# nSolver™ 4.0

## Analysis Software User Manual

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

**www.nanostring.com**

T: 206.378.6266
888.358.6266

E: info@nanostring.com

MAN-C0019-08, January 2018

## Intellectual Property Rights

This nSolver™ Analysis Software user manual and its contents are the property of NanoString Technologies, Inc. ("NanoString"), and are 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 www.nanostring.com/company/patents.

## 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, the NanoString logo, nCounter, nSolver, PlexSet and Plex$^2$ 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

©2018 NanoString Technologies, Inc. All rights reserved.

# Contents

Introduction ..... 5
nSolver 4.0 Basics ..... 5
Workflow ..... 6
3D Example Dataset ..... 6
Analyte Types ..... 7
Installation ..... 9
nSolver 4.0 Quick Start Guide ..... 13
Import ..... 17
RCC files ..... 17
RLFs ..... 19
Import Example ..... 20
Quality Control (QC) ..... 21
QC Parameters for All Analytes ..... 22
Positive Control QC Parameters ..... 23
Additional Analyte-Specific QC Parameters ..... 24
QC Example ..... 28
Exploring Raw Data ..... 29
Checking for QC Flags ..... 29
What to Do If You Have a Flag ..... 30
Re-running QC ..... 30
Exploring Raw Data Example ..... 31
Viewing & Customizing Samples ..... 32
Using Table Options Buttons ..... 33
Creating Experiments ..... 35
Studies ..... 35
Experiments ..... 36
Creating Experiments Example ..... 37
Annotations ..... 38
Annotations Example ..... 39
Background Subtraction & Thresholding ..... 40

Background Thresholding ........ 40
Background Subtraction ........ 41
Background Example ........ 42
Normalization ........ 43
Steps of Normalization ........ 43
Normalization Settings ........ 44
Normalization Example ........ 45
What to Do If You Have a Content Normalization QC Flag ........ 46
Ratios ........ 47
Data Output & Export ........ 50
Types of Data ........ 50
Viewing Tables ........ 51
Raw Data Table ........ 51
Normalized Data Table ........ 52
Grouped Data Table ........ 53
Ratio Data Table ........ 53
Experiment Reports ........ 56
Exporting Data ........ 57
Analysis & Advanced Analysis ........ 59
Analyte- & Assay-Specific Notes ........ 78
Plex$^2$ Data Analysis ........ 78
PlexSet Data Analysis ........ 81
SNV Data Analysis ........ 86
CNV Data Analysis ........ 90
Multi-RLF Experiments & Batch Calibration ........ 91
Appendix A: 3D Bio Data Example ........ 98
Glossary ........ 110

# Introduction

## nSolver 4.0 Basics

NanoString Technologies' nCounter® assays are designed to provide a sensitive, reproducible, and highly-multiplexed method for detecting proteins, RNA and DNA targets. These assays provide direct detection of targets using molecular barcodes, without the necessity of reverse transcription and most without the necessity of amplification. NanoString data is obtained utilizing the fully-automated Prep Station followed by data collection on a nCounter MAX or FLEX Digital Analyzer; alternatively, processing and data collection may be accomplished together using the SPRINT platform. The resulting data can then be imported into, and analyzed by, the nSolver 4.0 Analysis Software System. This user manual describes quality control, normalization, experimental setup, data analysis, and visualization using nSolver 4.0.

The nSolver 4.0 Analysis Software is designed to manage and analyze nCounter instrument data of all analyte types and combinations. The seven steps of a basic analysis are featured in the workflow image (see the *Workflow* section), as are the four wizards that assist in the process. In addition, they are all listed below.

When you initiate the *Data Import* step, the ***RCC Import Wizard*** launches and guides you through importing data from nCounter instruments as well as through the selection of *Quality Control (QC)* parameters.

You must then create experiments, at which point the ***Experiment Wizard*** opens and prompts you to *Normalize* the data and create *Ratios*.

You can explore the various table formats, and utilize the ***Export Wizard*** for *Data Export* when ready. Multiple export formats are available to facilitate integration with other statistical, analysis and visualization programs, if desired.

Graphical *Analysis* is assisted by the ***Analysis Wizard***.

### Changes from nSolver 3.0 to nSolver 4.0

nSolver 4.0 is keeping pace with the rapidly expanding nCounter technology. In this version, data analysis becomes more **data-focused** and less analyte-restricted. Single Nucleotide Variance (**SNV**) and PlexSet analyses are supported, and new methods of **background thresholding and normalization** are available, allowing users more flexibility in calculating these metrics.

## Workflow

The nSolver 4.0 workflow is shown in the image below (Figure 1). The seven steps of the process are shown in black blocks and the four wizards which assist are in green blocks. See the *Quick Start Guide* section for an overview of these basic steps.

For more details on a subject:

- Follow the hyperlinks in the Quick Start Guide, or
- Navigate the manual using the *Table of Contents* and relevant links.

Figure 1: nSolver 4.0 workflow

## 3D Example Dataset

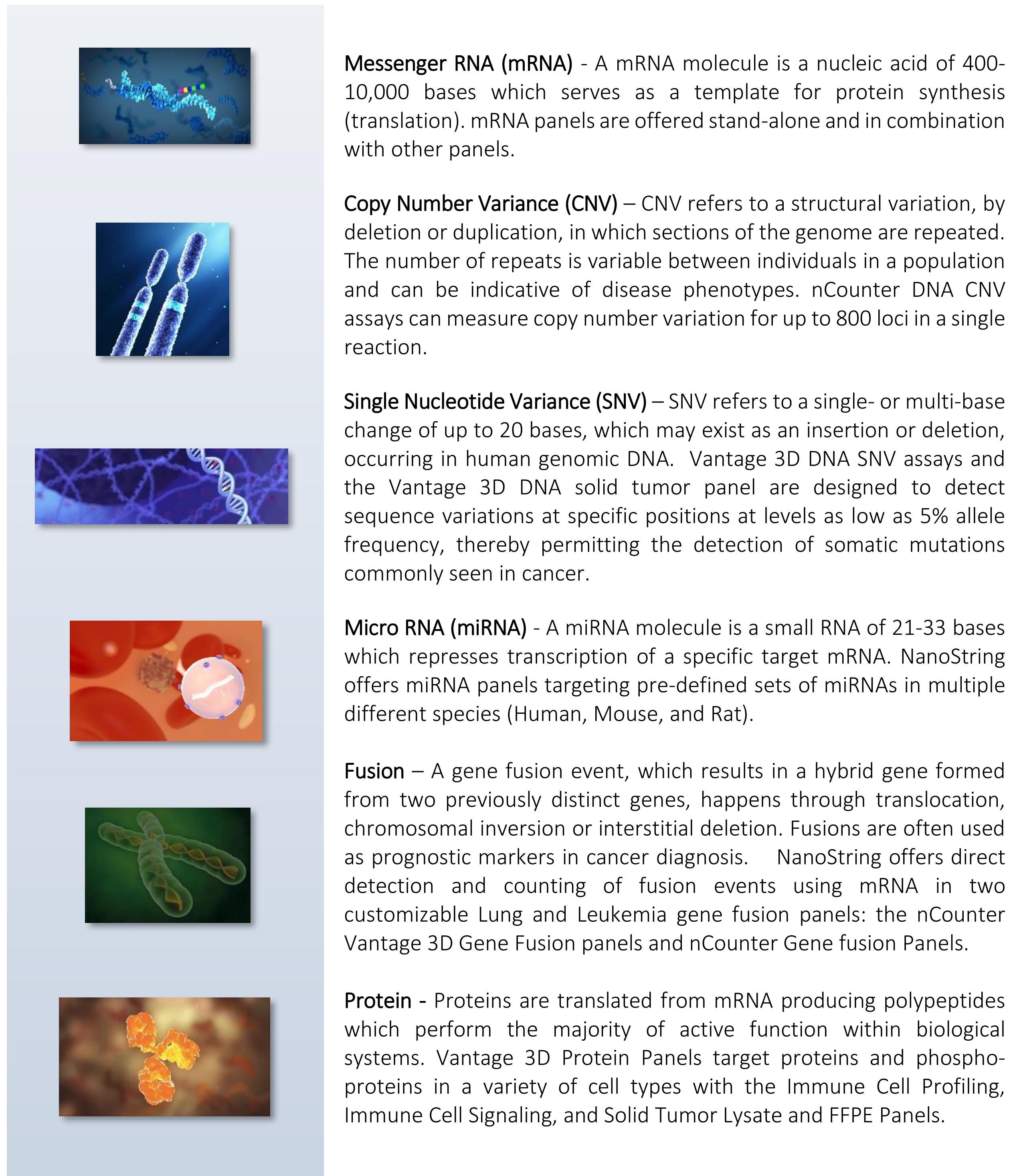
The dataset, **3D Bio Data**, is included when you download the nSolver 4.0 Analysis Software. This data is a result of three biological replicates of two different melanoma cell lines, SKMEL28, which has a known mutation (c.1799T>A; p.V600E) in both copies of the BRAF gene, and SKMEL2, which has two normal copies of the BRAF gene. Both cell lines were treated with either DMSO (vehicle) or vemurafenib (a specific inhibitor of the V600E mutant BRAF protein) dissolved in DMSO for 8 hours.

Throughout the nSolver 4.0 User Manual, you will find excerpts of this dataset's analysis.

The entire 3D Bio Data Example is attached as *Appendix A*.

## Analyte Types

nSolver 4.0 is designed to identify and support the analyte types listed below (Figure 2), whether alone or together in any multi-omic combination of NanoString 3D Biology™ Technology.



**Messenger RNA (mRNA)** - A mRNA molecule is a nucleic acid of 400-10,000 bases which serves as a template for protein synthesis (translation). mRNA panels are offered stand-alone and in combination with other panels.

**Copy Number Variance (CNV)** – CNV refers to a structural variation, by deletion or duplication, in which sections of the genome are repeated. The number of repeats is variable between individuals in a population and can be indicative of disease phenotypes. nCounter DNA CNV assays can measure copy number variation for up to 800 loci in a single reaction.

**Single Nucleotide Variance (SNV)** – SNV refers to a single- or multi-base change of up to 20 bases, which may exist as an insertion or deletion, occurring in human genomic DNA. Vantage 3D DNA SNV assays and the Vantage 3D DNA solid tumor panel are designed to detect sequence variations at specific positions at levels as low as 5% allele frequency, thereby permitting the detection of somatic mutations commonly seen in cancer.

