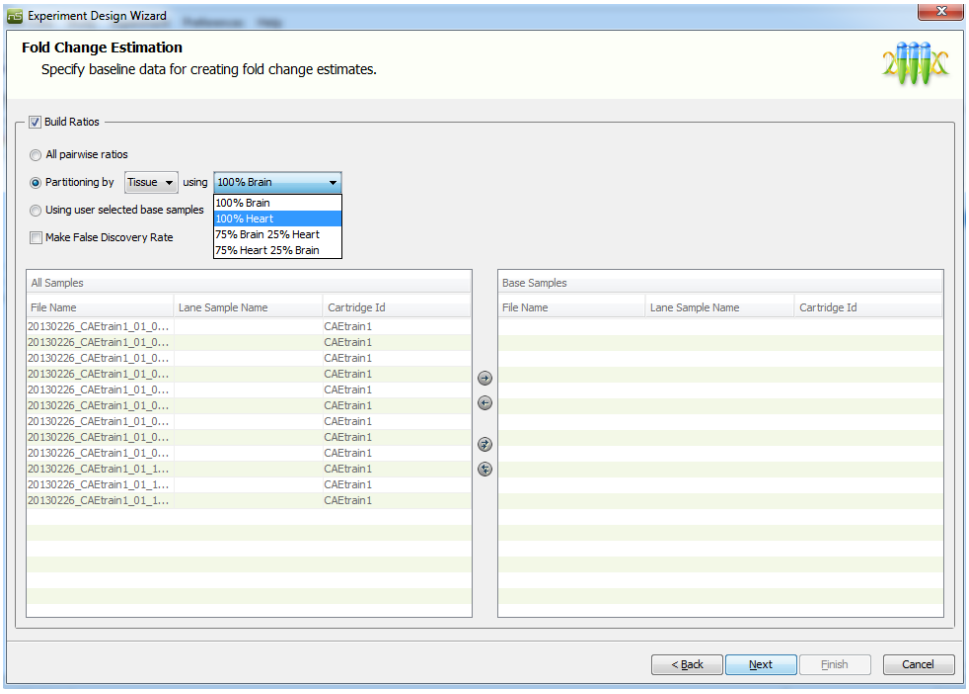
Choosing a Baseline

Ratios simply represent the division of one Grouped Data item divided by another. Where A/B is the ratio, the baseline B is defined in the Fold Change parameters screen of the experimental design wizard as described above.

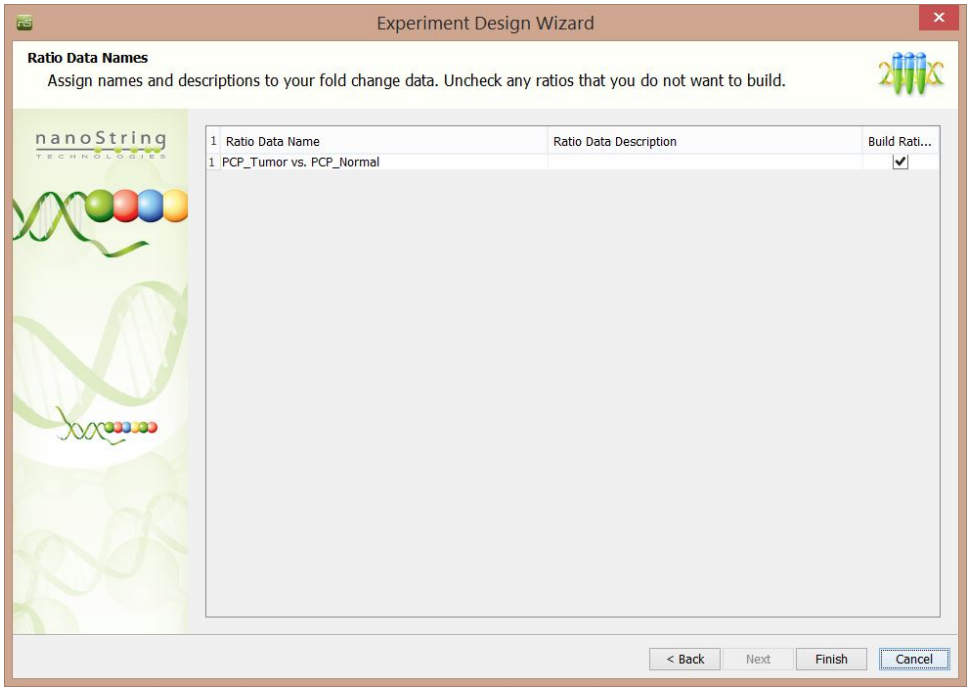
By default, the data are displayed by default in terms of fold change for GX, miRNA, miRGE, Plex² GX and Other. Fold change is equal to A/B when that ratio is greater than 1. When A/B is less than 1, then the fold change displays as the negative reciprocal [-1/(A/B) or –B/A]. You can display and export data as Fold Change, Ratio, or Log_2 Ratio.

NOTE: A $\log_2$ ratio of 1 is equivalent to a fold change of 2, and a $\log_2$ ratio of -1 is equivalent to a fold change of -2.

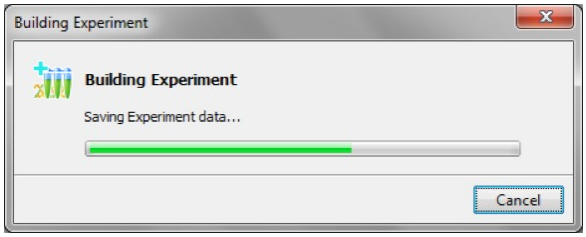
Select the appropriate baseline and click **Next** to proceed to the Ratio Data Names page.



Unselect any ratios that you do not want to create by clicking on the checkbox under the Build Ratio Data? column. In addition, you can add a more descriptive name to any Ratios as desired.

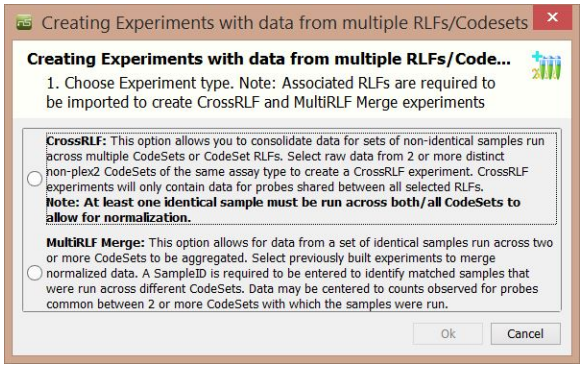


Click **Finish** to complete setting up your experimental design. A progress bar appears. For very large datasets, this may take a few minutes.



Creating a MultiRLF Experiment

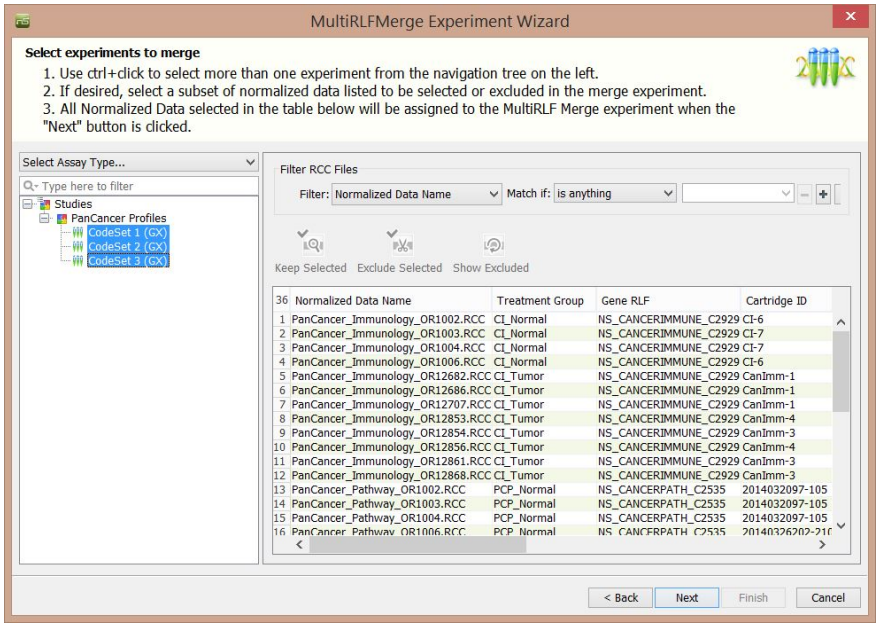
A MultiRLF Merge uses previously-built experiments and allows you to merge and compare data from a set of identical samples across multiple CodeSets. MultiRLF Merge allows normalized data for identical samples to be scaled based on the expression of overlapping probes between CodeSets and then combine the scaled data across all CodeSets for the samples.



Use the radio button to select MultiRLF and click **OK**. You will see the Experiment Properties dialog box. Enter a unique experiment name and choose the study associated with your experiment. Owner, protocol, and descriptions are optional but recommended. Click **Next** to proceed to the next screen of the wizard.

Selecting Experiments to Merge in MultiRLF Experiments

When creating a new MultiRLF Merge experiment, you must select the experiments you want to merge. After you select MultiRLF Experiment, the Select Experiments to Merge dialog box will appear. Choose the experiments you want to merge. There are several ways to search for the data you are looking for and select it for use in your experiment.



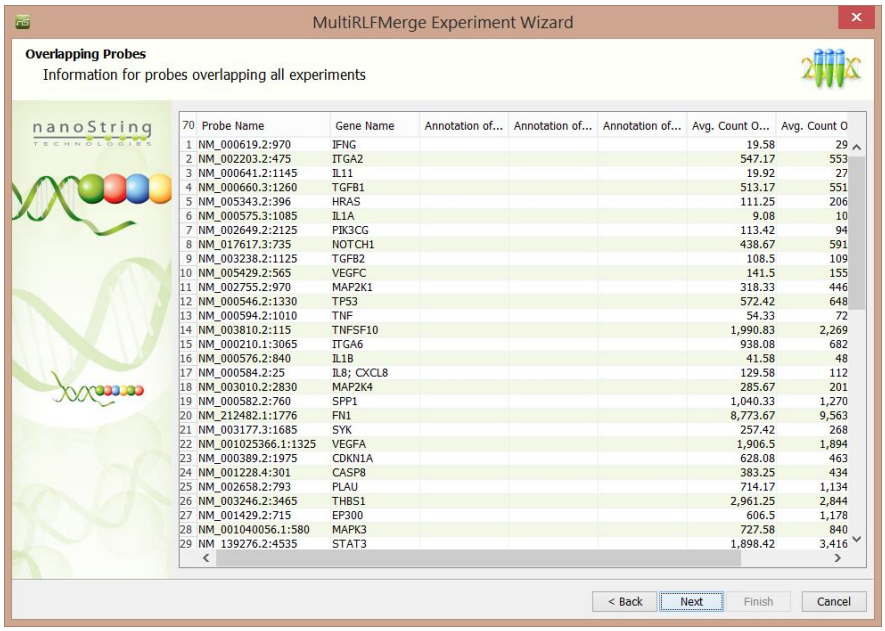
To create a MultiRLF Merge experiment using normalized data from previously created experiments in nSolver all desired experiment files from the navigation tree on the left need to be selected using Ctrl-click (or Command-click for Macintosh uses). When the additional experiments are chosen, the table on the right will expand accordingly.

NOTE: All RLF files used for the experiment must be imported in order to proceed to the Overlapping Probes dialog

box.

Overlapping Probes

The Overlapping Probes dialog box displays information about the overlapping genes across selected RLFs and experiments respectively.

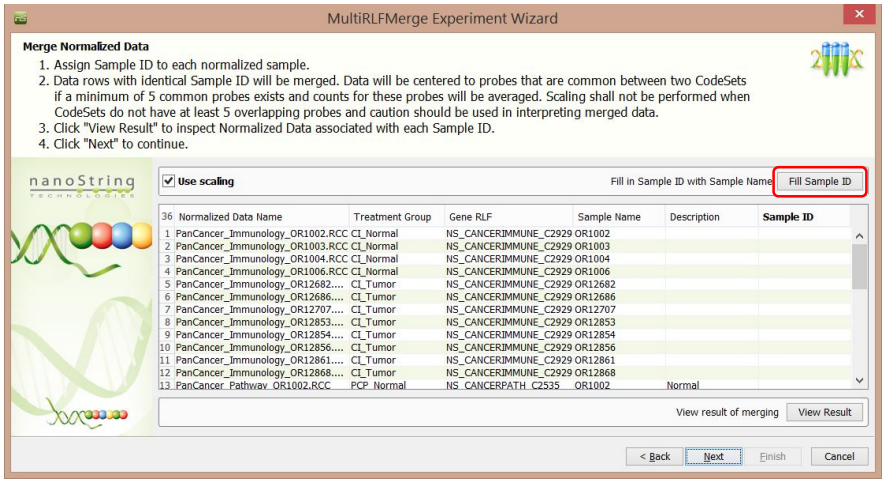


	Probe Name	Gene Name	Annotation of...	Annotation of...	Annotation of...	Avg. Count O...	Avg. Count O
1	NM_000619.2:970	IFNG				19.58	29
2	NM_002203.2:475	ITGA2				547.17	553
3	NM_000641.2:1145	IL11				19.92	27
4	NM_000660.3:1260	TGFB1				513.17	551
5	NM_005343.2:396	HRAS				111.25	206
6	NM_000575.3:1085	IL1A				9.08	10
7	NM_002649.2:2125	PIK3CG				113.42	94
8	NM_017617.3:735	NOTCH1				438.67	591
9	NM_003238.2:1125	TGFB2				108.5	109
10	NM_005429.2:565	VEGFC				141.5	155
11	NM_002755.2:970	MAP2K1				318.33	446
12	NM_000546.2:1330	TP53				572.42	648
13	NM_000594.2:1010	TNF				54.33	72
14	NM_003810.2:115	TNFSF10				1,990.83	2,269
15	NM_000210.1:3065	ITGA6				938.08	682
16	NM_000576.2:840	IL1B				41.58	48
17	NM_000584.2:25	IL8; CXCL8				129.58	112
18	NM_003010.2:2830	MAP2K4				285.67	201
19	NM_000582.2:760	SPP1				1,040.33	1,270
20	NM_212482.1:1776	FN1				8,773.67	9,563
21	NM_003177.3:1685	SYK				257.42	268
22	NM_001025366.1:1325	VEGFA				1,906.5	1,894
23	NM_000389.2:1975	CDKN1A				628.08	463
24	NM_001228.4:301	CASP8				383.25	434
25	NM_002658.2:793	PLAU				714.17	1,134
26	NM_003246.2:3465	THBS1				2,961.25	2,844
27	NM_001429.2:715	EP300				606.5	1,178
28	NM_001040056.1:580	MAPK3				727.58	840
29	NM_139276.2:4535	STAT3				1,898.42	3,416

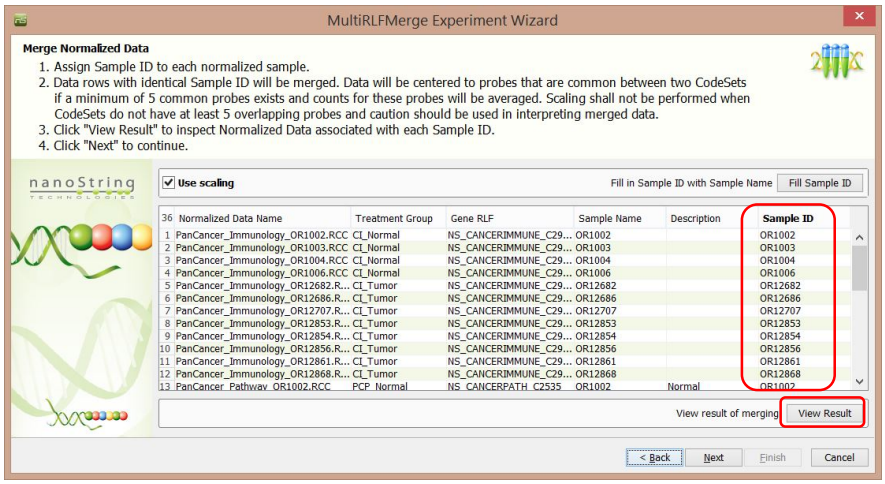
Merging Normalized Data for MultiRLF Experiments

For MultiRLF Merge experiments, the next step is to merge the normalized data. Normalized data of samples run on multiple CodeSets that are to be merged are recognized by their Sample ID. Sample ID have the following characteristics.

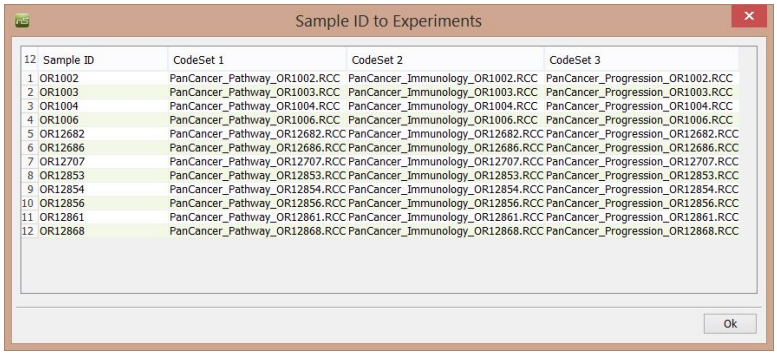
- Each sample lane from a CodeSet is identified by a unique Sample ID
- Identical Sample IDs are used to recognize samples that were run on 2 or more experiments from different CodeSets.
- Sample ID can be entered by entering text manually, pasting information copied from a spreadsheet or by clicking ‘Fill Sample ID’ if to automatically parse and fill the Sample ID from the data files.
- Sample IDs associated with Normalized data file name can be previewed by clicking the View Result button.



Click the **Fill Sample ID** button on the right-hand side above the data table to populate Sample ID column from the data files.



To view the result of the merging before proceeding, click the **View Result** button to bring up the Sample ID to Experiments dialog box.

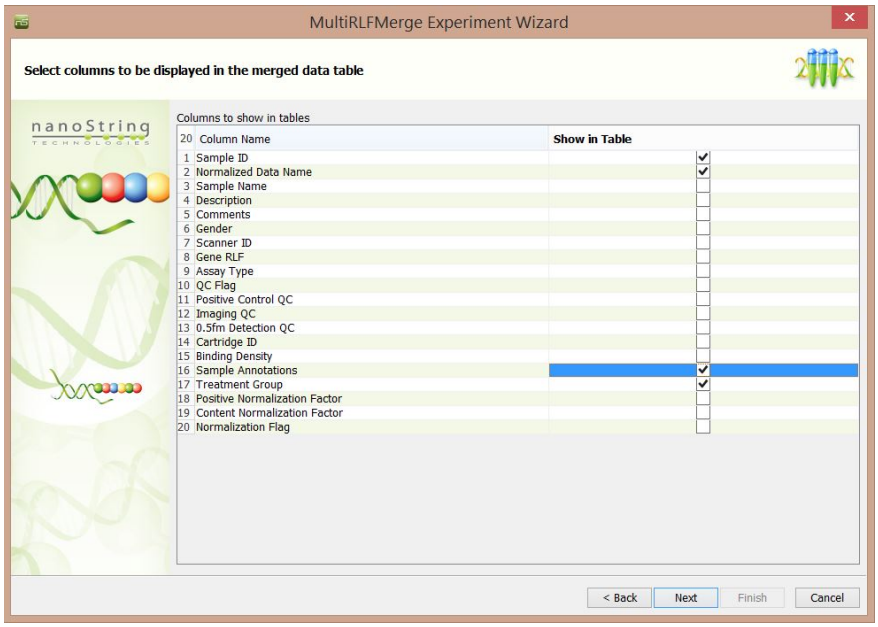


When you are finished reviewing the results, click **OK** to return to the Merge Normalized Data dialog box. Click **Next** to continue.

Selecting Columns for the Merged Data Table

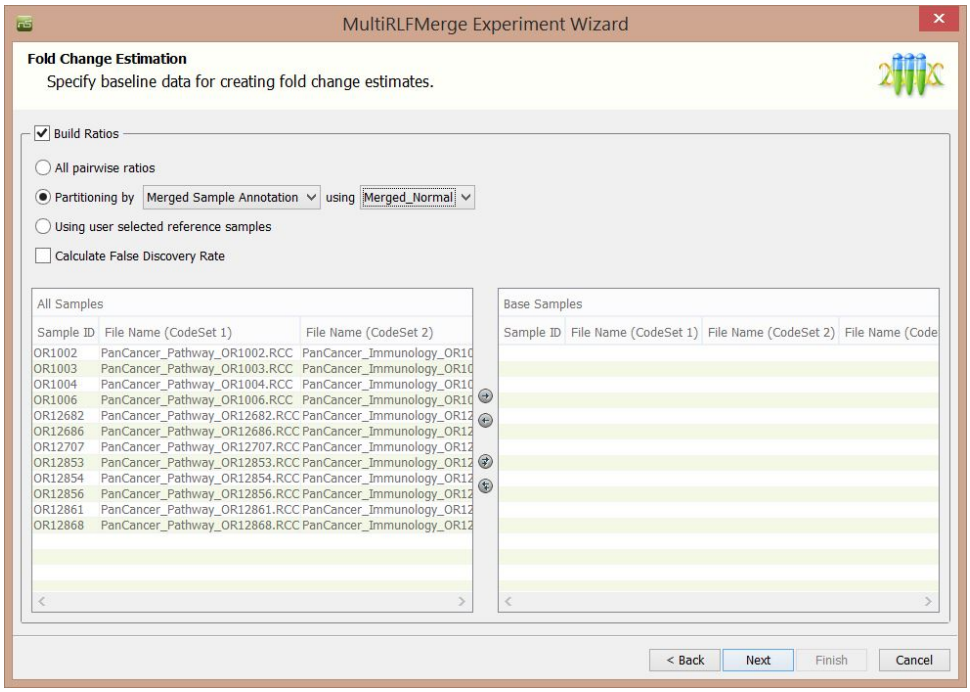
The next step for creating a MultiRLF Merge experiment is to select the columns that you want to display on the merged

data table. Use the checkboxes to select the columns and click **Next**.



Estimating Fold Changes

After selecting the normalization parameters click **Next** to proceed to the Fold Change Estimation Dialog Box.



Fold change is computed from the ratio of the geometric mean of the normalized lanes associated with the same treatment group from the Annotation Parameters screen to the geometric mean of the baseline sample lanes or treatment group specified on this screen. There are three ways to choose the baseline for computing ratios and fold

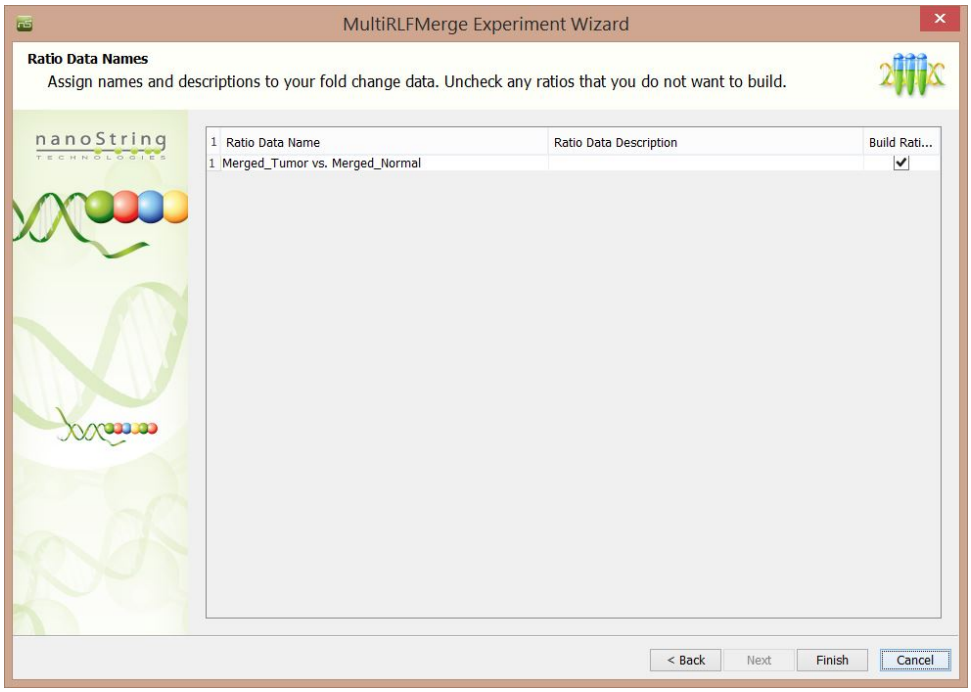
change estimates.

- **All pairwise ratios**—This method uses each treatment group as a baseline and creates ratios in all possible unique combinations. Note, if more than 200 ratios will be created, this option is disabled.
- **Partitioning by**—This is the most commonly-used method. When using **Partitioning by**, sample lanes belonging to an annotation group are the baseline for all ratios.
- **User selected reference samples**—This method allows you to select any sample, including those already used in treatment groups, to be used as a baseline in building ratios. Select the sample/s from the table on the left hand side and then move the ones that you want to use in the baseline to the right hand side using the arrow button. You can move any sample into and out of the box using the single arrow buttons. Use the double arrow buttons to move all lanes/samples back and forth. The baseline is the mean of all lanes you select. Using this option results in a total number of experiments equal to the total number of treatment groups (one ratio experiment for each treatment group using the same baseline).

If sample annotations columns were added previously on the Add Sample Annotations page of the Experiment Wizard, and if more than one sample is added to the treatment groups, then a t-test will be automatically run as the experiment is built. If you would also like to calculate the False Discovery Rate, choose that option on this page by selecting the checkbox.

When you have finished your selections, click Next.

Use the Ratio Data Names dialog box to assign names and descriptions to your Fold Change data. You can unselect any ratios that you do not want to create by clicking on the checkbox under the Build Ratio Data column. In addition, you can add a more descriptive name to any Ratios as desired.

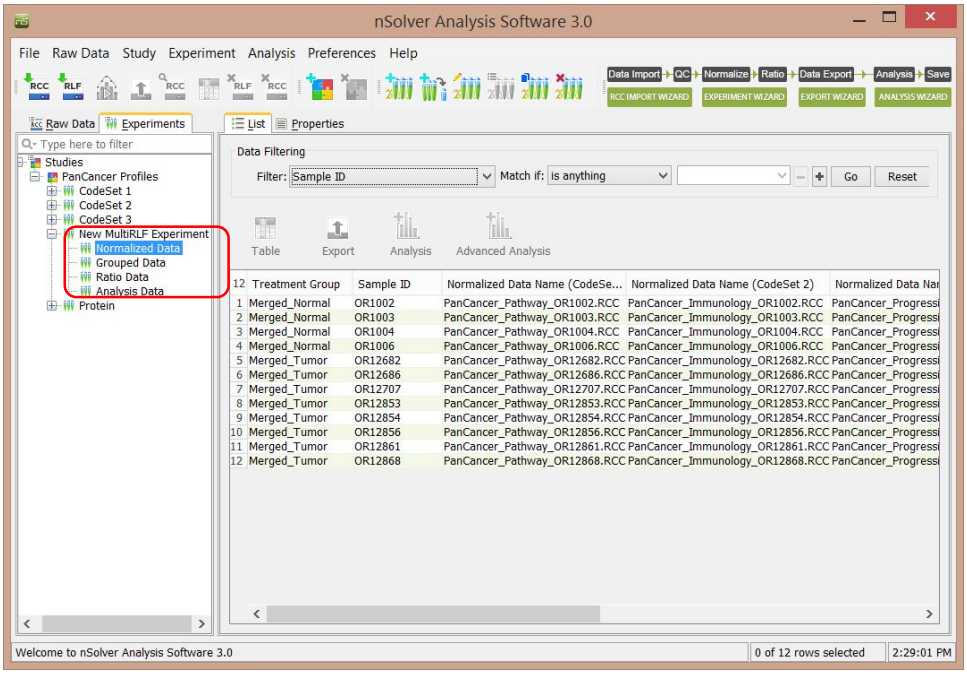


Click **Finish** to complete setting up your experimental design. A progress bar shows the status of the merge. For very large datasets, this may take a few minutes.

A MultiRLF Merge experiment consists of the following tables:

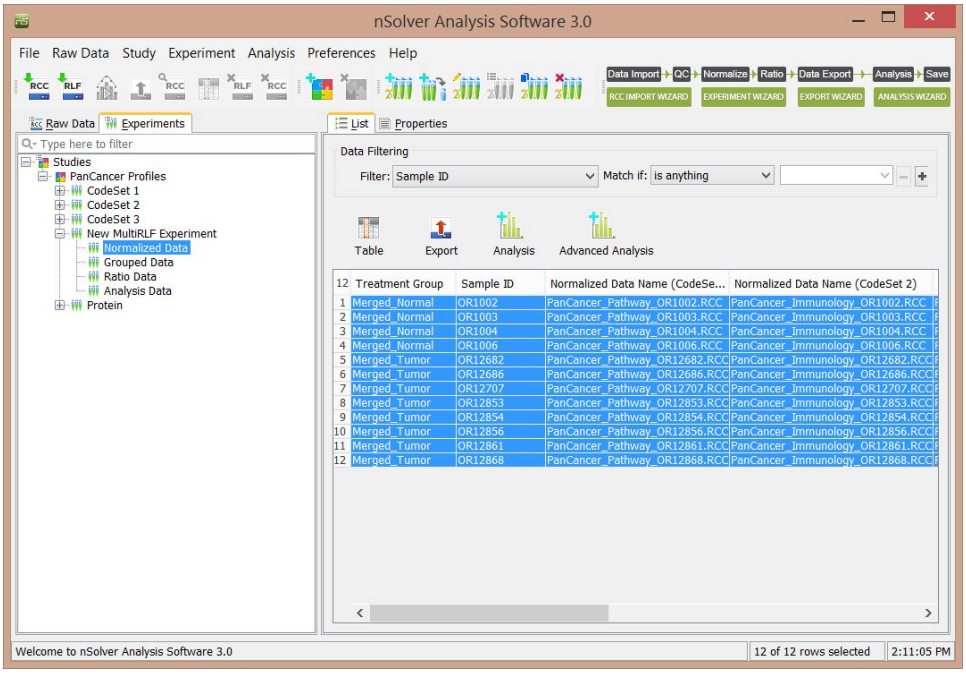
- Normalized data
- Grouped Data
- Ratio Data
- Analysis Data

It is important to note that MultiRLF Merge experiments do not contain a table for Raw Data, since they are created with normalized data from previously created experiments.



Select the desired data using the mouse or the Data Filtering options. When data is selected, the icons above the data table are active. Use them to:

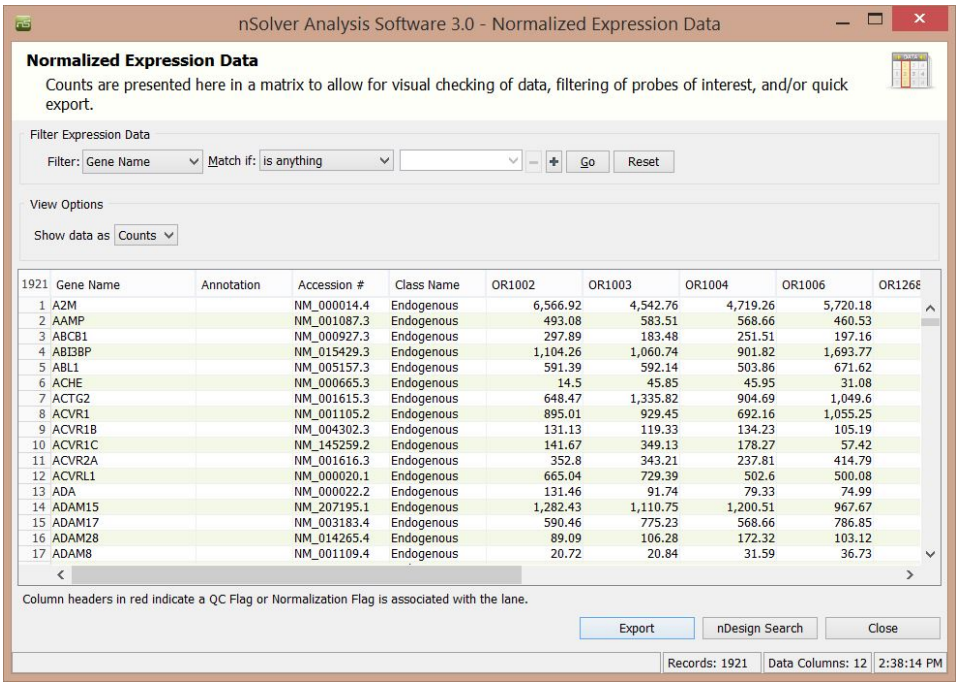
- View a table of the data, which you can export or do an nDesign search
- Export the file
- Analyze the data
- Use the Advanced Analysis module to analyze the data

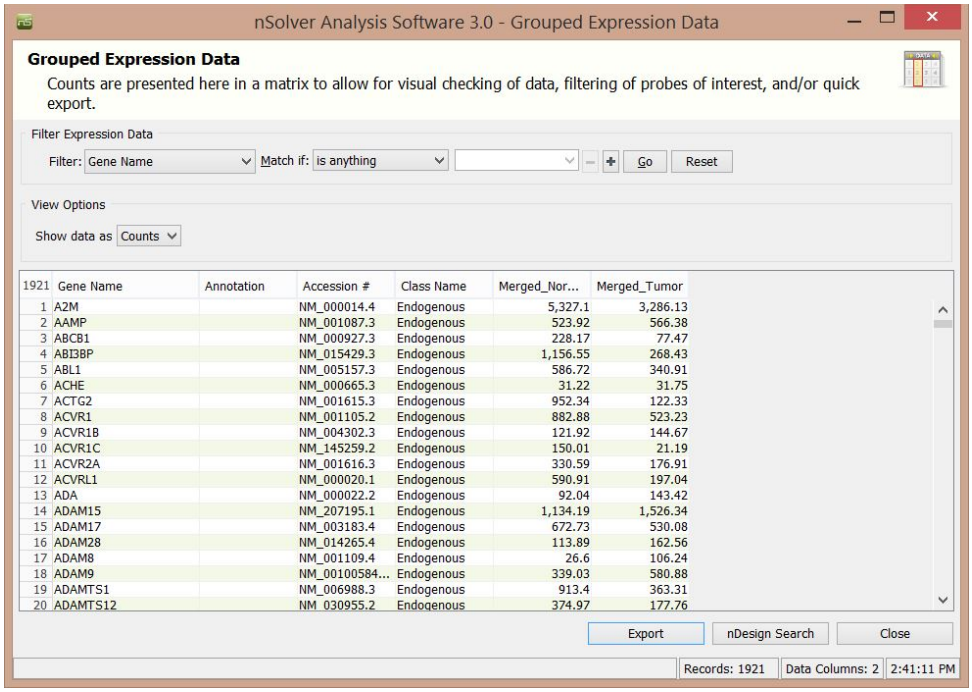


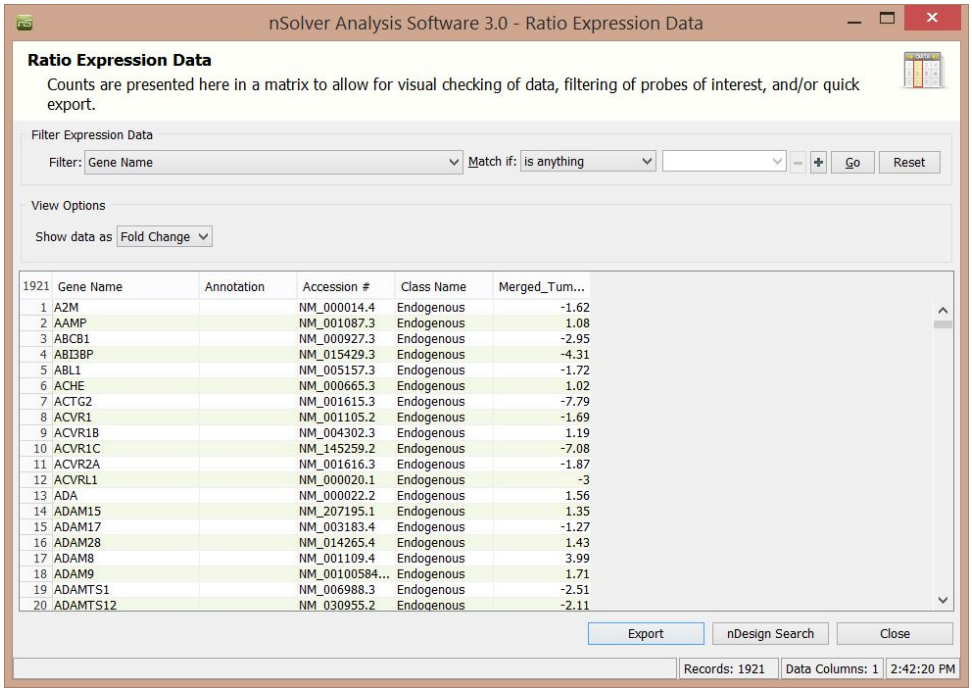
Select the Table icon to view the data tables. The data tables for a MultiRLF Merge experiment are displayed by Sample ID and contain data for all the probes from the selected RLFs for the CodeSets that the samples were run across, excluding the Positive and Negative ERCC control probes. The tables also include the probe annotations assigned to each probe in the context of the original CodeSet/RLF.

The Normalized, Grouped, and Ratio data tables are shown in the following figures.

The Normalized data table displays all Sample ID and any other columns preferentially selected in the MultiRLF Experiment Wizard in the “Select columns to be displayed in the merged data table” screen.

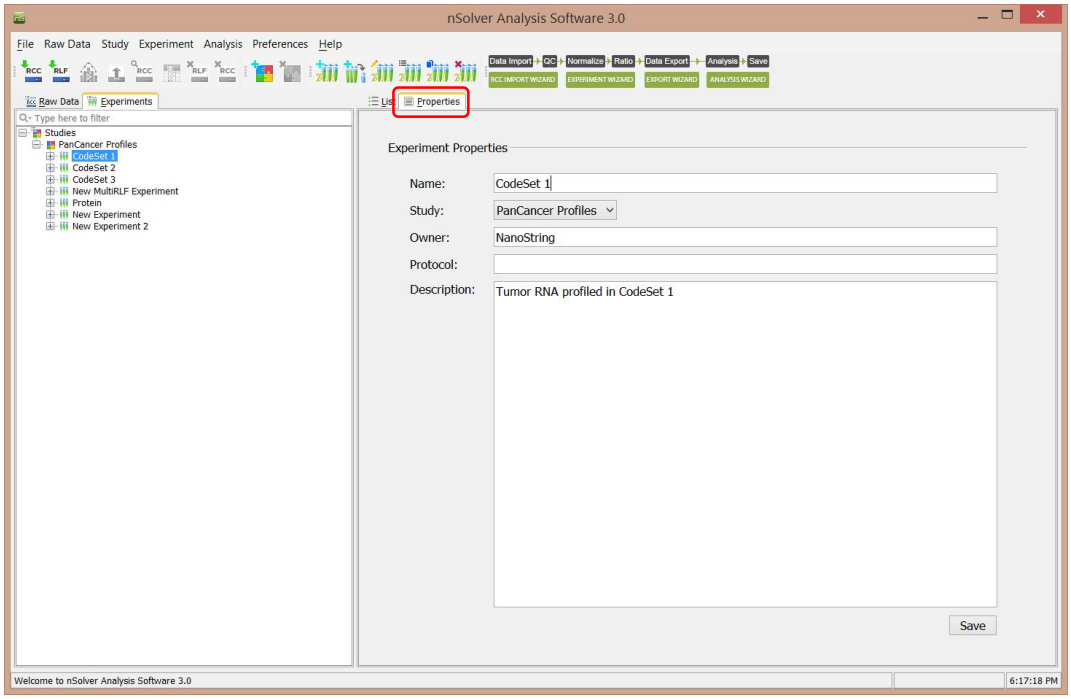






Editing an Experiment

An experiment can be edited in two different ways. If you only need to change the Experimental Properties, then select the experiment name from the navigation tree on the left side and click on the **Properties** tab.