**Micro RNA (miRNA)** - A miRNA molecule is a small RNA of 21-33 bases which represses transcription of a specific target mRNA. NanoString offers miRNA panels targeting pre-defined sets of miRNAs in multiple different species (Human, Mouse, and Rat).

**Fusion** – A gene fusion event, which results in a hybrid gene formed from two previously distinct genes, happens through translocation, chromosomal inversion or interstitial deletion. Fusions are often used as prognostic markers in cancer diagnosis. NanoString offers direct detection and counting of fusion events using mRNA in two customizable Lung and Leukemia gene fusion panels: the nCounter Vantage 3D Gene Fusion panels and nCounter Gene fusion Panels.

**Protein -** Proteins are translated from mRNA producing polypeptides which perform the majority of active function within biological systems. Vantage 3D Protein Panels target proteins and phospho-proteins in a variety of cell types with the Immune Cell Profiling, Immune Cell Signaling, and Solid Tumor Lysate and FFPE Panels.

Figure 2: Analyte types supported by nSolver 4.0

## System Requirements

The nSolver 4.0 Analysis Software is compatible with both Macintosh (10.10–10.11) and Windows (8.1 and 10) operating systems. 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, contact the nSolver support team at support@nanostring.com.

- Windows users will find that nSolver installs and runs its own Java runtime environment.
- Mac users must have Java 1.7 or higher installed separately

Advanced Analyses are available in nSolver versions 2.5 and higher, and require the use of R, a statistical software package freely available to the public.

- Windows users will find that there is an option to install R version 3.3.2 automatically during the installation of nSolver 4.0. *R v.3.3.2 is required for Advanced Analysis 2.0.*
- Mac users must install R separately before using the Advanced Analyses 2.0 plug-in module. Customers using Macintosh 10.9 and later must also install XQuartz. Links are provided to download R and XQuartz on the NanoString website.

## Support Options and Contacts

In addition to this manual, visit https://www.nanostring.com/products/analysis-software/nsolver or click on the links below for other documentation and guidance, including:

- The Advanced Analysis User Manual (MAN-10030).
- The nCounter Gene Expression Data Analysis Guide (MAN-C0011).
- The All About SNV Analysis in nSolver and Advanced Analysis (MAN-10075) guide.
- The All About Fusion Analysis in nSolver and Advanced Analysis (MAN-10076) guide.
- The All About PlexSet™ Technology Data Analysis in nSolver Software (MAN-10044) guide.
- Training webinars and videos.

Please contact support@nanostring.com for any issues or requests.

## Installation

The nSolver 4.0 Data Analysis Software is available to NanoString customers, as is the Advanced Analysis 2.0 plug-in. The program R 3.3.2 is required for Advanced Analysis 2.0 users but not for those who will only use nSolver for analysis.

Advanced Analysis must be separately downloaded from the NanoString website and imported into the nSolver 4.0 application. All Advanced Analysis plug-ins distributed by NanoString depend on a specific version of R version. Refer to the instruction manual of the specific Advanced Analysis plug-in you intend to use to ensure you have the correct version of R installed.

Instructions for the following software downloads are listed individually below: **nSolver 4.0 Analysis Software**, **R 3.3.2**, and the **Advanced Analysis 2.0** plug-in.

## Downloading nSolver 4.0 Analysis Software

If you have been using another version of **nSolver 4.0 alpha**, you will need to back up your database and start with a **clean or blank nSolver 4.0 database**. Then, download and install the software.

**Windows** users:

- Navigate to c:\users\<username>\appdata\roaming\. Rename your nSolver4 folder to *nSolver4_old* (or similar). You may need to *show hidden files* to see the *appdata* folder.
- **Download** and extract **nSolver 4.0** from https://www.nanostring.com/products/analysis-software/nsolver. **Install** the nSolver 4.0 application.
- When prompted, check the box to **Install R** (see next section).

**Mac** users:

- From your home directory, make sure your hidden files are shown so you can see your nSolver4 folder. Rename it *nSolver4_old* (or similar).
- **Download** and extract **nSolver 4.0** from https://www.nanostring.com/products/analysis-software/nsolver. **Install** the nSolver 4.0 application.

## Downloading R 3.3.2

***R 3.3.2 is required for version 2.0 of Advanced Analysis.***

**Windows** users:

- You will be given the option to download R 3.3.2 when you install **nSolver 4.0**. If you did not, go to https://cran.r-project.org/bin/windows/base/old/3.3.2/.
- If you've previously used a different version of R with Advanced Analysis and are updating to a new version of R, you will need to **change the R home path in nSolver**. Select **Analysis** on the top toolbar in nSolver and select **Change R Home Path** to the R 3.3.2 installation folder. Browse to the desired directory and then select **Ok**.

**Mac** users:

- Install R separately. Go to https://cran.r-project.org/bin/macosx/old/R-3.3.2.pkg.
- Install XQuartz if you use Mac OS X 10.10 or higher. Go to https://www.xquartz.org/.
- You may need to download **R Switch** or a similar app to replace your current version of R with 3.3.2. Alternatively, you may uninstall all other R versions.

When initiating an analysis in Advanced Analysis 2.0, nSolver 4.0 will check the version of R you have installed and will issue a warning if it is a version incompatible with the program.

## Downloading Advanced Analysis 2.0

You will find the most recent version of Advanced Analysis on https://www.nanostring.com/products/analysis-software/nsolver. Save this to your computer as a compressed .zip file. *Do not extract this file before uploading it to nSolver.*

In **nSolver 4.0,** select **Analysis** on the top toolbar (see Figure 3) and select **Advanced Analysis Manager**. Any previously-installed versions of Advanced Analysis will be displayed. You can **Remove** them or simply **Import** the current version. To import, select the **Import New Advanced Analysis** button and navigate to the .zip file with the current Advanced Analysis version. This version will be added to the list within the Advanced Analysis Manager. Select **OK**.

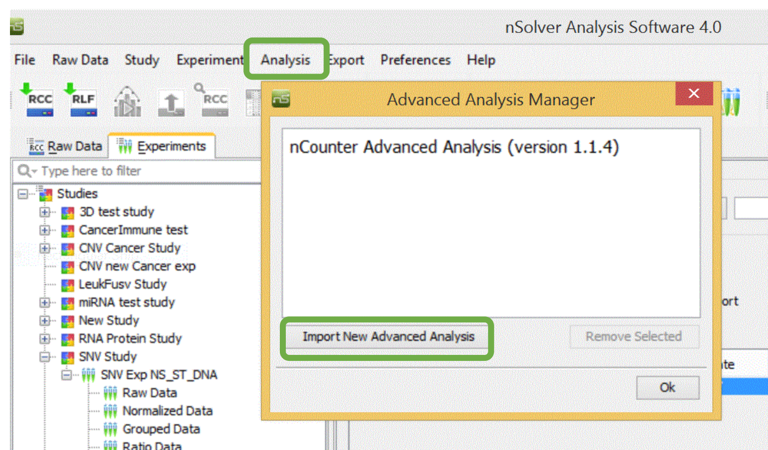



Figure 3: Importing Advanced Analysis or changing the analysis version

## Migration of Data & Analyses from nSolver 3.0 to nSolver 4.0

When installing nSolver 4.0, please note that you have only ***one opportunity to migrate any nSolver 3.0 data*** you may have to your new program (see Figure 4). Migrating your data at this time gives you the option to preserve any analyses previously performed.

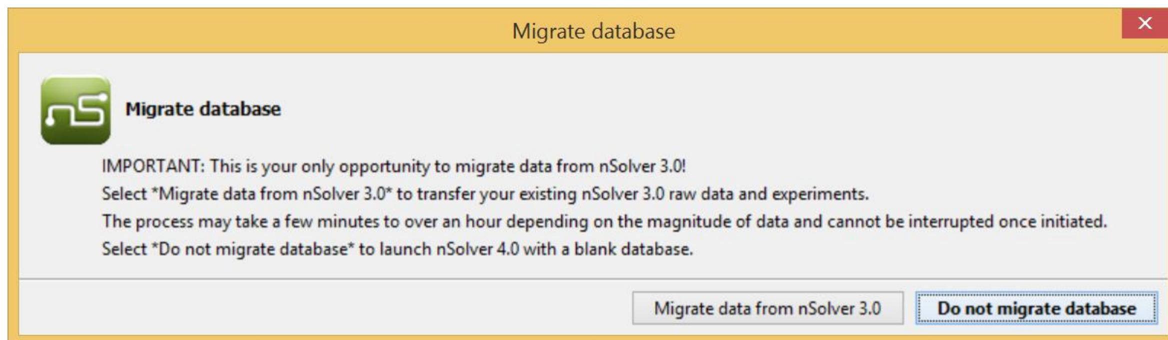



Figure 4: Database migration alert

Select **Migrate data from nSolver 3.0**. Confirm the path that the program suggests or **Browse** to the location of your **nSolver3** database. Select **Next**.

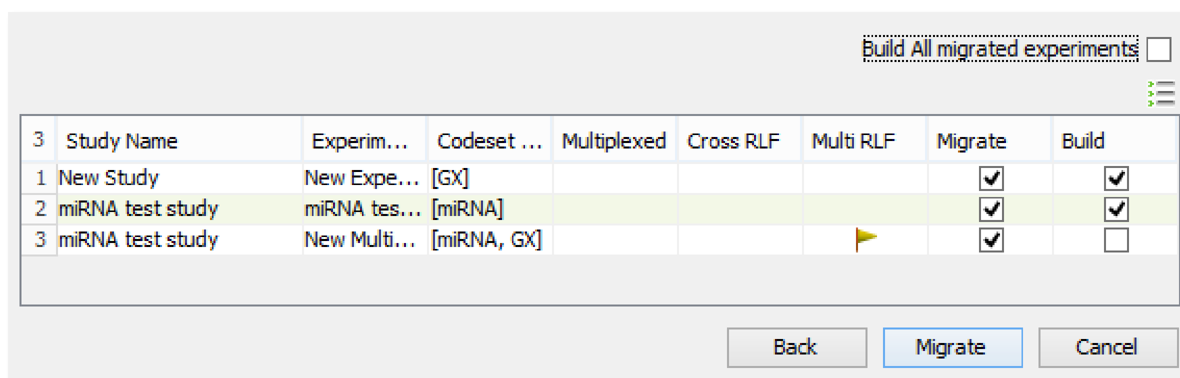



Figure 5: Experiment migration window

You may choose to **Build All migrated experiments** using the checkbox at the top of the window (see Figure 5) or address each study individually:

- Checking the **Migrate** box for a study will result in the transfer of all levels of data ***except for the analyses*** performed on that data. This is a faster option than the Build option.
- Checking the **Build** box will result in the transfer of all data, as in Migrate option, but will also ***include any analyses*** previously performed in nSolver 3.0. This is a more time-consuming option than the Migrate option.
- NOTE: migrating a multi-RLF Merge experiment requires that the parent experiments which comprised that multi-RLF Merge be migrated and built, as well. If you select a multi-RLF Merge experiment, the software will automatically select, migrate, and build whatever parent experiments were used to comprise that multi-RLF experiment.

Once you've made your selections, select **Migrate**. The data will migrate.

Upon opening nSolver 4.0, you will need to select migrated experiments and select the **Re-Build** button to reinstate them (see Figure 6). Data which was moved over using the **Build** feature does not require this additional step.

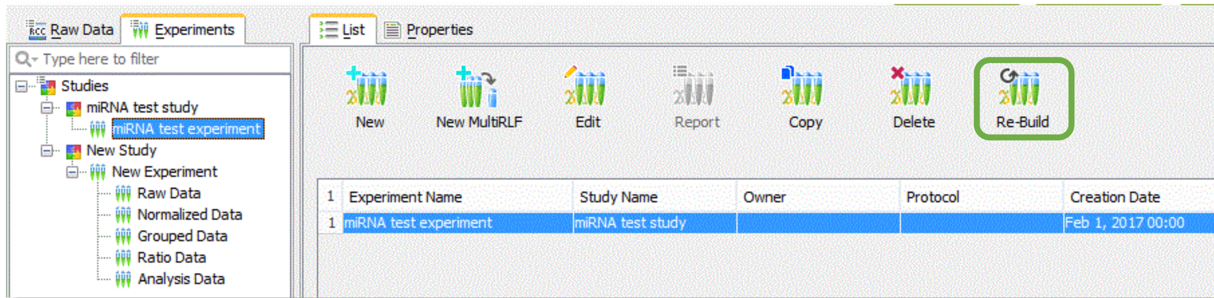



Figure 6: Rebuilding a migrated experiment.

# nSolver 4.0 Quick Start Guide

This quick-start guide is designed to direct you through the steps of importing, processing, exporting, and analyzing your nCounter data in nSolver 4.0. While this guide is appropriate for all analyte types and combinations, more detailed information may be needed and is available elsewhere in the nSolver 4.0 User Manual; section headings are hyperlinked for convenience. *All About SNV Analysis* (MAN-10075), *All About Fusion Analysis* (MAN-10076), *and All About PlexSet Analysis* (MAN-10044) guides are also available.

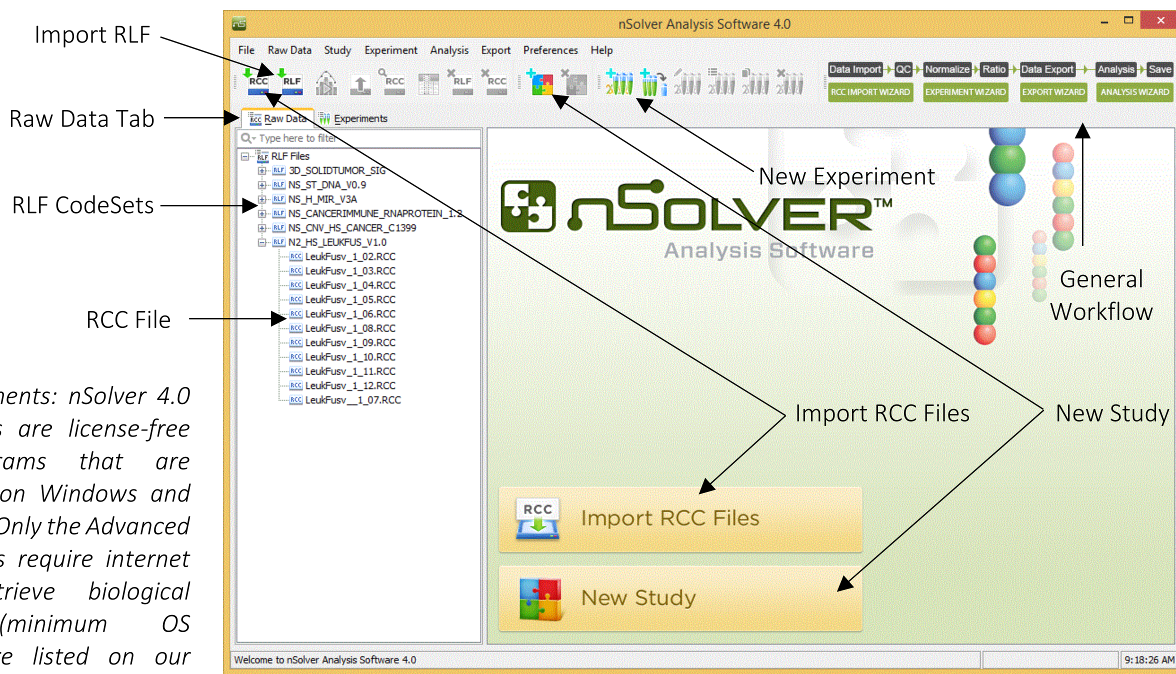
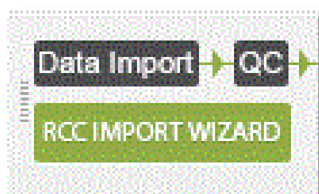



Figure 7: nSolver dashboard - raw data tab

*System Requirements: nSolver 4.0 and its modules are license-free software programs that are installed locally on Windows and Mac computers. Only the Advanced Analysis modules require internet access to retrieve biological annotations (minimum OS requirements are listed on our website).*

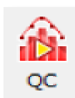



## Data Import

Open your data folder and unzip RCC data files using right click and **Extract All**. *Note: Most operating systems have built-in unzipping functions but a freeware program (e.g. 7-Zip) can achieve the same thing*.

Open **nSolver 4.0** and select **Import RLF,** then **Import RCC Files**. Follow the prompts in each process to navigate to your unzipped data folder and select files. Select **Next**. For SNV and Fusion data, refer to the *All About SNV Analysis* (MAN-10075) and *All About Fusion Analysis* (MAN-10076) guides.

## QC



Choose the RLF, then the QC parameters (see Figure 8). If hidden, select the double arrow at the right side of the screen to reveal the System QC parameters.

- The **Imaging QC** is a measure of the percentage of requested fields of view successfully scanned in each cartridge lane.
- The **Binding Density QC** is a measure of reporter probe density on the cartridge surface within each sample lane.
- The **Positive Control Linearity QC** is a measure of correlation between the counts observed for the Positive ERCC probes and the concentrations of the spike-in synthetic target nucleic acids.
- The **Positive Control Limit of Detection QC** indicates whether the counts for the POS_E control probe and target sequence, spiked in at 0.5fM (assumed to be the system's limit of detection) are significantly above the counts of the Negative control probes.

## QC (continued)

Additional QC parameters will appear, depending on the analyte types detected in your data. nSolver displays and applies the QC parameters recommended by NanoString; it is usually not necessary to adjust the default settings. Select **Import**.

Figure 8: QC window

## Exploring Raw Data

Your RCC data files will now be stored under the corresponding RLF CodeSet on the **Raw Data tab** (see Figure 7). Selecting the RLF name allows you to view all RCC files in a table format. Scroll to check for QC flags (see Figure 9). Selecting samples and clicking the **Table** button allows you to review the raw data in more detail. Examine the data to ensure that counts of POS/NEG controls and Housekeeping/Endogenous genes meet expectations, especially for samples with QC flags.

Figure 9: Table with QC flags

Figure 10: nSolver dashboard – experiments tab

## Creating Experiments

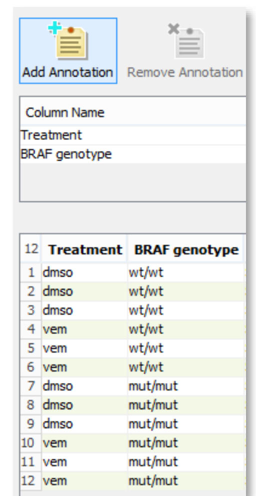
Within nSolver, any studies and experiments you create will be visible on the **Experiments tab** (see Figure 10).

A **Study** is an organizational folder used to store experiments; select the **New Study** button to create a study.

An **Experiment** is a collection of samples that have been analyzed together to allow comparisons between samples or samples grouped in conditions; select the **New Experiment** button to create an experiment under your study of choice. Follow the prompts to select the samples to include in your experiment.


nanoString

## Creating Experiments (continued)




| 12 | Treatment | BRAF genotype |
|---|---|---|
| 1 | dmso | wt/wt |
| 2 | dmso | wt/wt |
| 3 | dmso | wt/wt |
| 4 | vem | wt/wt |
| 5 | vem | wt/wt |
| 6 | vem | wt/wt |
| 7 | dmso | mut/mut |
| 8 | dmso | mut/mut |
| 9 | dmso | mut/mut |
| 10 | vem | mut/mut |
| 11 | vem | mut/mut |
| 12 | vem | mut/mut |

Figure 11: Annotations

**Annotations** to define sample groups should be assigned for experiments in which fold-change estimates and their statistical significance will be studied. These annotations can also be used in Advanced Analysis (see Figure 11).

The recommended method for adjusting for **Background** noise in data will appear by default for most analyte types. Confirm/select an option below (see wizard steps) and select **Next.**

- No background calculation (option clicked off or greyed out).
- Background thresholding, which uses a user-defined threshold count value; all raw counts below this value will be adjusted to it. This is recommended over subtraction.
- Background subtraction, which can be calculated by using the blank lane (if loaded) counts, by assigning a defined value (any negative value after background correction will be set to 1), or by taking the mean/geometric mean/median/max of the Negative Control counts.