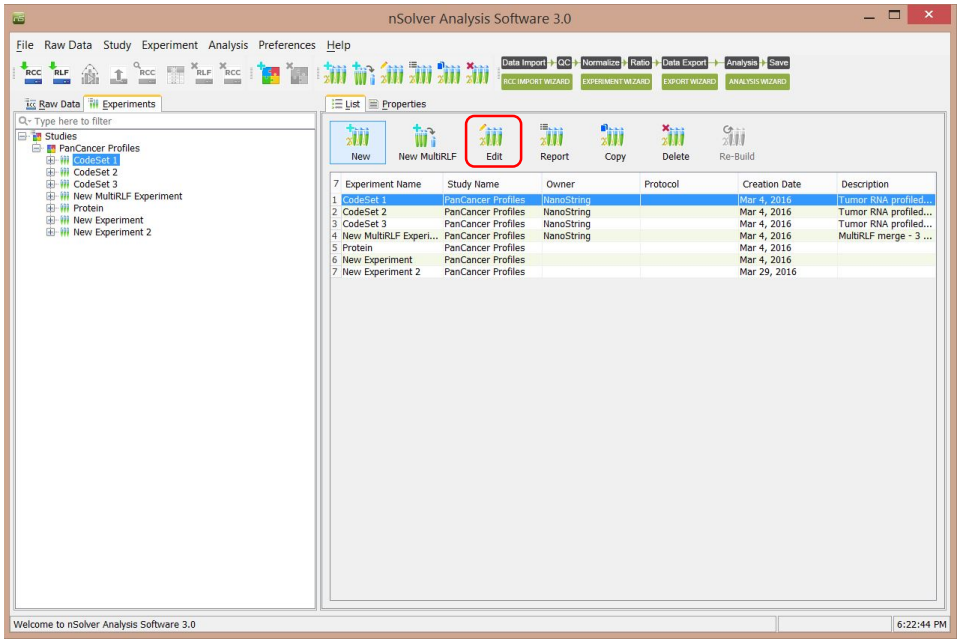


Update the Name, Study, Owner, Protocol and Description and click **Save**.

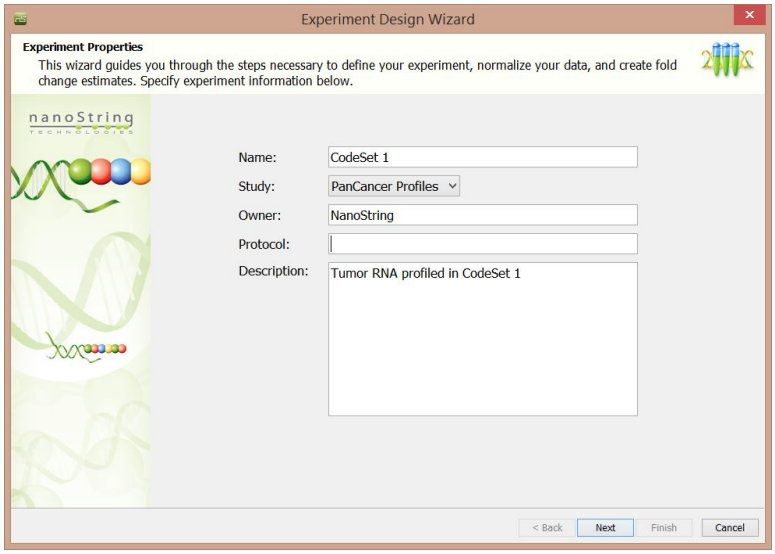
The second way to edit an experiment is to use the **Edit** icon (see next page). This allows you to open up the experimental wizard and change any parameters or values that were entered when the experiment was first set up.

NOTE: All results will be updated when you click Finish to complete the wizard and save your changes, and the previously saved data will be replaced with the new data. Also, all analysis data from the previous experiment will be discarded.

To edit an experiment, select the experiment from the navigation tree on the left side and click the **Edit** icon

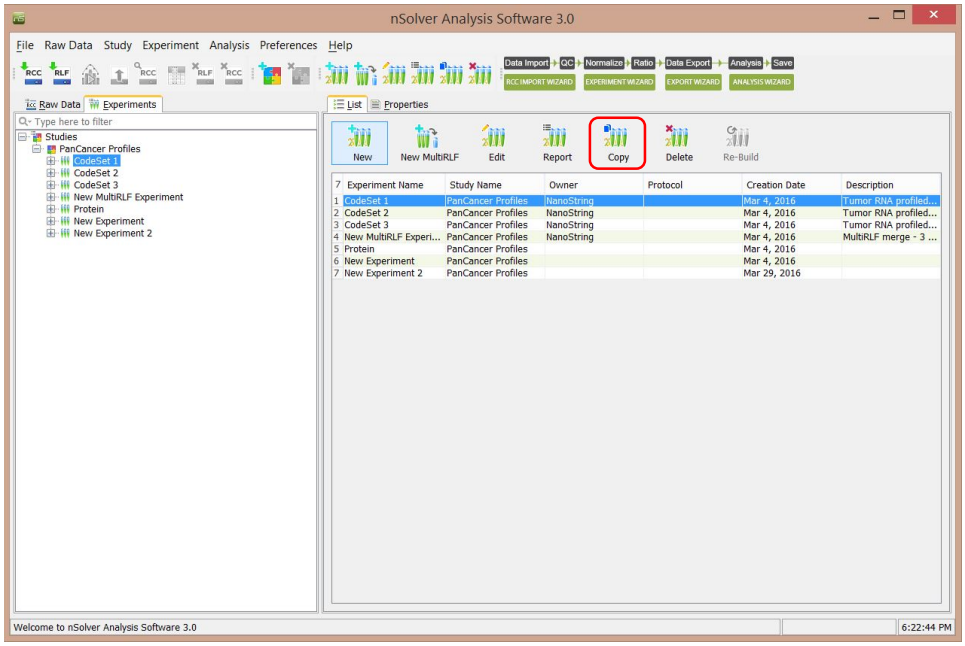


The Experiment Wizard will open. Click **Next** and update data on each screen as desired. Click **Finish** when complete to re-build your experiment.

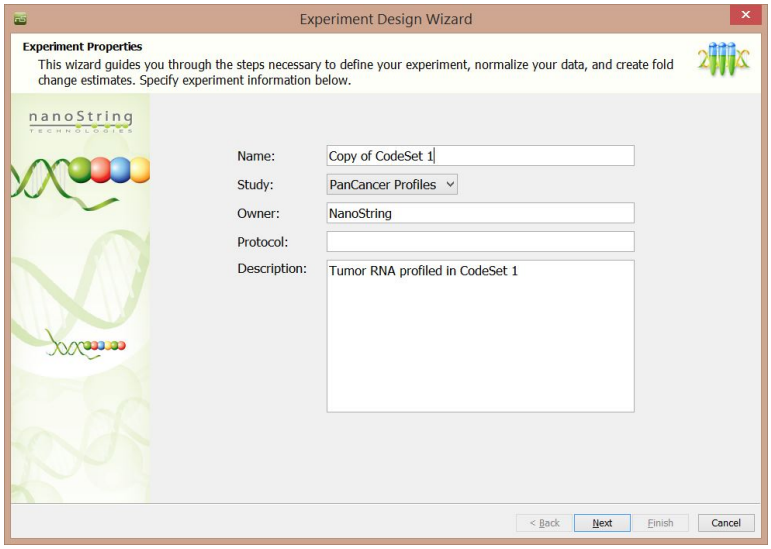


Experiments may be copied to quickly and easily create an experiment similar to one that already exists. Modifications can be made on all pages of the Experiment Wizard. This is also a simple way to compare two different methods of analysis or normalization without deleting any results.

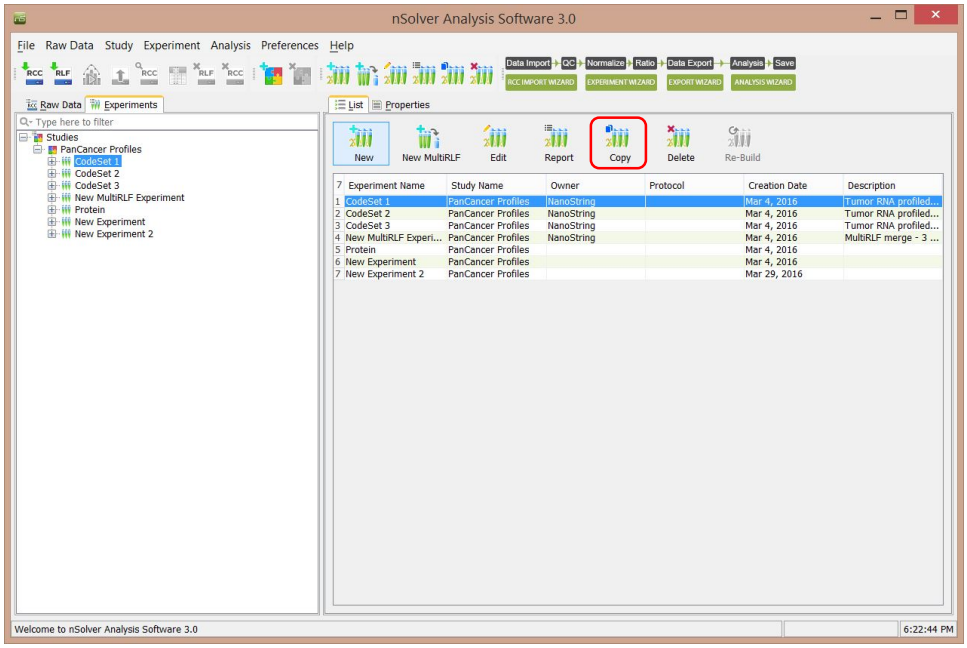
To create a copy of an experiment, select the experiment to be copied from the navigation tree and click on the **Copy** icon.



The Experiment Wizard opens and all data from the original experiment are contained. Step through the wizard and make changes where desired. Click **Finish** to build the new experimental design.



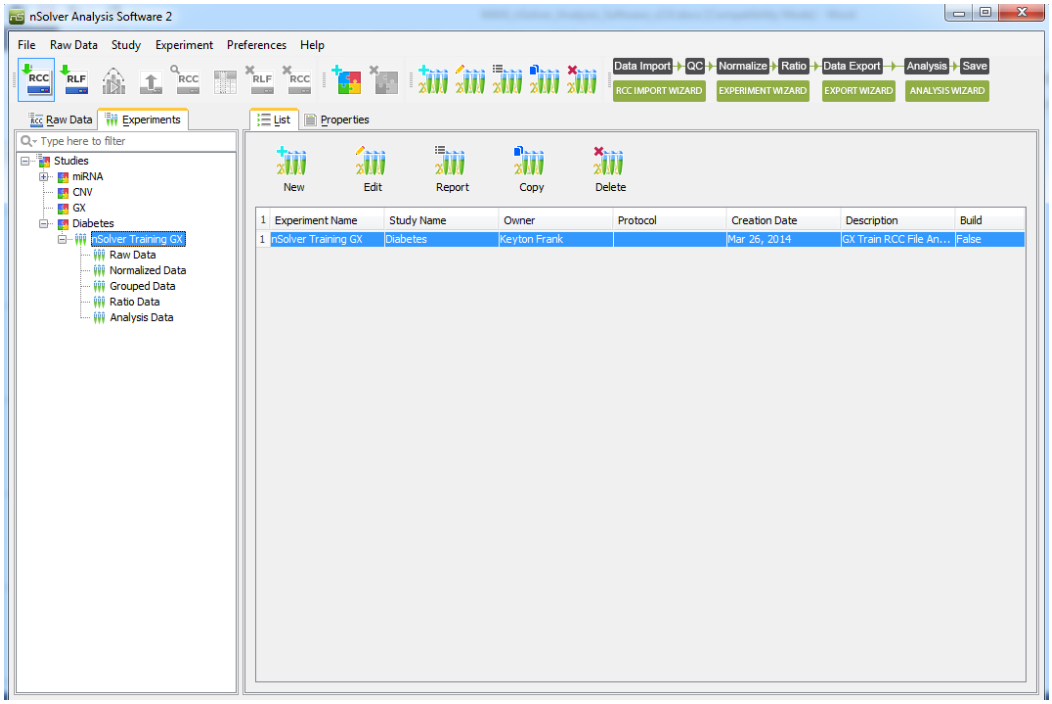
Experiments that are no longer wanted can be completely deleted from the system. To delete an experiment, select the experiment name from the navigation tree on the left side and click on the **Delete** icon. You will need to confirm that your experiment and all analyzed data contained within it should be deleted from the system.



9 Experiment Output—Data Tables and Reports

Viewing Analyzed Data

After completing your experiment, the experiment name appears under the appropriate study in the navigation tree on the left side. Click on the '**+**' button next to the experiment name to expand and view different levels of data that are now available as a result of 'building' the experiment.



nanoString TECHNOLOGIES

After clicking the '**+**' button, the following 4 levels of data will be displayed:

- **Raw Data**—the raw data counts that are collected on the Digital Analyzer.

NOTE: For miRNA and miRGE assays the background subtraction denoted with '+++' in the RCC and RLF files is automatically performed during import and displayed as raw data in the raw data table.

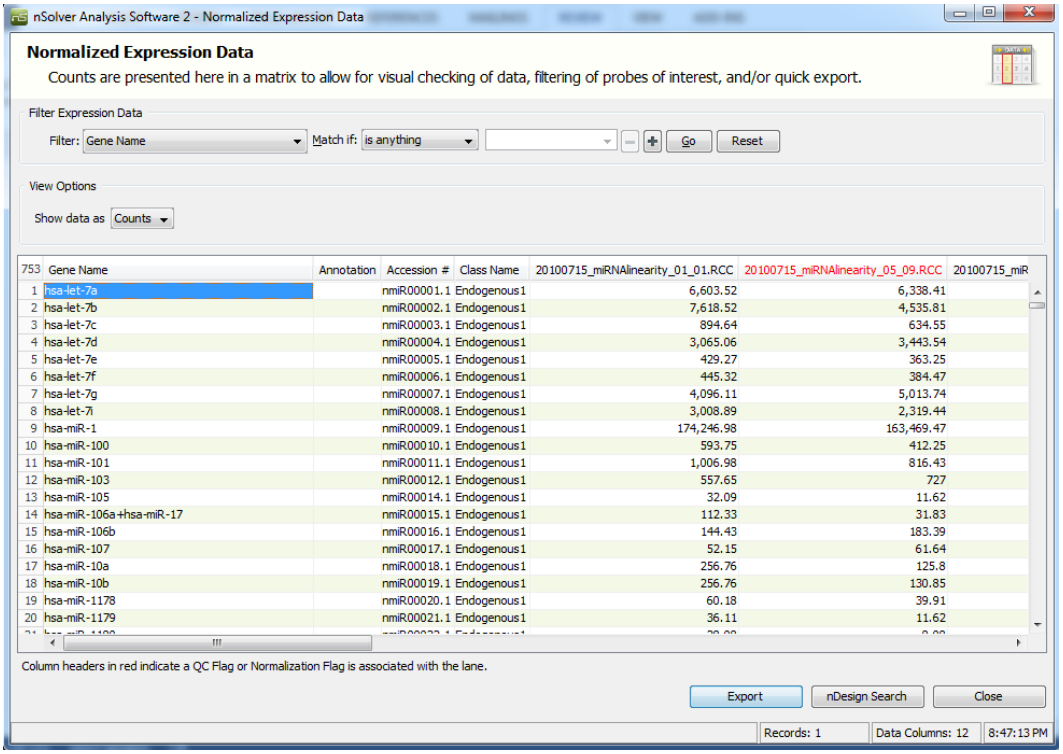
- **Normalized Data**—each lane/sample after it has been adjusted as previously designated on the normalization parameters page of the Experiment Wizard. For details on actual calculations see NanoString's *nCounter Expression Data Analysis Guide* and *nCounter Data Analysis Guidelines for Copy Number Variation (CNV)*.
- **Grouped Data**—the geometric mean of the lanes/samples that were assigned to the same treatment group (e.g., geometric mean of all lanes/samples assigned to Compound A @ Time 1, etc.). If no annotations are added in the Experiment Wizard, the grouped data are the same as the normalized data.
- **Ratio Data**—the division of one Grouped Data item by another. Where A/B is the ratio, the baseline B was defined in the Fold Change parameters page of the Experiment Wizard.
- **Analysis**—the results of any analysis carried out as described in *Chapter 12, Analyses and Visualizations*.

By default, data for GX, miRNA, miRGE, Protein and Other are displayed in terms of fold change. Fold change is the same as A/B when A/B is greater than 1. When A/B is less than 1, then the fold change displays as the negative reciprocal [-1/(A/B) or –B/A]. Data can be displayed and exported in terms of both Fold Change and Ratio formats.

NOTE: For CNV, only ratios and copy number estimates are displayed. For autosomal probes, copy number estimates are the ratio A/B times 2, where B is the baseline as specified in the Experiment Wizard. For samples specified as male, the copy number estimates for X and Y chromosome probes will be the ratio A/B times 1. For samples specified as female, the copy number estimates for X chromosome probes will be the ratio A/B times 2 and no estimates will be displayed for Y chromosome probes. For reference samples of unknown gender, the copy number estimates are not displayed for chromosome probes.

To open any data table, click on the appropriate level of data from the navigation tree on the left side and select one or more rows in the main viewer corresponding to the data that you want to view in a table. Then click **Table**.

Normalized Data Table



NOTE: Column headers in **red** indicate a QC or Normalization Flag is associated with the lane.

NOTE: You can view normalized data in counts, $\log_{10}$, or $\log_2$ format by selecting the desired option as displayed above.

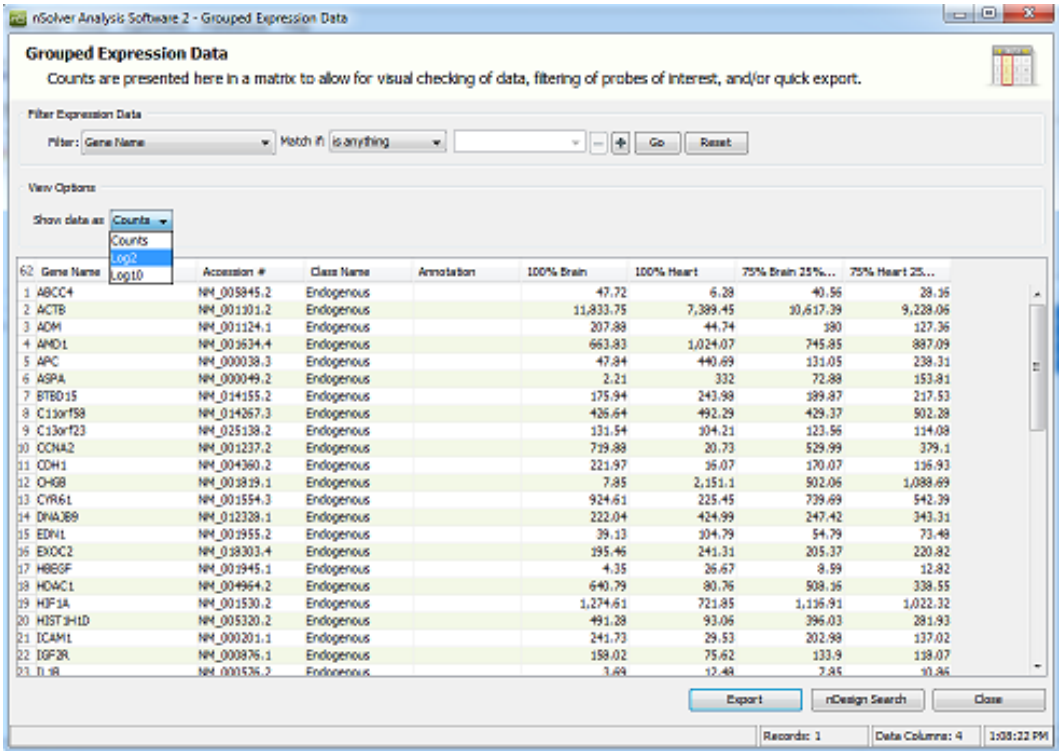
Columns available in the Normalized Data table are as follows:

- Gene Name
- Annotation (user-defined)
- Accession #
- Class NameSpecies Name (hidden by default; RLF import required)
- Probe Name (hidden by default; RLF import required)
- Target Sequence (hidden by default; RLF import required)
- Comments (hidden by default; RLF import required)
- Class Source (hidden by default; RLF import required)
- Positive Flag (hidden by default; RLF import required)
- Class Active (hidden by default; RLF import required)
- Class Preparer (hidden by default; RLF import required)
- Class Date (hidden by default; RLF import required)
- Avg Count (hidden by default)—Can be used to obtain the average count for the gene across the samples (very

low counts may indicate lower confidence in expression levels observed)

- Min Count (hidden by default)—Can be used in conjunction with Avg Count to obtain a sense for the distribution of counts for the gene across the samples
- Max Count (hidden by default)—Can be used in conjunction with Avg Count to obtain a sense for the distribution of counts for the gene across the samples
- % CV (hidden by default)—%Coefficient of Variation—For housekeeping genes, this parameter can be used to assess the effectiveness of the normalization performed or to obtain a sense of the distribution of counts for a particular gene across the samples
- File Name(s)

Grouped Data Table



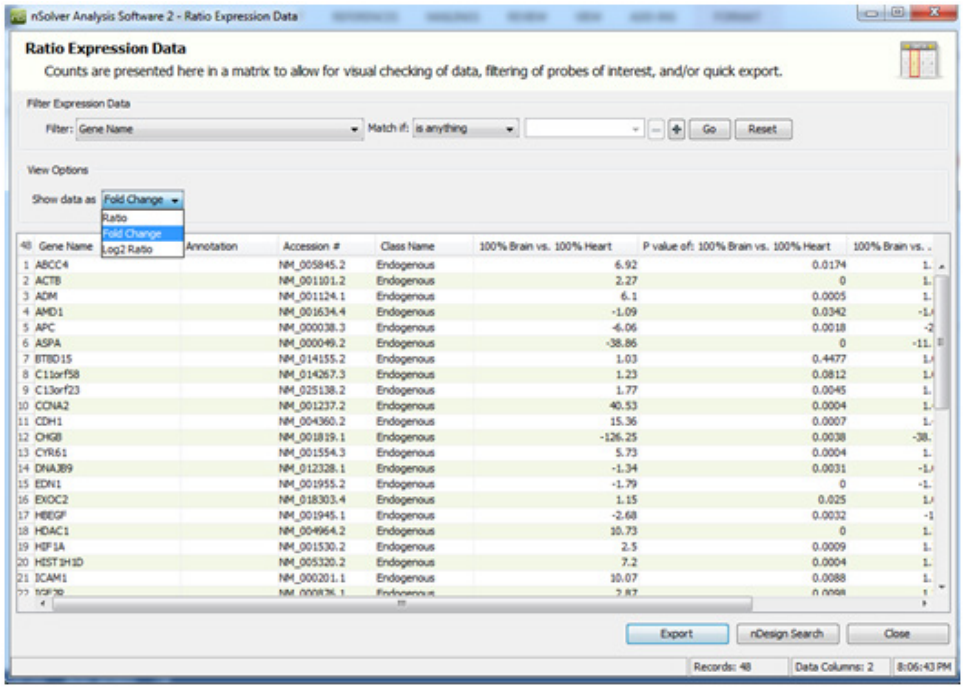
NOTE: You can view normalized data in counts, $\log_{10}$, or $\log_2$ format by selecting the desired option as displayed above.

Columns available in Grouped Data table are as follows:

- Gene Name
- Accession #
- Class Name
- Annotation (user-defined)
- Species Name (hidden by default; RLF import required)
- Probe Name (hidden by default; RLF import required)
- Target Sequence (hidden by default; RLF import required)

- Comments (hidden by default; RLF import required)
- Class Source (hidden by default; RLF import required)
- Positive Flag (hidden by default; RLF import required)
- Class Active (hidden by default; RLF import required)
- Class Preparer (hidden by default; RLF import required)
- Class Date (hidden by default; RLF import required)
- Avg Count (hidden by default) – Used to obtain the average count for the gene across the samples (very low counts may indicate lower confidence in expression levels observed)
- Min Count (hidden by default) – Used in conjunction with Avg Count to obtain a sense of the distribution of counts for a gene across samples
- Max Count (hidden by default) - Used in conjunction with Avg Count to obtain a sense of the distribution of counts for a gene across samples
- % CV (hidden by default) – % Coefficient of Variation - Used to identify the genes that have the most variation and to obtain a sense of the distribution of counts for a gene across samples
- Group Name(s)
- Standard Deviation of each Group Name (hidden by default) – Used to assess the variation of the specified replicates for each Group

Ratio Data Table

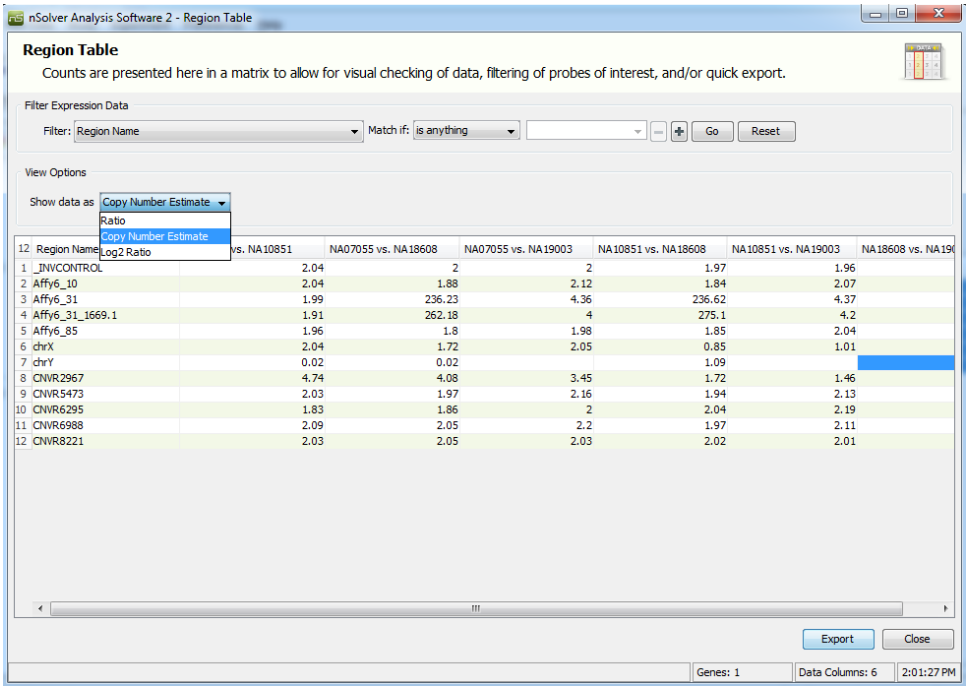


NOTE: You can view ratio data in ratio, fold change, or $\log_2$ ratio format by selecting the desired option as displayed above. Standard deviations for both numerator and denominator are available.

Columns available in Ratio Data table are as follows:

- Gene Name

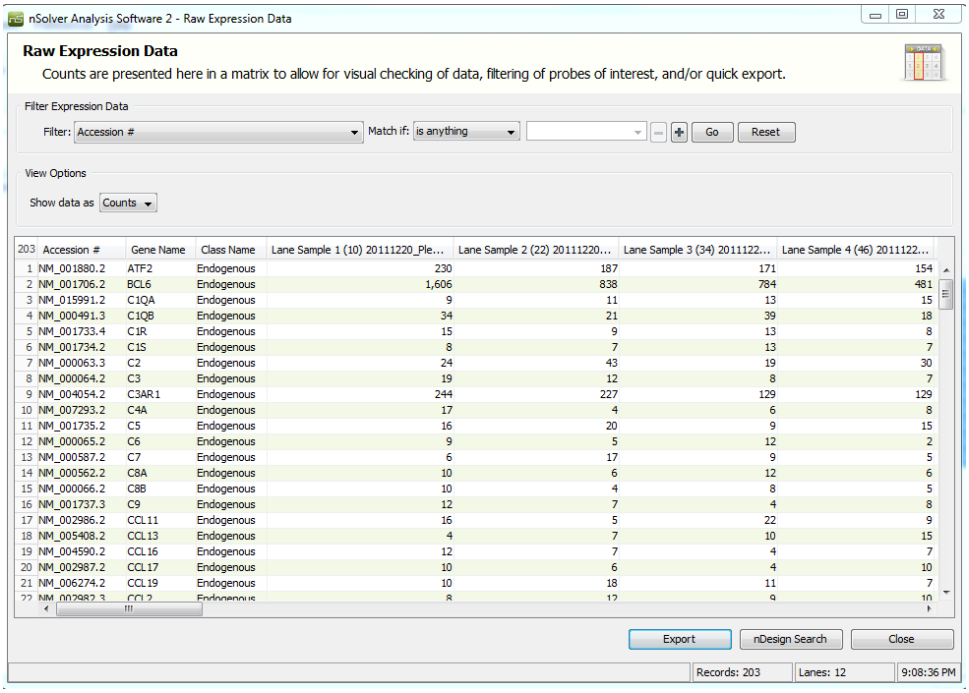
- Annotation (user-defined)
- Accession #
- Class NameSpecies Name (hidden by default; RLF import required)
- Probe Name (hidden by default; RLF import required)
- Target Sequence (hidden by default; RLF import required)
- Comments (hidden by default; RLF import required)
- Class Source (hidden by default; RLF import required)
- Positive Flag (hidden by default; RLF import required)
- Class Active (hidden by default; RLF import required)
- Class Preparer (hidden by default; RLF import required)
- Class Date (hidden by default; RLF import required)
- Avg Count (hidden by default)—Can be used to obtain the average count for the gene across the samples in order to assist in Fold Change significance estimations (higher counts -> significant fold change limit decreases)
- Min Count (hidden by default)—Minimum count for all numerator Groups—Can be used in conjunction with Avg count to assist in estimating Fold Change significance
- Max Count (hidden by default)—Maximum count for all numerator Groups—Can be used in conjunction with Avg count to assist in estimating Fold Change significance
- % CV (hidden by default)—% Coefficient of Variation—Can be used to identify the genes that have the most variation
- Numerator Name(s)—Group Name of Numerator for ratio
- Standard Deviation of each Numerator Name (hidden by default)—Can be used to assess the variation of the specified replicates for each Group
- Denominator Name—Group Name of Denominator for ratio
- Standard Deviation of each Denominator Name (hidden by default)—Can be used to assess the variation of the specified replicates for each Group
- Ratio Name(s)
- Differential Expression Call for each ratio—DE Call can be used to determine whether a probe differs between two samples. It can be used when no replicates exist. See the *Significance Testing* section below for additional information.
- P-value for each ratio—-p-value can be used to determine the evidence for differential expression between two groups and is a measure of the probability that the observed difference in expression can occur by chance alone
- 95% Confidence interval is comprised of the upper 95% confidence limit and lower 95% confidence limit .This is a range of values that has a 95% probability of containing the true difference in the mean expression values between the selected sample annotation groups.
- T-statistic quantifies the differential expression values between the groups relative to the degree of variability of the samples.
- FDR for each ratio—FDR can also be used to determine the confidence of differential expression between two groups. It is more conservative than the t-test alone. Similar to the t-test, it requires replicates. See the *Significance Testing* section below for additional information.



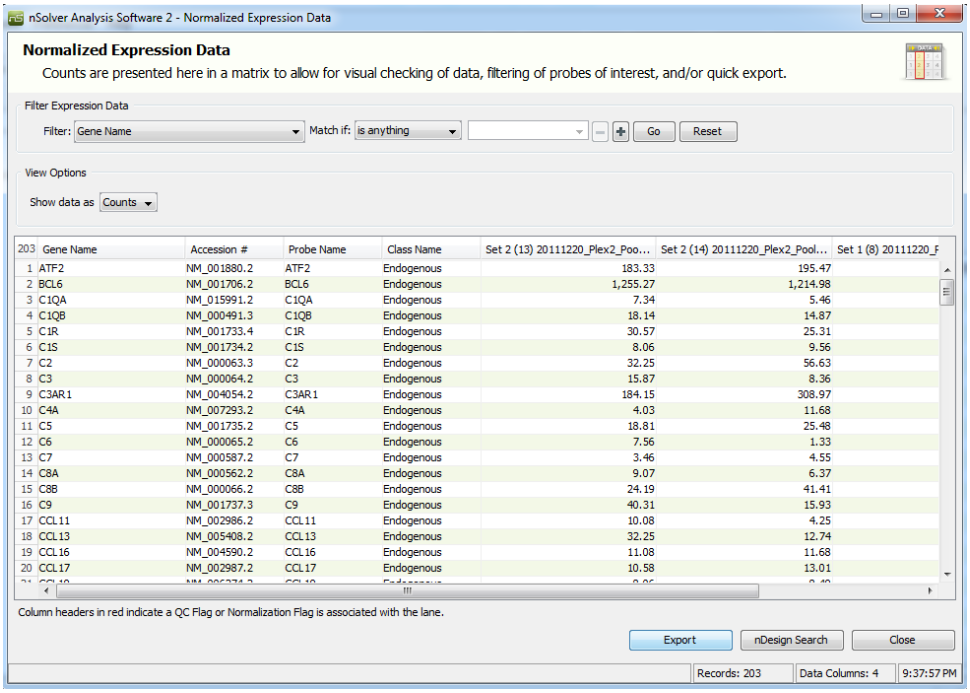
NOTE: For CNV, both probe and region tables/exports are available. For the region table, the mean of all probes for the same region is calculated.

Plex^2 Assay Data Tables

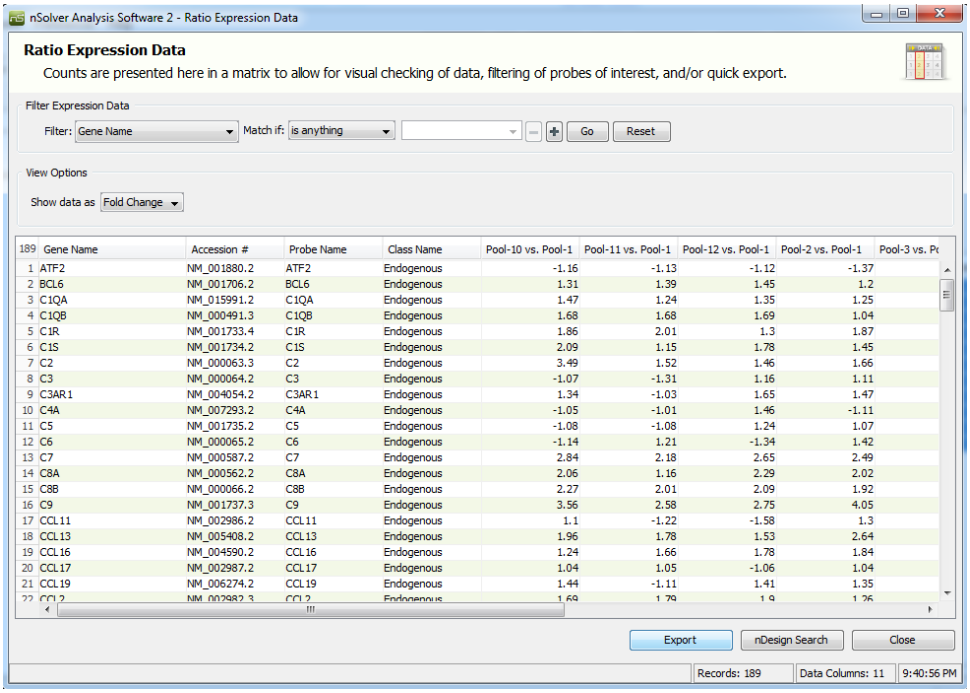
Due to the nature of the Plex^2 Expression Assay RCC files, each file is broken out into multiple columns, one for each sample (or sub-CodeSet). When building an experiment, the calibration sample is used to normalize the differences between the sub-CodeSets. This allows for the normalized and grouped data to look the same as other assay types and for ratios to be created similarly. The figures below show examples of the raw, normalized, and ratio data tables.



203	Accession #	Gene Name	Class Name	Lane Sample 1 (10) 20111220_Ple...	Lane Sample 2 (22) 20111220...	Lane Sample 3 (34) 2011122...	Lane Sample 4 (46) 2011122...
1	NM_001880.2	ATF2	Endogenous	230	187	171	154
2	NM_001706.2	BCL6	Endogenous	1,606	838	784	481
3	NM_015991.2	C1QA	Endogenous	9	11	13	15
4	NM_000491.3	C1QB	Endogenous	34	21	39	18
5	NM_001733.4	C1R	Endogenous	15	9	13	8
6	NM_001734.2	C1S	Endogenous	8	7	13	7
7	NM_000063.3	C2	Endogenous	24	43	19	30
8	NM_000064.2	C3	Endogenous	19	12	8	7
9	NM_004054.2	C3AR1	Endogenous	244	227	129	129
10	NM_007293.2	C4A	Endogenous	17	4	6	8
11	NM_001735.2	C5	Endogenous	16	20	9	15
12	NM_000065.2	C6	Endogenous	9	5	12	2
13	NM_000587.2	C7	Endogenous	6	17	9	5
14	NM_000562.2	C8A	Endogenous	10	6	12	6
15	NM_000066.2	C8B	Endogenous	10	4	8	5
16	NM_001737.3	C9	Endogenous	12	7	4	8
17	NM_002986.2	CCL11	Endogenous	16	5	22	9
18	NM_005408.2	CCL13	Endogenous	4	7	10	15
19	NM_004590.2	CCL16	Endogenous	12	7	4	7
20	NM_002987.2	CCL17	Endogenous	10	6	4	10
21	NM_006274.2	CCL19	Endogenous	10	18	11	7
22	NM_002982.3	CCL2	Endogenous	8	12	9	10

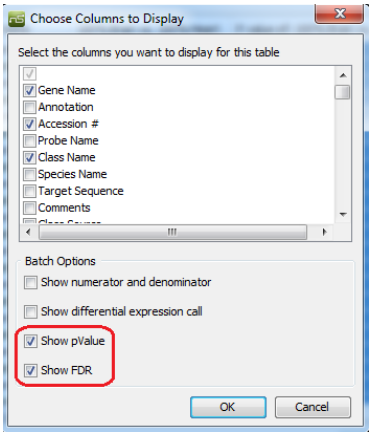


203	Gene Name	Accession #	Probe Name	Class Name	Set 2 (13) 20111220_Plex2_Poo...	Set 2 (14) 20111220_Plex2_Pool...	Set 1 (8) 20111220_F
1	ATF2	NM_001880.2	ATF2	Endogenous	183.33	195.47	
2	BCL6	NM_001706.2	BCL6	Endogenous	1,255.27	1,214.98	
3	C1QA	NM_015991.2	C1QA	Endogenous	7.34	5.46	
4	C1QB	NM_000491.3	C1QB	Endogenous	18.14	14.87	
5	C1R	NM_001733.4	C1R	Endogenous	30.57	25.31	
6	C1S	NM_001734.2	C1S	Endogenous	8.06	9.56	
7	C2	NM_000063.3	C2	Endogenous	32.25	56.63	
8	C3	NM_000064.2	C3	Endogenous	15.87	8.36	
9	C3AR1	NM_004054.2	C3AR1	Endogenous	184.15	308.97	
10	C4A	NM_007293.2	C4A	Endogenous	4.03	11.68	
11	C5	NM_001735.2	C5	Endogenous	18.81	25.48	
12	C6	NM_000065.2	C6	Endogenous	7.56	1.33	
13	C7	NM_000587.2	C7	Endogenous	3.46	4.55	
14	C8A	NM_000562.2	C8A	Endogenous	9.07	6.37	
15	C8B	NM_000066.2	C8B	Endogenous	24.19	41.41	
16	C9	NM_001737.3	C9	Endogenous	40.31	15.93	
17	CCL11	NM_002986.2	CCL11	Endogenous	10.08	4.25	
18	CCL13	NM_005408.2	CCL13	Endogenous	32.25	12.74	
19	CCL16	NM_004590.2	CCL16	Endogenous	11.08	11.68	
20	CCL17	NM_002987.2	CCL17	Endogenous	10.58	13.01	
21	CCL19	NM_006274.2	CCL19	Endogenous	9.06	9.49	

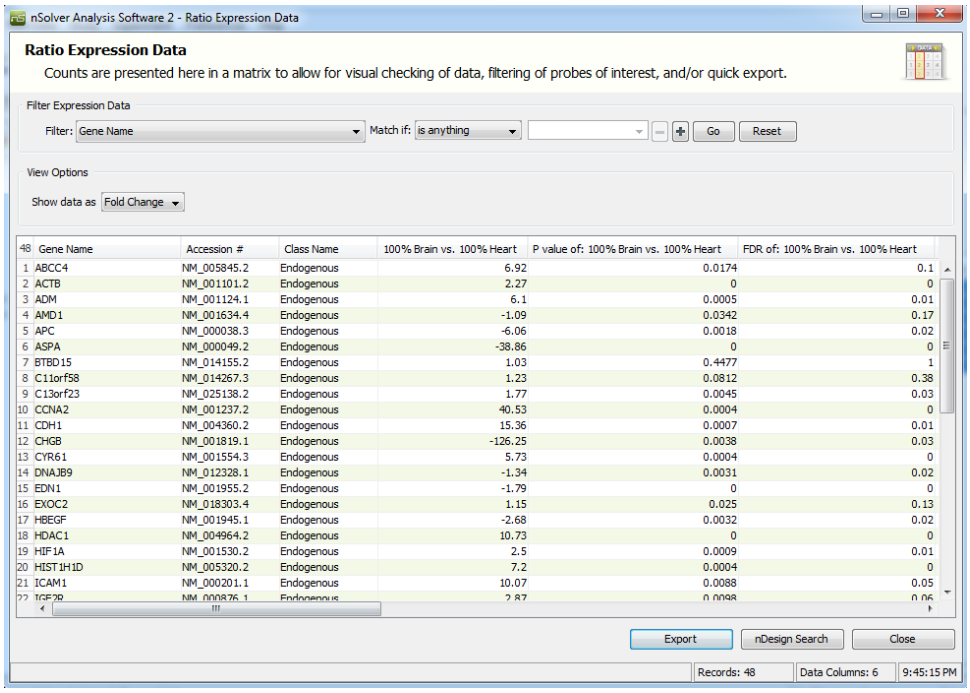


Significance Testing

nSolver provides three different methods to determine significance of ratios: error model, t-test, and false discovery rate (FDR). To view these values, including DE call, p-value, or FDR, right click on any existing column header and choose **Select Columns**. The following parameters box will appear. Select the columns you want to display.