### Normalize

**Normalization** can be accomplished by using the geometric mean of the Positive Control counts and by selecting normalization genes in the CodeSet Content. The settings recommended for most analyte types will appear by default. For SNV and Fusion data, refer to the *All About SNV Analysis* (MAN-10075) and *All About Fusion Analysis* (MAN-10076) guides. Review defaults, set preferences, and select **Next**.

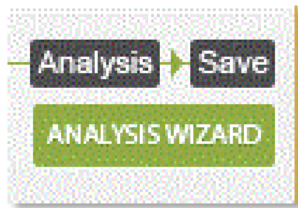
### Ratio

Fold Changes (**Ratios**) can be calculated by specifying the sample(s) that represent the baseline of your experiment. *All pairwise ratios* compares all groups to one another, while *Partitioning by* allows you to choose a group as the reference. Use the **Calculate False Discovery Rate** box (if active) to calculate FDR; output for this will be in the Ratio Table. Select **Next**, confirm the ratios you wish to calculate, and select **Finish**.

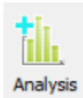
### Data Export

Your experiment will now be visible under your study on the **Experiments tab**. Expand the navigation tree. Select the desired data table, highlight samples of interest in the central window, and use the **Table** button to examine your data or the **Export** button to export results. Highlight an experiment and select the **Report** button (not shown) to run a report.

- The **Raw Data** table contains unprocessed data for all samples in this experiment.
- The **Normalized Data** table contains the processed data for all samples. Samples with unusually low counts for POS controls or Housekeeping genes may receive Normalization flags, which can be seen by scrolling to the far right in the central window. Paying particular attention to any flagged samples, review this data to ensure that counts of POS/NEG controls and Housekeeping/Endogenous genes meet expectations (refer to the respective *All About…* guide for SNV (MAN-10075), Fusion (MAN-10076), or PlexSet (MAN-10044) data).
- The **Grouped Data** table contains the geometric mean of expression levels for all samples from each group (as defined by the sample annotations).
- The **Ratio Data** table contains the fold-change results, as well as any statistical inferences surrounding those estimates. You may need to view hidden columns of data by right clicking any column header or using the column options icon to view all columns.
- The **Analysis Data** table contains any analysis you have run.

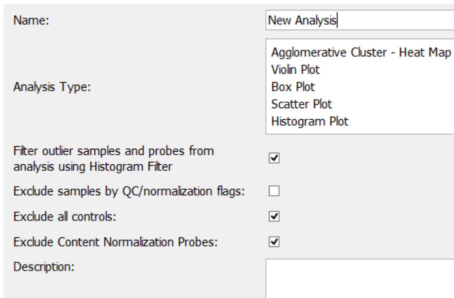



## Analysis & Advanced Analysis

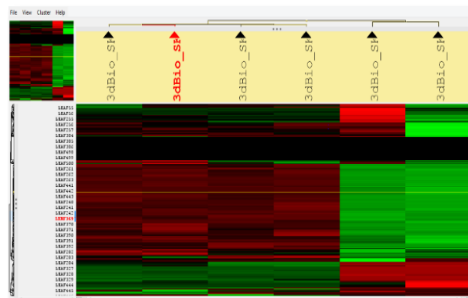



Highlight your data table and select **Analysis**. Select the plot desired (see Figure 12), then select **Next**. Select the samples, then the probes you would like included in your analysis and select **Next**/**Finish**. If creating a heatmap, you will be asked to set **Clustering Parameters**. Once your data is plotted, you can fine tune the settings.

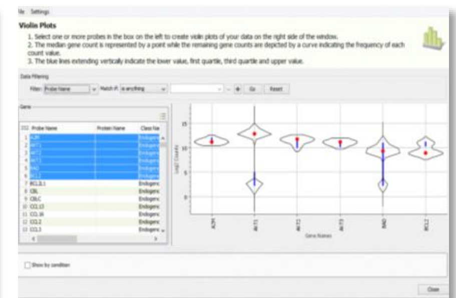
- **File** allows you to save and print the plot image.
- The tables to the left of the **Violin**, **Box**, and **Histogram** plots allow you to select the probes you would like included in the plots.
- The **Heatmap** is interactive and provides several customization options.
- The options to the left of the **Scatter Plot** allow you to select the sample(s) you would like included in the plot, as well as the color designations of the data points. Additional customization (such as the axis and legend settings) is possible in **Settings**.



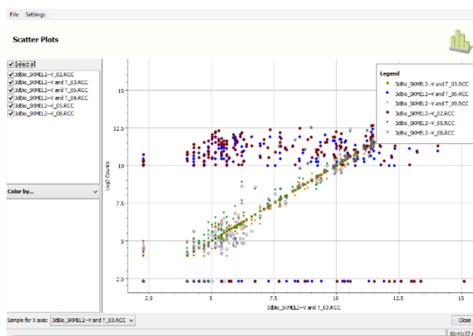

*Analysis Wizard menu*



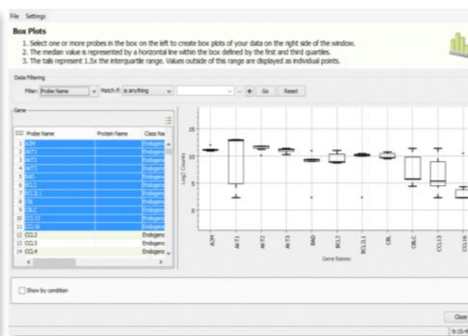
*Heat Map*



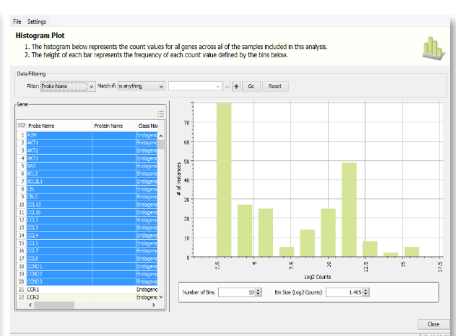
*Violin Plot*



*Scatter Plot*

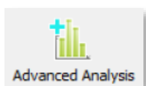


*Box Plot*



*Histogram*

Figure 12: Visualizations available in basic Analysis in nSolver 4.0




**Advanced Analysis** is available for raw and normalized SNV, Fusion, mRNA, and protein data (not for miRNA, CNV, Plex2™, or PlexSet™ assays) and uses the R program. This plug-in gives you more options and flexibility in statistically analyzing and visualizing your data. You can choose between a **Quick** and **Custom Analysis.**

Help

For troubleshooting and/or guidance through your analysis, contact support@nanostring.com

# Import

## RCC files

A Reporter Code Count (RCC) file is an output file generated by nCounter instruments. One RCC file is produced for each lane tested; this one file contains the barcode counts from each gene and control in the CodeSet.

After the completion of the instrument run, the data folder containing all RCC files from that run can be saved as a zipped file to a location of your choice on your computer or USB drive. Right-click or command-click on the zipped file and select **Extract All** or run your preferred file extraction program.  Once the extraction is complete, you should be able to see one RCC file for every cartridge lane.

Open **nSolver 4.0** and select either the **Import RCC Files** button on the central dashboard or the **Import RCC File** icon on the toolbar at the top of the page (see Figure 13). This will automatically open the ***RCC Import Wizard***, which will provide prompts to guide you in importing your data.  **Browse** to navigate to the folder in which your data was saved and select **Open**. Your RCC files will automatically populate the Import Wizard window in a table format (see Figure 14). You may use the checkboxes in the **Import** column to select which files to import.

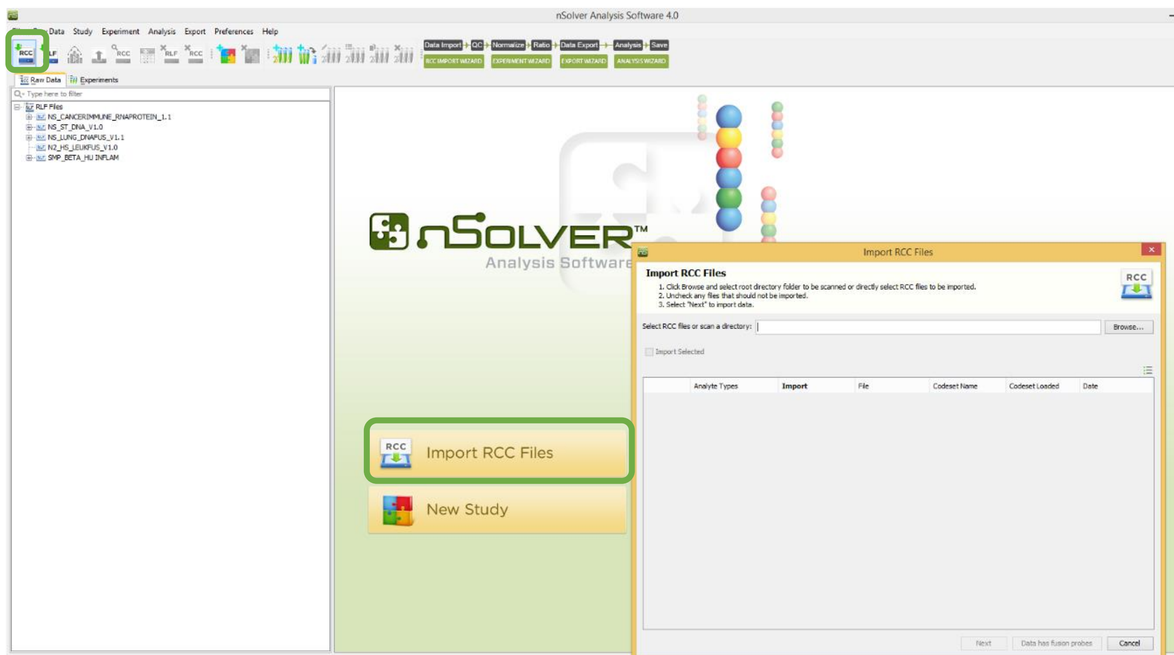



Figure 13: Importing RCC files

Checked boxes in the **CodeSet Loaded** column indicate that the RLF for this dataset has already been imported to nSolver. Conversely, unchecked boxes indicate that the RLF has not yet been imported.

Use the Column Options icon to select columns to hide or unhide. If processing fusion data, select **Data has fusion probes** (see *A Note About Fusion Assays* box, below).

Select **Next** to choose QC parameters (see the *QC* section of this manual for more information on these parameters).