The ratio table changes correspondingly to display the p-value and FDR.



48	Gene Name	Accession #	Class Name	100% Brain vs. 100% Heart	P value of: 100% Brain vs. 100% Heart	FDR of: 100% Brain vs. 100% Heart
1	ABCC4	NM_005845.2	Endogenous	6.92	0.0174	0.1
2	ACTB	NM_001101.2	Endogenous	2.27	0	0
3	ADM	NM_001124.1	Endogenous	6.1	0.0005	0.01
4	AMD1	NM_001634.4	Endogenous	-1.09	0.0342	0.17
5	APC	NM_000038.3	Endogenous	-6.06	0.0018	0.02
6	ASPA	NM_000049.2	Endogenous	-38.86	0	0
7	BTBD15	NM_014155.2	Endogenous	1.03	0.4477	1
8	C11orf58	NM_014267.3	Endogenous	1.23	0.0812	0.38
9	C13orf23	NM_025138.2	Endogenous	1.77	0.0045	0.03
10	CCNA2	NM_001237.2	Endogenous	40.53	0.0004	0
11	CDH1	NM_004360.2	Endogenous	15.36	0.0007	0.01
12	CHGB	NM_001819.1	Endogenous	-126.25	0.0038	0.03
13	CYR61	NM_001554.3	Endogenous	5.73	0.0004	0
14	DNAJB9	NM_012328.1	Endogenous	-1.34	0.0031	0.02
15	EDN1	NM_001955.2	Endogenous	-1.79	0	0
16	EXOC2	NM_018303.4	Endogenous	1.15	0.025	0.13
17	HBEGF	NM_001945.1	Endogenous	-2.68	0.0032	0.02
18	HDAC1	NM_004964.2	Endogenous	10.73	0	0
19	HIF1A	NM_001530.2	Endogenous	2.5	0.0009	0.01
20	HIST1H1D	NM_005320.2	Endogenous	7.2	0.0004	0
21	ICAM1	NM_000201.1	Endogenous	10.07	0.0088	0.05
22	IGF2R	NM_000876.1	Endogenous	2.87	0.0098	0.06

Error Model—(DE Call)

The NanoString bioinformatics team developed an error model to assist in determining confidence of ratios when no replicates exist. The output of this is called the DE call (differential expression call). The DE call can be viewed in the ratio data table. The output is either “Yes,” which means the difference between two groups (numerator and denominator) is significant beyond technical noise, or “No,” which means the ratio can be explained by technical noise.

At any given number of counts, the nCounter platform is subject to a certain degree of technical variability. The nCounter error model is a mapping of raw count values to estimated levels of technical variability. NanoString derived the error model from a large experiment in which multiple technical replicates were run on numerous samples and CodeSets. This experiment allowed for measurement of the average expression and technical variability (on a log scale) of thousands of gene/sample combinations. These observations were then used to create a rule mapping any given raw expression value to the 95th percentile of technical variability of genes at that expression level. Thus the DE call provides conservative estimates of any gene’s technical variability.

nSolver uses the DE call to provide guidance on whether two counts are within the technical noise of each other or whether they’re demonstrably different from each other. For any gene in any given pair of samples, nSolver can report whether the gene’s expression level is due to bona fide differential expression or able to be explained by technical noise.

The DE call is NOT designed to provide conclusions about the populations from which samples were drawn. For example, if the error model finds that the gene PDCD1 is higher in a treatment sample than in a control sample and that this difference is greater than technical noise can explain, it does not support any conclusions about whether PDCD1 is generally higher in treatment samples than in control samples; biological replicates would be needed for that broader conclusion. In other words, the DE call can be thought of as a rough replacement for technical replicates but not for biological replicates.

T-test

When replicates have been assigned as defined on the Add Sample Annotations page of the Experiment Wizard, then a t-test runs automatically and a p-value is calculated and viewable in the ratio data table.

The Student’s t-test is a statistical comparison of the mean expression between two groups of samples. In nSolver, it compares the numerator samples to the denominator samples in order to assist in determining if the fold change calculated is statistically significant or not. nSolver performs a two-tailed t-test on the log-transformed normalized data that assumes unequal variance. It is important to note that nSolver performs this using 7 digits for each data point. Therefore, the t-test output may be slightly different than if you were to perform the t-test on normalized data exported from nSolver because nSolver exports normalized data with 2 decimal places.

The t-statistic is calculated as a ratio of difference in the means of the log-transformed normalized data to the square root of the sum of the variances of samples in the two groups.

The distribution of the t-statistic is calculated by using the Welch-Satterthwaite equation for the degrees of freedom and in the estimation of the 95% confidence limits for observed differential expression between groups.

The output of the t-test is a p-value. The lower the p-value, the stronger the evidence that the two different groups have different expression levels.

False Discovery Rate

When replicates have been assigned as defined on the Add Sample Annotations page of the Experiment Wizard, the **Calculate FDR** option is enabled on the fold Change Estimations screen. Selecting this prompts nSolver to use the Benjamini-Yekutieli procedure to calculate False Discovery Rates (FDR) from the p-values returned by the t-test. FDR can be used as a more conservative and informative alternative to p-values. A gene’s FDR is the proportion of genes with p-values at least as low as the gene in question that are expected to be false discoveries.

The original paper describing the Benjamini-Yekutieli procedure is:

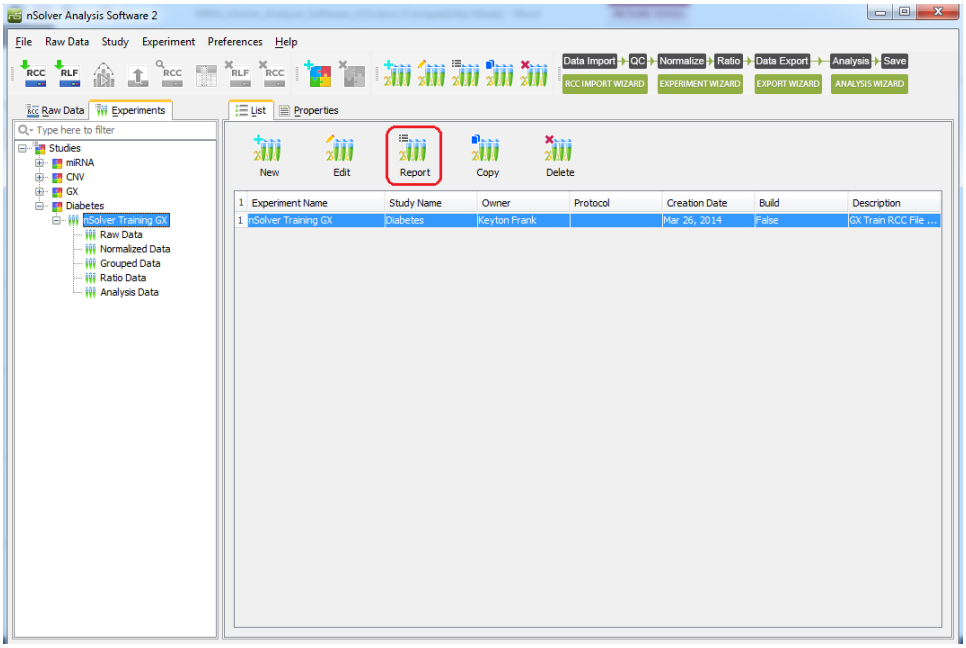
Benjamini, Y, and Yekutieli, D. (2001) "The control of the false discovery rate in multiple testing under dependency." Annals of Statistics. 29(4):1165-1188.

Experiment Reports

Experiment Reports are available for all experiments built using the Experiment Wizard. These reports contain information such as the following:

- Assay Type
- RLF Name
- Study Name
- Experiment Name
- Experiment Owner
- Experiment Protocol
- Experiment Description
- Names of All RCC Files in Experiment
- Sample ID for each RCC File
- Annotations/treatment group for each RCC File
- Description field for each RCC File
- QC Tests Performed
    - Values From Each Performed QC Test
    - QC Test Thresholds (for applied QCs)
    - QC Test Flags
- Background Subtraction type performed
    - Probes utilized
    - Calculations performed (mean, median, max, etc.)
- Normalization types performed (Positive Control Probes, Content)
    - Probes used for each normalization type
    - Mean vs. Geometric Mean for each normalization type
    - Normalization Factors
    - Normalization Thresholds
    - Normalization Flags (QC or Content)
- Calibration samples used (for Plex2 and Cross RLF experiments only)
- Ratio Method (All pairwise, Partition by, User Selected Reference Lanes)
    - List of Treatment groups and the associated RCC files
    - Ratio Descriptions
- Average, Median, and St Dev of Positive Control Probes (e.g., Avg POS_A, Avg POS_B, etc. across all samples)
- Average, Median, and St Dev of Negative Control Probes
- Average, Median, and St Dev of Each of the content probes used for normalization

To view or print the experiment report, select the experiment from the navigation tree on the left side and click on the **Report** icon.



The report opens in HTML format. (It will open an external window in the program defined as the default HTML viewer for your computer, typically a web browser.)

10 Exporting Data

This chapter contains information on how to export your nSolver data.

Throughout the application, you can export data in several file formats for use in other applications or analysis tools.



NanoString has formed partnerships with several software visualization and analysis companies. Some partners' products, such as GeneData's Analyst™, support direct import of RCC files. Other exports are available from the nSolver system and are listed below:

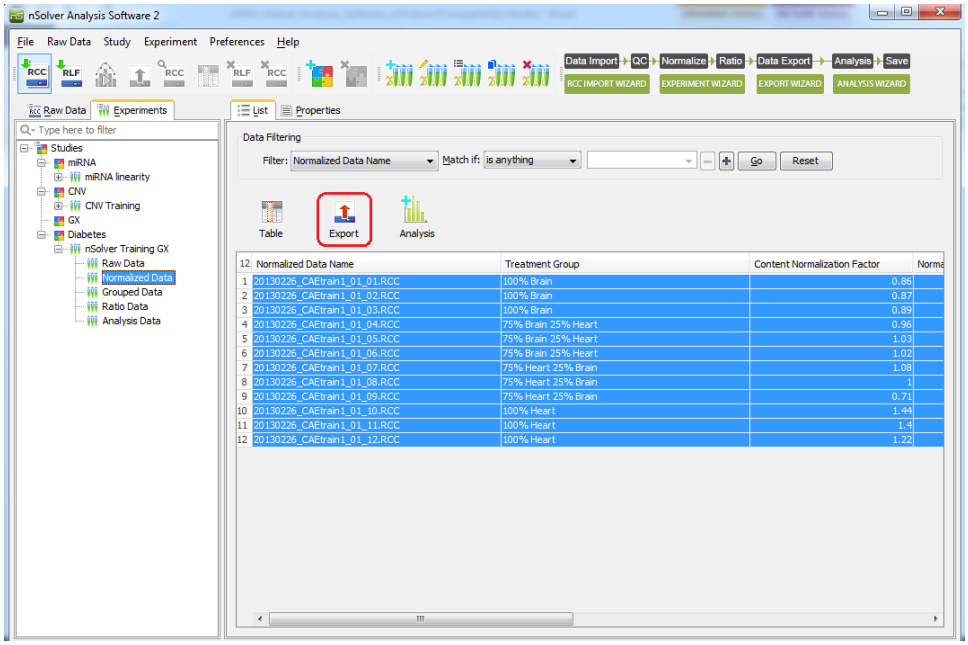
- **Custom Text Format—**This is the most widely used export option due to its flexibility. This option allows for direct import into several visualization packages, such as Gene Cluster/Java Tree View (open source), MeV (open source), FDA ArrayTrack (open source), and the Rosetta Resolver® System. You can export separate tables into separate files per lane/sample or combine all data into one file. You can customize columns for export by automatically excluding files failing QC and output ratio data as either fold changes or ratios (Copy Number Estimates or Ratios for CNV data).
- **RCC Collector Tool Format—**For customers that prefer the RCC Collector file format, the RCC Collector Tool Format Export provides raw and normalized data outputs in exactly the same format. For those not familiar with the RCC Collector Tool, it was an Excel® template which combines individual RCC data files produced by the nCounter Digital Analyzer into a single spreadsheet for analysis. It is similar to the Custom Text Format Export but includes additional cartridge lane statistics information
- **Upload MultiSet to Ingenuity®—**For those with an Ingenuity IPA® license, you can upload datasets with their associated calculations to the IPA system. Integration is seamless and the IPA system automatically launches and prompts you to login before uploading.
- **Partek Genomics Suite™—**nSolver provides an export format that is directly compatible for import into the Partek Genomics Suite system.
- **BioDiscovery Nexus Copy Number™—**For CNV users, the nSolver system exports in a format directly compatible for import into BioDiscovery's Nexus Copy Number software. Nexus Copy Number requires log ratios, and hence BioDiscovery export is only available at the ratio data level.
- **iPathwayguide™ Support—**For Advaita Bioinformatics iPathwayguide™ users, the nSolver system exports in a format directly compatible for import into the iPathwayguide™ software and is available for ratio data and associated p values for the fold change estimates in GX assays only

NOTE: The order of rows in the main table viewer will be consistent with the order of columns when opening a table as well as exporting. To view a table or export with columns in a specific order, sort the rows as desired prior to exporting or launching the table. The Ratio Description field in the Experiment Wizard can be used to assist in sorting of ratio data columns in tables and for export.

NOTE: NanoString provides support for the export function only. Interpretation of data outputs for all partner software packages are up to individual users. For information on purchasing or details on partner software packages, contact the appropriate company directly.

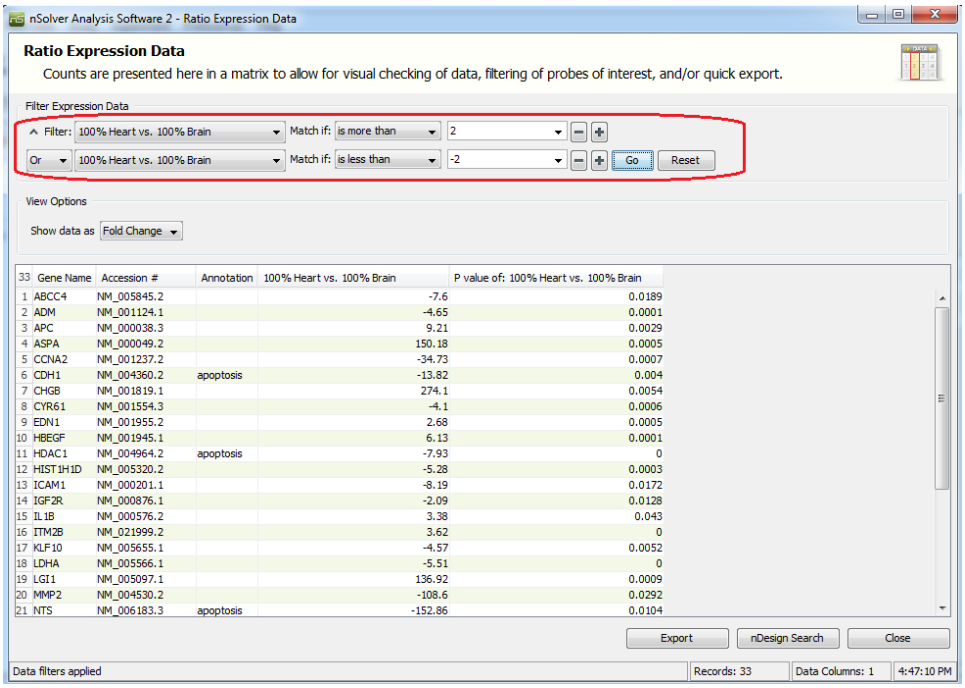
Exporting All Data

When exporting from the main table viewer, use the **Export** icon to export all contents of the selected rows.



Exporting a Subset of a Data Table

To select and export only genes of interest, open a table and use the filter to display the desired rows. Then click **Export**.

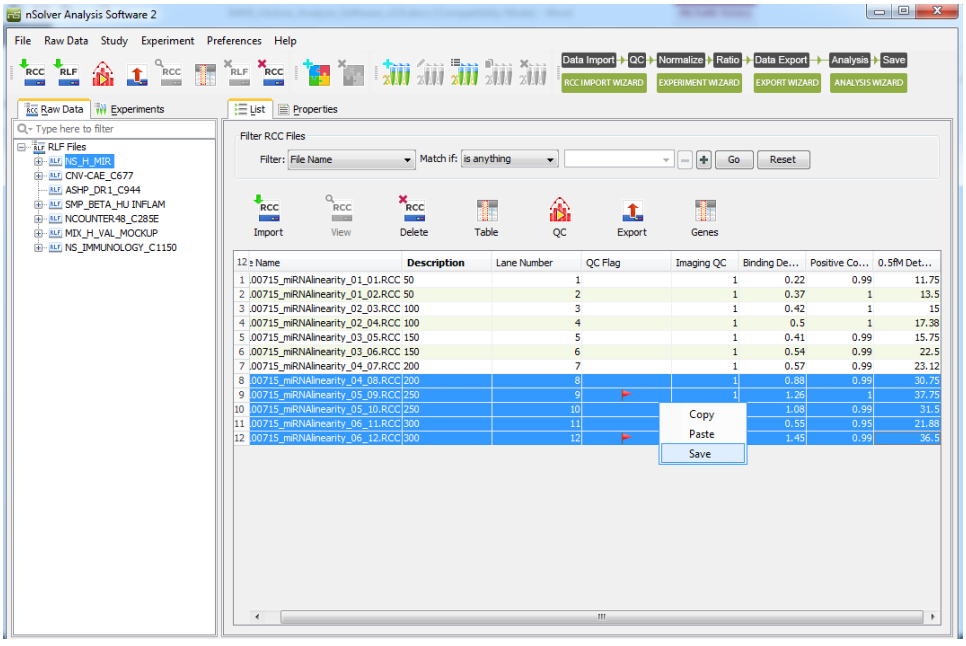


33	Gene Name	Accession #	Annotation	100% Heart vs. 100% Brain	P value of: 100% Heart vs. 100% Brain
1	ABCC4	NM_005845.2		-7.6	0.0189
2	ADM	NM_001124.1		-4.65	0.0001
3	APC	NM_000038.3		9.21	0.0029
4	ASPA	NM_000049.2		150.18	0.0005
5	CCNA2	NM_001237.2		-34.73	0.0007
6	CDH1	NM_004360.2	apoptosis	-13.82	0.004
7	CHGB	NM_001819.1		274.1	0.0054
8	CYR61	NM_001554.3		-4.1	0.0006
9	EDN1	NM_001955.2		2.68	0.0005
10	HBEGF	NM_001945.1		6.13	0.0001
11	HDAC1	NM_004964.2	apoptosis	-7.93	0
12	HIST1H1D	NM_005320.2		-5.28	0.0003
13	ICAM1	NM_000201.1		-8.19	0.0172
14	IGF2R	NM_000876.1		-2.09	0.0128
15	IL1B	NM_000576.2		3.38	0.043
16	ITM2B	NM_021999.2		3.62	0
17	KLF10	NM_005655.1		-4.57	0.0052
18	LDHA	NM_005566.1		-5.51	0
19	LGI1	NM_005097.1		136.92	0.0009
20	MMP2	NM_004530.2		-108.6	0.0292
21	NTS	NM_006183.3	apoptosis	-152.86	0.0104

Clicking **Export** launches the Export Wizard and allows you to export only the selected genes to a text file or other application.

It is also possible to export specific genes for export within a data table without using the filter function. Simply select the genes that you wish to export for all samples and click the **Export** button. The Export Wizard will then allow you to export only the selected genes to a text file or other application.

Alternatively, you may select items of interest for export, right click, and choose **Save**. You can save the selected data as a comma-separated value (*.CSV) file. The Export Wizard will not launch. This is particularly useful for exporting QC values from the main viewers.



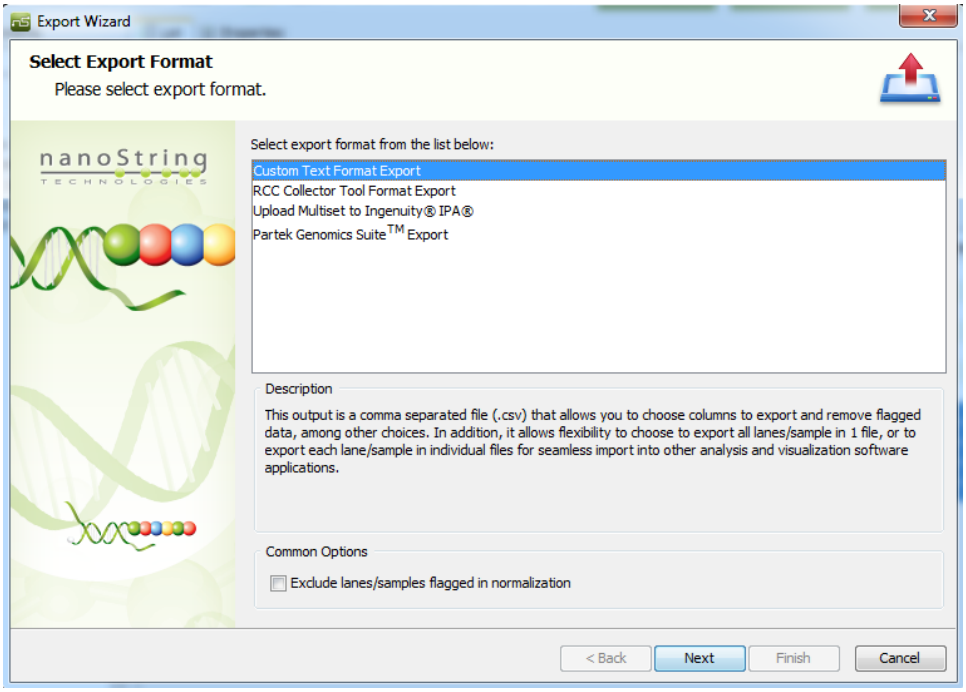
12	Name	Description	Lane Number	QC Flag	Imaging QC	Binding De...	Positive Co...	0.5fM Det...
1	00715_miRNAlinearity_01_01.RCC	50	1		1	0.22	0.99	11.75
2	00715_miRNAlinearity_01_02.RCC	50	2		1	0.37	1	13.5
3	00715_miRNAlinearity_02_03.RCC	100	3		1	0.42	1	15
4	00715_miRNAlinearity_02_04.RCC	100	4		1	0.5	1	17.38
5	00715_miRNAlinearity_03_05.RCC	150	5		1	0.41	0.99	15.75
6	00715_miRNAlinearity_03_06.RCC	150	6		1	0.54	0.99	22.5
7	00715_miRNAlinearity_04_07.RCC	200	7		1	0.57	0.99	23.12
8	00715_miRNAlinearity_04_08.RCC	200	8		1	0.88	0.99	30.75
9	00715_miRNAlinearity_05_09.RCC	250	9		1	1.26	1	37.75
10	00715_miRNAlinearity_05_10.RCC	250	10			1.08	0.99	31.5
11	00715_miRNAlinearity_06_11.RCC	300	11			0.55	0.95	21.88
12	00715_miRNAlinearity_06_12.RCC	300	12			1.45	0.99	36.5

Navigate to the desired folder to save your data and click **Save**. A prompt will display asking you if you want to open the file right now. Click **OK** to open the file.

Custom Export

The nSolver system provides customized export options. The output of the **Custom Text Format Export** is a comma-separated value file (*.CSV) that allows you to choose columns to export and remove flagged data, among other choices. In addition, it has the flexibility to export all lanes/samples in a single file, or to export each lane/sample into individual files. This allows seamless import into other analysis and visualization software applications. Some of these analysis software packages are Cluster Tree View, Rosetta Resolver®, Spotfire®, and GeneSpring®. For convenience, you can export files that are pre-formatted for direct import into various partner software platforms.

To export using the nSolver custom export, first select the data to be exported as described previously and then choose **Custom Text Format Export**. To automatically exclude from export all data flagged in QC or Normalization, select that checkbox under **Common Options**. Click **Next** to proceed.

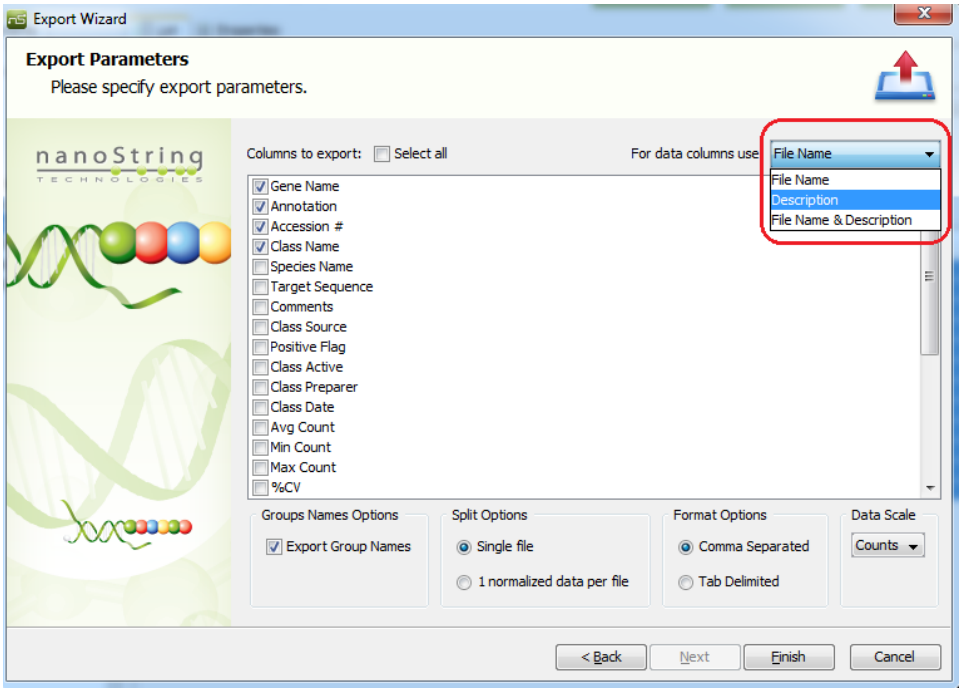


You can now customize the format and columns for your exported data. Select the check box next to only the columns that are desired for export. In addition, you can choose the header that you want displayed for the data columns. Choosing **Description** from the drop down menu replaces the RCC file name with the description for the file that you entered in ***Chapter 5 – Adding Lane/Sample Descriptions and Gender to Raw Data***.

To export all data into a single file, select **Single file** within the **Split Options** menu. To export each lane/sample separately, choose **1 normalized data per file** or similar depending on what level of data you are exporting.

You can choose to export data as either a **Comma Separated** file (*.CSV) or **Tab Delimited** file (*.TXT).

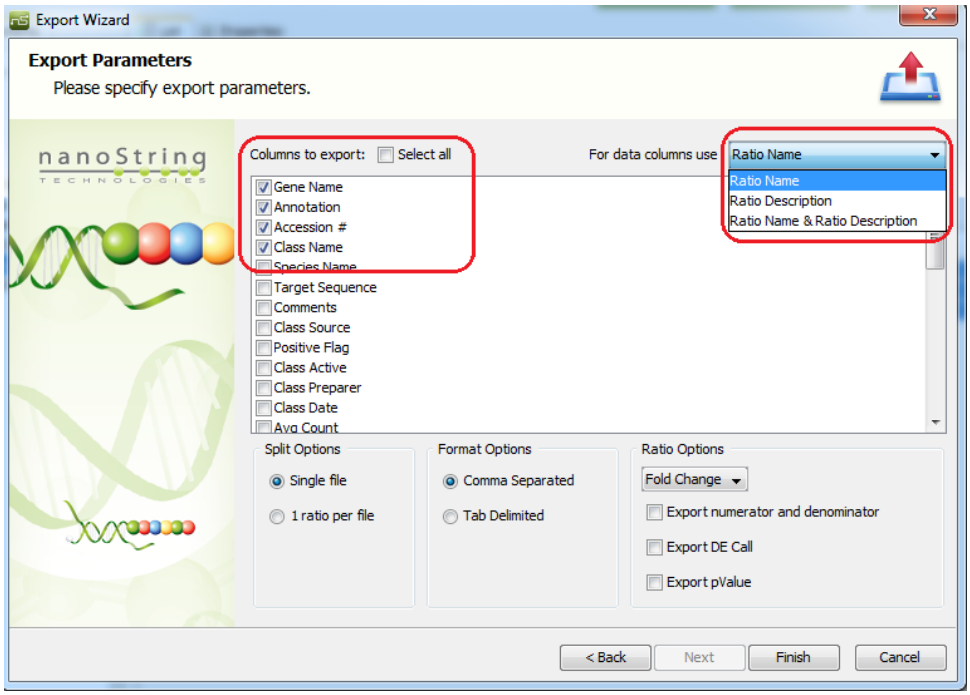
For Raw and Normalized Data, select **Export Group Names** to export the annotation group names (treatment groups) that were added on the **Add Annotation Columns** screen of the Experiment Wizard.



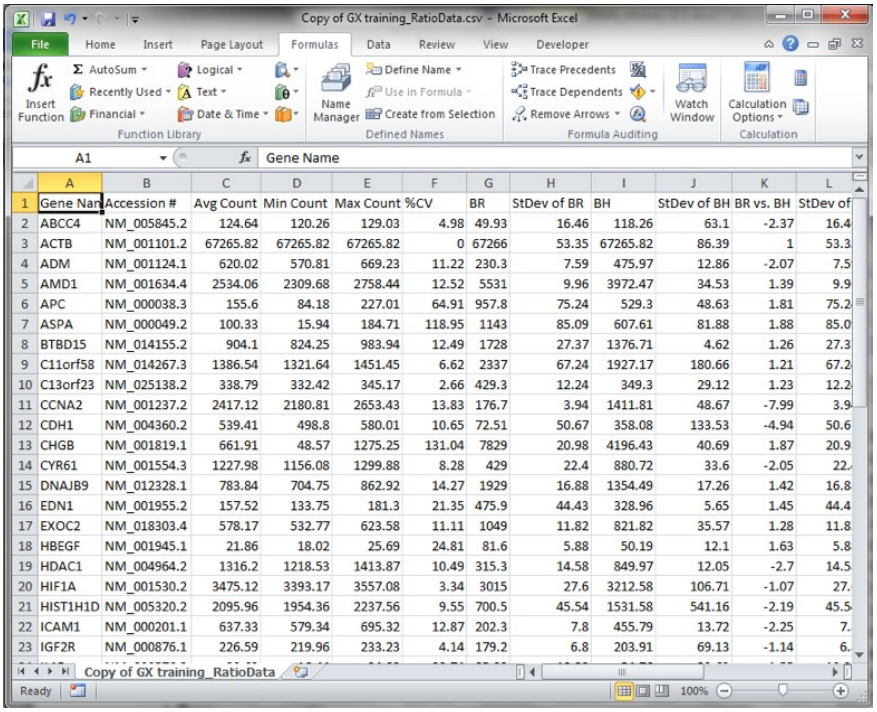
For ratio data the export wizard is similar, but you have an additional option to choose to export data as fold change, ratio, or $\log_2$ ratio. For ratio data you can customize your ratio data column headers in the export wizard, similar to the other data levels, although the choices are slightly different. Choosing to export the Ratio Data Description overwrites the default ratio name in the exported file with the description entered on the final screen of the Experiment Wizard as described in *Chapter 9 – Experiments*.

MultiRLF Merge experiment data have the option to choose Sample ID as the column header.

NOTE: For Plex^2 Assays, you also have the option to use the Sample Description associated with each sub-CodeSet for the ratio column headers. This allows columns to be displayed with the sample name or in combination with the other options (Ratio Name and Ratio Description).



Select **Finish** to complete the export and open the file. See an example file opened below:

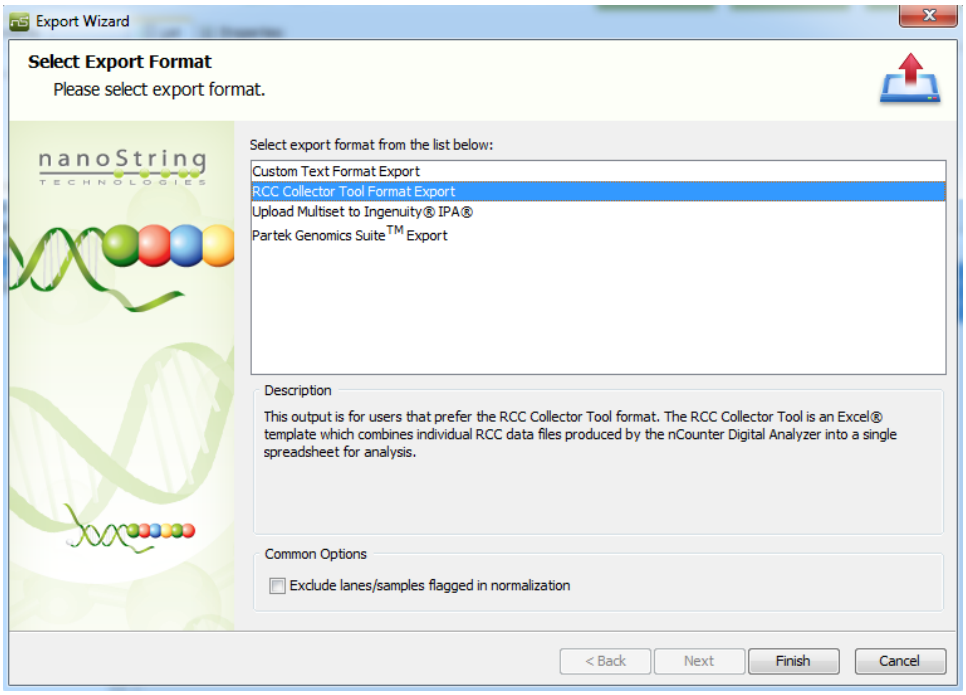


	A	B	C	D	E	F	G	H	I	J	K	L
1	Gene Nan	Accession #	Avg Count	Min Count	Max Count	%CV	BR	StDev of BR	BH	StDev of BH	BR vs. BH	StDev of
2	ABCC4	NM_005845.2	124.64	120.26	129.03	4.98	49.93	16.46	118.26	63.1	-2.37	16.4
3	ACTB	NM_001101.2	67265.82	67265.82	67265.82	0	67266	53.35	67265.82	86.39	1	53.3
4	ADM	NM_001124.1	620.02	570.81	669.23	11.22	230.3	7.59	475.97	12.86	-2.07	7.5
5	AMD1	NM_001634.4	2534.06	2309.68	2758.44	12.52	5531	9.96	3972.47	34.53	1.39	9.9
6	APC	NM_000038.3	155.6	84.18	227.01	64.91	957.8	75.24	529.3	48.63	1.81	75.2
7	ASPA	NM_000049.2	100.33	15.94	184.71	118.95	1143	85.09	607.61	81.88	1.88	85.0
8	BTBD15	NM_014155.2	904.1	824.25	983.94	12.49	1728	27.37	1376.71	4.62	1.26	27.3
9	C11orf58	NM_014267.3	1386.54	1321.64	1451.45	6.62	2337	67.24	1927.17	180.66	1.21	67.2
10	C13orf23	NM_025138.2	338.79	332.42	345.17	2.66	429.3	12.24	349.3	29.12	1.23	12.2
11	CCNA2	NM_001237.2	2417.12	2180.81	2653.43	13.83	176.7	3.94	1411.81	48.67	-7.99	3.9
12	CDH1	NM_004360.2	539.41	498.8	580.01	10.65	72.51	50.67	358.08	133.53	-4.94	50.6
13	CHGB	NM_001819.1	661.91	48.57	1275.25	131.04	7829	20.98	4196.43	40.69	1.87	20.9
14	CYR61	NM_001554.3	1227.98	1156.08	1299.88	8.28	429	22.4	880.72	33.6	-2.05	22.
15	DNAJB9	NM_012328.1	783.84	704.75	862.92	14.27	1929	16.88	1354.49	17.26	1.42	16.8
16	EDN1	NM_001955.2	157.52	133.75	181.3	21.35	475.9	44.43	328.96	5.65	1.45	44.4
17	EXOC2	NM_018303.4	578.17	532.77	623.58	11.11	1049	11.82	821.82	35.57	1.28	11.8
18	HBEGF	NM_001945.1	21.86	18.02	25.69	24.81	81.6	5.88	50.19	12.1	1.63	5.8
19	HDAC1	NM_004964.2	1316.2	1218.53	1413.87	10.49	315.3	14.58	849.97	12.05	-2.7	14.5
20	HIF1A	NM_001530.2	3475.12	3393.17	3557.08	3.34	3015	27.6	3212.58	106.71	-1.07	27.
21	HIST1H1D	NM_005320.2	2095.96	1954.36	2237.56	9.55	700.5	45.54	1531.58	541.16	-2.19	45.5
22	ICAM1	NM_000201.1	637.33	579.34	695.32	12.87	202.3	7.8	455.79	13.72	-2.25	7.
23	IGF2R	NM_000876.1	226.59	219.96	233.23	4.14	179.2	6.8	203.91	69.13	-1.14	6.

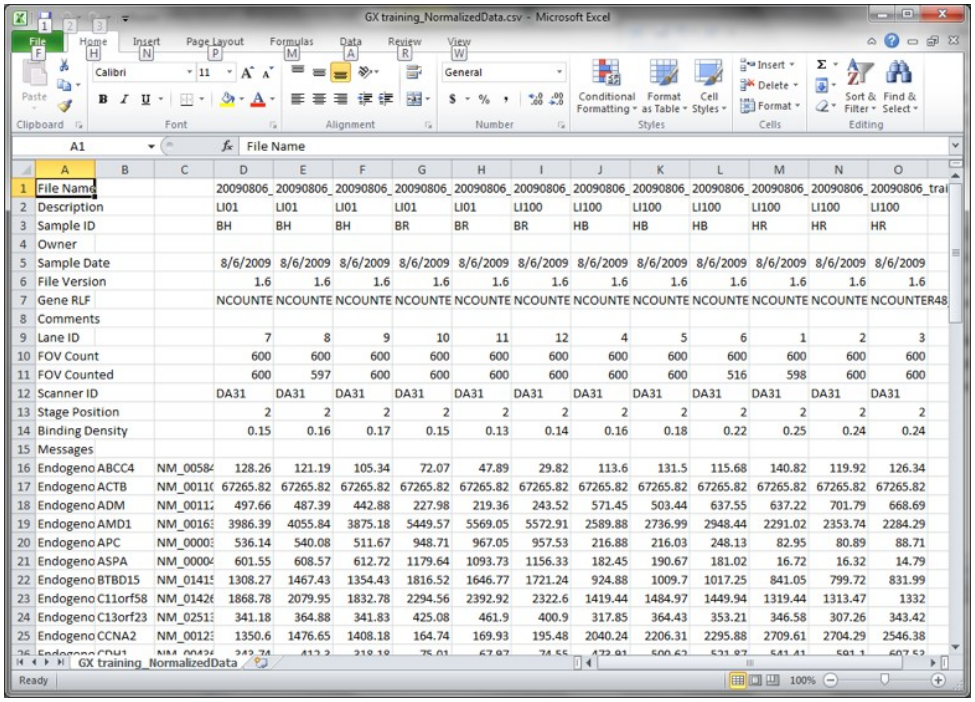
RCC Collector Format Export

For those that are familiar and prefer the RCC Collector tool export, this option is available for all raw and normalized data. The RCC Collector Tool is an Excel® template that combines the individual RCC files produced by the nCounter Digital Analyzer into a single spreadsheet for analysis. It also includes specific header information that is different than the Custom Text Format Export option.

To export in the RCC Collector Tool Format, select the desired data for export as described previously and select the **RCC Collector Tool Format Export** option. To automatically exclude from export all data that was flagged in QC or Normalization, select that checkbox under **Common Options**. Click **Finish** to proceed.



The file will open in the RCC Collector format.