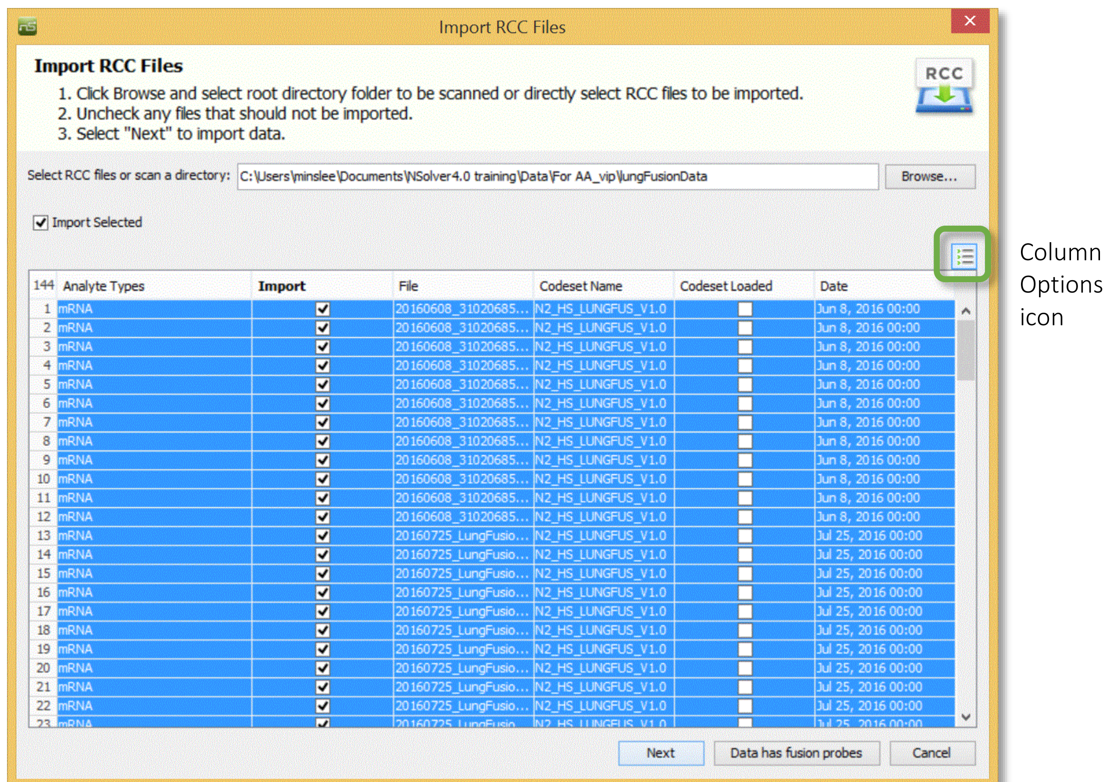



Figure 14: Importing RCC files - choosing files in the Import Wizard

A Note About Fusion Assays

If working with fusion data, the **Data has fusion probes** button allows you to designate fusion probes in your CodeSets. Use the **Select CodeSet** dropdown menu and the **Apply Analyte Type for selected Probes** dropdown menu to change the analyte type from the default *mRNA* to *Fusion*. If not designated, fusion probes will be treated as mRNA probes. If you neglected to select the Data has fusion probes button in error, you will need to delete the RLF and RCC files and re-import. Most fusion probes will have “FUS” (junction probes) or “3P” or “5P” (end probes) associated. See the *All About Fusion Analysis* guide (MAN-10076) for more information.

## RLFs

A Reporter Library File (RLF) is a file specific to your CodeSet. It provides necessary information to the nCounter instruments and the nSolver 4.0 Software Application about the CodeSet, such as the assignment of probe to gene or protein.

The MAX and FLEX instruments only scan for the barcodes in the RLF assigned at the onset of the run; data for other barcodes is not collected. For this reason, if the wrong RLF is applied on a MAX or FLEX instrument, the cartridges will need to be re-scanned with the correct RLF to ensure collection of data for all barcodes present in the CodeSet.

The SPRINT instrument scans cartridge lanes for all barcodes, regardless of the RLF that was assigned at the onset of the run. For this reason, you can apply a new RLF post-run and retrieve information on any set of probes without having to rerun the instrument.

Depending on what type of analysis you will run, nSolver may be able to pull all the information it needs from the RCC files, without importing the corresponding RLF.

Circumstances in which **you *may not need to* import an RLF** include:

- If you will only use **basic Analysis.**
- If you will only run a **single-RLF** experiment.
- If you **don't have any SNV reference samples** in the set.

Circumstances in which **you *must* import an RLF** include:

- If you plan on using the **Advanced Analysis** plug-in.
- If you plan on running **a multi-RLF experiment.**
- If you plan on using any of the **samples as SNV references**.

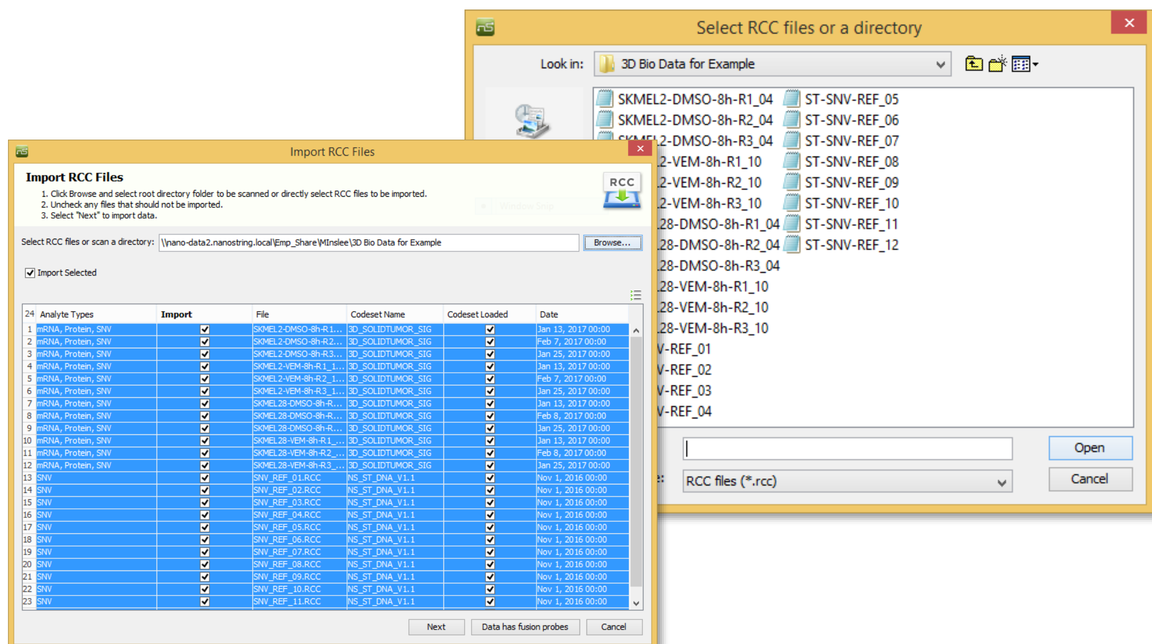
## Importing an RLF

Once you have determined that you need to import an RLF (see sections above), select the **Import RLF File** icon on the nSolver toolbar. **Browse** to navigate to the folder in which your RLF is stored and select **Import**.

## Import Example

When importing the files contained in the data folder, **3D Bio Data,** note the following:

- The **Analyte Types** column lists mRNA, Protein, and SNV for the samples (and only SNV for the SNV reference samples).
- All boxes in the **Import** column are checked, indicating that you want to import each sample.
- The **CodeSet Name** column; your data will ultimately be stored under this name on the nSolver raw data tab.
- The boxes in the **CodeSet Loaded** column are checked, indicating that you have imported the RLF.




.

Figure 15: Windows associated with RCC import



nanoString™

## Quality Control (QC)

The Quality Control (QC) Parameters window (Figure 16) appears when importing files for analysis and can reappear by selecting the QC button (see the *Using Table Options Buttons* section). The parameters listed will depend on the analyte types detected in the dataset. Select the analyte icon from the analyte list on the left side of the window to see the associated QC parameters. You can select or de-select individual QC parameters to indicate whether your samples should be assessed for each, however, it is recommended that an overall QC check be performed on all imported data. See additional details in the respective assay-specific sections for *Plex*[2] and *PlexSet* assays.

There are two quality control parameters common to all nCounter assays (Imaging QC and Binding Density QC) and numerous others that are specific to particular assays. The default settings for each parameter represent the typical levels of performance. If a sample performs outside of one of these ranges, the software system will produce a flag in that category of the data table (see the *Exploring Raw Data* section).

***In most cases, it is not necessary nor recommended that you change the default QC settings***. They are customizable, however, and altering these parameters may be appropriate to prevent flagging of useful data. See sections below for details on each of the QC parameters. Once you have selected or confirmed the QC settings, select **Import** (or **Run QC** if you are rerunning QC) and proceed to the *Exploring Raw Data* section of this manual.