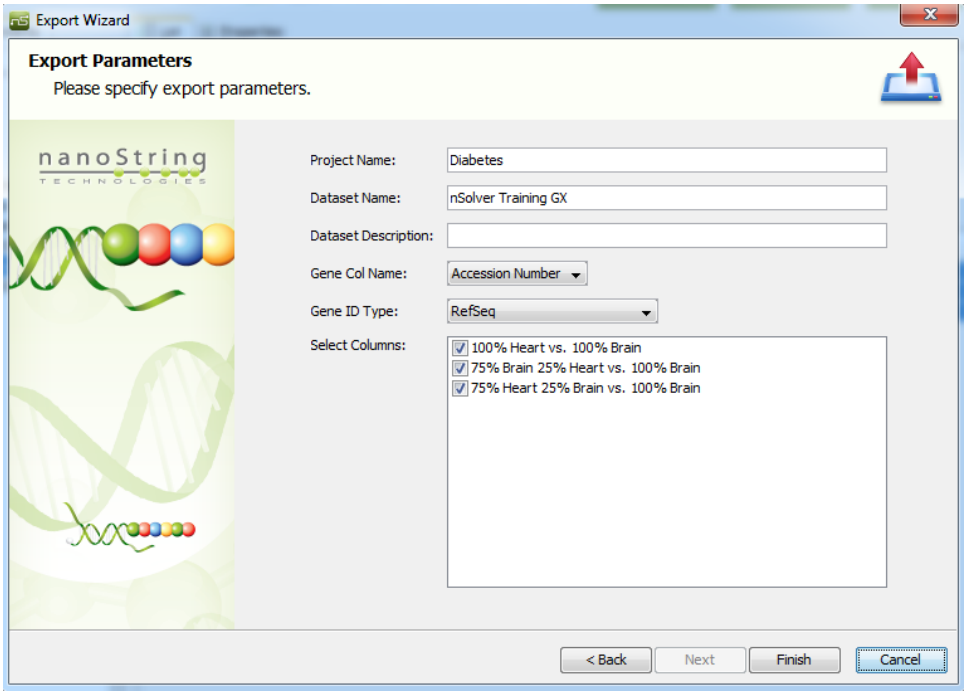
Upload MultiSet to Ingenuity®

nSolver contains seamless integration with the Ingenuity® Pathway Analysis software (IPA®). The IPA® system is a web-based application used to better understand the biology underlying experimental results. Scientists can concurrently analyze multiple datasets across different experimentation platforms and identify key functions and pathways that distinguish biological states. This technology is especially useful for dose response assessment, time course experiments, and simultaneous analysis of gene lists and protein lists. The web-delivered application makes use of the Ingenuity® Pathways Knowledge Base, a database consisting of millions of individually modeled relationships between proteins, genes, complexes, cells, tissues, drugs, and diseases.

If you have an Ingenuity® license, gene sets of interest can be uploaded with their associated fold change values to the IPA system for further analysis and investigation.

To upload data to IPA® using the nSolver system, first select the data to be uploaded as described above and then choose **Upload MultiSet to Ingenuity® IPA®**. To automatically exclude from export all data that was flagged in QC or Normalization, select the checkbox under **Common Options**. Click **Next** to proceed.

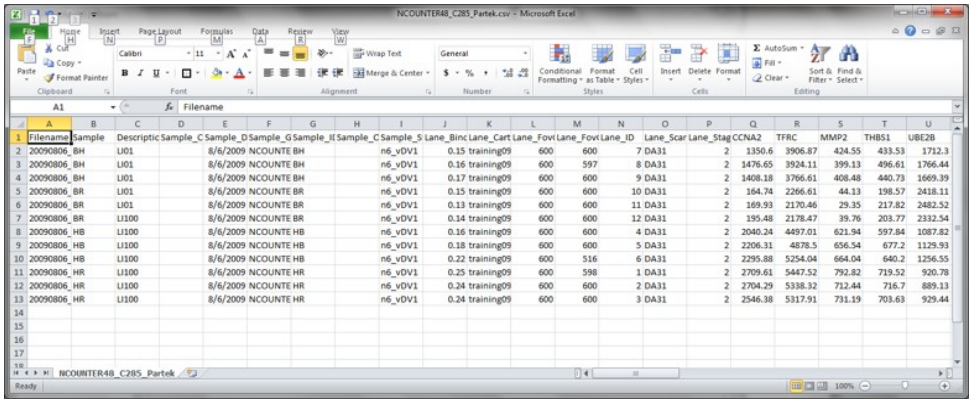
Enter a unique project name and dataset name. Choose the correct Gene Column name and Gene ID type. For nCounter data, this should typically be Accession Number/RefSeq for Gene Expression experiments and Gene Name/miRBase for miRNA.



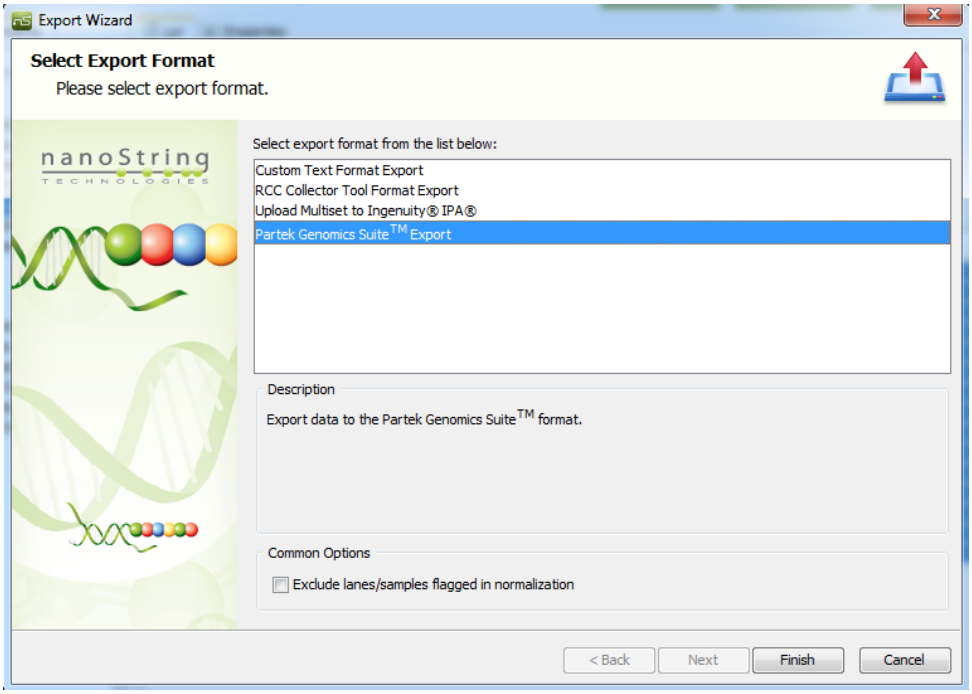
Click **Finish** to proceed and launch the Ingenuity® Pathway Analysis system. You may be prompted to log in before IPA® is launched.

Partek® Genomics Suite™ Export

nSolver also provides an export format that is directly compatible for import into the Partek® Genomics Suite™ system. To export your data to perform further analysis and visualization using Genomics Suite™, first select the data to be uploaded as described previously and then choose **Partek Genomics Suite™ Export**. To automatically exclude from export all data that was flagged in QC or Normalization, select the checkbox under **Common Options**.



Click **Finish** to proceed. You will be prompted to name your data file and can choose to open it upon export.

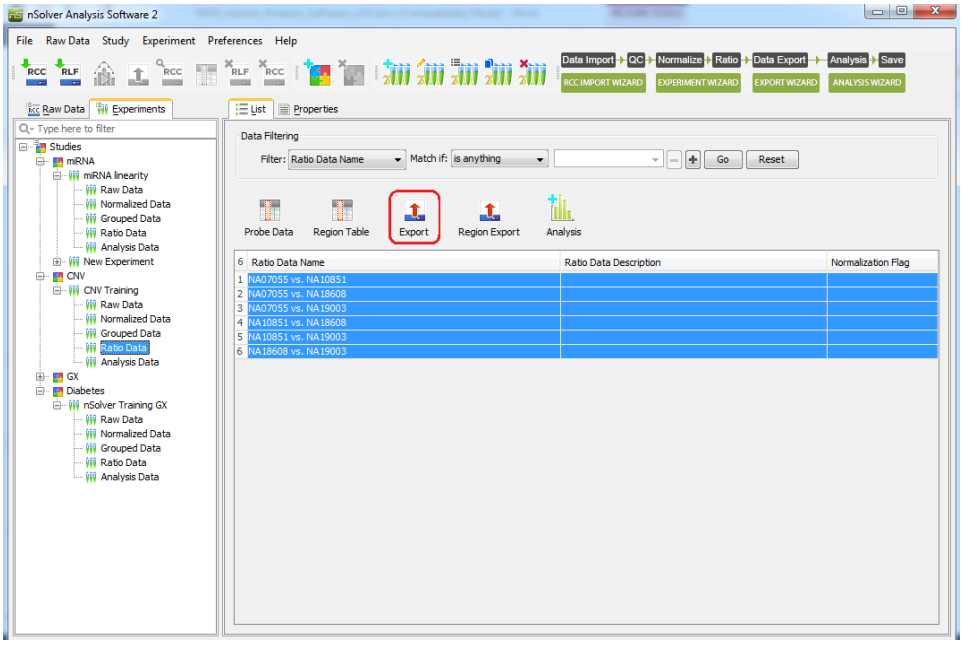


The Partek® Genomics Suite™ export format is inverted with respect to other nSolver exports: genes are displayed in columns rather than rows.

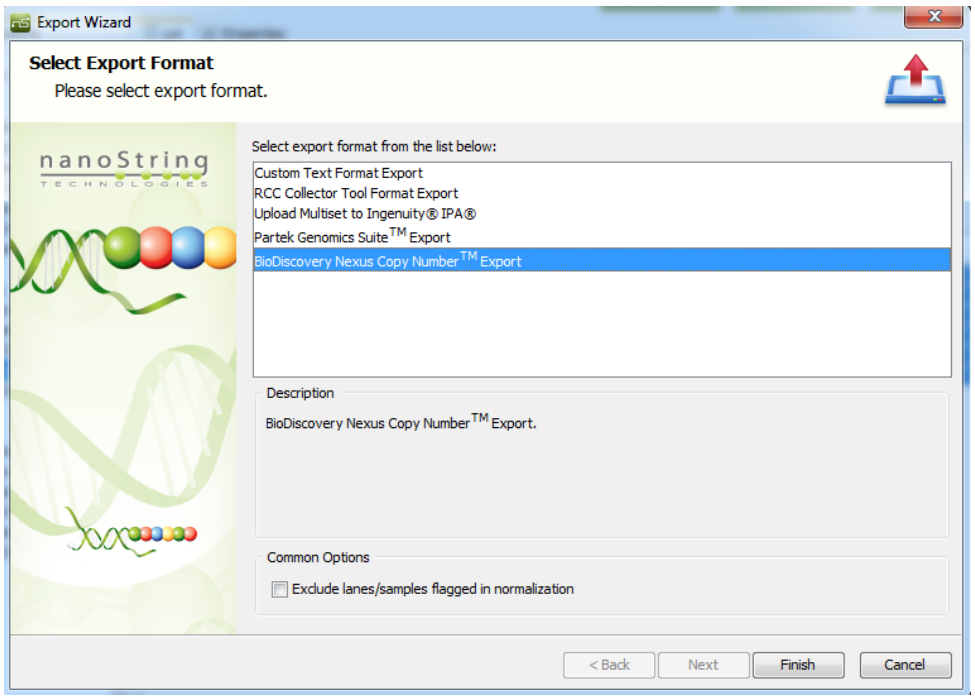
BioDiscovery Nexus Copy Number™ Export

For CNV users, the nSolver system exports in a format directly compatible for import into BioDiscovery's Nexus Copy Number™ software. BioDiscovery requires log ratio, thus BioDiscovery export is only available at the ratio data level.

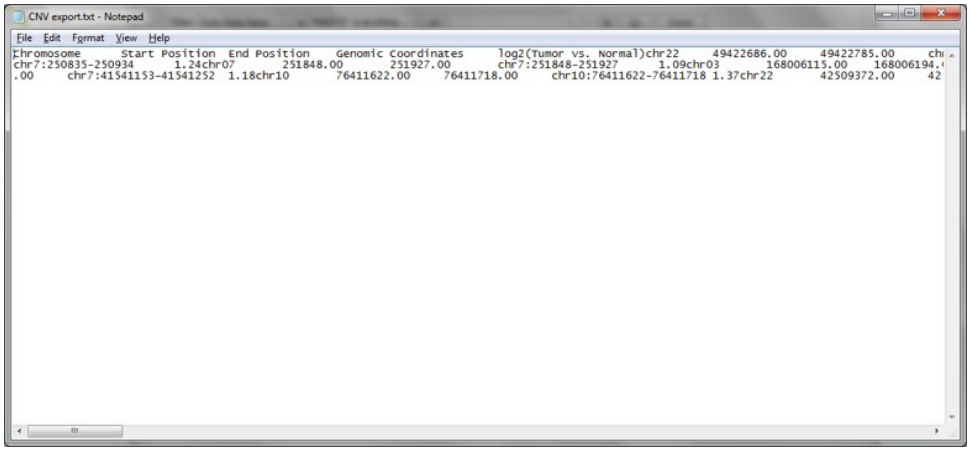
To export data in the format compatible with the Nexus Copy Number™ system, first select the data to be uploaded as described previously, click the **Export** button, and then choose **BioDiscovery Nexus Copy Number™ Export**.



NOTE: You can also export in BioDiscovery's required format from inside a CNV ratio table.



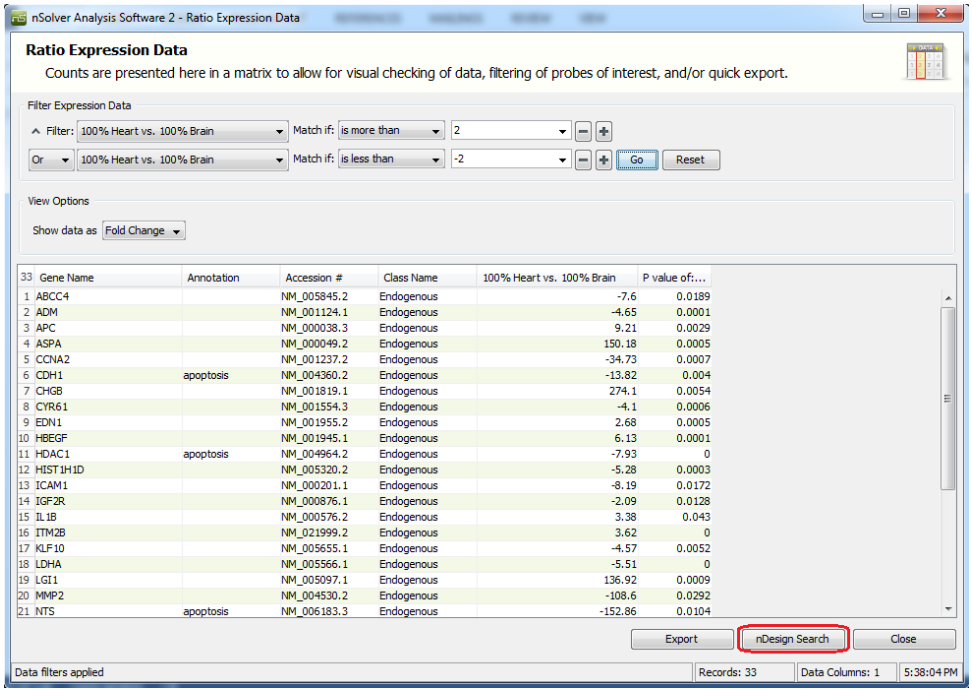
To automatically exclude from export all data that was flagged in QC or Normalization, select that checkbox under **Common Options**. Click **Finish** and you will be prompted to name your data file. You can choose to open it upon export. The BioDiscovery export is a tab delimited (*.TXT) file.



nDesign™ Gateway Integration

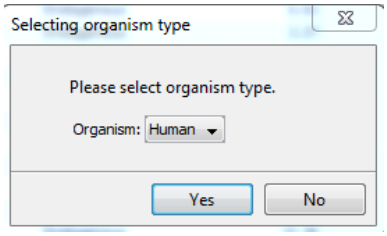
nSolver allows you to export genes to the nDesign Gateway web application, a tool for helping NanoString customers create CodeSets based on pathways, gene ontologies, and diseases. In addition, it can take a list of genes and allow you to view pertinent pathways and NanoString panels. nSolver includes support for sending a set or subset of your experiment to search nDesign and view pertinent pathways or create a custom CodeSet.

To send a list of genes to nDesign, filter from any nSolver table and then click the **nDesign Search** button.

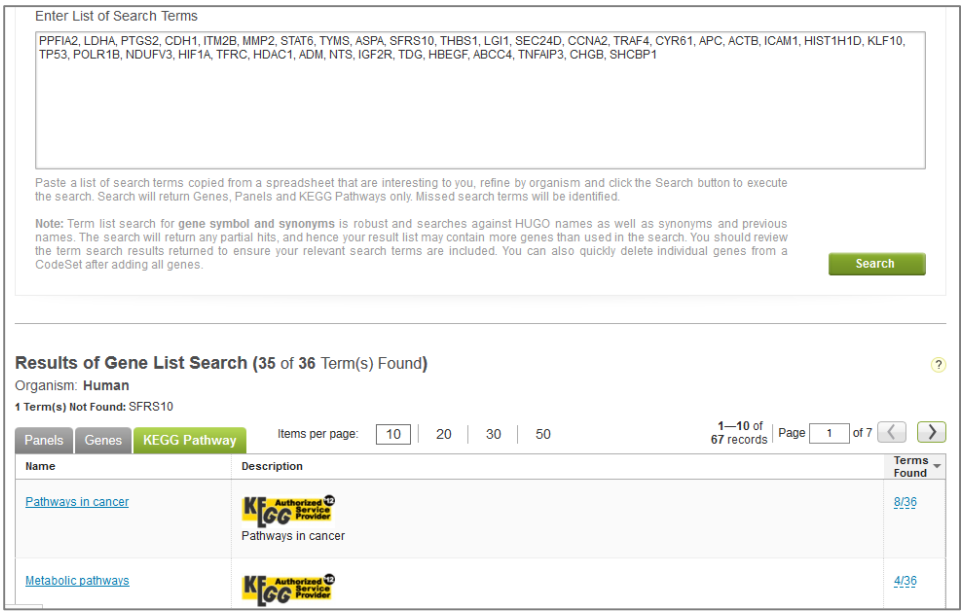


33	Gene Name	Annotation	Accession #	Class Name	100% Heart vs. 100% Brain	P value of:...
1	ABCC4		NM_005845.2	Endogenous	-7.6	0.0189
2	ADM		NM_001124.1	Endogenous	-4.65	0.0001
3	APC		NM_000038.3	Endogenous	9.21	0.0029
4	ASPA		NM_000049.2	Endogenous	150.18	0.0005
5	CCNA2		NM_001237.2	Endogenous	-34.73	0.0007
6	CDH1	apoptosis	NM_004360.2	Endogenous	-13.82	0.004
7	CHGB		NM_001819.1	Endogenous	274.1	0.0054
8	CYR61		NM_001554.3	Endogenous	-4.1	0.0006
9	EDN1		NM_001955.2	Endogenous	2.68	0.0005
10	HBEGF		NM_001945.1	Endogenous	6.13	0.0001
11	HDAC1	apoptosis	NM_004964.2	Endogenous	-7.93	0
12	HIST1H1D		NM_005320.2	Endogenous	-5.28	0.0003
13	ICAM1		NM_000201.1	Endogenous	-8.19	0.0172
14	IGF2R		NM_000876.1	Endogenous	-2.09	0.0128
15	IL1B		NM_000576.2	Endogenous	3.38	0.043
16	ITM2B		NM_021999.2	Endogenous	3.62	0
17	KLF10		NM_005655.1	Endogenous	-4.57	0.0052
18	LDHA		NM_005566.1	Endogenous	-5.51	0
19	LGI1		NM_005097.1	Endogenous	136.92	0.0009
20	MMP2		NM_004530.2	Endogenous	-108.6	0.0292
21	NTS	apoptosis	NM_006183.3	Endogenous	-152.86	0.0104

Select the desired organism for searching nDesign and click **Yes**.



nDesign will load in the computer's default web browser. Click the pathways tab if you want to view the various pathways associated with your gene set.



11 Analyses and Visualizations

This chapter contains information on how to perform analysis and visualizations in nSolver.



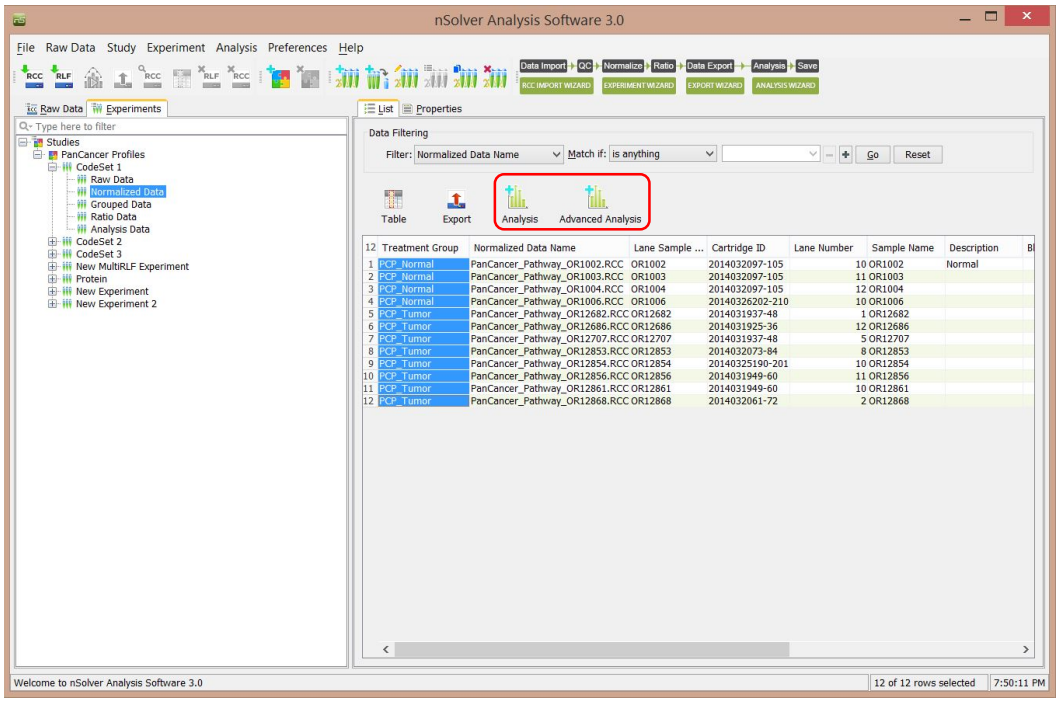
nSolver has a variety of analyses and visualizations:

- Agglomerative Cluster (Heat Map)
- Violin Plot
- Box Plot
- Scatter Plot
- Histogram

These five visualizations are accomplished using the Analysis Wizard, which may be launched from any of the raw, normalized, grouped, or ratio data levels. In addition, nSolver versions 2.5 and higher permit Advanced Analyses that make use of the open source R statistical software and may be particularly useful for pathway-based research. Advanced Analyses may only be launched when viewing the raw or normalized data levels and will be discussed at the end of this chapter.

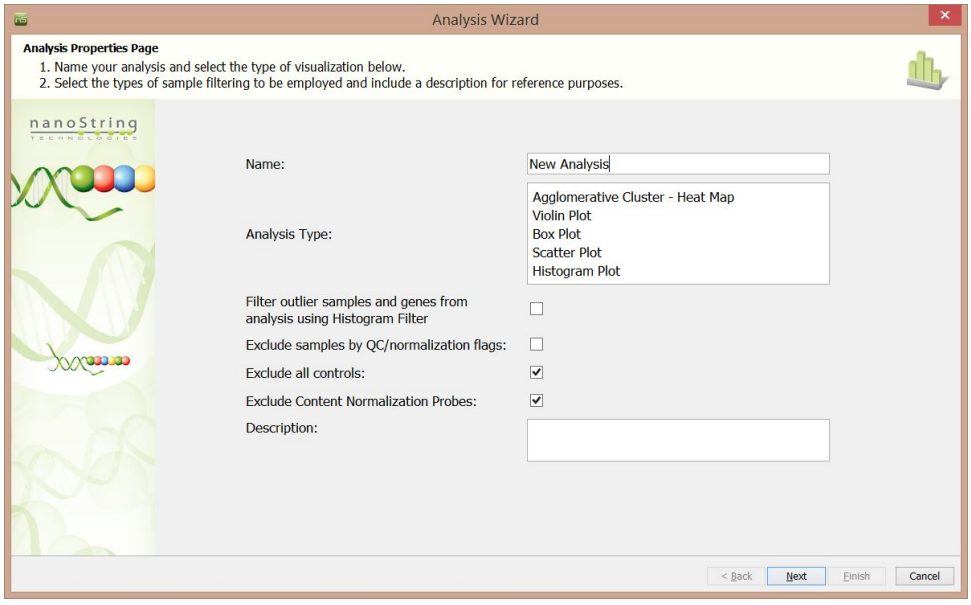
Analysis Wizard

To run an analysis, select the data level and rows and click the **Analysis** or **Advanced Analysis** icons.



Analysis Properties

The Analysis Wizard will launch, and the first screen will be the Analysis Properties Page.



Enter the following properties and click **Next**.

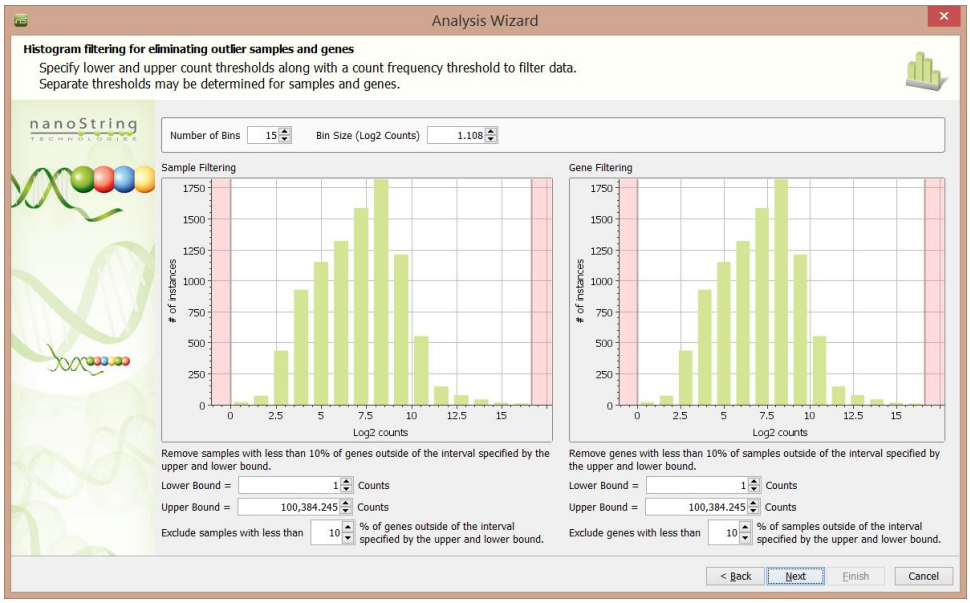
- **Name**—Name the analysis. Use a unique and meaningful name for easier retrieval.
- **Analysis Type**—Choose the desired analysis type.
- **Filter outlier samples and genes from analysis using Histogram Filter**—Check this box to display the Histogram Filter on the next page of the wizard. The Histogram Filter assists with filtering outliers from the

analysis. It is particularly valuable for large sample sets. Uncheck this option if you want to bypass the Histogram Filter page.

- **Exclude samples by QC/normalization flags**—Check this option to exclude any samples that have a QC or Normalization Flag. You can make adjustments later as you proceed through the wizard.
- **Exclude all controls**—Check this option to exclude all probes that have a Code Class of Positive, Negative, Spike In, or Invariant.
- **Exclude Content Normalization Probes**—Check this option to exclude all probes used for Content (Housekeeping) Normalization.
- **Description**—Add an optional text description for the analysis.

Histogram Filter

If you selected the checkbox to filter samples and genes using the Histogram Filter, then the Histogram Filter will be the next page of the Analysis Wizard. Otherwise the Histogram Filter will be skipped and the wizard will continue on to the include/exclude samples page.



The Histogram Filter is generally designed to allow removal of poor quality samples or genes from downstream analysis and visualization. At the top of the page, the histogram can be adjusted to appear more or less granular by changing the **Number of Bins** or the **Bin Size**.

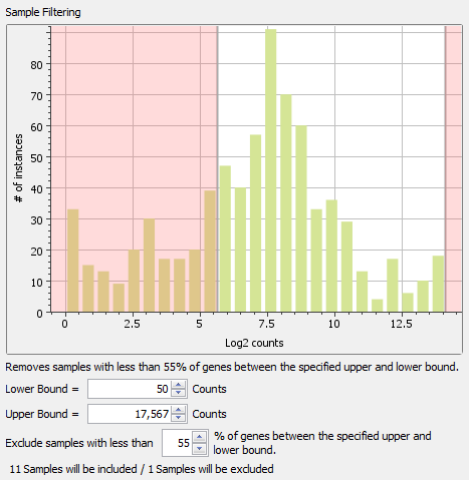
The histogram on the left is for Sample Filtering, displaying $\log_2$ counts on the x-axis and the number of instances (samples) on the y-axis. Use the following steps to remove outliers:

1. Move the sliding bars to the right and left to select the genes per sample you want to include. As you move the bars, the lower and upper bound count levels will adjust automatically. The genes per sample between the slider settings will be included in the analysis.
2. Alternately, you can edit the **Lower Bound** and **Upper Bound** settings below the histogram, which adjusts the sliding bars to match the upper and lower count levels specified.
3. For data points that are outside of the bounds specified, samples will be excluded for further analysis only if the specified percentage of the genes is outside the upper and lower bounds.

A summary of the samples to be included and excluded is always displayed dynamically as adjustments are made to the graph. If you selected Exclude QC and Normalization Flags on the Analysis Properties page and there were any flagged samples, then these are displayed as excluded by default on this page.

For example, in the histogram below, if the **Lower Bound** count is set to 50, and the **% of genes** between the specified

bounds is set to 55%, then one sample would be excluded from analysis. This is shown in the following figure.

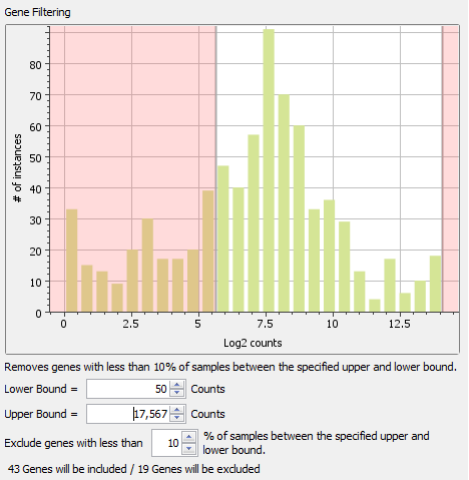


The histogram on the right side is for gene/probe filtering, displaying $\log_2$ counts on the x-axis and the number of instances (genes) on the y-axis. Similar to the sample filtering graph, use the following functions to removing outliers:

1. Move the sliding bars to the right or left to select which genes per sample to be included. As you move slider bars, the lower and upper bound count levels are automatically adjusted. The genes per sample in between the slider settings are included.
2. Alternately, you can edit the Lower and Upper Bounds, which moves the slider bars automatically to match the upper and lower count levels specified.
3. For data points that are outside of the bounds specified, genes are excluded for further analysis only if the specified percentage of the samples is outside the upper and lower bounds.

A summary of the genes to be included and excluded is displayed dynamically as adjustments are made to the histogram. Seventeen genes/probes were excluded by default in the example above because the Content Normalization and Controls were specified to be excluded by default on the Analysis Properties page.

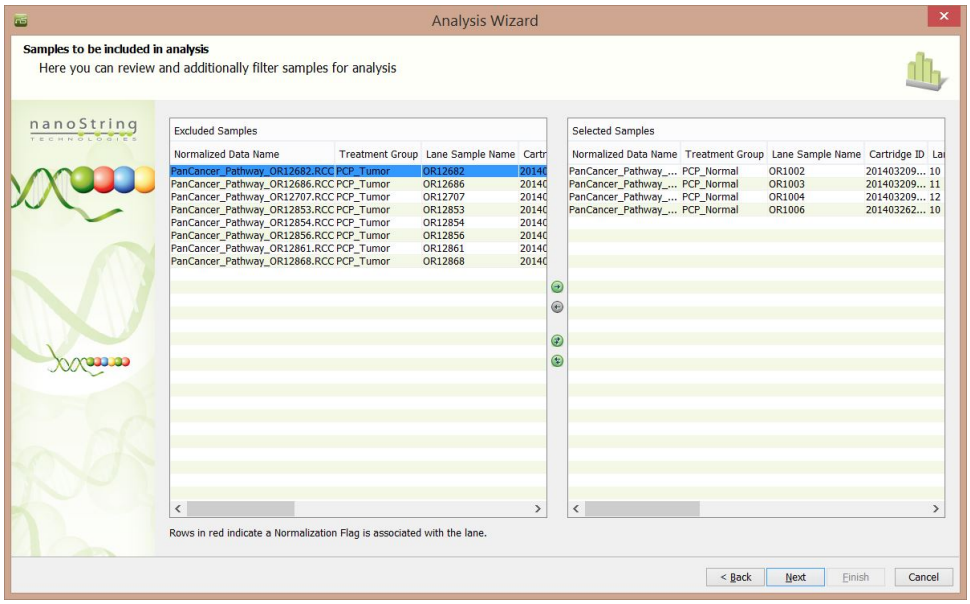
For example, in the histogram below, if the lower bound count is set to 50, and the % of samples between the specified bounds is set to 55%, an additional two genes will be excluded from analysis (19 total) as shown below.



After specifying desired sample and filtering options, click the **Next** button. The samples and genes specified by the histogram filter will be carried over to the next two pages of the Analysis Wizard and classified as either “included” or “excluded” from the analysis/visualization.

Samples to Be Included in Analysis

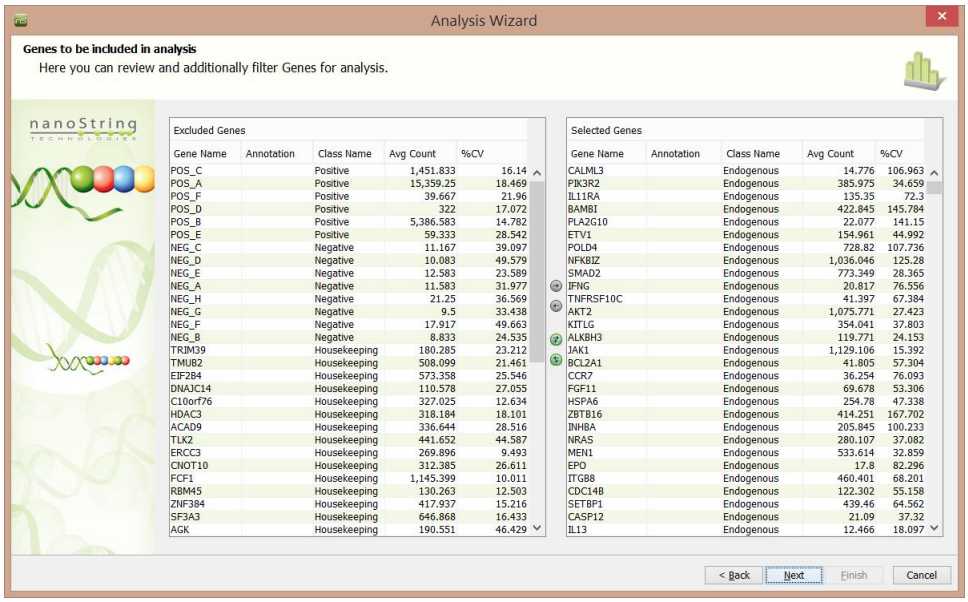
The next dialog box allows you to view and edit the samples to include and exclude from analysis. Adjust the Samples lists by selecting the samples and using the right and left arrows.



NOTE: QC Flags are displayed as a column in the table. Samples with a Normalization Flag are denoted in **red**.

Genes/Probes to Be Included in Analysis

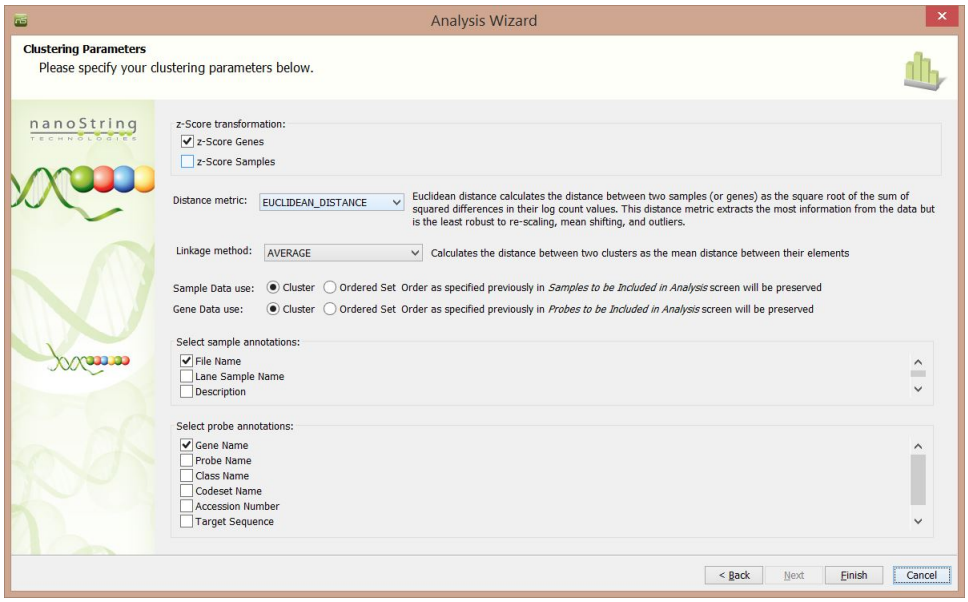
The next dialog box allows you to view and edit the genes/probes to include and exclude from the analysis. Adjust these by selecting the probes and using the right and left arrows.



If you are creating an Agglomerative Cluster (Heat Map) plot, click **Next** to enter the visualization-specific parameters. For all other plots, click **Finish** to launch the visualization.

Agglomerative Cluster (Heat Map)-Specific Parameters

If you select the Agglomerative Cluster analysis, then you will enter Clustering Parameters on the final dialog box of the Analysis Wizard.

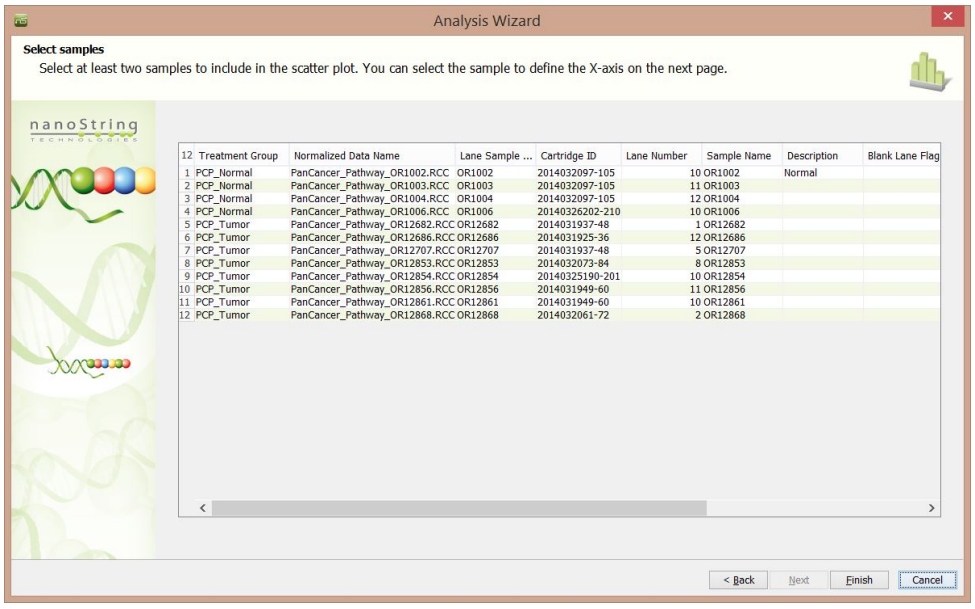


The following functions are available for cluster analysis.

- If the assay type is CNV, then an option will be available to run the cluster on probes or on regions.
- **Z-score transformation (Genes/probes)**—Check this if you want to perform z-score transformation (centering and scaling) on genes/probes. Note that this is recommended to produce the most easily interpreted heat maps.
- **Z-score transformation (Samples)**—Check this if you want to perform z-score transformation (centering and scaling) on samples
- **Distance metric**—Choose between Spearman correlation, Euclidean distance, and Pearson correlation. A short definition explaining each metric is dynamically displayed.
- **Linkage method**—Choose between Average, Median, Complete, and Centroid linkage methods. A short definition explaining each method is dynamically displayed.
- **Sample data use**—Agglomerative clustering can be performed on samples, genes, both, or neither. If you want to simply plot the heat map in a specified sample order, make sure the rows are sorted in the 'Samples to be Included in Analysis' page of the Analysis Wizard as you desire them to appear in the heat map. Then choose the Ordered Set option here. Otherwise agglomerative clustering will be performed on samples.
- **Gene data use**—If you want to simply plot the heat map in a specified gene order, make sure the rows are sorted in the 'Genes to be Included in Analysis' page of the analysis wizard as desired to appear in the Heat map. Then choose the Ordered Set option here. Otherwise the agglomerative cluster will be performed on genes.
- **Select sample annotations**—Check any of the fields that you want to be displayed as a sample label on the heat map. Items will be separated by a comma.
- **Select probe annotations**—Check any of the fields that you want to be displayed as gene/probe labels on the heat map. Items will be separated by a comma.

Scatter Plot Parameters

If you selected scatter plot analysis, specify the samples to include in the scatter plot on the final dialog box of the Analysis Wizard.



	Treatment Group	Normalized Data Name	Lane Sample ...	Cartridge ID	Lane Number	Sample Name	Description	Blank Lane Flag
1	PCP_Normal	PanCancer_Pathway_OR1002.RCC	OR1002	2014032097-105	10	OR1002	Normal	
2	PCP_Normal	PanCancer_Pathway_OR1003.RCC	OR1003	2014032097-105	11	OR1003		
3	PCP_Normal	PanCancer_Pathway_OR1004.RCC	OR1004	2014032097-105	12	OR1004		
4	PCP_Normal	PanCancer_Pathway_OR1006.RCC	OR1006	20140326202-210	10	OR1006		
5	PCP_Tumor	PanCancer_Pathway_OR12682.RCC	OR12682	2014031937-48	1	OR12682		
6	PCP_Tumor	PanCancer_Pathway_OR12686.RCC	OR12686	2014031925-36	12	OR12686		
7	PCP_Tumor	PanCancer_Pathway_OR12707.RCC	OR12707	2014031937-48	5	OR12707		
8	PCP_Tumor	PanCancer_Pathway_OR12853.RCC	OR12853	2014032073-84	8	OR12853		
9	PCP_Tumor	PanCancer_Pathway_OR12854.RCC	OR12854	20140325190-201	10	OR12854		
10	PCP_Tumor	PanCancer_Pathway_OR12856.RCC	OR12856	2014031949-60	11	OR12856		
11	PCP_Tumor	PanCancer_Pathway_OR12861.RCC	OR12861	2014031949-60	10	OR12861		
12	PCP_Tumor	PanCancer_Pathway_OR12868.RCC	OR12868	2014032061-72	2	OR12868		

If you selected raw, normalized, or grouped data for analysis, you will be prompted to select a minimum of two samples to use in the scatter plot. If you selected ratio data for analysis and if p-values were calculated (replicates were specified), then you must select at least one ratio to be used in scatter plot. Otherwise, if no p-values exist, then you must select at least two ratios to use in the scatter plot.

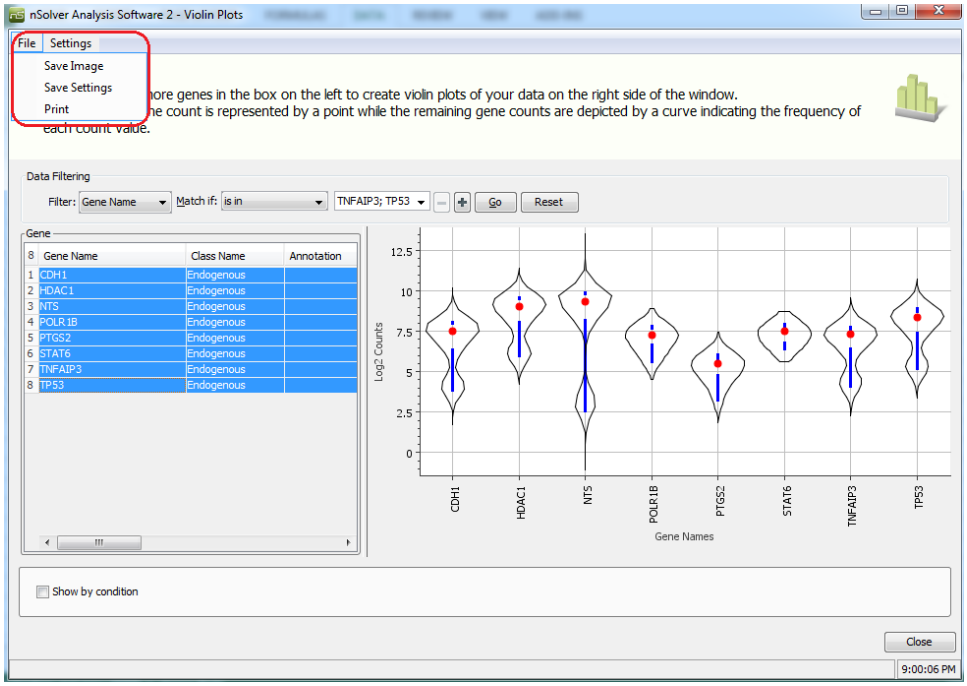
Analyses/Visualizations

After completing the Analysis Wizard, the visualizations will appear. For box plots, violin plots, and scatter plots, there are some common features you can use to customize the plots.

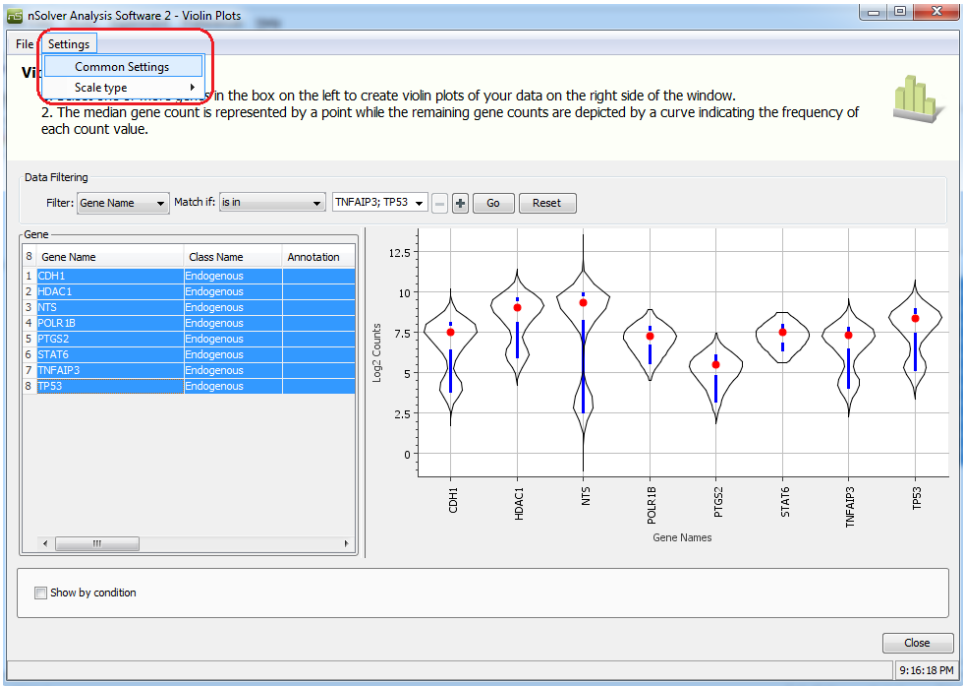
Common Settings (Excludes Heat Maps)

The **File** menu is available from the menu bar at the top. You can perform the following functions:

- **Save Image**—Save the plot as an image (either *.JPG, *.PNG, or *.GIF).
- **Save Settings**—Save the settings that were made as the visualization was adjusted using various options provided to users. If a user selects Save Settings, when they bring up the analysis in the future, it should already have these parameters set as when the visualization was previously opened.
- **Print**—Send the image to a designated printer.

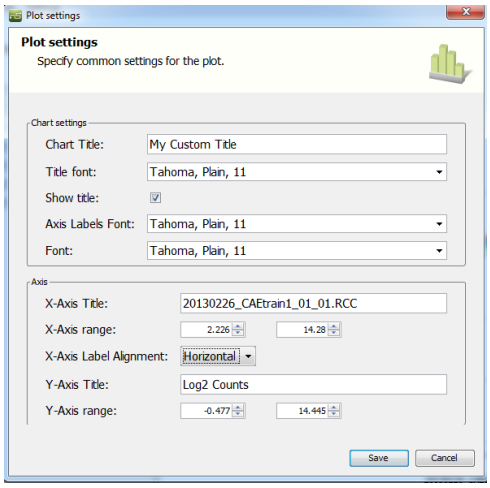

nSolver Analysis Software 2 - Violin Plots
File
Settings
Save Image
Save Settings
Print
Data Filtering
Filter: Gene Name
Match if: is in
TNFAIP3; TP53
Go
Reset
Gene
Gene Name
Class Name
Annotation
1 CDH1 Endogenous
2 HDAC1 Endogenous
3 NTS Endogenous
4 POLR1B Endogenous
5 PTGS2 Endogenous
6 STAT6 Endogenous
7 TNFAIP3 Endogenous
8 TP53 Endogenous
Log2 Counts
12.5
10
7.5
5
2.5
0
CDH1
HDAC1
NTS
POLR1B
PTGS2
STAT6
TNFAIP3
TP53
Gene Names
Show by condition
Close
9:00:06 PM

Common Settings are also available from the top menu bar under the **Settings** menu.

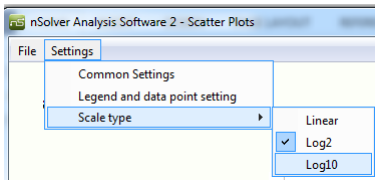


Open **Common Settings** to adjust the visualization in the following ways:

- Create a chart title that will display above the visualization.
- Choose the chart title font.
- Decide if the chart title shall be displayed or not.
- Choose the font for the axis titles.
- Choose the font for the axis labels.
- Create the x-axis title.
- Adjust the x-axis range (scatter plot only).
- Show regression (scatter plot only).
    - If Show Regression is selected, choose to display the R^2 value, the equation of the line slope, or both. The corresponding values display in the legend.
- Choose the x-axis label alignment (histogram, box plots, and violin plots only). This option determines whether the labels on the x-axis are displayed vertically or horizontally.
- Create the y-axis title.
- Adjust the y-axis range.



Select **Scale type** from the **Settings** menu to change the scale of the plot between log_2, linear, and log_{10}. The default is log_2 for all visualizations.



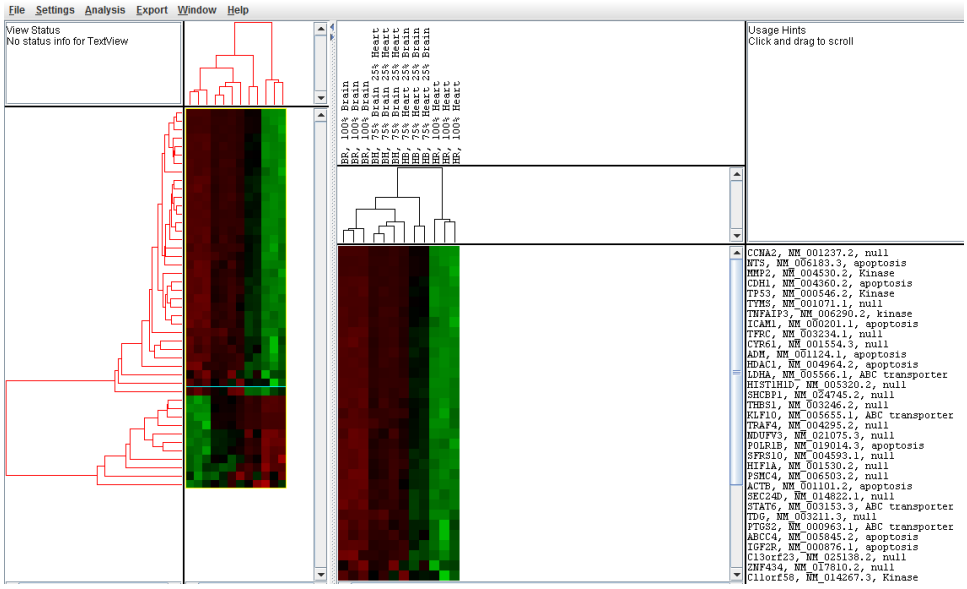
Plots are dynamically adjusted after changing the option as shown above.

Agglomerative Cluster (Heat Map)

Agglomerative clustering is a bottom up form of hierarchical clustering. Hierarchical clustering makes inferences about all data sets relationships to each other. Specifically, agglomerative clustering yields an extensive hierarchy of clusters that merge with each other at certain distances. First, the two closest objects are paired and their values are averaged. After the initial pairing, the next two most similar objects are paired and averaged. The pairing continues until all objects have been compared.

The output of the agglomerative cluster is a heat map with a dendrogram tree. Data sets belonging to the same branch of a cluster are similar to each other at some level; data sets in separate branches are less similar at some level.

nSolver uses the open source application Java TreeView to view heat maps. nSolver automatically launches Java TreeView and exports the clustered data to create a heat map. No changes to Java TreeView code were made. All documentation, including technical documentation on Java TreeView, can be found at **http://jtreeview.sourceforge.net/manual.html**.



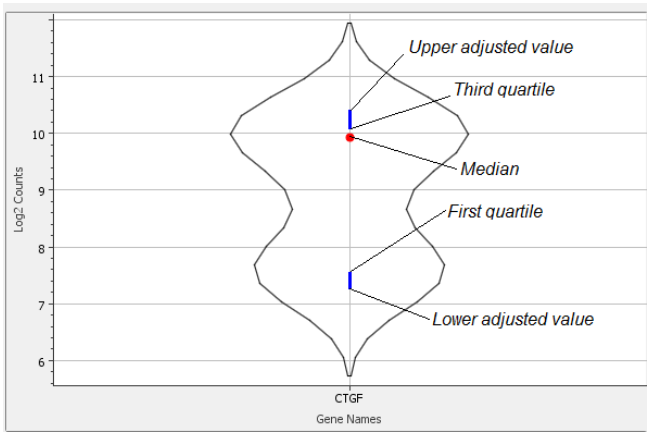
NOTE: The methods used for computations can be found at http://sape.inf.usi.ch/hac.

Violin Plots

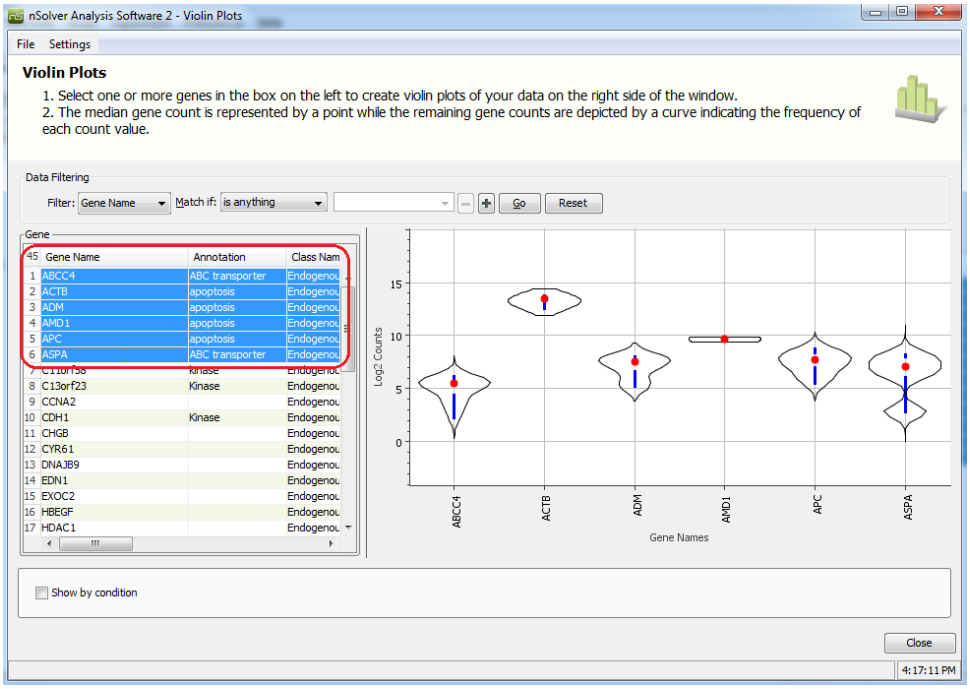
Violin plots are a convenient way of depicting subsets of your experiment with a curve around the plot to show the rounded distribution of data. Violin plots have blue lines extending vertically to indicate the lower adjusted value, first quartile, third quartile, and upper adjusted value. The lower adjusted value and upper adjusted value are the minimum and maximum values in the data series. The median is depicted with a red dot.

- The first quartile (designated Q1) also called the lower quartile or the 25th percentile (separates the lowest 25% of data from the highest 75%)
- The second quartile (designated Q2) also called the median or the 50th percentile (separates the data set in two halves)
- The third quartile (designated Q3) also called the upper quartile or the 75th percentile (separates the highest 25% of data from the lowest 75%)

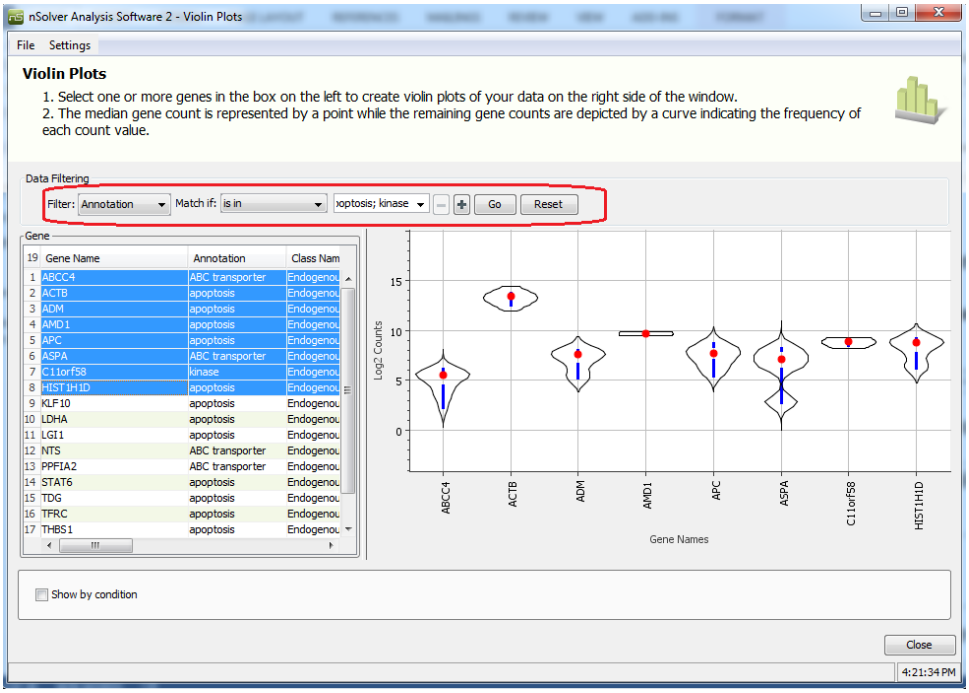
An example violin plot with annotations is displayed below.



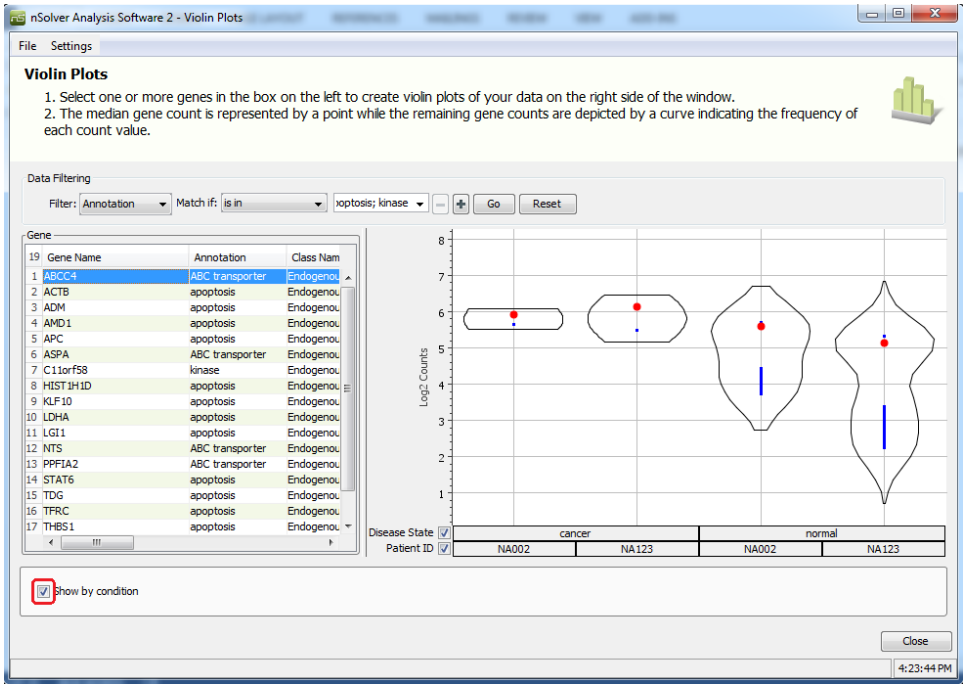
Selecting genes of interest from the left side of the Violin Plots page dynamically builds the plot on the right side.



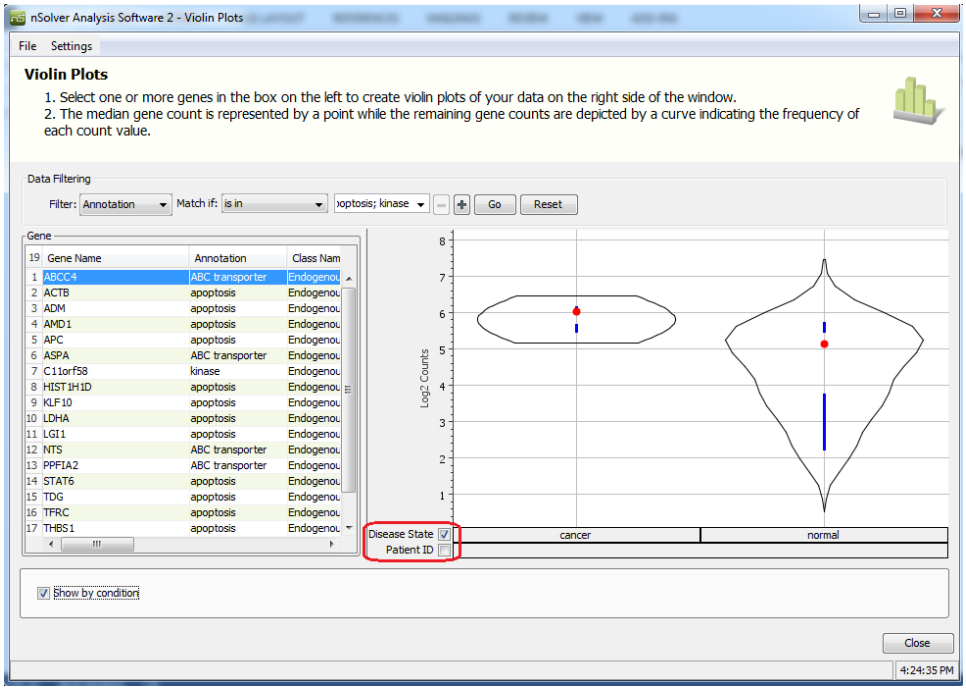
The Violin Plots page contains several features you can use to quickly identify the genes you want to plot. The first is the Data Filtering function at the top of the page that is available throughout nSolver. You can always sort on the various columns by clicking the column header. In addition, filter based on a list of genes, or if gene annotations are available as described in *Chapter 6 – Adding Custom Annotations to Genes*, filter based on a list of these as shown below.



nSolver also provides the capability to display the violin plot by condition. If experiment annotation columns were added to create treatment groups of data, this option allows for a useful visual of a single gene or small group of genes plotted across the various experimental annotations. Select the checkbox for **Show by condition** in the lower left corner, as shown below.



Notice that a row will appear for each annotation column that was added when setting up the experiment. Therefore, if there is a two-by-two factor experiment of disease and patient, the default view will show each of the four treatment groups. However, by selecting the checkbox as displayed below on one of the treatment groups, the plot changes to display only one of the factors. This allows the ability to see the expression between all conditions or just for a specified condition.

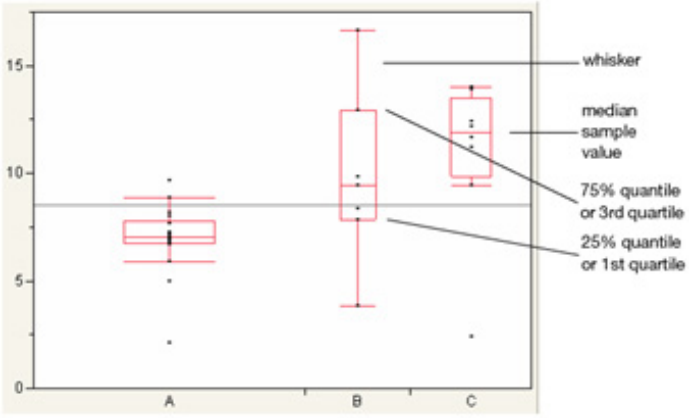


NOTE: When viewing the plot by condition, if multiple genes are selected, they will each be plotted on the same x-axis. Identify which gene belongs to which plot by hovering over each one with the cursor.

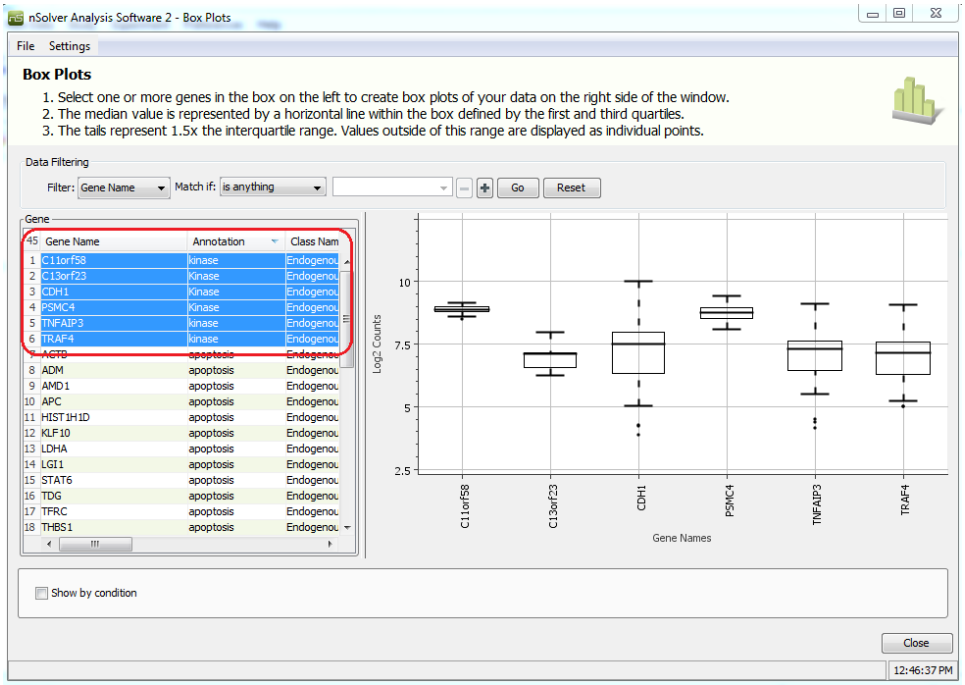
Box Plots

Box plots are a convenient way of depicting subsets of your experiment through their quartiles. Box plots have lines extending vertically from the boxes (whiskers) that indicate variability outside the upper and lower quartiles. Outliers may be plotted as individual points. Box plots display differences between subsets of an experiment without making any assumptions about the underlying statistical distribution; they are non-parametric.

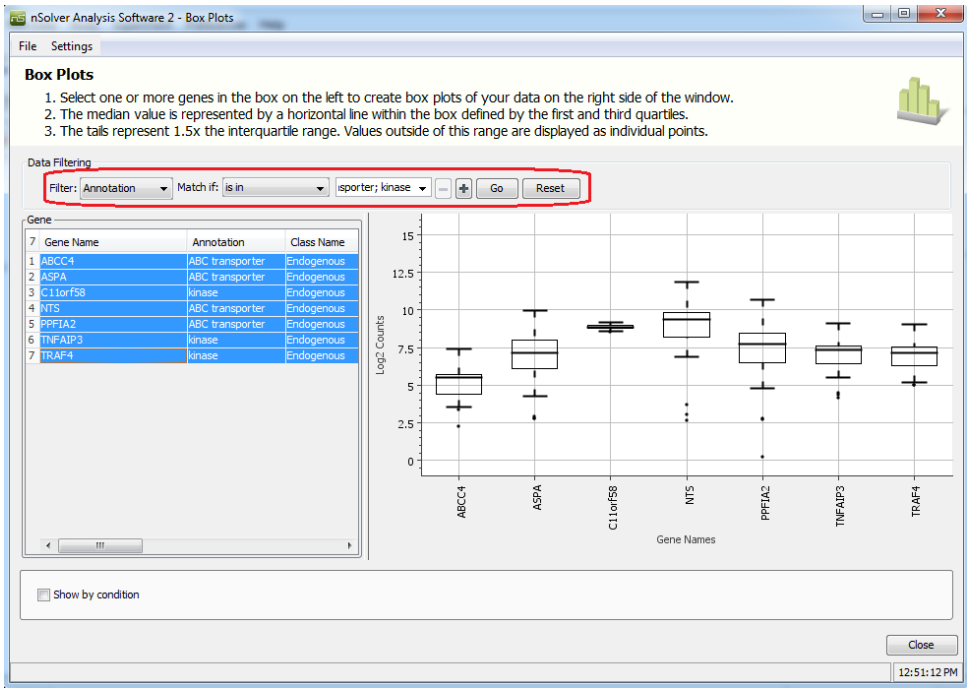
An example of a box plot with annotations is displayed below.



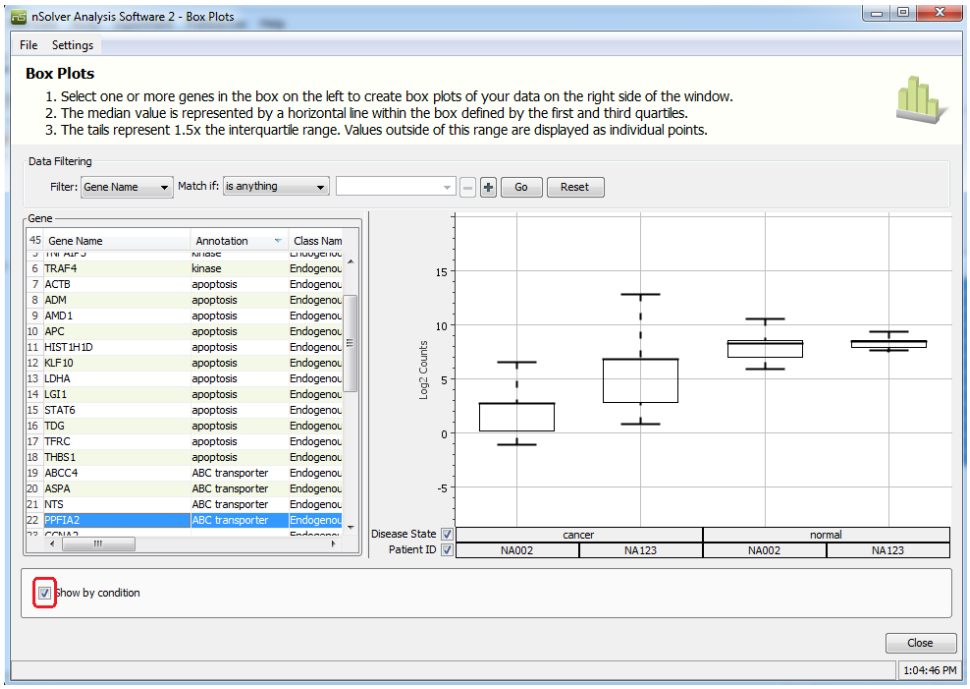
Box plots function similarly to violin plots. Select genes of interest from the left side to dynamically build the plot on the right side.



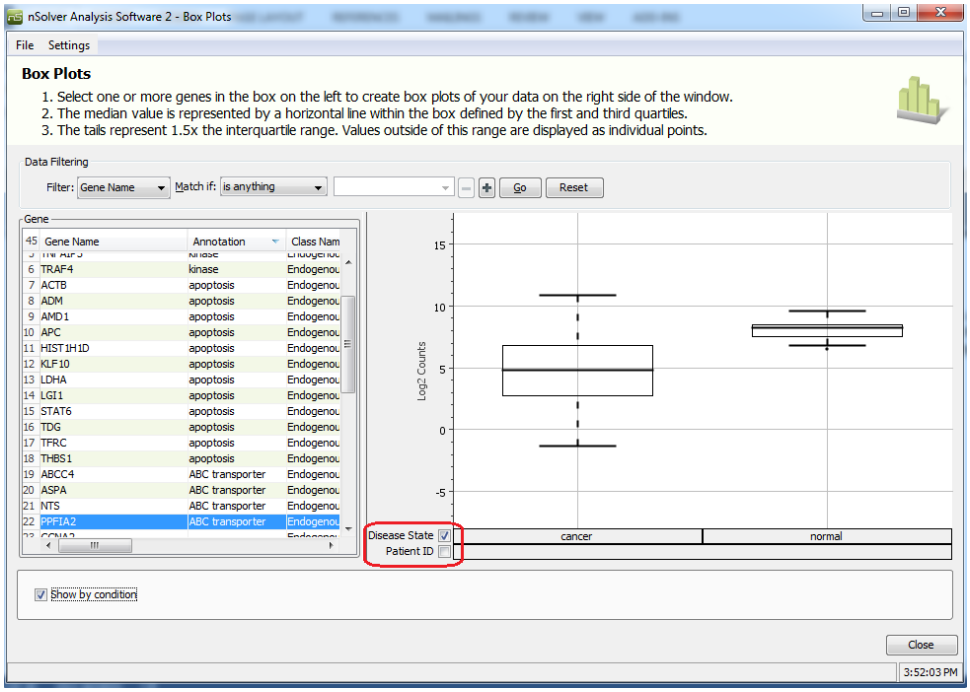
The box plot has several features you can use to quickly identify the genes you want to plot. The first is the Filter function at the top that is available throughout nSolver. Filter based on a list of genes, or if gene annotations have been added as described in *Chapter 6 – Adding Custom Annotations to Genes*, filter based on a list of these as shown below.



You can also display the box plot by condition. If you added experiment annotation columns to create treatment groups of data, you can use this option to create a useful visual of a gene or small group of genes plotted across the various conditions. Select the checkbox for **Show by condition** in the lower left corner, as shown below.

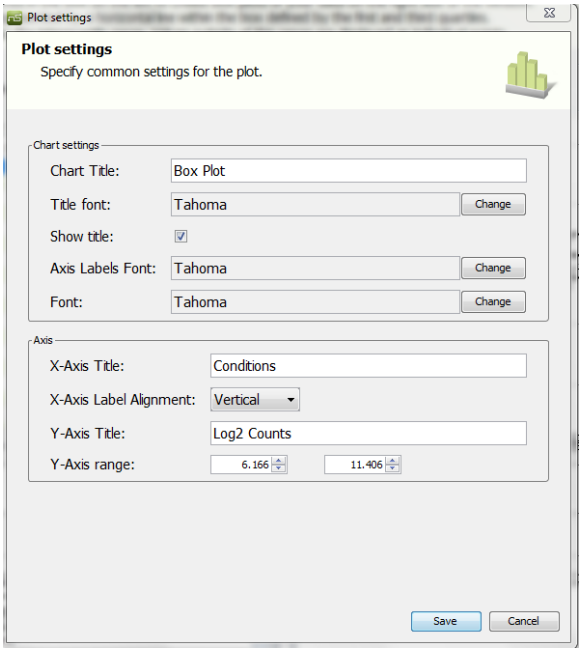


Notice that a row appears for each annotation column that was added when setting up the experiment. As such, if there is a two-by-two factor experiment of disease and patient, the default view shows each of the four treatment groups. However, by selecting the checkbox as displayed below on one of the treatment groups, the plot changes to display only one of the factors, allowing you to see the expression between all conditions or just for a specified condition.

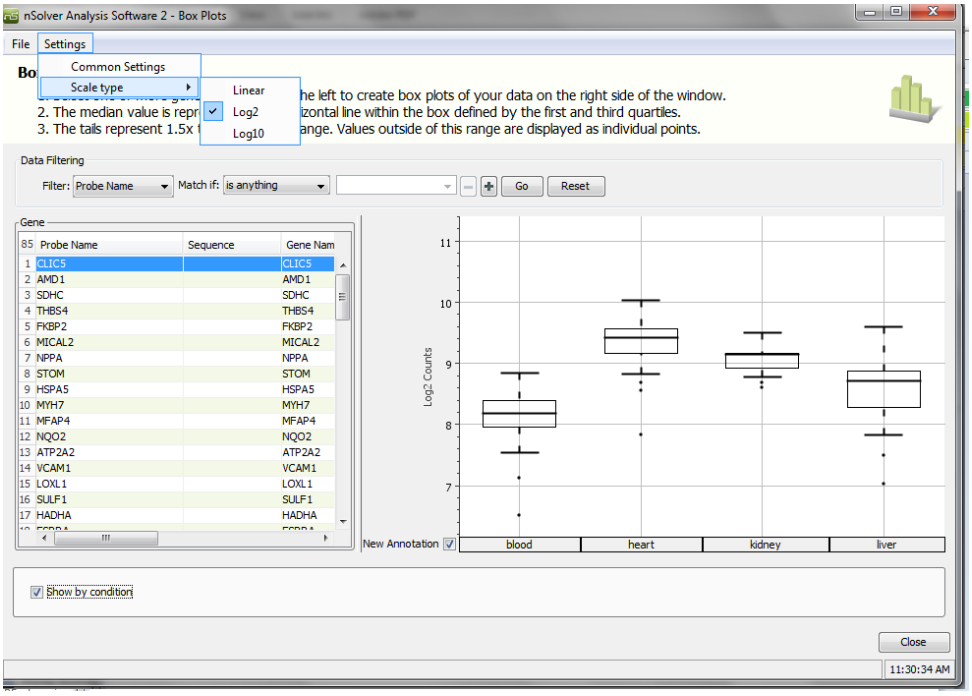


NOTE: When viewing the plot by condition, if you select multiple genes, they will each be plotted on the same x-axis. Identify which gene belongs to which plot by hovering over each one with the cursor.

The **Settings** menu on the top left field of the screen allows users to specify plot settings, including chart title, axis font, and axis label alignments.

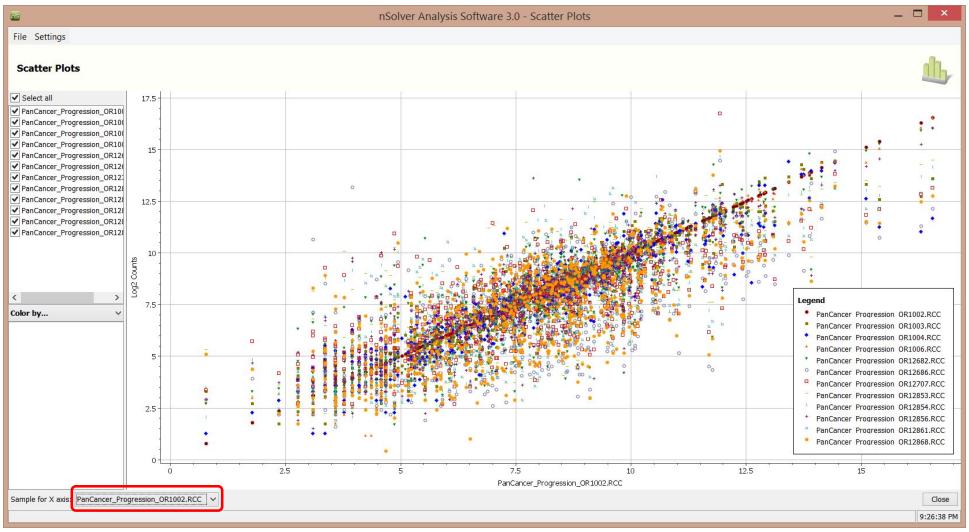


The **Settings** menu also allows users to specify **Scale type**. Users can choose to display box plots using linear count values, $\log_2$ values, or $\log_{10}$ values of the data for the selected genes.