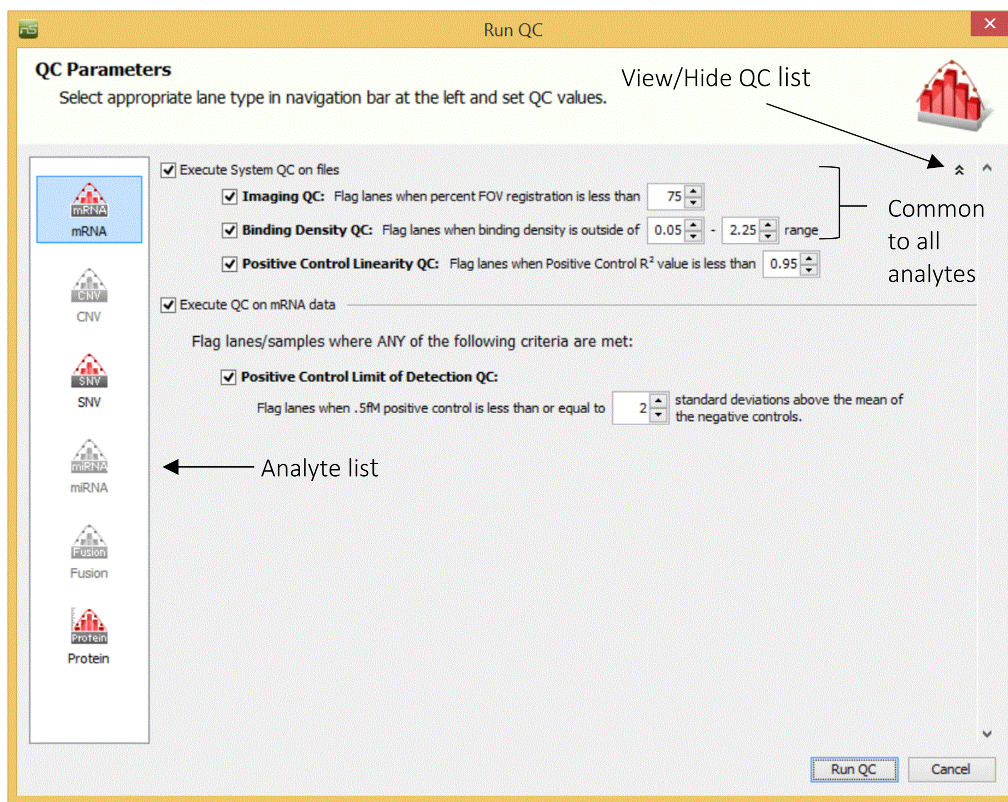



Figure 16: QC parameters window

## QC Parameters for All Analytes

The **Binding Density** and **Image Quality** (**Imaging**) are QC parameters measured for every assay. The Imaging QC measures how much of the lane the instrument was able to see and the Binding Density reports how dense the barcode count was in the scan.

### Binding Density QC

*mRNA | CNV | SNV | miRNA | Fusion | Protein*

This metric is a measurement (in spots per square micron) of the concentration of barcodes seen by the instrument. The Digital Analyzer may not be able to distinguish each probe from the others if too many are present. The ideal range **0.1 - 2.25** spots per square micron has been established for assays run on an **nCounter *MAX* or *FLEX*** system and **0.1 - 1.8** spots per square micron on the **nCounter *SPRINT*** system.

Measurements outside of these ideal ranges will be flagged, but should be checked to see how much they deviate from the ideal range. If they are *only slightly outside of range, they do not indicate a problem* in the data and can be bypassed. If they deviate a great deal, troubleshooting should continue since this may indicate reduced resolution. See the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section.

Troubleshooting a Binding Density Flag

- **Check the input amount.** More sample input will result in an increased Binding Density.
- **Consider the expression level** of the targets in the CodeSet. If the targets in the CodeSet are highly expressed, Binding Density will go up simply because more mRNA molecules are being targeted in your samples.
- **Consider the size of the CodeSet**. If a CodeSet contains probes for more targets, then Binding Density will usually be higher.

### Imaging QC

*mRNA | CNV | SNV | miRNA | Fusion | Protein*

This metric reports the percentage of fields of view (FOVs) the Digital Analyzer or Sprint was able to capture. At least **75%** of FOVs should be successfully counted to obtain robust data. A flag in this area may indicate something as simple as a crooked or smeared cartridge, which can be remedied by rescanning (MAX or FLEX systems only; ideally within one week). Consistently reduced percentages, however, can be indicative of an issue associated with the instrumentation (see the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section).

Troubleshooting an Imaging Flag

- **Check alignment of the cartridge** in the instrument – if crooked, reload and rescan.
- **Check the bottom of the cartridge** - clean with 70% EtOH and a lint-free wipe, reload, and rescan.

## Positive Control QC Parameters

Positive controls assess three general QC purposes:

- **Overall assay efficiency**. nSolver raises a warning flag when the geometric mean of positive controls is more than three-fold different from the geometric mean of all samples.
- **Assay linearity**. Decreasing linear counts are expected from POS_A to POS_E (POS_F is considered below the limit of detection).
- **Limit of detection (LOD).** It is expected that counts for POS_E will be higher than background, which is represented by the mean of the negative controls plus two standard deviations (for most assays) or simply the mean of the negative controls (for miRNA assays).

## Positive Control Linearity QC

*mRNA | CNV | miRNA | Fusion | Protein*

This metric performs a correlation analysis in $log_2$ space between the known concentrations of positive control target molecules added by NanoString and the resulting counts. Correlation values lower than **0.95** may indicate an issue with the hybridization reaction and/or assay performance. See the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section.

### Troubleshooting a Positive Control Linearity Flag

If all the following are true, it is safe to include the flagged sample:

- r-squared values are ~ 0.90
- only a single POS control is abnormal
- the raw counts for the sample appear normal

If one or more of the above is *not* true, contact support@nanostring.com

### Positive Control Limit of Detection QC

*mRNA | CNV | miRNA | Fusion | Protein*

This measures the limit of detection of the assay by comparing the results from the positive control probes and those from the negative control probes. Specifically, it is expected that the 0.5 fM positive control probe (*Pos_E*) will produce raw counts at least **two standard deviations higher than the mean of the negative control** probes. You can modify the number of standard deviations used to estimate significance. In the event of a flag, see the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section.

Troubleshooting a Positive Control Limit of Detection Flag

Too high background may be due to:

- Mixing of reporter and capture probes in hybridization buffer master mix.
- Too much time elapsed between adding capture probe and loading in thermal cycler.
- High counts in one of the negative controls – NEG control could be elevated due to cross hybridization with targets in the sample.

Too low POS counts may be due to:

- Sub-optimal hybridization – check thermal cycler temperature and consider whether sample impurities (chaotropic salts, for example) may have been introduced.

## Additional Analyte-Specific QC Parameters

Depending on the nature of the assay and the analyte tested, additional QC parameters may be offered. nCounter Custom assays that target DNA and provide copy number estimates, such as CNV and SNV, require additional controls and QC parameters to ensure high quality results.

### Ligation QC

*miRNA*

Each miRNA assay contains six short synthetic RNA constructs. Three of these are subjected to ligation and each release a miRNA tag as a positive ligation control. The other three are not subject to ligation and serve as ligation-negative controls. Ideally, the ligation-negative controls will yield counts in the negative control range (background level) and the ligation-positive control counts will be significantly higher, increasing from LIG_POS_C to LIG_POS_A. In the event of a flag, see the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section.

Troubleshooting a Ligation Flag

- If counts for LIG_POS controls are at background level (similar to those for LIG_NEG controls), consider whether the ligase enzyme may have been omitted from the reaction entirely or indirectly (added to the side of the tube, for example).
- Consider whether sample impurities (chaotropic salts, for example) may have been introduced.


nanoString

### Fragmentation QC

*CNV*

Four fragmentation control probes (two positive and two negative) are included with each Custom CNV assay. Their target oligonucleotides are 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. The Fragmentation QC measures this. See the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section.

Troubleshooting a Fragmentation Flag

- If you used alternative methods to Alu1 digestion with NanoString's CNV Sample Preparation reagents for DNA fragmentation, disregard this QC flag.
- Consider whether sample impurities (chaotropic salts, for example) may have been introduced.

### Invariant Control QC

*CNV*

The Invariant Control QC measures the level at which the invariant control probes contained in each custom CNV assay bound to their target regions. A mean invariant count level of **over 100 counts** typically ensures normalization will be performed within the linear range of the system, whereas mean invariant count levels of less than 100 counts may result in less accurate normalization and/or copy number estimates. See the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section.

Troubleshooting an Invariant Control Flag

- Check whether counts for the general positive controls are within the expected range to ensure hybridization, itself, was optimal.
- Relate Invariant Control counts to counts of other CNV targets.
- Check sample input amounts.

### CarryOver Contamination QC

*SNV*

The CarryOver Contamination QC is an internal control consisting of a control template which contains dU's and a probe set. The absence of a flag in this column implies that the UDG digest prep step was successful and the probe set found no template with which to hybridize. A flag in this column implies that the probe set did, indeed, bind to the template, which indicates that it was not digested properly. See the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section.

Troubleshooting a CarryOver Contamination Flag

- Consider factors which may have allowed PCR product contamination from previous experiments.
- SNV calls from lanes with this flag should be interpreted with caution, as they may result from amplified sample carried-over from a previous sample.

### Amplification Control QC

*SNV*

The Amplification Control is a set of internal controls consisting of templates and probe sets. A flag in this column appears when this signal's yield is lower than expected, indicating that overall amplification was not ideal. The calling algorithm may fail due to this sample being an outlier. Including a lane with an Amplification Control flag may increase overall noise of the data group and affect calling of other samples. See the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section.

Troubleshooting an Amplification Control Flag

- Consider factors impacting amplification efficiency, such as contamination with inhibitors, pipetting error, or poor mixing during PCR prep.
- SNV calls from lanes with this flag should be interpreted with caution. Data quality may improve by excluding this lane and rerunning the sample.

### Sample Amplifiability QC

*SNV*

The Sample Amplifiability QC consists of probe sets for three largely invariant human genes. A flag in this column indicates that these genes did not amplify as efficiently as expected and may indicate overall low quality in the genomic sample. **If Sample Amplifiability *is* flagged, while Amplification Control QC is *not*,** this indicates that either the gDNA sample, itself, is of poor quality or too little of it was added. Inclusion of this lane may increase overall noise of the data group and affect the calling of other samples. See the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section.

Troubleshooting a Sample Amplifiability Flag

- Consider factors affecting overall genomic quality, such as degradation.
- Consider whether genomic DNA input may have been low.
- SNV calls from this lane can be used, but should be interpreted with caution.
- Data quality may improve by excluding this lane. Rerun the sample, adding an additional cycle of PCR.

### Lane Temperature QC

*SNV*

The Lane Temperature QC consists of both positive and negative controls whose probes have very similar sequences. A flag in this column indicates that during the Prep Station sample processing steps, the instrument's temperature was not controlled precisely enough to produce the appropriate count ratios for these controls. This will increase the likelihood of False Positive and False Negative calls. See the *Troubleshooting* box, here, and the *What to Do If You Have a Flag* section.

Troubleshooting a Lane Temperature Flag

- Consider factors that would contribute to the instrument running outside the most optimal temperature range.
- Ensure that all current instrument software updates have been installed.
- Ensure that you are running on a Qualified Prep Station. If you are unsure if your Prep Station has been qualified, contact SNVQualify@nanostring.com
- SNV calls from this lane can be used, but should be interpreted with caution.
- If this flag is seen in multiple lanes and/or multiple runs (>2 lanes per occasional run), contact NanoString Support.