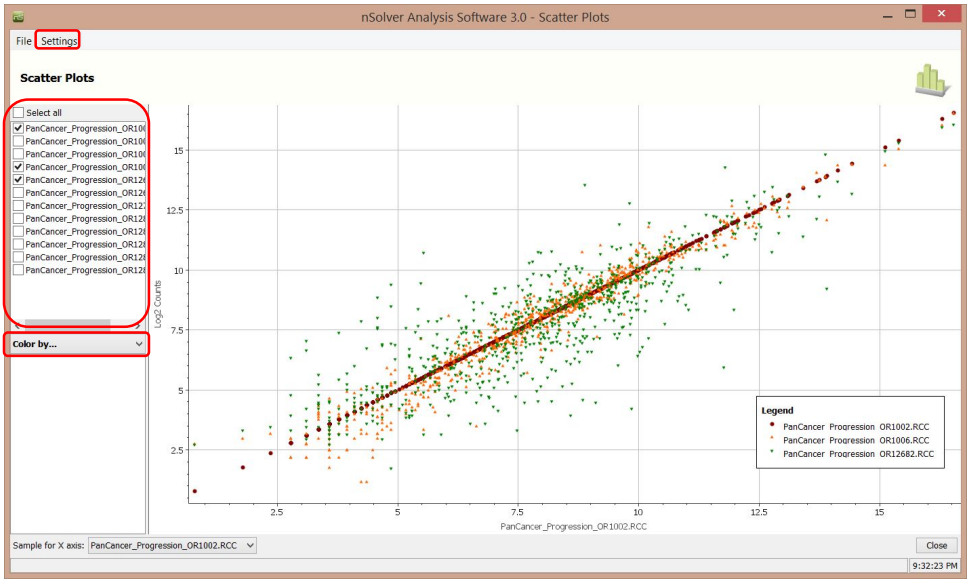


Scatter Plot

The scatter plot for raw, normalized, and grouped data is a visualization that plots one of the samples on the x-axis and the remaining samples on the y-axis. Choose a new x-axis sample by selecting from the drop down list as shown below.



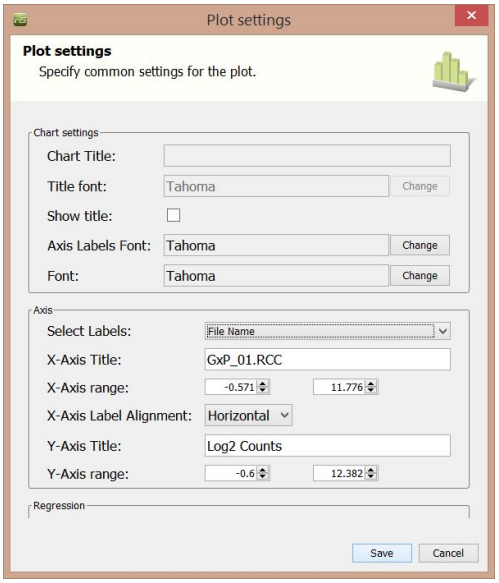
Adjust the plot dynamically to add or remove samples by selecting the checkboxes on the y-axis. This scatter plot has three samples selected:



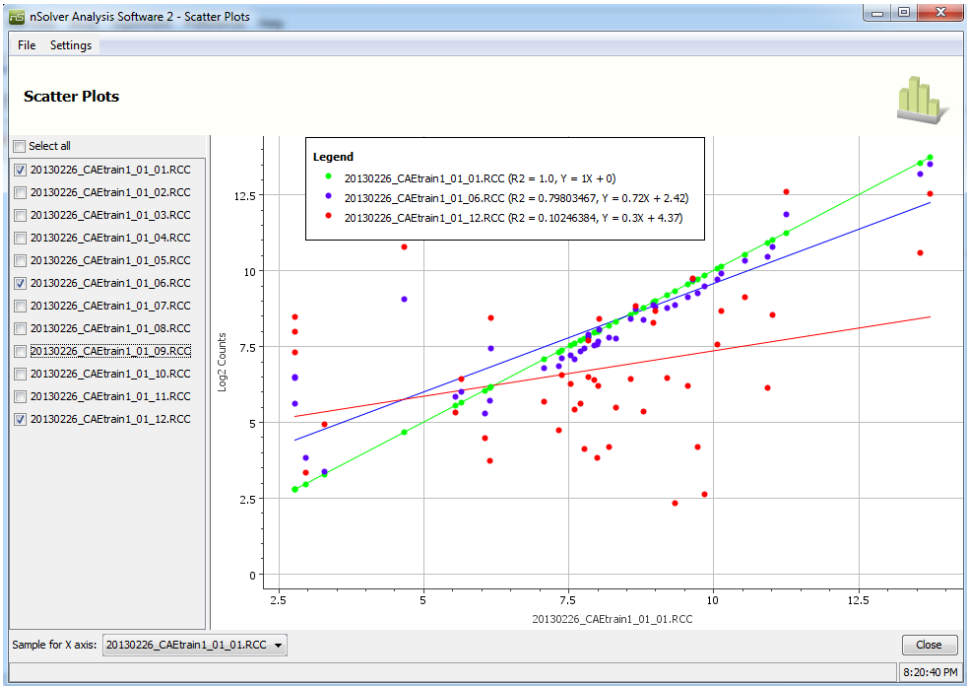
Use the **Color by…** dropdown to color the scatter plot by Code Class, Gene Annotation, or Normalization Reference. When you select a **Color by…**option, it will affect your options for the **Legend and data point settings**.

The mouse scroll wheel allows you to zoom in and out of the scatter plot.

Use the Settings dropdown to customize the appearance; change the common settings, change the legend and data point settings, and select the scale type.

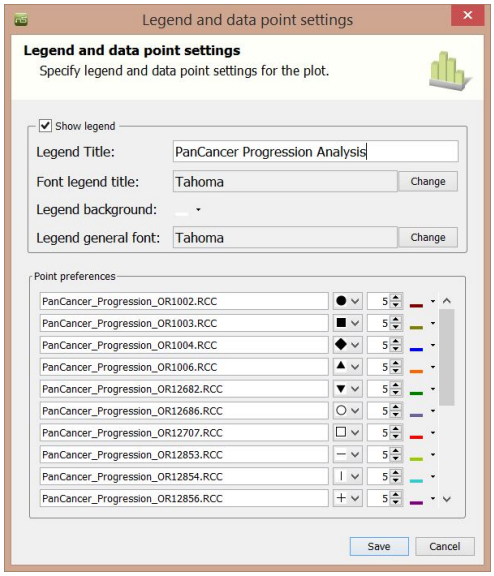


For scatter plots, the Common Settings option includes **Chart settings** to specify and customize chart titles, axis labels and axis titles. Display the linear regression by selecting the **Show regression** checkbox at the bottom of the Common Settings window. When this is checked, the plot will change to display the linear regression line. Also, one may choose to display the R^2 value and the equation of the line on the plot.

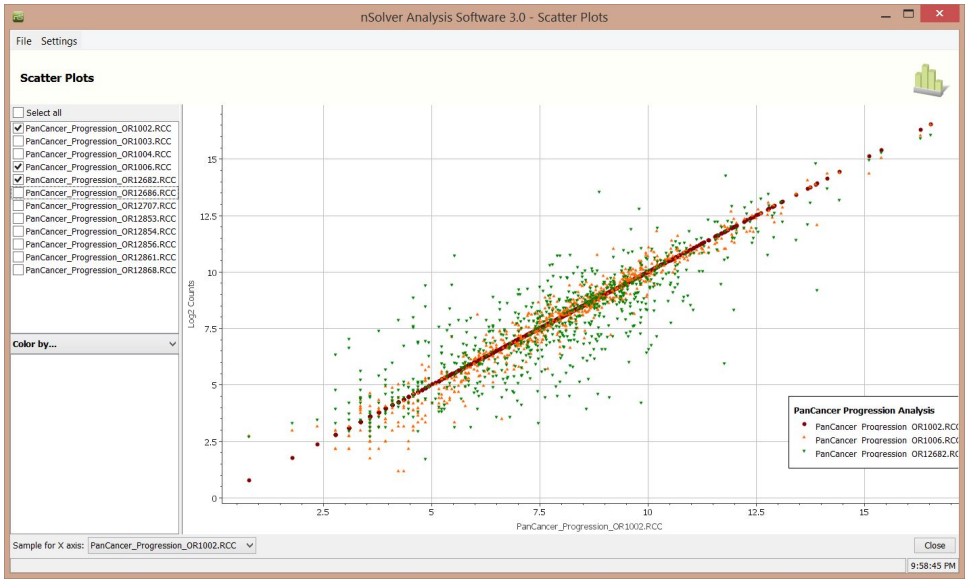


You can also use the Common Settings option to display your samples using labels other than file name. Choices include Description, Treatment Group, and Sample Name.

Scatter plots have a legend that indicates the samples or ratios plotted in the analysis. To customize this legend, select **Settings** from the top menu bar and choosing **Legend and data point setting**. The **Legend and data point settings** dialog box allows you to enter a title for the legend and customize the font style font and color, and background color for the scatter plot legend. It also allows you to specify the shape and color of data points to visually distinguish between samples. This allows data points to be distinguished by Code Class, whether or not they were used for Normalization, or if they have custom annotations.



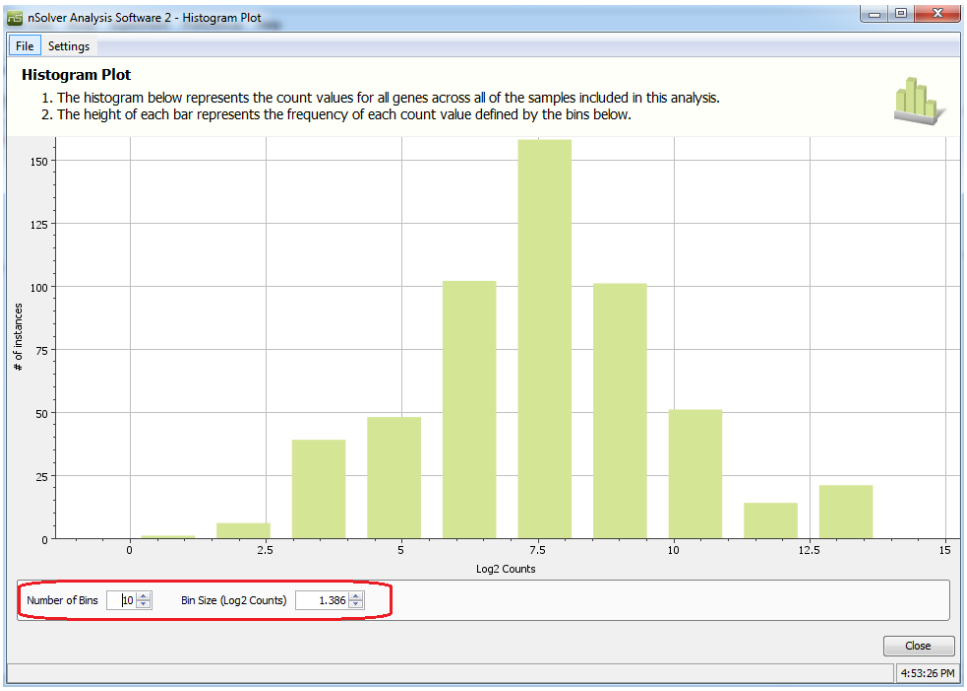
After changing the legend and data point settings, click the **Save** button to display the updated plot.



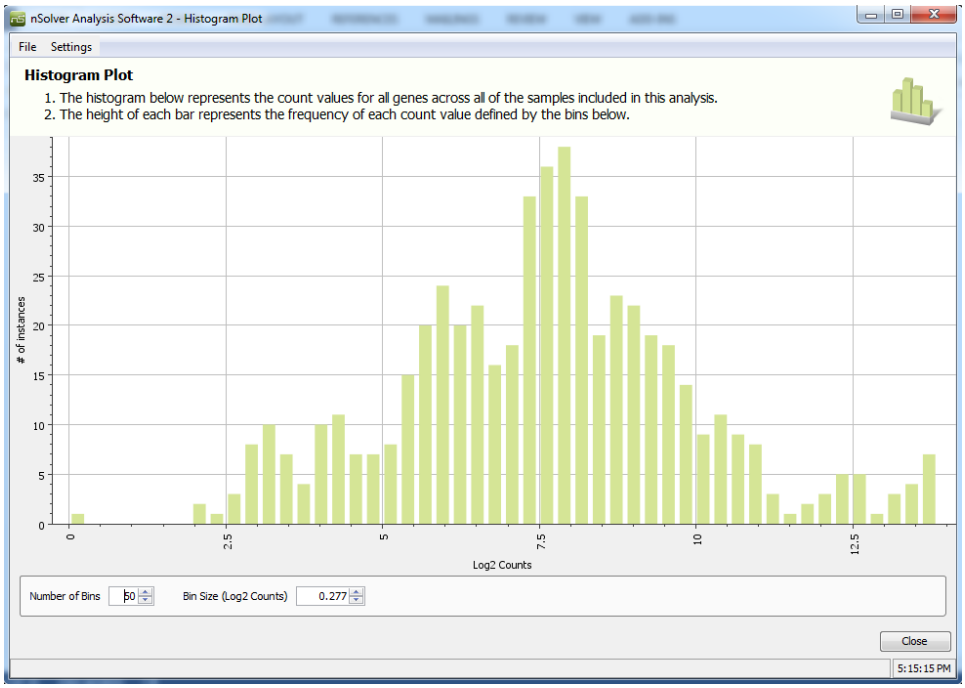
As noted earlier, that it is important to Save Settings to return to the plot later with the same settings intact.

Histogram Plot

The Histogram Plot represents the count values of all genes across all samples included in the analysis. The height of each bar represents the frequency of each count defined by the bins. Similar to all visualizations in nSolver, this plot is also available at all data levels. Below is an example of a Histogram plot on raw data.



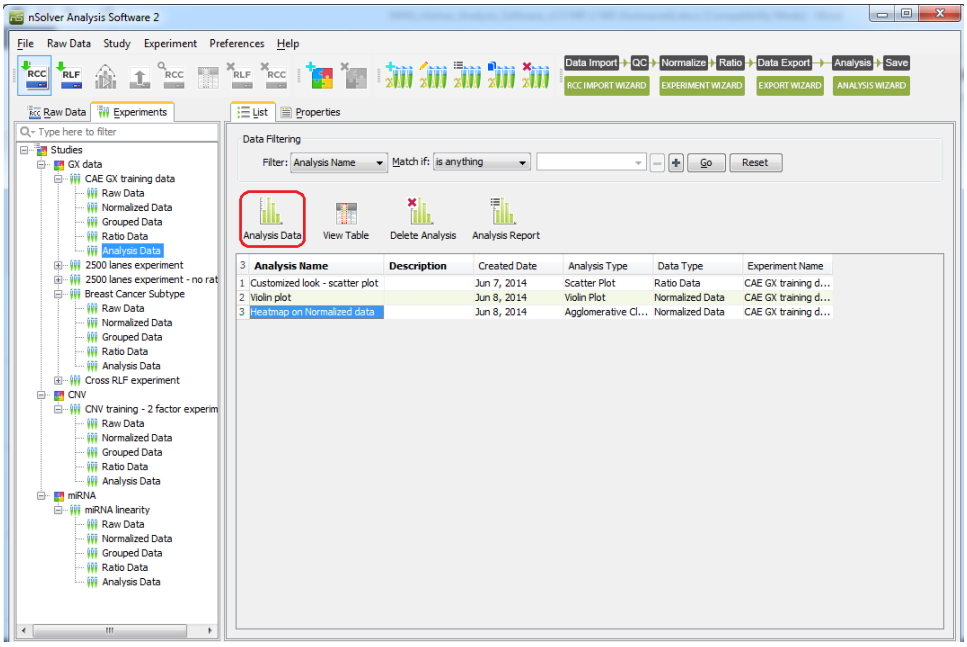
You can display a more granular view by adjusting the **Number of Bins** or the **Bin Size** at the bottom.



Managing Saved Analyses/Visualizations

Viewing Analyses

All analyses are saved automatically upon completion of the Analysis Wizard. Analyses are saved per experiment and can be accessed from the navigation tree at the Analysis Data level. To view a previously created analysis, select the desired analysis and click the **Analysis Data** icon as shown below. The saved visualization will appear.



Editing Analyses

The name and description of an analysis can be edited in the main view after selecting Analysis Data form the navigation tree. Simply double-click on the name or description of the analysis (bolded headings indicate values in that column are editable) and type in the new text. These fields will automatically be saved.

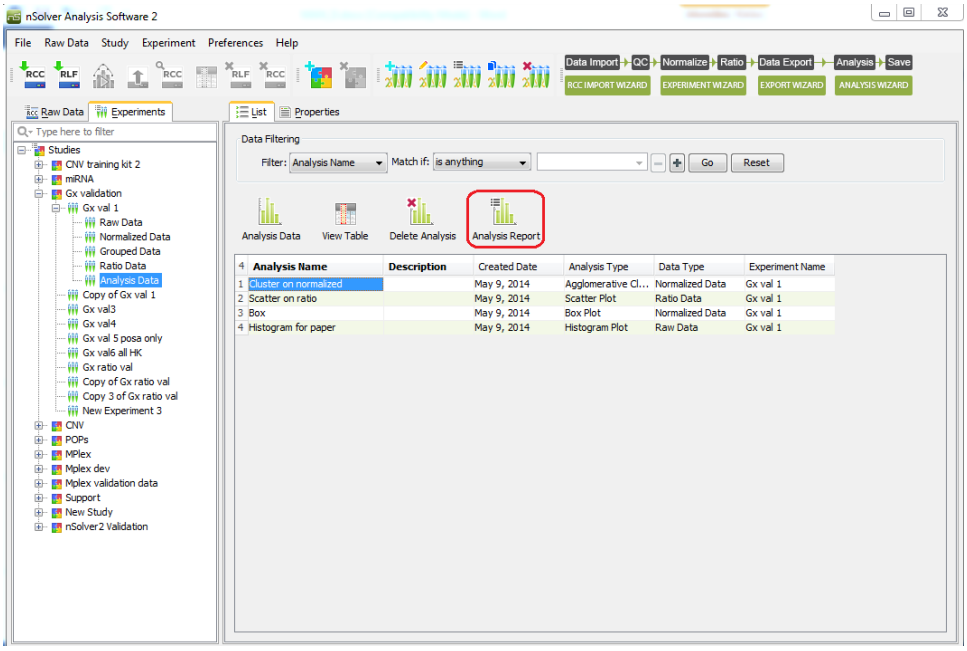
Analyses can also be deleted. To delete an analysis, select that analysis and then click the **Delete Analysis** icon. The analysis will be deleted.

Analysis Report

An Analysis Report is available with formatting and content similar to the Experiment Report. This report contains all information about the data used for the analysis as well as the parameters used in the analysis.

To view or print the Analysis Report, select the experiment from the navigation tree on the left side, navigate to the Analysis Data view, and click on the **Analysis Report** icon.

The report opens in HTML format. Information in the report varies depending on the analysis.



Advanced Analyses

Advanced analyses are a new kind of sample analysis introduced in nSolver 2.5. Rather than perform the analysis and visualization of sample data within nSolver, the data are automatically exported to an open source statistical environment called "R." nSolver is used to adjust the settings for each advanced analysis as well as to manage and launch the resulting visualization(s). No familiarity with R is necessary to use the advanced analyses supplied by NanoString.

Third-party Software Requirements

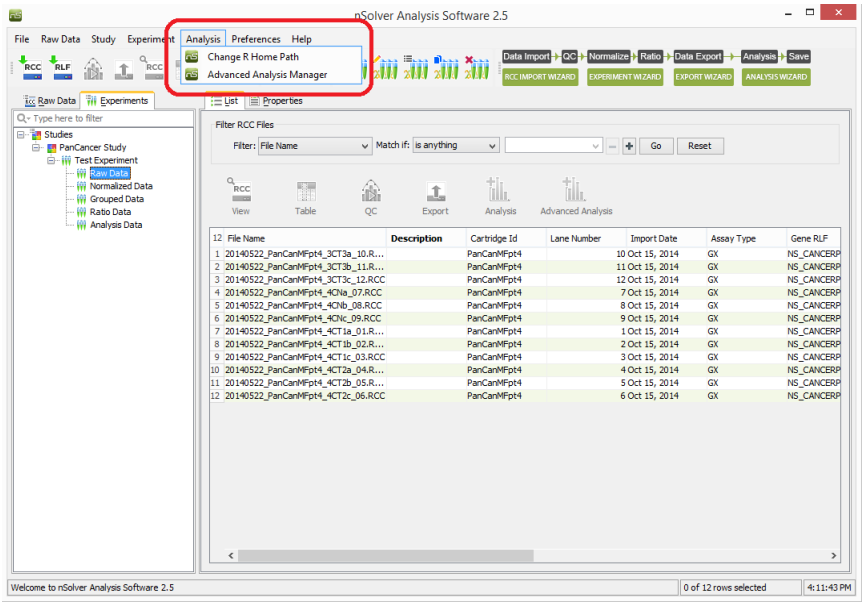
Before performing any advanced analysis, R must be installed on the computer. Users with a Windows PC will be prompted to install R during the installation of nSolver 2.5 or higher; the two applications are packaged within the same file. Users who skip this step or who operate a Macintosh computer will need to install R separately. Information about R and links to download the application for either Windows or Macintosh are available at http://www.r-project.org/.

Most Macintosh users will also have to install XQuartz, a second open source application that provides libraries required for certain R packages to function. Recent versions of the Macintosh OS X operating system do not include these libraries, so NanoString recommends installing XQuartz on systems operating OS X v10.9 and newer (this includes Mavericks and Yosemite). Information about XQuartz and a link to download the application for Macintosh are available at http://xquartz.macosforge.org/landing/.

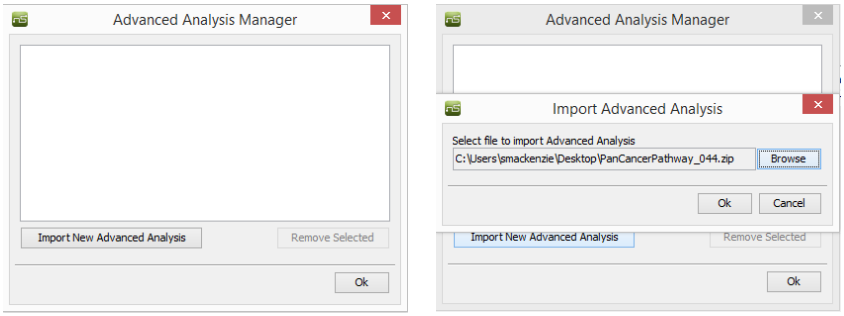
Installing New Advanced Analyses

The instructions (analysis module) for each type of advanced analysis must also be uploaded to nSolver. These are provided by NanoString as a compressed *.ZIP file. Do not extract the files before uploading them to nSolver.

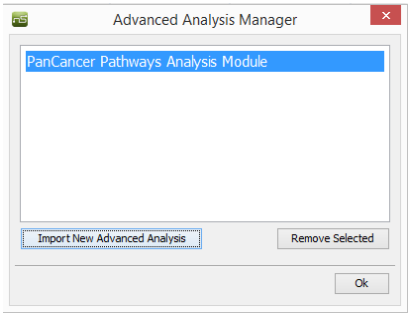
1. Open the **Analysis** menu at the top of the window and select **Advanced Analysis Manager**.



2. The Advanced Analysis Manager will open, displaying a list of existing analysis modules. Click the **Import New Advanced Analysis** button and use the file manager to select the *.ZIP file that contains the new analysis module.



3. The analysis module will be added to the list within the Advanced Analysis Manager. Click **Ok** to return to nSolver.

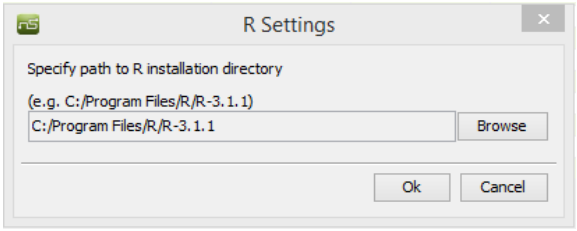


4. If you wish to remove an analysis module, open the Advanced Analysis Manager as before, select that analysis module, and click **Remove Selected**.

Changing the R Home Path

Some users may choose to change the R home path, the directory where the R application is installed. This is an optional feature but can be useful in some circumstances. For example, some users may wish to install the application in an unusual location, or they may wish to install multiple versions.

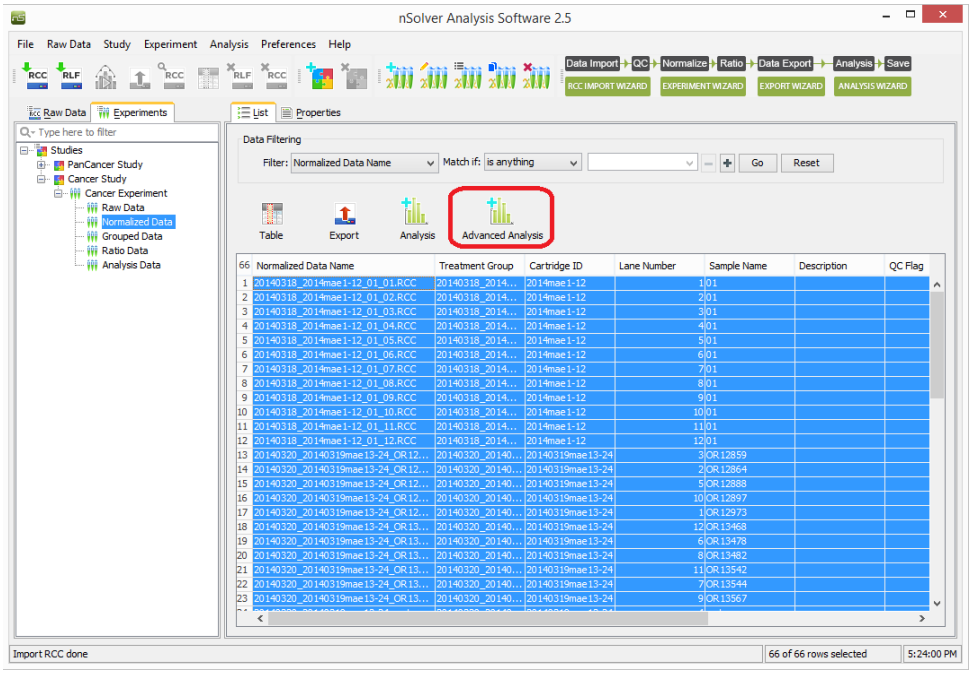
To change the R home path, open the Analysis menu at the top of the window and select Change R Home Path. A new window will appear.



Browse to the desired directory and then select **Ok**.

Advanced Analysis Wizard

Advanced analyses can only be performed at the raw or normalized data level. Other data levels (Grouped Data or Ratio Data) cannot be used. To run an advanced analysis, navigate to the raw or normalized data level using the navigation tree on the left. Then select the desired rows from the table and click the **Advanced Analysis** icon.



NOTE: If you have selected data that share the same annotation column heading more than once, then these column headings may be renamed in the analysis results. Duplicate annotations will be appended with a period and a number to distinguish them, e.g., “Group” will become “Group.1”. However, do not create your own column headings that use this period-number format.

NOTE: If a column name begins with a number, an underscore, or a period and then a number, R will prepend “X” to that column name, e.g., “1Group” will become “X1Group”.

NOTE: The following illegal characters may not be used for column names:

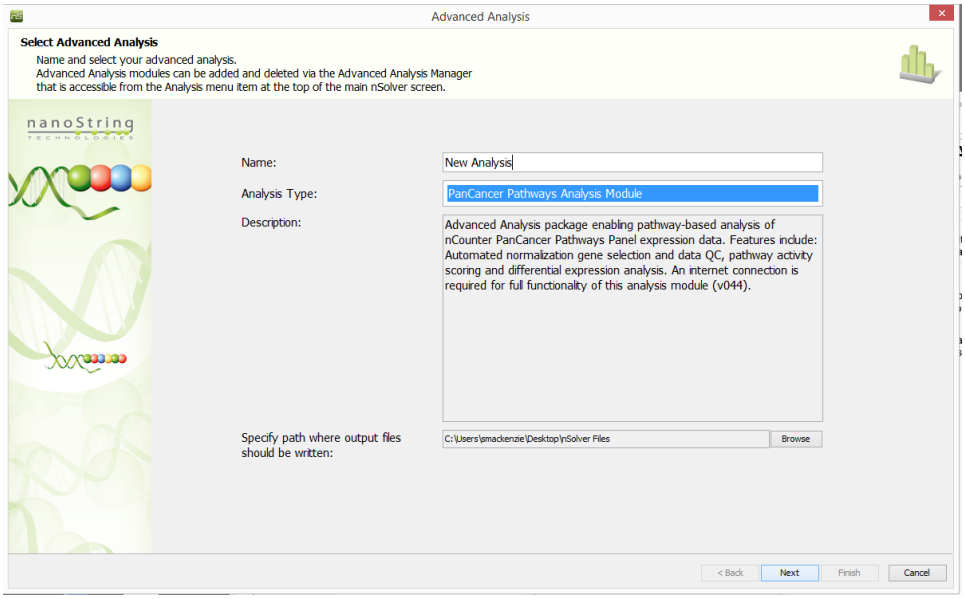
! @ # $ % ^ & * ( ) - + = ` ~ < > , . : ; ‘ | [ ] { }

If these characters are found, they will be automatically converted to a period, e.g., “sample%1” will become “sample.1”.

However, multiple illegal characters in a series will be converted to only a single period.

Advanced Analysis Properties

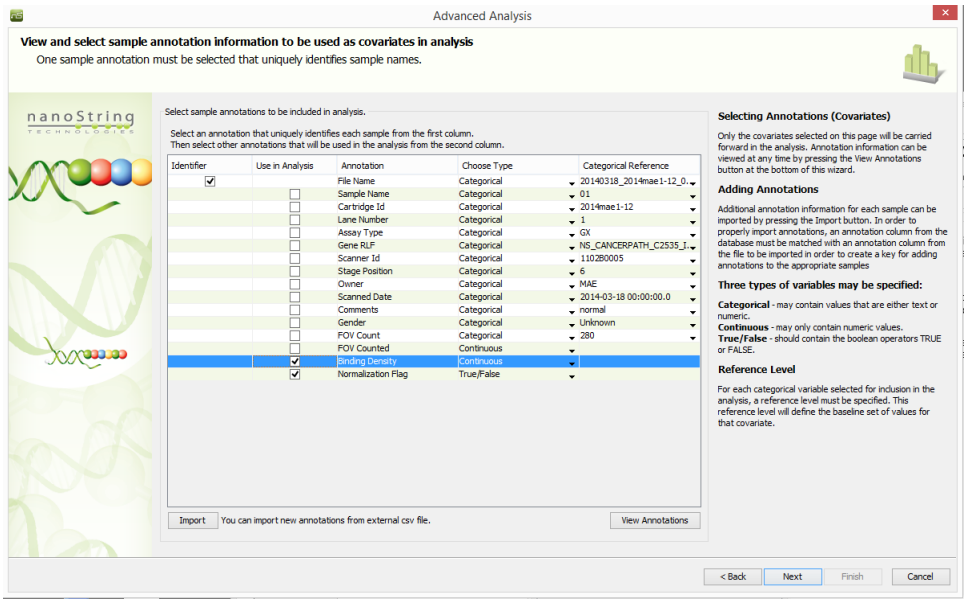
The Advanced Analysis Wizard will launch. The first screen will ask you to name this analysis and select a type of advanced analysis from the **Analysis Type** menu. A description of that advanced analysis will be displayed below but cannot be edited (this description is included in the advanced analysis module before uploading it to nSolver). Select an output path for the files generated by the advanced analysis by clicking the **Browse** button. When ready, click **Next**.



NOTE: Some advanced analyses are only compatible with certain CodeSets. Incompatible advanced analyses will appear in grey text and cannot be selected. Choose an advanced analysis from the remaining options. If no advanced analyses are available for the selected data, an error message will appear. Select different data, or confirm that you created your experiment using the correct CodeSet.

Sample Annotations

The next screen requires selecting sample annotations to be used as covariates during the advanced analysis.



A checkbox will appear in the **Identifier** column only if a unique annotation exists for every lane/sample in the data selected for this advanced analysis. At least one annotation must be selected as a unique identifier to track data.

The file name can be used as the identifier since RCC data for each lane/sample is saved in its own file. Other annotations, such as description or sample ID, may be available depending on the nature of the experiment. If you wish to add your own custom annotations, click the **Import** button to add new annotations from a *.CSV file. The next section, *Import Annotations*, provides more detailed instructions.

After selecting an identifier, choose which annotations will be used in the advanced analysis by selecting the checkboxes in the **Use in Analysis** column. Some advanced analyses may require that one or more annotations be selected before continuing. nSolver will automatically classify each of the default annotations as continuous, categorical, or true/false, but you may change these settings using the drop down menu next to each annotation included in the analysis. If one of these is changed to a categorical annotation, nSolver requires that you also specify a Categorical Reference to be used as a baseline for some analyses.

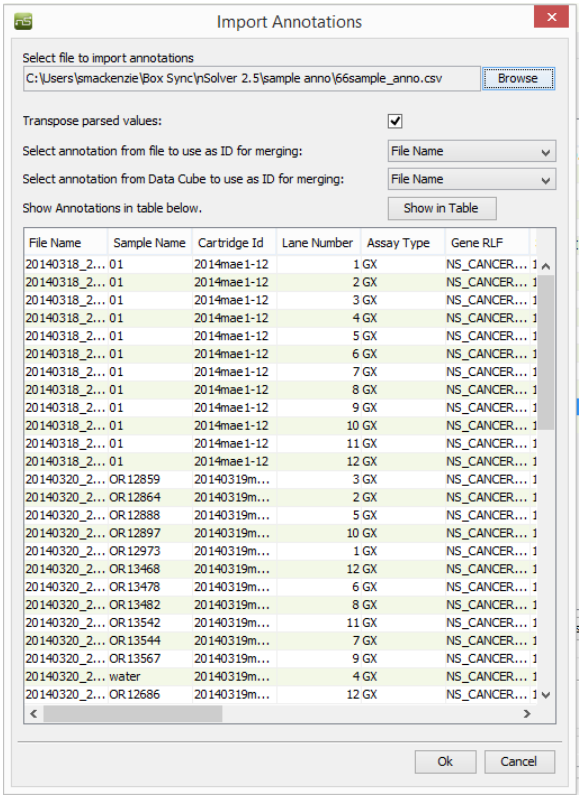
Although this window displays only the annotation heading, the **Preview Annotations** button at the bottom of the page will display all data so that you can see the annotations for each lane/sample.

Click **Next** when ready to proceed.

Import Annotations

If you wish to add custom annotations to an advanced analysis, click the **Import** button on the Sample Annotations screen from the previous section.

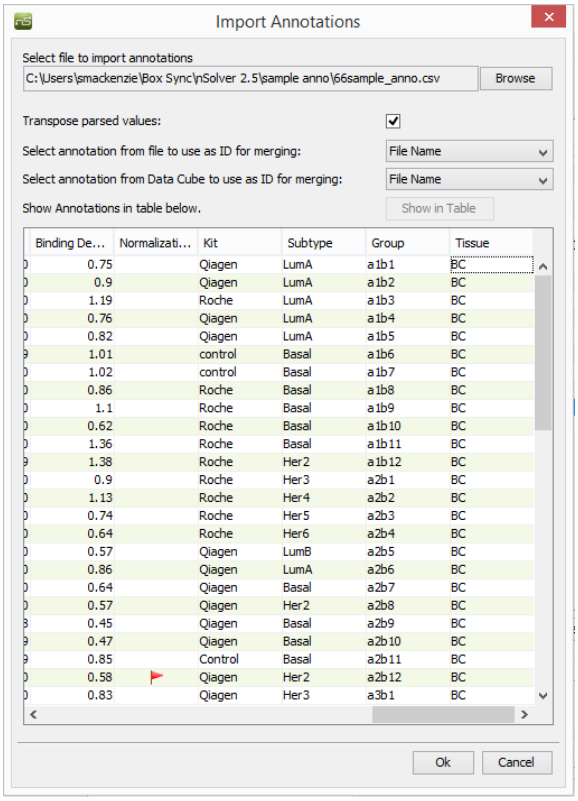
Click **Browse** and navigate to the *.CSV file that contains the desired annotations.



Additional sample annotations may be as simple as numbering lanes/samples but can also provide more specific information about the assay. In the example below (opened in Microsoft Excel®), the “File Name” and “Group” annotations are unique for each sample. However, the “Kit,” “Subtype,” and “Tissue” annotations contain some variables that are repeated for multiple lanes/samples.

	A	B	C	D	E
1	File Name	Kit	Subtype	Group	Tissue
2	20140318_2014mae1-12_01_01.RCC	Qiagen	LumA	a1b1	BC
3	20140318_2014mae1-12_01_02.RCC	Qiagen	LumA	a1b2	BC
4	20140318_2014mae1-12_01_03.RCC	Roche	LumA	a1b3	BC
5	20140318_2014mae1-12_01_04.RCC	Qiagen	LumA	a1b4	BC
6	20140318_2014mae1-12_01_05.RCC	Qiagen	LumA	a1b5	BC
7	20140318_2014mae1-12_01_06.RCC	control	Basal	a1b6	BC
8	20140318_2014mae1-12_01_07.RCC	control	Basal	a1b7	BC
9	20140318_2014mae1-12_01_08.RCC	Roche	Basal	a1b8	BC
10	20140318_2014mae1-12_01_09.RCC	Roche	Basal	a1b9	BC
11	20140318_2014mae1-12_01_10.RCC	Roche	Basal	a1b10	BC
12	20140318_2014mae1-12_01_11.RCC	Roche	Basal	a1b11	BC
13	20140318_2014mae1-12_01_12.RCC	Roche	Her2	a1b12	BC
14	20140320_20140319mae13-24_OR12859_03.RCC	Roche	Her3	a2b1	BC
15	20140320_20140319mae13-24_OR12864_02.RCC	Roche	Her4	a2b2	BC

After selecting the desired annotations file, use the drop down menus to choose one annotation from the imported data and one annotation from the existing data. These two annotations must be identical (for example, “File Name”) and will be used to integrate the new annotations into the analysis. Click **Show in Table** to confirm that this matching process is successful and display which new annotations will be matched with which lanes/samples.

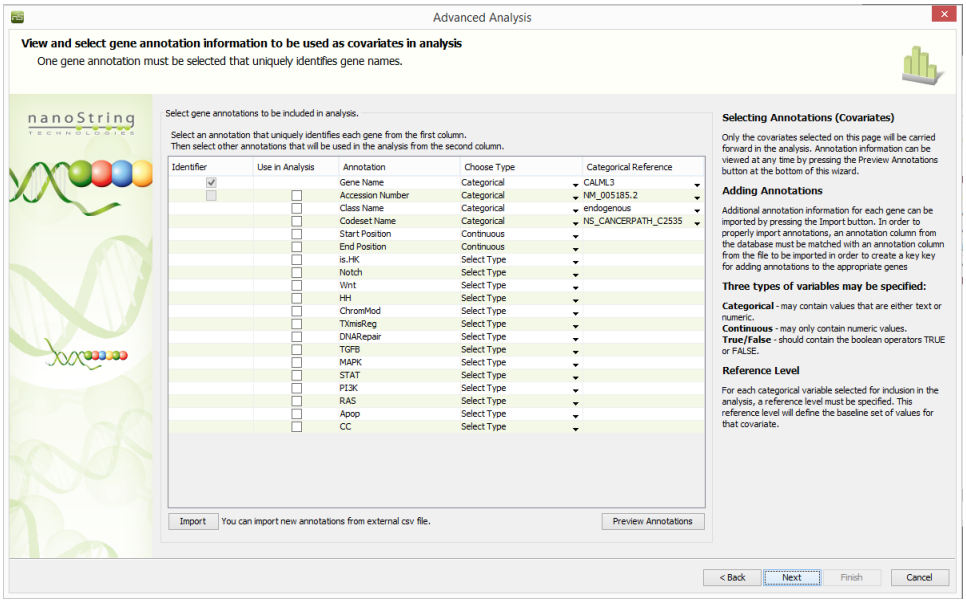


The checkbox next to **Transpose parsed values** is selected by default. If an error occurs when you click **Show in Table**, this is sometimes because of a transposition error (turning columns into rows and vice versa). Try unselecting the checkbox and try again. If the error persists, open the *.CSV file in an editor like Microsoft Excel® to verify the new annotations are correct and do not contain any omissions or typos.

When annotations have been successfully added, click **Ok** to return to the Advanced Analysis Wizard. You will be able to select these new annotations to be included in the analysis as described in the previous section. You will also need to choose the annotation type – continuous, categorical, or true/false – using the drop down menu next to each annotation included in the analysis. If one of these is a categorical annotation, nSolver requires that you also specify a Categorical Reference to be used as a baseline for some analyses.

Gene Annotations

Some advanced analyses include an additional screen that contains gene annotations associated with that analysis. (If the gene annotations screen does not appear, disregard this step and continue to the next section.) In this example, the gene annotations screen will enable you to select annotations such as specific pathway names.

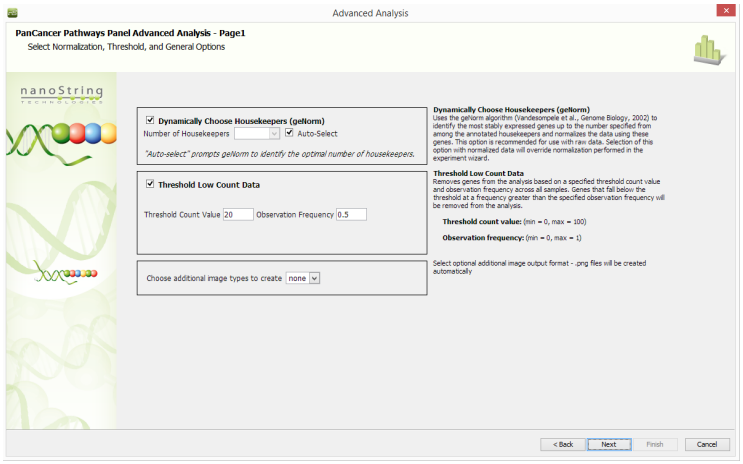


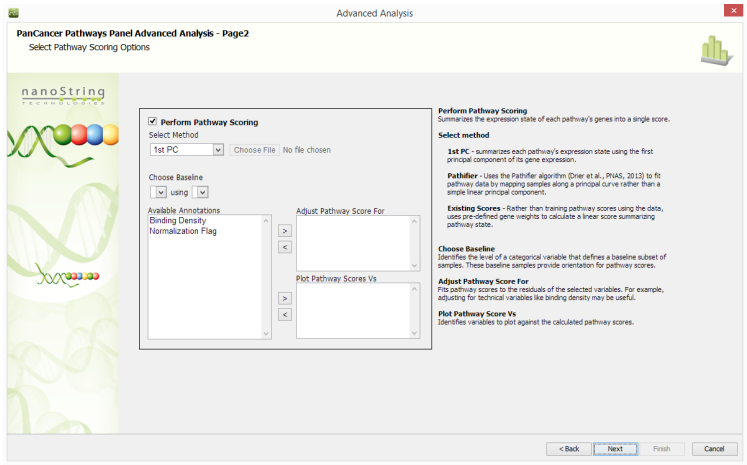
Choose which annotations will be used in the advanced analysis. Similar to the sample annotations screen, some default annotations may already be classified as specific annotation types. However, any gene annotations imported by the advanced analysis module will not have a defined type. You will need to choose the annotation type – continuous, categorical, or true/false – using the drop down menu next to each annotation included in the analysis. If one of these is a categorical annotation, nSolver requires that you also specify a Categorical Reference to be used as a baseline for some analyses.

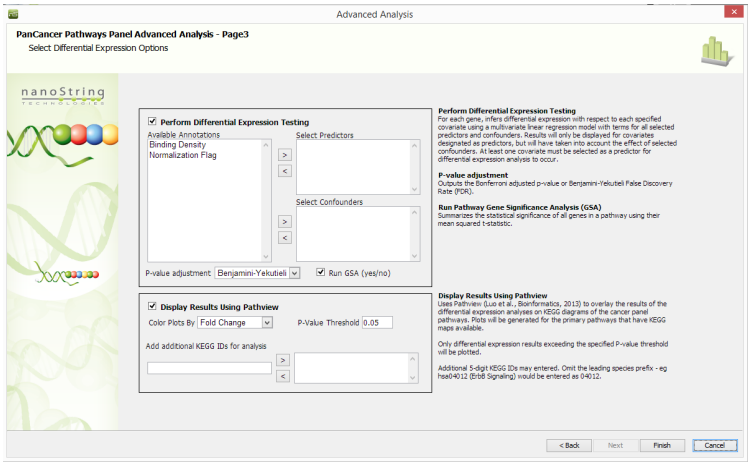
You may import additional gene annotations if desired. Click on **Import** at the bottom of the screen and select a *.CSV file following the same procedure described in the previous section, Import Annotations. Click **Next** when ready.

Custom Parameters for Advanced Analyses

Once the desired annotations are selected, the remaining pages contain parameters unique to each advanced analysis. These pages are in HTML format and the content varies depending on the instructions provided by the advanced analysis module for that analysis. Instructions and supplementary documentation for adjusting these settings will be provided separately with each module. Example screens are provided below for the PanCancer Pathways advanced analysis.





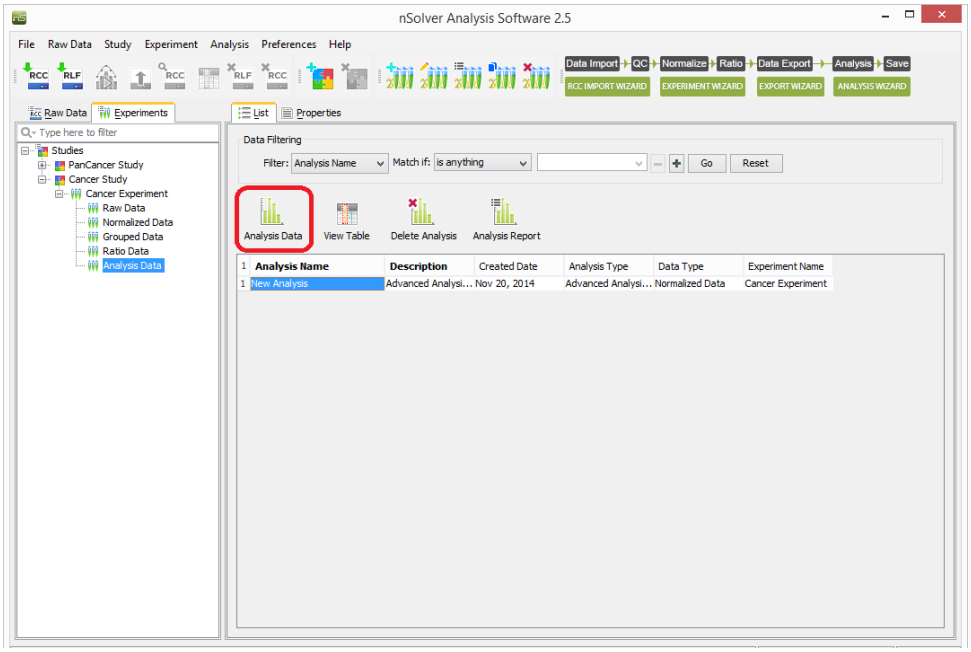


Once all the parameters have been selected, click **Finish**.

Managing Advanced Analyses

Viewing Analyses

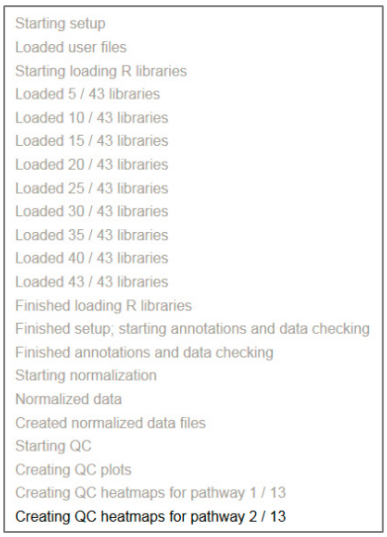
Each advanced analysis is organized within its parent experiment and can be accessed from the navigation tree at the Analysis Data level. Upon finishing the Advanced Analysis Wizard, nSolver will take you to this page automatically. Select the desired analysis and click on the **Analysis Data** icon as shown below.



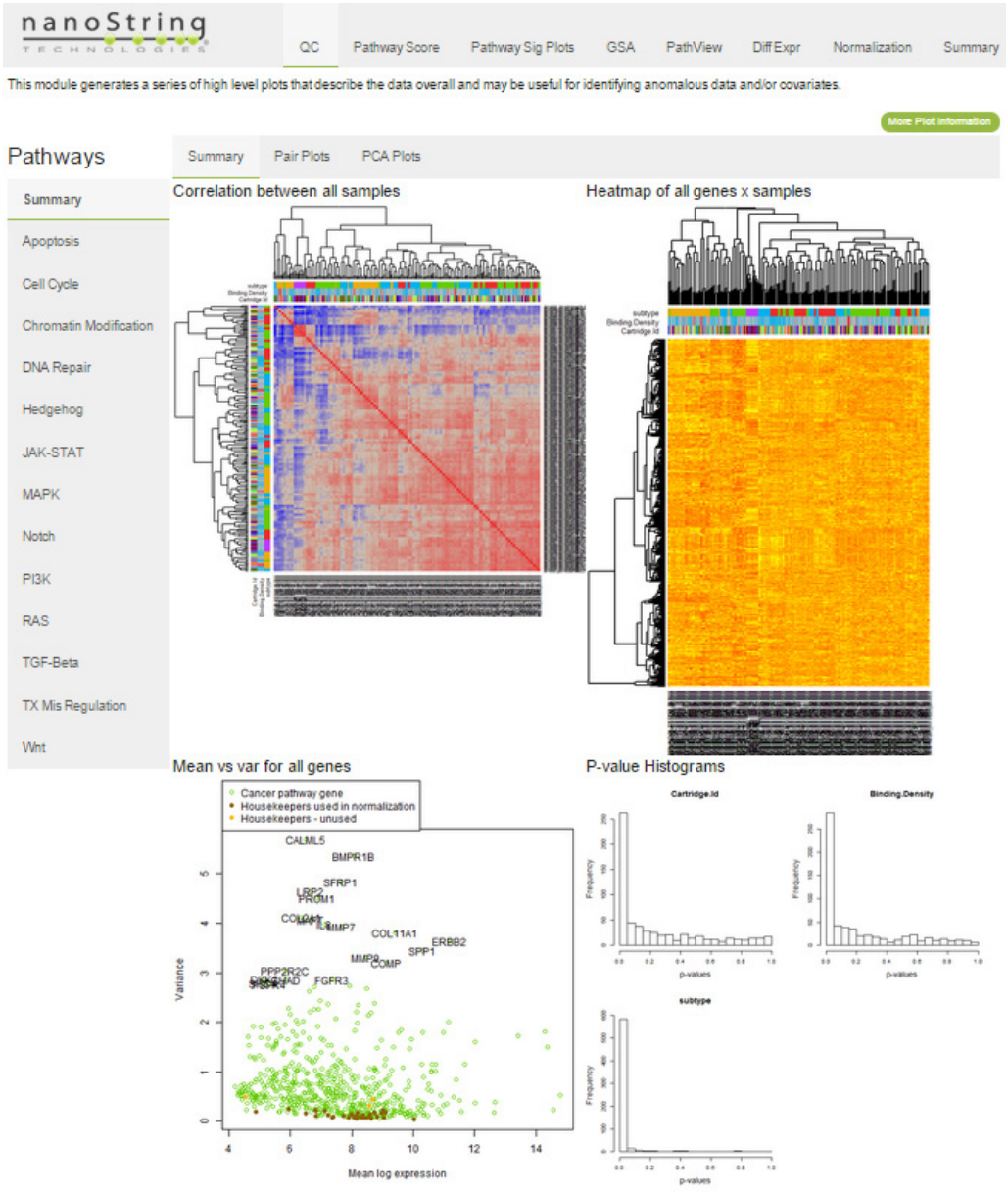
All advanced analyses are saved automatically in the file path specified at the beginning of the Advanced Analysis Wizard. However, the Analysis Data screen on nSolver is used only to launch the advanced analysis. Data files are generated by R and displayed in your default web browser.

NOTE: At a minimum, you must be connected to the Internet when the advanced analysis is initially run. Some analyses may require that you be connected to the Internet each time it is viewed, while others may cache files so that the completed analysis can be viewed offline in the future.

If the analysis has not yet completed running in the background, a progress report will be displayed.



If the analysis is complete, the saved visualizations will appear with links on the page that can be used to access different parts of the report.



Editing Analyses

The name and description of an analysis can be edited in the main view after selecting Analysis Data form the navigation tree. Simply double-click on the name or description of the analysis (bolded headings indicate values in that column are editable) and type in the new text. These fields will automatically be saved.

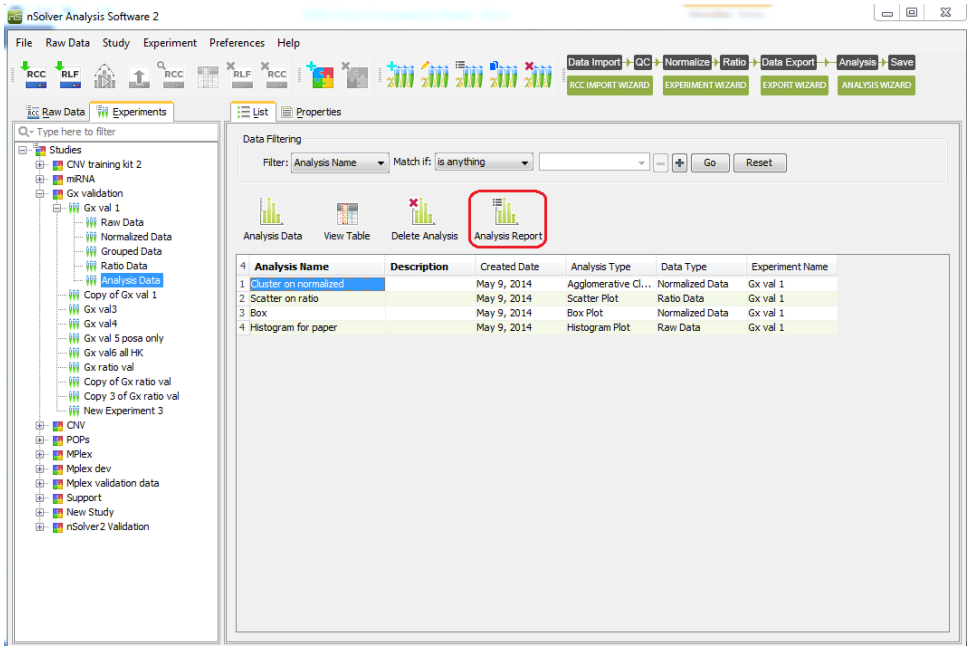
Analyses can also be deleted. To delete an analysis, select that analysis and then click the **Delete Analysis** icon. The analysis will be deleted.

Analysis Report

An Analysis Report is available with formatting and content similar to the Experiment Report. This report contains all information about the data used for the analysis as well as the parameters used in the analysis.

To view or print the Analysis Report, select the experiment from the navigation tree on the left side, navigate to the

Analysis Data view, and click on the **Analysis Report** icon.



The Analysis Report will open in HTML format in your default web browser. Information in the report varies depending on the analysis and contains all sample data, similar to reports for regular analyses. However, it does not contain all of the parameters entered in the Advanced Analysis Wizard. Those parameters are captured by the module that creates the advanced analysis and are typically available on a summary tab within the Analysis Data file.

Analysis Properties

Analysis Name:	New Analysis
Description:	Advanced Analysis package enabling pathway-based analysis of nCounter PanCancer Pathways Panel expression data. Features include: Automated normalization gene selection and data QC, pathway activity scoring and differential expression analysis. An internet connection is required for full functionality of this analysis module (v044).
Experiment Name:	Cancer Experiment
Analysis Type:	Advanced Analysis. Plugin name: 'PanCancer Pathways Analysis Module'
Analysis Plugin:	PanCancer Pathways Analysis Module
Data Level:	Normalized Data

Sample Annotations to use in Analysis

	File Name	Sample Name	Cartridge Id	Lane Number	Assay Type	Gene RLF	Scanner Id	F
1	20140318_2014mae1-12_01_01.RCC	01	2014mae1-12	1	GX	NS_CANCERPATH_C2535_INT	1102B0005	6
2	20140318_2014mae1-12_01_02.RCC	01	2014mae1-12	2	GX	NS_CANCERPATH_C2535_INT	1102B0005	6
3	20140318_2014mae1-12_01_03.RCC	01	2014mae1-12	3	GX	NS_CANCERPATH_C2535_INT	1102B0005	6
4	20140318_2014mae1-12_01_04.RCC	01	2014mae1-12	4	GX	NS_CANCERPATH_C2535_INT	1102B0005	6
5	20140318_2014mae1-12_01_05.RCC	01	2014mae1-12	5	GX	NS_CANCERPATH_C2535_INT	1102B0005	6
6	20140318_2014mae1-12_01_06.RCC	01	2014mae1-12	6	GX	NS_CANCERPATH_C2535_INT	1102B0005	6
7	20140318_2014mae1-12_01_07.RCC	01	2014mae1-12	7	GX	NS_CANCERPATH_C2535_INT	1102B0005	6
8	20140318_2014mae1-12_01_08.RCC	01	2014mae1-12	8	GX	NS_CANCERPATH_C2535_INT	1102B0005	6
9	20140318_2014mae1-12_01_09.RCC	01	2014mae1-12	9	GX	NS_CANCERPATH_C2535_INT	1102B0005	6
10	20140318_2014mae1-12_01_10.RCC	01	2014mae1-12	10	GX	NS_CANCERPATH_C2535_INT	1102B0005	6

12 Data Management

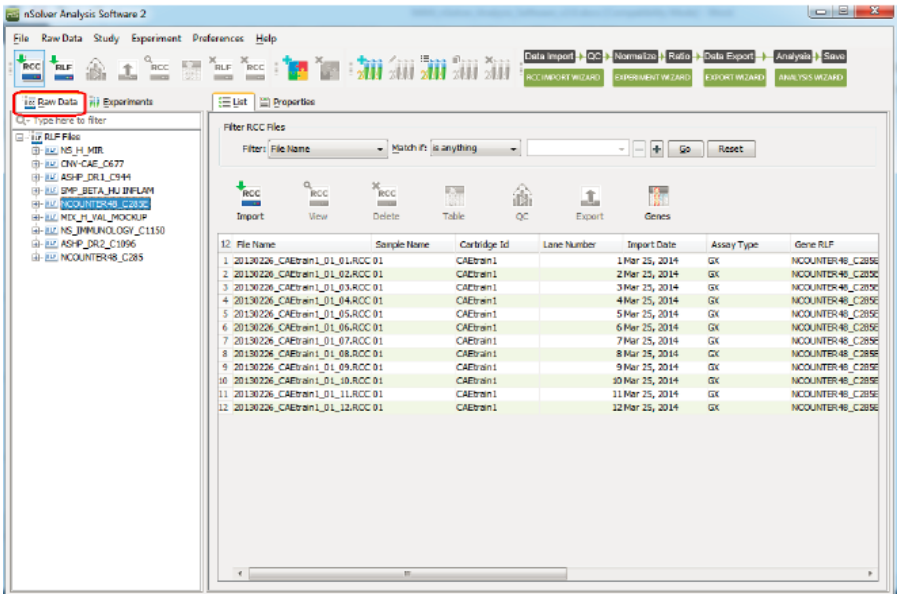
This chapter contains general information about basic features related to data management. It explains the different views/tabs and how to search for anything throughout the application. In addition, it provides some useful functions that can be used within tables such as sorting, customizing columns, displaying hidden fields, and copying or pasting to and from Excel®.

Views/Tabs

You can choose from two main views when working in the nSolver application (Raw Data and Experiment) and two minor views within each main view (List and Properties). You can toggle between the two main views by clicking on the tab above the tree on the left hand side.

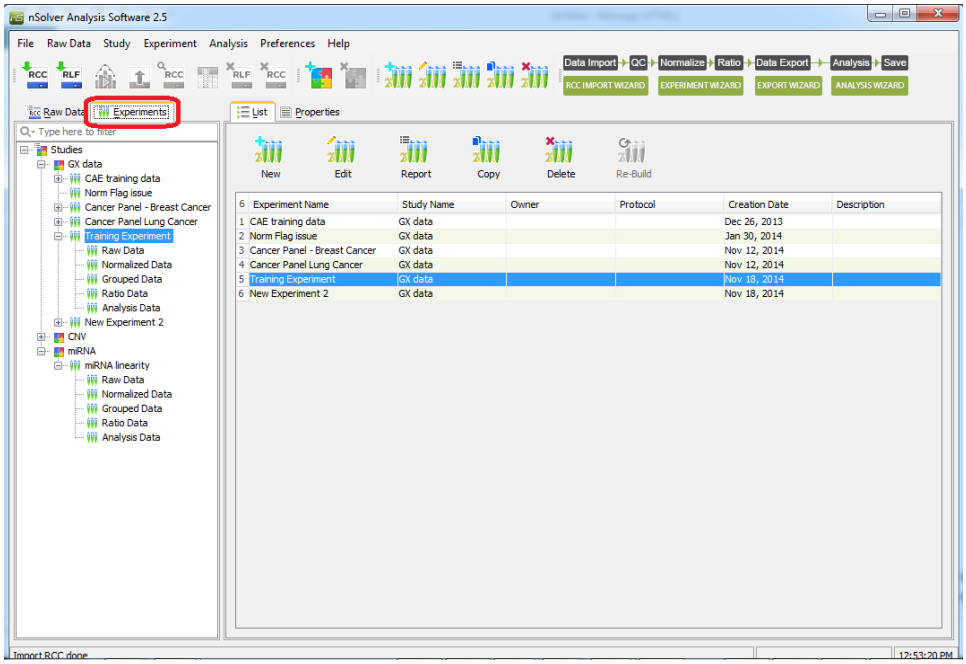
Raw Data Main View/Tab

The Raw Data Main View is the default view when the application opens. It contains all of the raw data files organized by CodeSet in the left navigation tree. When you select an item such as raw data or one of the CodeSets, those data files will be displayed in the panel on the right side of the window. Relevant icons will become active in the toolbar and menu.



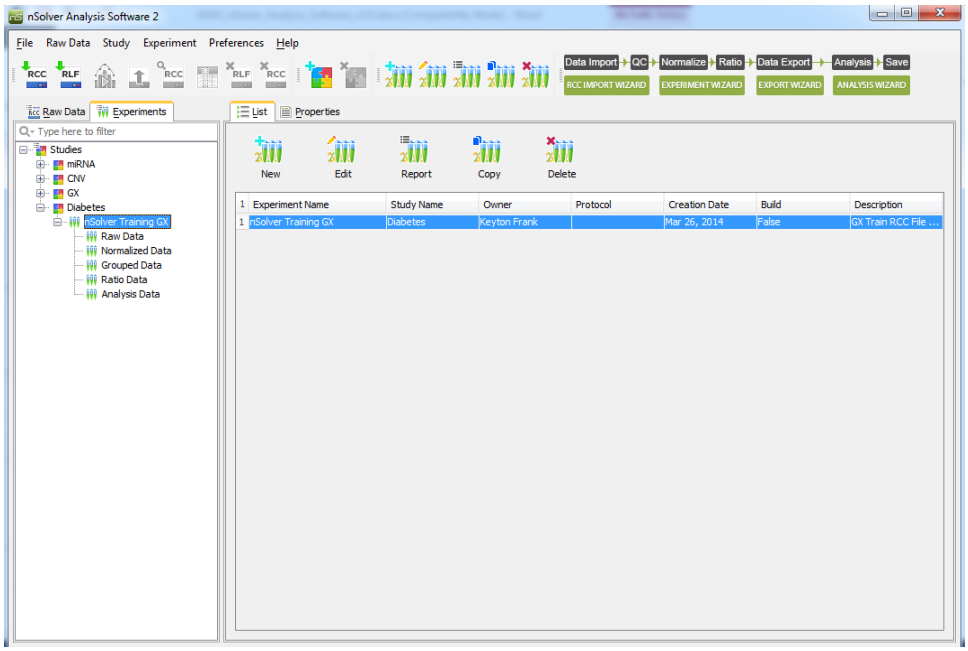
Experiment Main View/Tab

In the left navigation tree, the Experiments tab displays all of your experiments, organized by study. It also displays all studies that are in the application. The Experiment Main View is helpful for locating your analyzed data. When you select an item from the navigation tree, the associated list view or properties view will appear to the right.

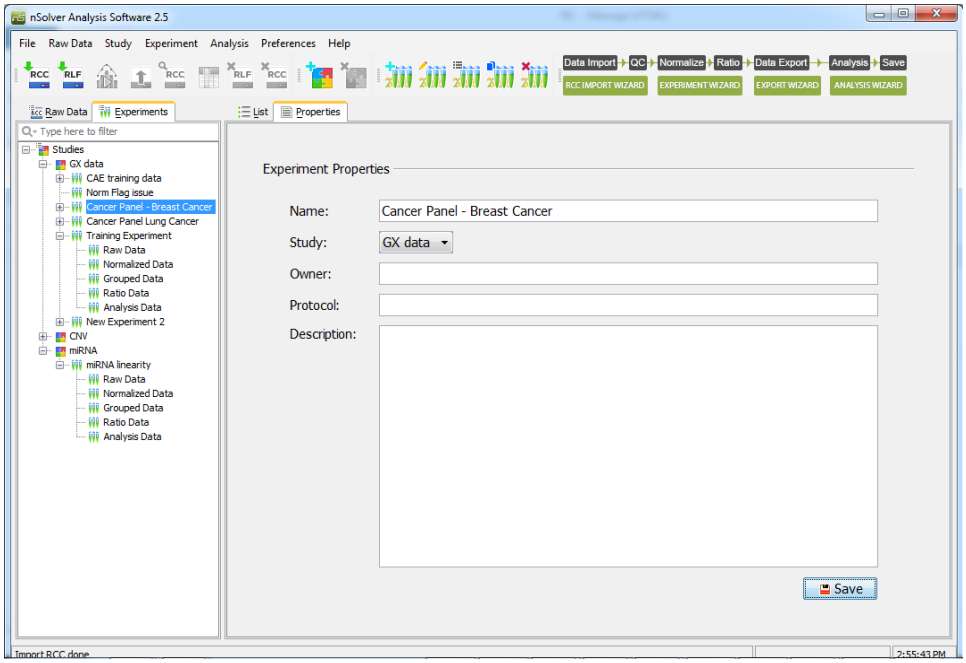


List and Properties Minor Views/Tabs

Within the Experiment and Raw Data Main Views are two minor views: the List View and the Properties View. Most of your time will be spent in the List View, which is where you will see a list of all of the items that are selected from the navigation tree on the left side. Everything in the List View is presented in a table format.



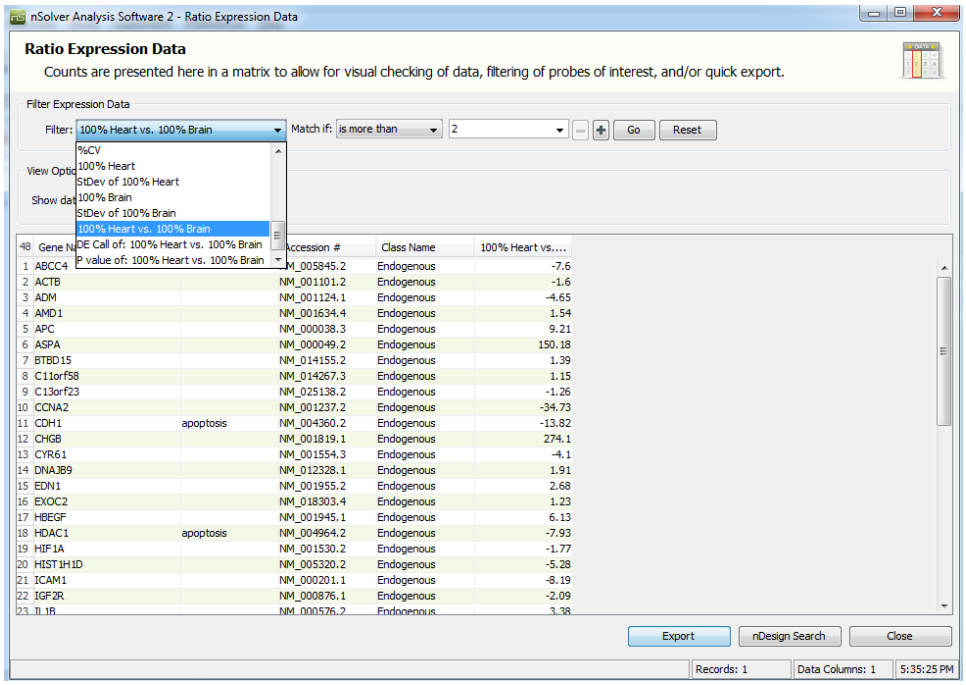
The Properties View is where you go to view or edit an individual item (study, experiment, etc.).



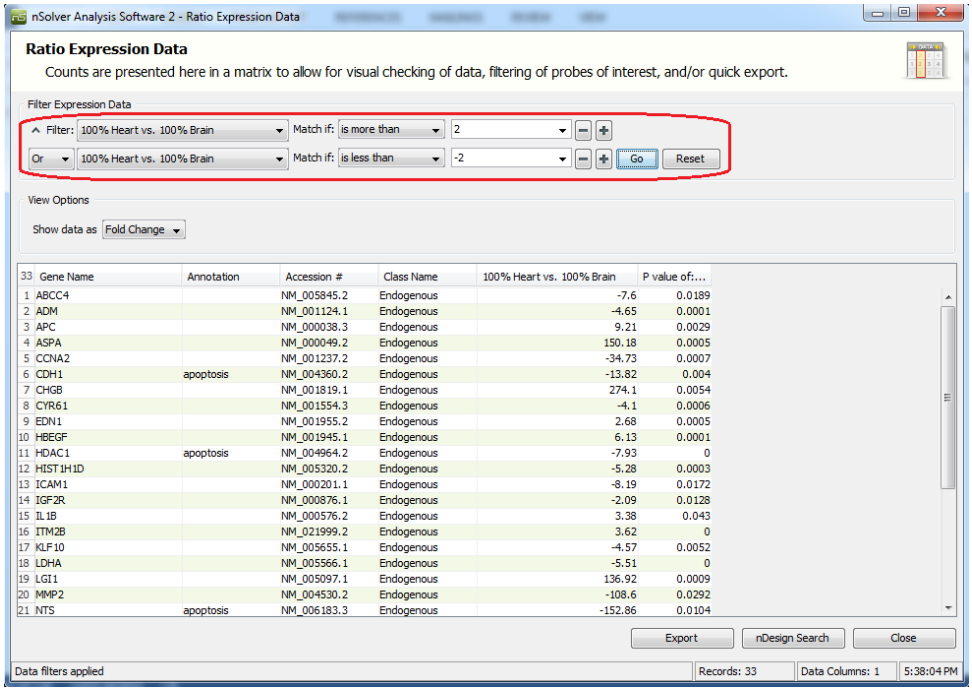
Filter/Search

Within a ratio table, you might want to view only genes for which the absolute value of the fold change is greater than 2. To do so, select the appropriate item from the **Filter** menu and add criteria.

Search Based on Fold Change



To add the Boolean function 'and/or' click on the '+' button and add your remaining filter criteria. When you are satisfied with your criteria, click **Go** to execute the filter.

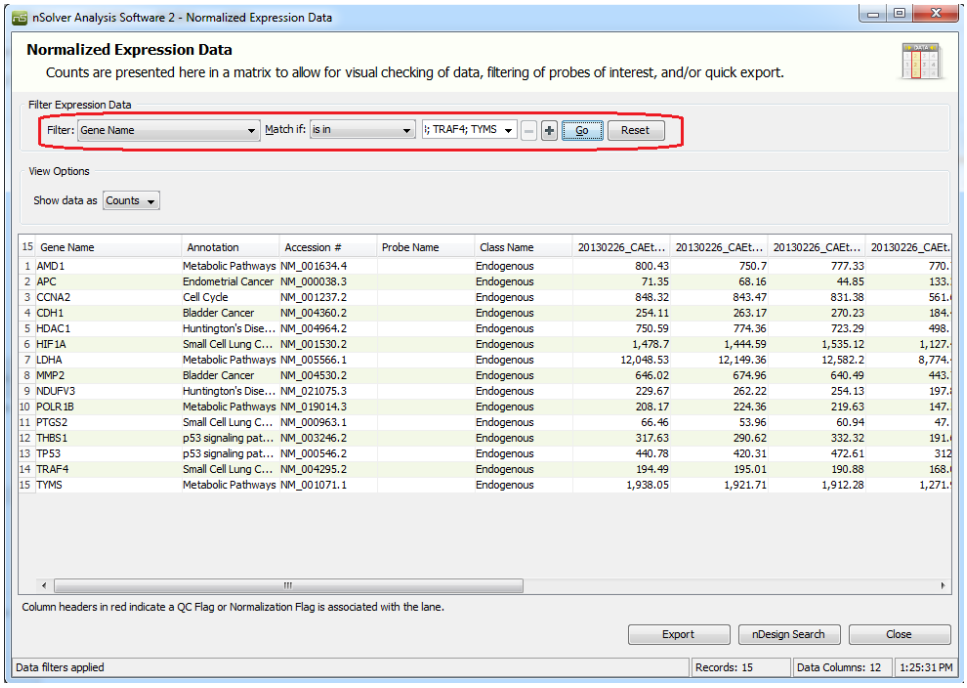


The result in this example contains only rows for which the absolute value of the fold change is greater than 2. Click **Reset** at any time to reset the filter and display all fields.

Search for a List (Such as Genes)

Often you have a list of genes or other items that you want to search for in a table. Use the '**is in**' option from the filter settings to copy a list (from Excel® or elsewhere) and search nSolver for items that are only in that list.

To search for a list of genes, first copy the list to your clipboard by pressing **Ctrl** + **C** (press **Command** + **C** on a Macintosh). Then select the **Gene Name** filter and the '**is in**' criterion. Paste the list from your clipboard into the search text box by pressing **Ctrl** + **V** (press **Command** + **V** on a Mac). The list will automatically populate the text box by delimiting each value with a semi-colon as shown below. Click **Go**, and the search will display only genes matching those in the list.



15	Gene Name	Annotation	Accession #	Probe Name	Class Name	20130226_CAEt...	20130226_CAEt...	20130226_CAEt...	20130226_CAEt.
1	AMD1	Metabolic Pathways	NM_001634.4		Endogenous	800.43	750.7	777.33	770.
2	APC	Endometrial Cancer	NM_000038.3		Endogenous	71.35	68.16	44.85	133.
3	CCNA2	Cell Cycle	NM_001237.2		Endogenous	848.32	843.47	831.38	561.
4	CDH1	Bladder Cancer	NM_004360.2		Endogenous	254.11	263.17	270.23	184.
5	HDAC1	Huntington's Dise...	NM_004964.2		Endogenous	750.59	774.36	723.29	498.
6	HIF1A	Small Cell Lung C...	NM_001530.2		Endogenous	1,478.7	1,444.59	1,535.12	1,127.
7	LDHA	Metabolic Pathways	NM_005566.1		Endogenous	12,048.53	12,149.36	12,582.2	8,774.
8	MMP2	Bladder Cancer	NM_004530.2		Endogenous	646.02	674.96	640.49	443.
9	NDUFV3	Huntington's Dise...	NM_021075.3		Endogenous	229.67	262.22	254.13	197.
10	POLR1B	Metabolic Pathways	NM_019014.3		Endogenous	208.17	224.36	219.63	147.
11	PTGS2	Small Cell Lung C...	NM_000963.1		Endogenous	66.46	53.96	60.94	47.
12	THBS1	p53 signaling pat...	NM_003246.2		Endogenous	317.63	290.62	332.32	191.
13	TP53	p53 signaling pat...	NM_000546.2		Endogenous	440.78	420.31	472.61	312
14	TRAF4	Small Cell Lung C...	NM_004295.2		Endogenous	194.49	195.01	190.88	168.
15	TYMS	Metabolic Pathways	NM_001071.1		Endogenous	1,938.05	1,921.71	1,912.28	1,271.

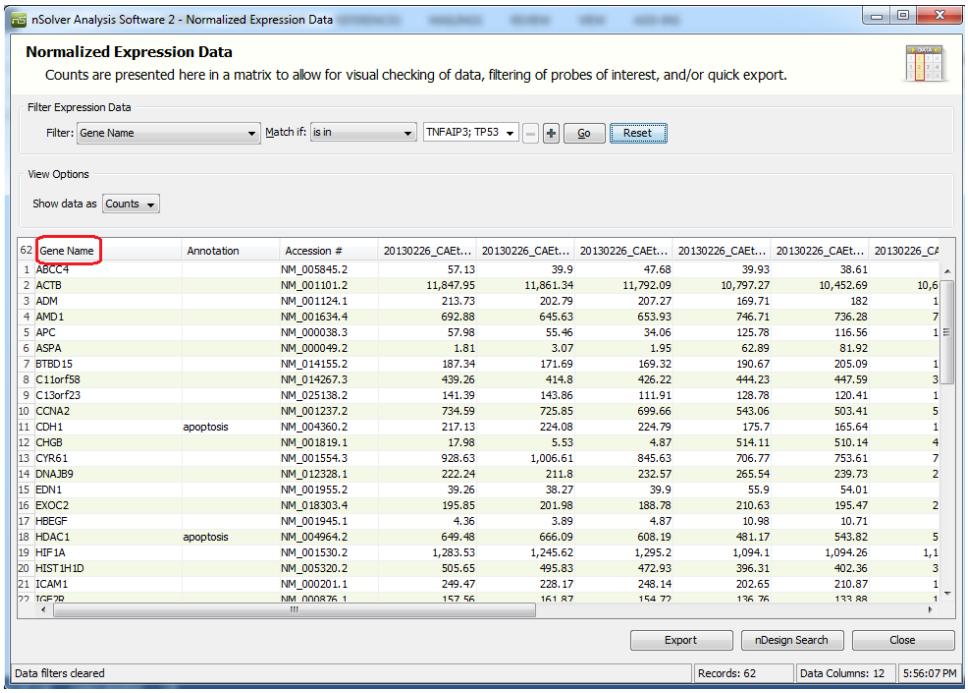
Table Functionality

There are several features that are useful to understand that are common across all tables in the nSolver system.

NOTE: The order of rows in the main viewer will be consistent with the order of columns when opening a table as well as exporting. To view a table or export with columns in a specific order, sort the rows as desired prior to exporting or launching the table. The Ratio Description field in the Experiment Wizard can be used to assist in sorting of ratio data columns in tables and for export.

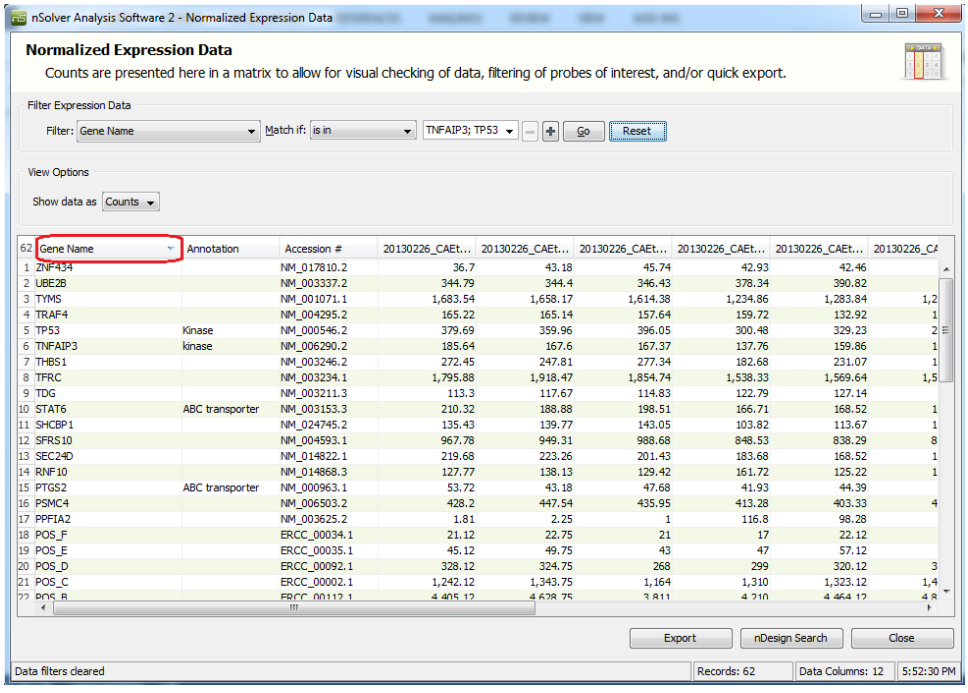
Sort

Data in tables can be sorted by any of the fields displayed in the table. To sort the data in a table in ascending order by a field, such as **Gene Name**, click the column header.



62	Gene Name	Annotation	Accession #	20130226_CAEt...	20130226_CAEt...	20130226_CAEt...	20130226_CAEt...	20130226_CAEt...	20130226_CA
1	ABCC4		NM_005845.2	57.13	39.9	47.68	39.93	38.61	
2	ACTB		NM_001101.2	11,847.95	11,861.34	11,792.09	10,797.27	10,452.69	10,6
3	ADM		NM_001124.1	213.73	202.79	207.27	169.71	182	1
4	AMD1		NM_001634.4	692.88	645.63	653.93	746.71	736.28	7
5	APC		NM_000038.3	57.98	55.46	34.06	125.78	116.56	1
6	ASPA		NM_000049.2	1.81	3.07	1.95	62.89	81.92	
7	BTBD15		NM_014155.2	187.34	171.69	169.32	190.67	205.09	1
8	C11orf58		NM_014267.3	439.26	414.8	426.22	444.23	447.59	3
9	C13orf23		NM_025138.2	141.39	143.86	111.91	128.78	120.41	1
10	CCNA2		NM_001237.2	734.59	725.85	699.66	543.06	503.41	5
11	CDH1	apoptosis	NM_004360.2	217.13	224.08	224.79	175.7	165.64	1
12	CHGB		NM_001819.1	17.98	5.53	4.87	514.11	510.14	4
13	CYR61		NM_001554.3	928.63	1,006.61	845.63	706.77	753.61	7
14	DNAJB9		NM_012328.1	222.24	211.8	232.57	265.54	239.73	2
15	EDN1		NM_001955.2	39.26	38.27	39.9	55.9	54.01	
16	EXOC2		NM_018303.4	195.85	201.98	188.78	210.63	195.47	2
17	HBEGF		NM_001945.1	4.36	3.89	4.87	10.98	10.71	
18	HDAC1	apoptosis	NM_004964.2	649.48	666.09	608.19	481.17	543.82	5
19	HIF1A		NM_001530.2	1,283.53	1,245.62	1,295.2	1,094.1	1,094.26	1,1
20	HIST1H1D		NM_005320.2	505.65	495.83	472.93	396.31	402.36	3
21	ICAM1		NM_000201.1	249.47	228.17	248.14	202.65	210.87	1
22	IGF2R		NM_000876.1	157.56	161.87	154.72	136.76	133.88	1

To sort in descending order, click on the column header again.