## QC Example

The following is an excerpt from the **3D Bio Data Example**, which can be found in *Appendix A.*

In choosing the QC parameters, note the following:

- You must use the **double arrow** icon in the right corner of the screen to reveal the top three (System QC) parameters.
- The activated buttons in the panel of **analytes** along the left side of the window represent the analytes detected in your data (mRNA, SNV, and Protein). Selecting an analyte reveals the default QC parameters associated with it.
- You may change the QC parameters, but this is not usually recommended nor necessary.

Select **Import**.

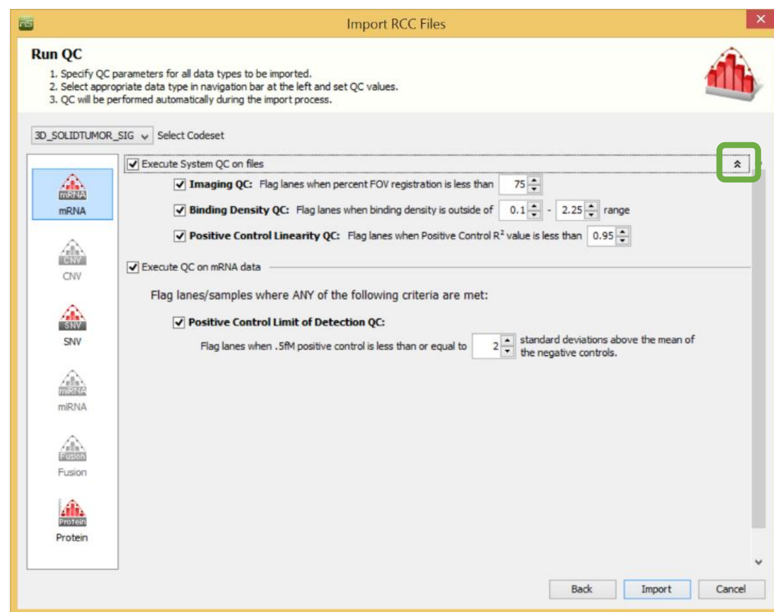



Figure 17: 3D Bio Data Example – QC

# Exploring Raw Data

Imported RCC files and RLFs should be visible on the **Raw Data tab**. Expanding the navigation tree (by clicking on the **+** sign) reveals the list of RCC files in the set and generates a central table of information. It is here that you can check for QC flags, customize your samples by adding tracking information to column fields (such as the **Description** or **Batch ID** fields), and delete or export files (see Figure 18). See additional details in the respective assay-specific sections for *Plex²* and *PlexSet* assays.

## Checking for QC Flags

Samples that fail to pass one of the QC parameters will display a flag in that parameter's column as well as one in the overall QC Flag column. After importing your data, check all samples for QC flags (you may need to scroll right to see these columns). Start by selecting the CodeSet (RLF) name on the Raw Data tab. You will see a table of your samples appear in the main screen. You can then search for QC-flagged samples in several ways:

- **Use the filter tool** above the data table to search for only flagged data. Do this by setting the filter to find files where the **QC Flag** is **YES** (see example in Figure 18).
- Click on the **column header labeled *QC Flag*** and sort by *QC Flag* to simply view all the QC flagged files on top. Note: you may click and drag column headers to move columns.
- **Create a table** (see the *Using Table Options Buttons* section), scroll through the samples in the data table and look for column headers in red (this indicates association with a QC flag).

Figure 18: Central table viewer orientation and depiction of checking for flags

## What to Do If You Have a Flag

QC flags presented in your data imply that closer examination of the affected lane is warranted prior to proceeding with analysis. Oftentimes, flags are raised due to intended attributes of a particular CodeSet or sample. QC flags do NOT necessarily mean that data is unreliable.

To determine whether a QC flag is indicating a critical problem, **examine the raw and normalized data** (normalized data is generated while making an Experiment; see the *Creating Experiments* section). Specifically, check low count transcripts to determine if flagged samples have a poorer limit of detection than non-flagged samples. There are several ways to do this:

- **A simple visual scan of the data** may suffice to detect problems in the flagged samples. This can be performed on raw data which have been background subtracted in nSolver to identify targets that are below the background.
- Review the results from the **positive and negative controls**. Positive controls with low counts or negative controls with counts significantly above background can trigger flags and should be checked to see if they indicate more serious issues with the data.
- **Outlier samples can be identified by generating a heat map** of normalized data from all samples to see if the flagged samples in question are strongly divergent from other samples with similar pathology. You will need to proceed through the steps to create an Experiment before doing an Analysis, at which point you can refer to the *Agglomerative Cluster (Heat Map)* section.
- You can **examine the calculated QC metrics**. Right-click or command-click on one of the table column headers in the raw data table or use the column options icon to *Show all Hidden Columns*. This will reveal a column next to each QC flag column containing the numbers on which the QC check was based. If these QC metrics only deviate from the threshold by a very small margin (i.e., the FOV registration is 74% instead of 75%), then the resultant data may be quite robust and usable. To review the parameter settings, select the QC button above the table.

In addition to these general guidelines, refer to the troubleshooting boxes in the individual *QC Parameters* sections, above, for more troubleshooting ideas on individual QC parameter flags. Normalization QC flags are addressed in the *Normalization* section.

In cases where there are no QC flags, it is still advised that you examine the calculated QC metrics and check whether there are outliers. For example, if the Imaging QC for 11 out of 12 samples is 95% or higher and one sample presents a value of only 76%, this would not trigger a QC flag. However, it is still important to investigate why this one sample may have had a significantly lower imaging quality score.

## Re-running QC

Because many nCounter CodeSets are custom designed, some cases may warrant the parameters for QC being adjusted to fit the needs of specific datasets or experimental designs. If you ran QC using the default parameters and observed flags that seemed to be set unnecessarily, you can either ignore the flags or re-run QC. To re-run QC, select the appropriate lanes/samples and select the **QC** button. This opens the QC Wizard for the assay type(s) selected. Change the thresholds, as appropriate, and select **Run QC**.

## Exploring Raw Data Example

The following is an excerpt from the **3D Bio Data Example**, which can be found in *Appendix A*.

- Selecting the CodeSet name, **3D_SOLIDTUMOR_SIG,** allows you to view all RCC files under it in a table format. You may edit fields in the Description and Batch ID columns, if desired.

- As you scroll right in this table, notice that the first QC column is labeled simply **QC Flag**. A flag in this column indicates that there is a QC flag of some sort associated with this sample. Since there are no flags in the QC Flag column, we know that there are no QC flags associated with this data, but we will still scroll farther to the right and note the individual QC columns.

| QC Flag | Imaging QC Flag | Imaging QC | Binding Density ... | Binding De... | Positive Control ... | Positive Co... | Limit of Detectio... | Limit of Detectio... | CarryOver Cont... | CarryOver... | Amplification Co... | Amplificati... |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0.96 | | 0.8 | | 0.98 | | 12.07 | | 252.45 | | 34.84 |
| | | 1 | | 1.03 | | 0.98 | | 11.19 | | 366.74 | | 25.07 |
| | | 1 | | 0.66 | | 0.98 | | 7 | | 299.41 | | 21.16 |
| | | 0.99 | | 0.8 | | 0.98 | | 9.5 | | 273.07 | | 24.96 |
| | | 0.99 | | 0.77 | | 0.98 | | 10.26 | | 221.98 | | 22.22 |
| | | 1 | | 0.56 | | 0.98 | | 6.83 | | 162.9 | | 24.51 |
| | | 0.99 | | 1 | | 0.98 | | 16.1 | | 376.76 | | 38.38 |
| | | 1 | | 0.98 | | 0.98 | | 16.36 | | 276.84 | | 19 |
| | | 0.98 | | 0.72 | | 0.98 | | 9.11 | | 352.96 | | 17.25 |
| | | 1 | | 0.96 | | 0.99 | | 11.83 | | 373.14 | | 48.35 |
| | | 1 | | 0.87 | | 0.98 | | 10.98 | | 304.11 | | 27.68 |
| | | 1 | | 0.64 | | 0.98 | | 11.56 | | 291.76 | | 22.51 |

Figure 19: 3D Bio Data Example - examining QC flags

- Select the **Column Options** icon in the upper right corner and select **Show All Hidden Columns** (this option will not be available if all columns are already being shown). This will reveal the QC columns with numerical data. Note that the values of the QC metrics vary only slightly from sample to sample. This is one indication of good data.

## Viewing & Customizing Samples

### Editing Columns and Selecting Samples

Columns can be hidden or revealed by right-clicking any column header and selecting **Select Columns** or by selecting the **Column Options** icon (Figure 20). Columns can be moved simply by clicking the column header and dragging it to the desired position. Clicking on the header of a column can also sort the samples in the ascending or descending order of the column's contents. Multiple samples can be highlighted using a mouse click combined with the control key, the shift key, or by clicking and dragging. The **number in the upper left corner** of the table highlights all samples.

### Description Column

The Description column is available to add a description of your sample that is informative for tracking. This information is not used for analysis, as annotations are when building an experiment (see the *Annotations* section). To add descriptions, click in the column and begin typing. You may also **Copy** and **Paste** this information from another source.

### Gender Column

The Gender column is typically only used in Copy Number Variance (CNV) assays. Use the drop-down menu in the Gender column to select **Male**, **Female**, or **Unknown** for each sample. You may also **Copy** and **Paste** this information from another source. The Gender Column may be hidden, depending on the analyte types in the data. See the *Editing Columns & Selecting Samples* section, above.