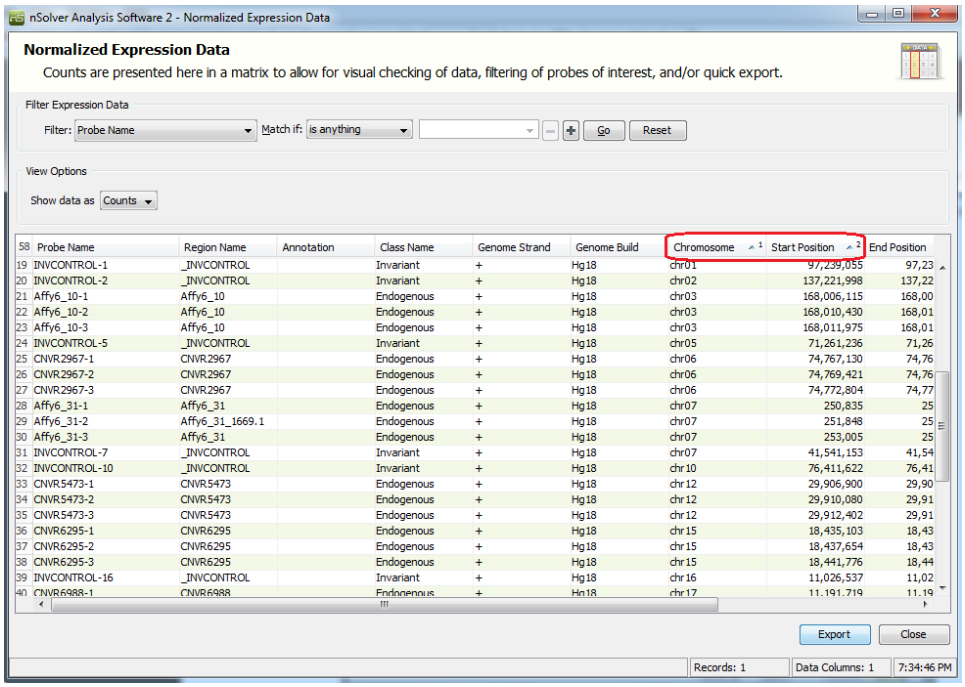


62	Gene Name	Annotation	Accession #	20130226_CAEt...	20130226_CAEt...	20130226_CAEt...	20130226_CAEt...	20130226_CAEt...	20130226_CA
1	ZNF434		NM_017810.2	36.7	43.18	45.74	42.93	42.46	
2	UBE2B		NM_003337.2	344.79	344.4	346.43	378.34	390.82	
3	TYMS		NM_001071.1	1,683.54	1,658.17	1,614.38	1,234.86	1,283.84	1,2
4	TRAF4		NM_004295.2	165.22	165.14	157.64	159.72	132.92	1
5	TP53	Kinase	NM_000546.2	379.69	359.96	396.05	300.48	329.23	2
6	TNFAIP3	kinase	NM_006290.2	185.64	167.6	167.37	137.76	159.86	1
7	THBS1		NM_003246.2	272.45	247.81	277.34	182.68	231.07	1
8	TFRC		NM_003234.1	1,795.88	1,918.47	1,854.74	1,538.33	1,569.64	1,5
9	TDG		NM_003211.3	113.3	117.67	114.83	122.79	127.14	
10	STAT6	ABC transporter	NM_003153.3	210.32	188.88	198.51	166.71	168.52	1
11	SHCBP1		NM_024745.2	135.43	139.77	143.05	103.82	113.67	1
12	SFRS10		NM_004593.1	967.78	949.31	988.68	848.53	838.29	8
13	SEC24D		NM_014822.1	219.68	223.26	201.43	183.68	168.52	1
14	RNF10		NM_014868.3	127.77	138.13	129.42	161.72	125.22	1
15	PTGS2	ABC transporter	NM_000963.1	53.72	43.18	47.68	41.93	44.39	
16	PSMC4		NM_006503.2	428.2	447.54	435.95	413.28	403.33	4
17	PPFIA2		NM_003625.2	1.81	2.25	1	116.8	98.28	
18	POS_F		ERCC_00034.1	21.12	22.75	21	17	22.12	
19	POS_E		ERCC_00035.1	45.12	49.75	43	47	57.12	
20	POS_D		ERCC_00092.1	328.12	324.75	268	299	320.12	3
21	POS_C		ERCC_00002.1	1,242.12	1,343.75	1,164	1,310	1,323.12	1,4
22	POS_B		ERCC_00112.1	4,405.12	4,628.75	3,811	4,210	4,464.12	4,8

To return to the original order, click the column header a third time.

NOTE: The order of rows in the main viewer will be consistent with the order of columns when opening a table as well as exporting. To view a table or export with columns in a specific order, sort the rows as desired prior to exporting or launching the table. The Ratio Description field in the Experiment Wizard can be used to assist in sorting of ratio data columns in tables and for export.

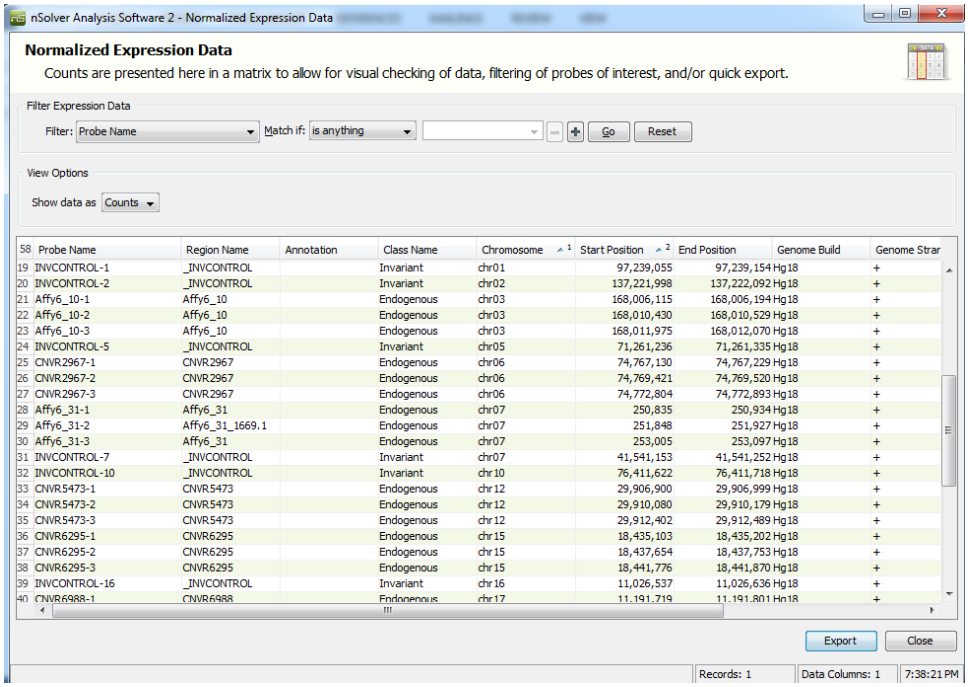
In addition to sorting rows by one column, you can perform second- and third-level sorting using the **Ctrl-click** function on a column heading (use **Command-click** on a Macintosh). For example, if you want to sort by Chromosome and then by Start Position, first click on the **Chromosome** column heading, then hold down the **Ctrl** key while clicking on the **Start Position** column header. Notice the superscript '1' and '2' that denote the order of sorting each column.



	Probe Name	Region Name	Annotation	Class Name	Genome Strand	Genome Build	Chromosome ¹	Start Position ²	End Position
19	INVCONTROL-1	_INVCONTROL		Invariant	+	Hg18	chr01	97,239,055	97,23
20	INVCONTROL-2	_INVCONTROL		Invariant	+	Hg18	chr02	137,221,998	137,22
21	Affy6_10-1	Affy6_10		Endogenous	+	Hg18	chr03	168,006,115	168,00
22	Affy6_10-2	Affy6_10		Endogenous	+	Hg18	chr03	168,010,430	168,01
23	Affy6_10-3	Affy6_10		Endogenous	+	Hg18	chr03	168,011,975	168,01
24	INVCONTROL-5	_INVCONTROL		Invariant	+	Hg18	chr05	71,261,236	71,26
25	CNVR2967-1	CNVR2967		Endogenous	+	Hg18	chr06	74,767,130	74,76
26	CNVR2967-2	CNVR2967		Endogenous	+	Hg18	chr06	74,769,421	74,76
27	CNVR2967-3	CNVR2967		Endogenous	+	Hg18	chr06	74,772,804	74,77
28	Affy6_31-1	Affy6_31		Endogenous	+	Hg18	chr07	250,835	25
29	Affy6_31-2	Affy6_31_1669.1		Endogenous	+	Hg18	chr07	251,848	25
30	Affy6_31-3	Affy6_31		Endogenous	+	Hg18	chr07	253,005	25
31	INVCONTROL-7	_INVCONTROL		Invariant	+	Hg18	chr07	41,541,153	41,54
32	INVCONTROL-10	_INVCONTROL		Invariant	+	Hg18	chr10	76,411,622	76,41
33	CNVR5473-1	CNVR5473		Endogenous	+	Hg18	chr12	29,906,900	29,90
34	CNVR5473-2	CNVR5473		Endogenous	+	Hg18	chr12	29,910,080	29,91
35	CNVR5473-3	CNVR5473		Endogenous	+	Hg18	chr12	29,912,402	29,91
36	CNVR6295-1	CNVR6295		Endogenous	+	Hg18	chr15	18,435,103	18,43
37	CNVR6295-2	CNVR6295		Endogenous	+	Hg18	chr15	18,437,654	18,43
38	CNVR6295-3	CNVR6295		Endogenous	+	Hg18	chr15	18,441,776	18,44
39	INVCONTROL-16	_INVCONTROL		Invariant	+	Hg18	chr16	11,026,537	11,02
40	CNVR6988-1	CNVR6988		Endogenous	+	Hg18	chr17	11,191,719	11,19

Continue to use **Ctrl-click** as desired for third- and fourth-level sorting.

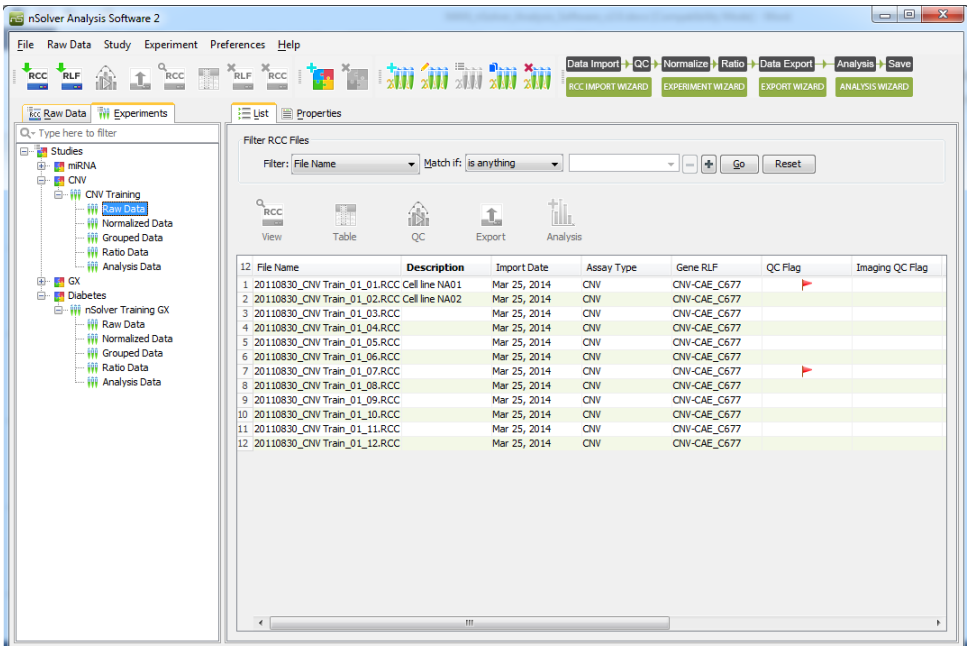
Columns as well as rows can be sorted in the nSolver application. To sort columns manually, click on and drag the column heading with your cursor to the desired spot.



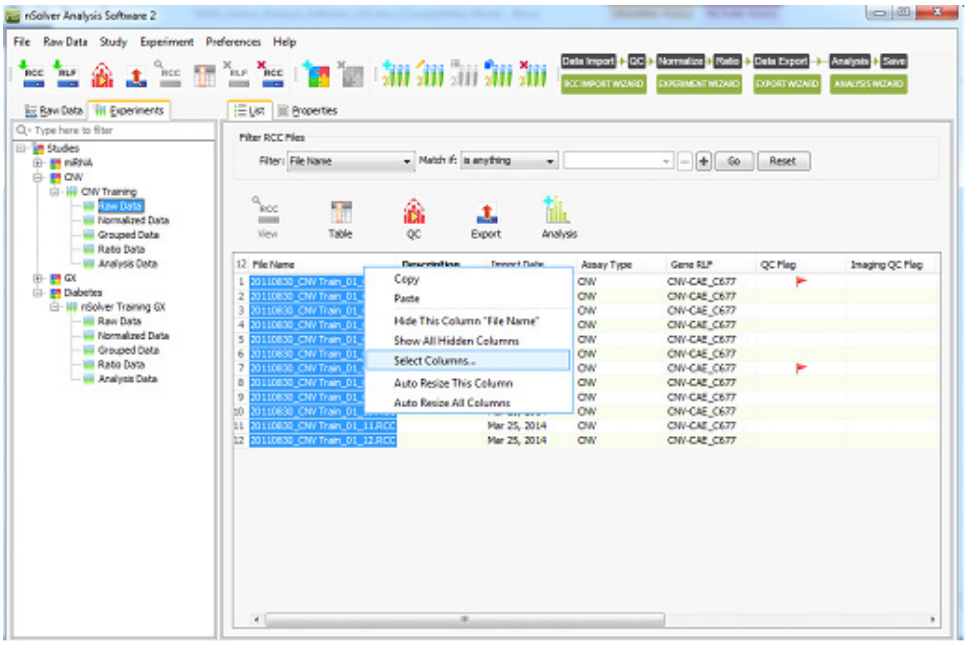
	Probe Name	Region Name	Annotation	Class Name	Chromosome ¹	Start Position ²	End Position	Genome Build	Genome Strar
19	INVCONTROL-1	_INVCONTROL		Invariant	chr01	97,239,055	97,239,154	Hg18	+
20	INVCONTROL-2	_INVCONTROL		Invariant	chr02	137,221,998	137,222,092	Hg18	+
21	Affy6_10-1	Affy6_10		Endogenous	chr03	168,006,115	168,006,194	Hg18	+
22	Affy6_10-2	Affy6_10		Endogenous	chr03	168,010,430	168,010,529	Hg18	+
23	Affy6_10-3	Affy6_10		Endogenous	chr03	168,011,975	168,012,070	Hg18	+
24	INVCONTROL-5	_INVCONTROL		Invariant	chr05	71,261,236	71,261,335	Hg18	+
25	CNVR2967-1	CNVR2967		Endogenous	chr06	74,767,130	74,767,229	Hg18	+
26	CNVR2967-2	CNVR2967		Endogenous	chr06	74,769,421	74,769,520	Hg18	+
27	CNVR2967-3	CNVR2967		Endogenous	chr06	74,772,804	74,772,893	Hg18	+
28	Affy6_31-1	Affy6_31		Endogenous	chr07	250,835	250,934	Hg18	+
29	Affy6_31-2	Affy6_31_1669.1		Endogenous	chr07	251,848	251,927	Hg18	+
30	Affy6_31-3	Affy6_31		Endogenous	chr07	253,005	253,097	Hg18	+
31	INVCONTROL-7	_INVCONTROL		Invariant	chr07	41,541,153	41,541,252	Hg18	+
32	INVCONTROL-10	_INVCONTROL		Invariant	chr10	76,411,622	76,411,718	Hg18	+
33	CNVR5473-1	CNVR5473		Endogenous	chr12	29,906,900	29,906,999	Hg18	+
34	CNVR5473-2	CNVR5473		Endogenous	chr12	29,910,080	29,910,179	Hg18	+
35	CNVR5473-3	CNVR5473		Endogenous	chr12	29,912,402	29,912,489	Hg18	+
36	CNVR6295-1	CNVR6295		Endogenous	chr15	18,435,103	18,435,202	Hg18	+
37	CNVR6295-2	CNVR6295		Endogenous	chr15	18,437,654	18,437,753	Hg18	+
38	CNVR6295-3	CNVR6295		Endogenous	chr15	18,441,776	18,441,870	Hg18	+
39	INVCONTROL-16	_INVCONTROL		Invariant	chr16	11,026,537	11,026,636	Hg18	+
40	CNVR6988-1	CNVR6988		Endogenous	chr17	11,191,719	11,191,801	Hg18	+

Customizing Columns Displayed

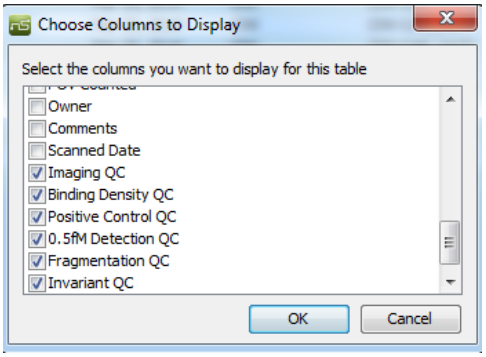
By default, not all columns are displayed in the different tables. For example, in the raw data view, all QC flags are displayed but the actual QC values calculated are not displayed by default. See below.

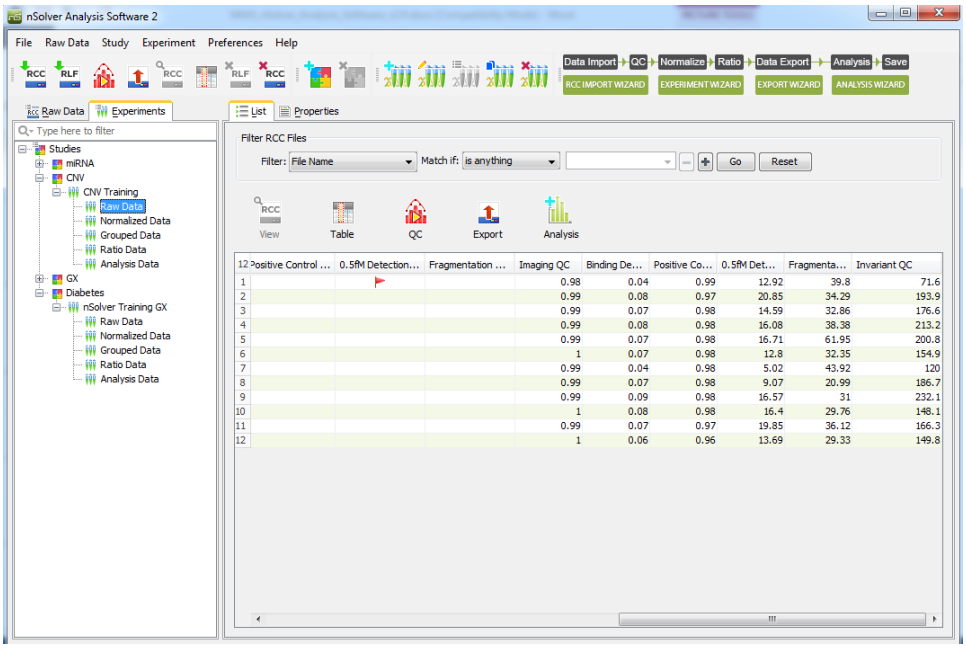


You can display these values by right-clicking on any column header and choosing **Select Columns…**



The following menu appears in which you will be able to select the columns to be displayed. In this example, some of the different QC values are chosen. Click **OK** and the table will include the selected columns.





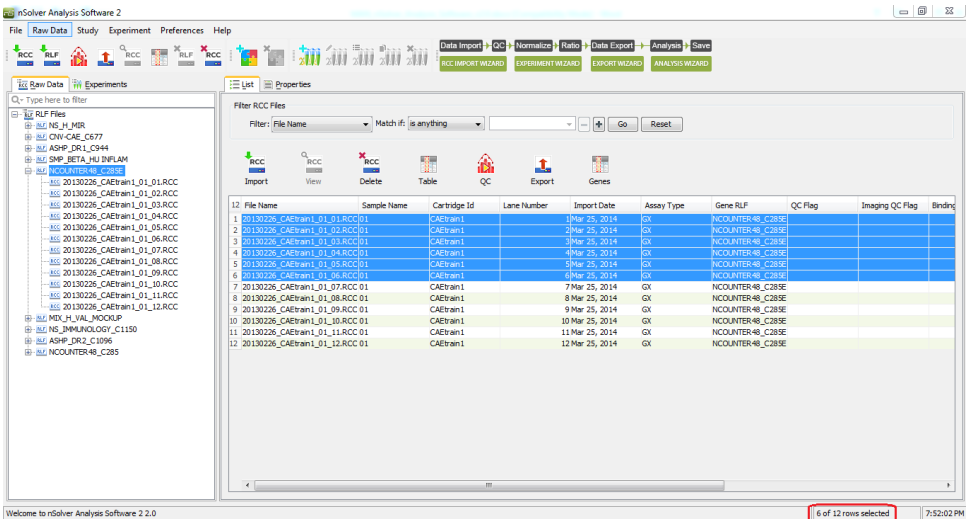
	Positive Control ...	0.5fM Detection...	Fragmentation ...	Imaging QC	Binding De...	Positive Co...	0.5fM Det...	Fragmenta...	Invariant QC
1				0.98	0.04	0.99	12.92	39.8	71.6
2				0.99	0.08	0.97	20.85	34.29	193.9
3				0.99	0.07	0.98	14.59	32.86	176.6
4				0.99	0.08	0.98	16.08	38.38	213.2
5				0.99	0.07	0.98	16.71	61.95	200.8
6				1	0.07	0.98	12.8	32.35	154.9
7				0.99	0.04	0.98	5.02	43.92	120
8				0.99	0.07	0.98	9.07	20.99	186.7
9				0.99	0.09	0.98	16.57	31	232.1
10				1	0.08	0.98	16.4	29.76	148.1
11				0.99	0.07	0.97	19.85	36.12	166.3
12				1	0.06	0.96	13.69	29.33	149.8

Other useful functions are available when you right-click on a column header:

- Copy—Copies all column values to the computer's clipboard. Works similar to Excel®.
- Paste—Pastes values that are copied to the clipboard into any editable field (for example, Description or annotation column). Works similar to Excel®.
- Hide this column—Hides the selected column so it will not be displayed in the table.
- Show all hidden columns –Displays all hidden columns in the table.
- Auto-resize this column—Automatically adjusts the width of the selected column to be large enough to see the column header and all values.
- Auto-resize all columns—Automatically adjusts the width of all columns to be large enough to see column headers and all values.

Row Counts

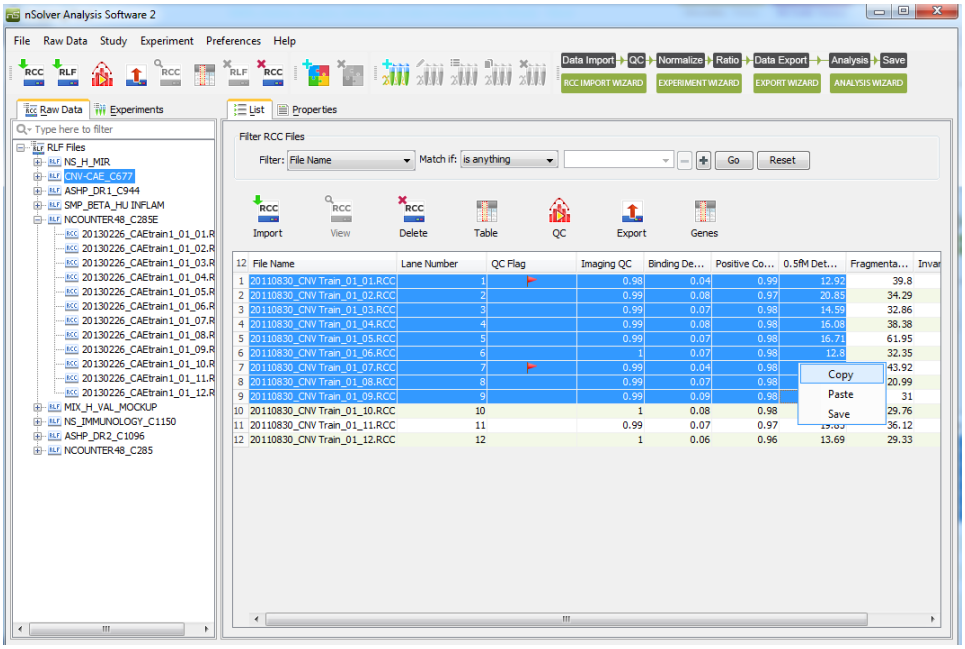
When working with nSolver tables, it is possible to count of the number of rows displayed as well as the number of rows selected. View these counts in the lower right corner of the table as shown below.



Copying Data in Tables

There are a number of different ways to copy data from tables. The first was described previously, in which performing a right-click on a column header copies the contents of an entire column.

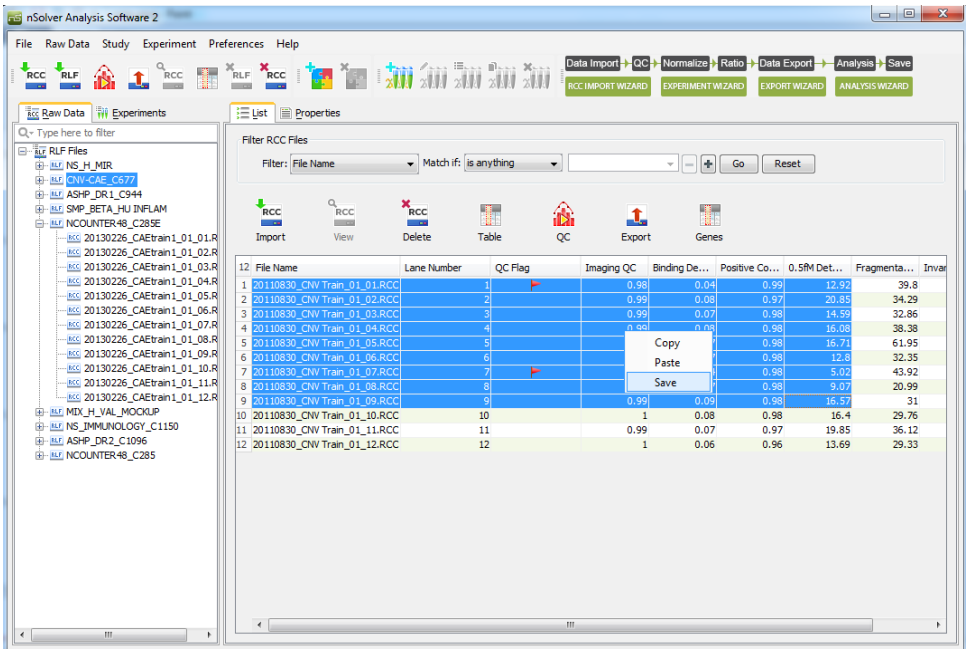
Another way is to simply select between one and many fields from a table and press **Ctrl** + **C** on the keyboard (press **Command** + **C** on a Mac) or right-click and choose **Copy**. This is shown below:



The values will be copied to the clipboard similar to copying a table in Excel® and can be pasted into any document including Excel®, Word®, PowerPoint®, or a plain text editor.

Export / Save

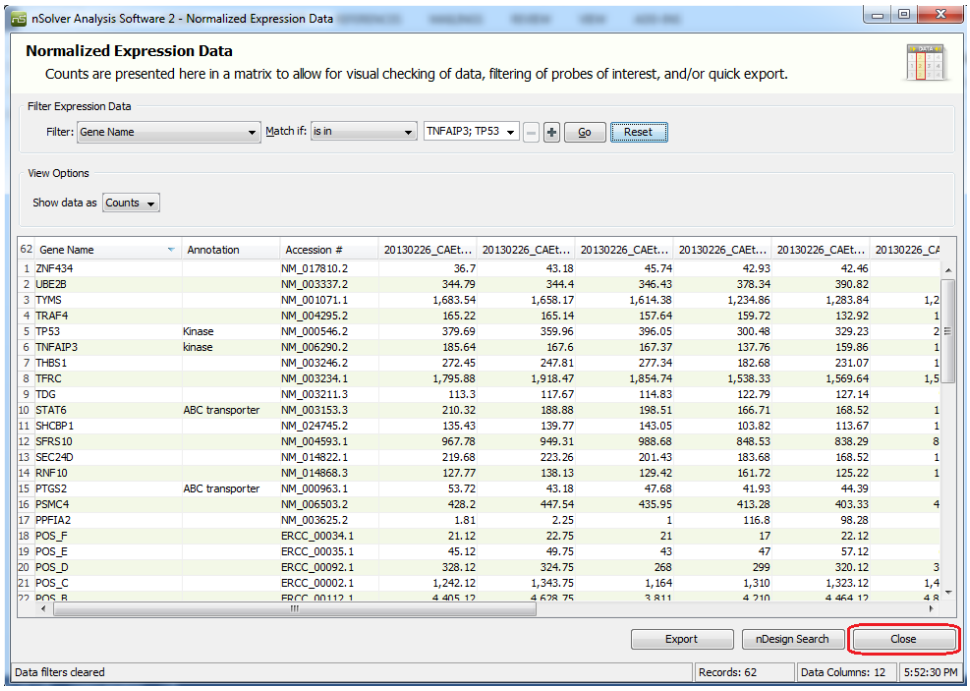
You can export either selected data or all data from any table to a text file. This is explained in detail in *Chapter 11 - Exporting Data*. Alternatively, to export only certain items of interest, right-click on that item and choose Save.



Navigate to the desired folder to save your data and click **Save**. A prompt will display asking if you want to open the file right now. Click **OK** to open the file.

Close

There are two ways to close a table in the nSolver system. Either click on the red **X** in the upper right corner of the table, or click on the **Close** button in the lower right corner.



62	Gene Name	Annotation	Accession #	20130226_CAEt...	20130226_CAEt...	20130226_CAEt...	20130226_CAEt...	20130226_CAEt...	20130226_CA
1	ZNF434		NM_017810.2	36.7	43.18	45.74	42.93	42.46	
2	UBE2B		NM_003337.2	344.79	344.4	346.43	378.34	390.82	
3	TYMS		NM_001071.1	1,683.54	1,658.17	1,614.38	1,234.86	1,283.84	1,2
4	TRAF4		NM_004295.2	165.22	165.14	157.64	159.72	132.92	1
5	TP53	Kinase	NM_000546.2	379.69	359.96	396.05	300.48	329.23	2
6	TNFAIP3	kinase	NM_006290.2	185.64	167.6	167.37	137.76	159.86	1
7	THBS1		NM_003246.2	272.45	247.81	277.34	182.68	231.07	1
8	TFRC		NM_003234.1	1,795.88	1,918.47	1,854.74	1,538.33	1,569.64	1,5
9	TDG		NM_003211.3	113.3	117.67	114.83	122.79	127.14	
10	STAT6	ABC transporter	NM_003153.3	210.32	188.88	198.51	166.71	168.52	1
11	SHCBP1		NM_024745.2	135.43	139.77	143.05	103.82	113.67	1
12	SFRS10		NM_004593.1	967.78	949.31	988.68	848.53	838.29	8
13	SEC24D		NM_014822.1	219.68	223.26	201.43	183.68	168.52	1
14	RNF10		NM_014868.3	127.77	138.13	129.42	161.72	125.22	1
15	PTGS2	ABC transporter	NM_000963.1	53.72	43.18	47.68	41.93	44.39	
16	PSMC4		NM_006503.2	428.2	447.54	435.95	413.28	403.33	4
17	PPFIA2		NM_003625.2	1.81	2.25	1	116.8	98.28	
18	POS_F		ERCC_00034.1	21.12	22.75	21	17	22.12	
19	POS_E		ERCC_00035.1	45.12	49.75	43	47	57.12	
20	POS_D		ERCC_00092.1	328.12	324.75	268	299	320.12	3
21	POS_C		ERCC_00002.1	1,242.12	1,343.75	1,164	1,310	1,323.12	1,4
22	POS_B		ERCC_00112.1	4,405.12	4,628.75	3,811	4,210	4,464.12	4,8

13 Glossary

This chapter contains definitions of terminology associated with the nCounter Analysis System and RCC file data fields.

Accession: Text field defining the gene accession information associated with an nCounter Reporter Probe pair.

Assay Type: A configuration of the nSolver Analysis procedure specifically designed to detect a specific type of molecule or perform a specific type of analysis. The nSolver application supports five assay types.

Assay: Each nCounter assay is capable of simultaneous, single-tube, multiplexed detection of up to 800 targets. With the exception of Plex^2, there is typically one assay per lane.

Binding Density: Number of fluorescent spots per µm^2 of the lane surface.

Calibration Sample: A reference sample that is used in Plex^2 assays to correct for intra-CodeSet probe variability. This sample is assayed across all Plex^2 sub-CodeSets.

Cartridge Definition File (CDF): A file used by the Digital Analyzer to associate user-input sample information and a RLF with a cartridge.

Cartridge: device containing 12 lanes that allows distribution and immobilization of nCounter probes on a solid surface for subsequent imaging and counting by the nCounter Digital Analyzer.

CNV: Custom copy number variation assay for the detection of user defined regions of a genome.

Code Class: Text field defining nCounter Reporter Probe pairs.

CodeSet Plus: Addition of up to 30 Reporter Probes customers can buy and add onto any custom CodeSet or Panel.

Comments: Text annotation field specified by the user in the CDF.

Digital Analyzer: The nCounter Digital Analyzer collects data by taking images of the immobilized fluorescent reporters in the sample cartridge with a CCD camera through a microscope objective lens. The Digital Analyzer may collect data at one of four resolutions: max, high, medium, and low, yielding data for hundreds of thousands of target molecules. The number of fields of view (FOV) taken correlates linearly with the number of reporters counted and this, in part, determines the dynamic range and level of sensitivity of the system. Images are processed internally by the Digital Analyzer and the results are exported as a zipped file that can be downloaded via memory stick.

Endogenous (1–4): Reporter probe pair for a target queried within a given sample. For Plex^2 assays, each sample is queried against unique reporter probe pairs for each target. These probe pairs are identified as Endogenous 1, Endogenous 2, Endogenous 3 and Endogenous 4.

Endogenous Probes: Probes designed to target biological content defined by the customer.

Experiment: An analysis method performed on a single CodeSet defined through use of the Experiment Wizard. Upon completion of the wizard, raw, normalized, grouped, ratio data tables and analyses are available for viewing and/or export.

Field of View (FOV): Area of the cartridge surface discretely imaged by the Digital Analyzer. A single image the Digital Analyzer uses to count molecules.

File Attributes: Data fields imported from CDF.

File Name: RCC file name. Unique for each lane.

File Version: RCC file version number.

FOV Count: Total number of FOVs successfully counted per lane.

Gene RLF: RLF used to associate gene names with counts specified by the user in the CDF.

GX: Custom gene expression assay for the detection of user defined mRNA transcripts.

Housekeeping Probes: Probes designed to target transcripts that vary minimally in expression level from sample to sample.

Invariant Probes: Probes containing probes designed to target regions of the genome that have been measured as two copies in a majority of individuals. Customers may specify the invariant probes for their custom CodeSet or utilize invariant probes pre-designed by NanoString. Invariant probes are utilized for normalization of input amounts between samples.

Lane Attributes: Data fields generated by the Digital Analyzer during a scan.

Lane ID: Numerical identifier of imaged lane (always a value from 1 to 12).

Lane: Area of the cartridge surface over which nCounter probes are immobilized. Each lane is 2600 µm x 8000 µm.

Messages: Text output field for Digital Analyzer QCs. A binding density greater than 2.25 will generate a warning indication.

miRGE: A custom combination GX and miRNA assay for the detection of user defined miRNA and mRNA transcripts.

miRNA: Panel assays targeting pre-defined sets of miRNAs. Panels are offered for different species (e.g., Human, Mouse, and Rat). miRNA assays controls undergo a ligation event and are fundamentally different from GX assay controls. See the nCounter® Expression Data Analysis Guide for more information on miRNA controls.

Multiplexing: This data field in nSolver is checked if the cartridge lane has been utilized for a Plex^2 assay.

Name: Text field defining the gene name associated with an nCounter Reporter Probe pair.

Negative Control Probes: Probes designed to target sequences defined by the External RNA Controls Consortium (ERCC) that are absent in the nCounter assay. The negative control probes may be used to infer information about the background level for a particular assay.

Other: A custom nCounter assay that requires unique analysis methods such as scaled total count normalization.

Plex^2: nCounter assay which allows for up to 4 samples to be profiled in a single lane.

Positive Control Probes: Probes designed to target sequences defined by the External RNA Controls Consortium (ERCC) that are included in each nCounter assay at varying concentrations. The positive control probes may be used to infer information about the technical performance of a particular assay.

Prep Station: An automated liquid-handling robot that utilizes magnetic bead-based purification to remove unbound CodeSet, non- target cellular transcripts and other cell debris after sample hybridization. In addition to purification of the hybridization, the Prep Station also automates immobilization of the sample onto an imaging surface for subsequent data collection.

Reporter Code Counts (RCC) File: A data file (*.CSV) output by the Digital Analyzer that contains sample information, probe information and probe counts for each lane scanned. Importing RCC files into the nSolver system is the first step in performing an analysis.

Reporter Counts: Data fields containing nCounter Reporter Probe information.

Reporter Library File (RLF): A file used by the Digital Analyzer to associate target name and annotation with a unique fluorescent barcode after imaging and counting. While optional, importing an RLF into the nSolver system provides additional annotations for the associated RCC files including unique probe IDs and target sequence information.

Sample Date: Date information specified by the user in the CDF.

Sample ID: Sample identifier specified by the user in the CDF.

Scanner ID: Numerical identifier of the Digital Analyzer.

Stage Position: Numerical identifier stage position / cartridge scanned (always a value from 1 to 6).

Sub-CodeSet: One of up to 4 CodeSets that are assayed in a single lane of a cartridge when using the Plex^2 assay. Each sub-CodeSet is referred to by number, *e.g.*, sub-CodeSet 1, sub-CodeSet 2, sub-CodeSet 3, and sub-CodeSet 4 if loading four samples in a Plex^2 assay.

Study: A set of experiments grouped for a user defined purpose (*e.g.,* research area, lab, tissue, etc.). A study may contain one to many experiments. The experiments may be derived from various CodeSets and assay types.

Wizard: A series of dialog boxes designed to aid the user in accomplishing one or more workflow steps.

14 Java TreeView Copyright Notice

Below is the copyright notice for Java TreeView.

This file is part of Java TreeView

Copyright © 2001–2003 Alok Saldanha, All Rights Reserved.

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details. You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 59 Temple Place, Suite 330, Boston, MA 02111-1307, USA

This software is provided under the GNU GPL Version 2. In particular,

- If you modify a source file, make a comment in it containing your name and the date.
- If you distribute a modified version, you must do it under the GPL 2.
- Developers are encouraged but not required to notify the Java TreeView maintainers at **alokito@users.sourceforge.net** when they make a useful addition. It would be nice if significant contributions could be merged into the main distribution.

A full copy of the license can be found in copy.txt or online at **http://www.gnu.org/licenses/gpl-2.0.html**.

15 R.app Copyright Notice

Below is the copyright notice for R.app.

The bulk of R.app code is copyrighted by members of the R Core Team. See the file [COPY.TEXT] for the exact conditions under which you may redistribute it.

Copyright © 2004–2013 The R Foundation

Written by Stefano M. Iacus and Simon Urbanek

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 59 Temple Place, Suite 330, Boston, MA 02111-1307, USA.

A full copy of the license can be found in the file copy.txt or online at **http://www.gnu.org/licenses/gpl-2.0.html.**

nSolver™ Data Analysis Software User Manual



NanoString Technologies, Inc.
530 Fairview Ave N
Seattle, Washington 98109

CONTACT US
info@nanostring.com
Tel: (888) 358-6266
Fax: (206) 378-6288
www.nanostring.com

SALES CONTACTS
United States: us.sales@nanostring.com
Europe: europe.sales@nanostring.com
Other Regions: info@nanostring.com

FOR RESEARCH USE ONLY. Not for use in diagnostic procedures.

MAN-C0019-06, April 2016