### Batch ID Column

The Batch ID column allows you to record CodeSet lot numbers or other identifying information of potential sources of variability in a run. You may also **Copy** and **Paste** this information from another source. When different Batch IDs are recorded across a set of samples, the software will give you the option to designate a reference sample to correct for any difference in performance between batches. You can also accomplish Batch Calibration by creating a Cross-RLF Experiment; this workflow will prompt you to designate a

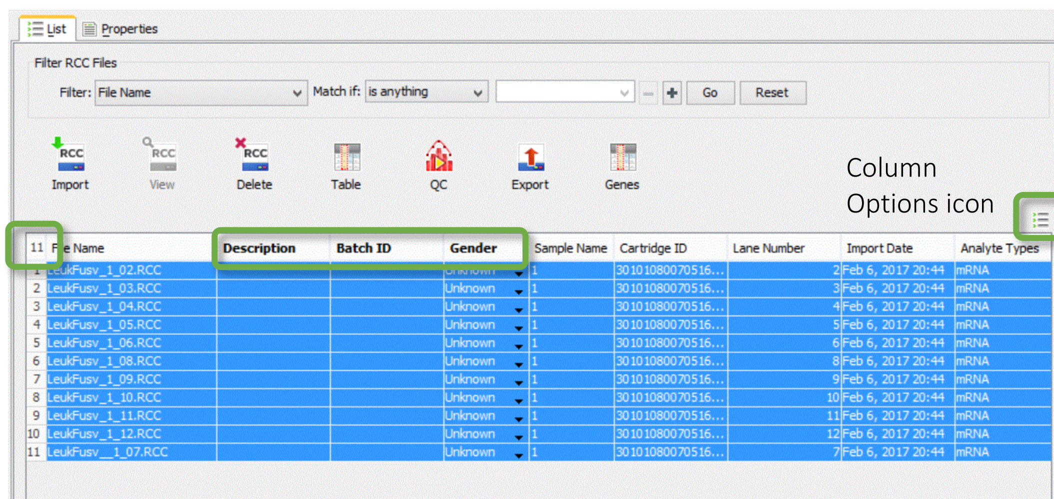



Figure 20: Central table viewer

reference sample that was run in both lots (or on both instruments) to calibrate the data. See the *Multi-RLF Experiments & Batch Calibration* section.

## Using Table Options Buttons

When you select one or more samples in the central data table, the **Table Options** buttons above it become active (Figure 21).

Figure 21: Active Table Options dashboard

### Import RCC Button

Select the **Import RCC** button to import new samples to the current set. See the *RCC Files* section of this manual for more information on importing RCC files.

### View RCC Button

Use the **View RCC** button to view the text file of your RCC data. Select one RCC file in the central data table viewer to activate this button.

### Delete RCC Button

If you choose to delete a sample, highlight it and select the **Delete RCC** button and confirm with **OK**. If the file is already incorporated into an experiment, you will need to first delete that experiment (see the *Experiments* section) and then return to delete the RCC file.

## Table Button

Select the **Table** button to view a data table (Figure 22) by probe. You can use the **Filter Expression Data** section to quickly find data of interest. Use the **View Options** section to change how your data is displayed. Red column headers indicate a sample which has a QC flag associated with it. Use the **Export** button to export the selection of your choice from this window. See the *Data Output & Export* section.

**Raw Expression Data**

Counts are presented here in a matrix to allow for visual checking of data, filtering of probes of in

Filter Expression Data

Filter: Probe Name | Match if: is anything | Go | Reset

View Options

Show data as Counts

| 54 | Accession # | Class Name | Annotation | LeukFusv_1_02... |
|---|---|---|---|---|
| 1 | NM_024812.2 | Endogenous | | 22,122 |
| 2 | tFUS_10002.1 | Endogenous | | 15,713 |
| 3 | tFUS_10003.1 | Endogenous | | 9,164 |
| 4 | tFUS_10006.1 | Endogenous | | 13,061 |
| 5 | tFUS_10007.1 | Endogenous | | 6,876 |
| 6 | tFUS_10008.1 | Endogenous | | 10,672 |
| 7 | tFUS_10004.1 | Endogenous | | 10,099 |
| 8 | tFUS_10005.1 | Endogenous | | 3,397 |
| 9 | tFUS_10009.1 | Endogenous | | 9,597 |
| 10 | tFUS_10010.1 | Endogenous | | 18,465 |
| 11 | tFUS_10011.1 | Endogenous | | 7,454 |
| 12 | tFUS_10012.1 | Endogenous | | 13,206 |

Figure 22: Table button displays a raw expression data table

## QC Button

You may revisit the QC parameters for your samples by selecting the **QC** button. You may simply review them or you may change them and rerun the QC check. See the *QC* section.

## Export Button

Select the **Export** button to export your entire data table to manipulate in the program of your choice. See the *Data Output & Export* section.

## Genes/Probes Button

Select the **Genes** (this may appear as **Probes**, depending on the assay) button to view a table of probe names and chromosomal locations (Figure 23). You can use the **Filter Data** section to quickly find your probe or region of interest. In addition, you can add annotations to each probe by editing or copying and pasting in the **Annotation** column cells. This is informative for exports, and particularly useful for filtering when creating visualizations.

**Gene Info**

Add custom Gene annotations

Filter Data

Filter: Probe Name | Match if: is anything | Go | Reset

| 54 | Probe Name | Analyte Type | **Annotation** | Class Name | Genomic Coordi... |
|---|---|---|---|---|---|
| 1 | BAALC | mRNA | | Endogenous | |
| 2 | BCR_13:ABL1_2 | mRNA | | Endogenous | |
| 3 | BCR_13:ABL1_3 | mRNA | | Endogenous | |
| 4 | BCR_14:ABL1_2 | mRNA | | Endogenous | |
| 5 | BCR_14:ABL1_3 | mRNA | | Endogenous | |
| 6 | BCR_19:ABL1_2 | mRNA | | Endogenous | |
| 7 | BCR_1:ABL1_2 | mRNA | | Endogenous | |
| 8 | BCR_1:ABL1_3 | mRNA | | Endogenous | |
| 9 | CBFB_5:MYH11_12 | mRNA | | Endogenous | |
| 10 | CBFB_5:MYH11_7 | mRNA | | Endogenous | |
| 11 | CBFB_5:MYH11_8 | mRNA | | Endogenous | |
| 12 | DEK_2:NUP214_6 | mRNA | | Endogenous | |
| 13 | ERG | mRNA | | Endogenous | |
| 14 | ETV6_5:RUNX1_2 | mRNA | | Endogenous | |
| 15 | ETV6_5:RUNX1_3 | mRNA | | Endogenous | |

Figure 23: Genes button displays a gene info table

## Creating Experiments

Experiments allow you to normalize, group, and create fold change estimates with your data. This then allows you to perform detailed analyses. A study must be created first; multiple experiments can be grouped under a study. The ***Experiment Wizard*** guides you through this process.

### 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 your data easy to find. A study can contain multiple experiments and span multiple CodeSets and assay types.

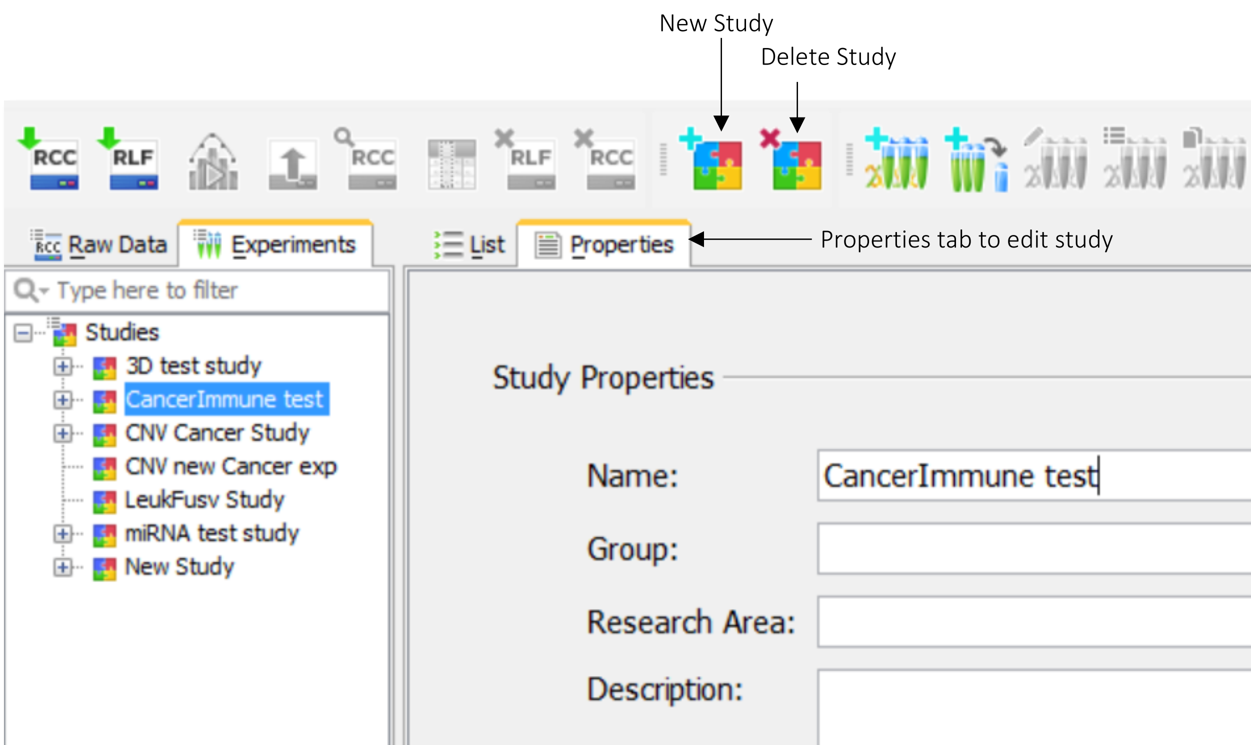



Figure 24: Buttons necessary to create, delete, and edit studies

To create a new study, select the **New Study** button on the main dashboard (Figure 24) and enter a unique study name. The Group, Research Area, and Description fields are optional. Select **Save**.

Your new study should now be visible on the **Experiments** tab. To **edit** the name or other details, select your study and then select the **Properties** tab. Edit your study and select **Save**.

To **delete** a study, select your study on the **Experiments tab** and select the **Delete Study** button. If a study contains experiments, you must delete the experiments first (see the *Experiments* section, below).

## Experiments

In creating an experiment, you will assign annotations to samples, set parameters for background thresholding, initiate normalization, and indicate the fold changes (ratios) you need to answer your biological questions. Creating **Single-RLF Experiments** and assigning **Annotations** are addressed here. **Multi-RLF Experiments** are created differently and are addressed in the *Multi-RLF Experiments* section.

### Single-RLF Experiments

The following steps will allow you to create a standard experiment using datasets run with a single RLF or CodeSet. If using the Batch ID column to calibrate across multiple reagent lots or instruments, see the *CrossRLF and Batch Calibration* section.

To create a new experiment, select a **New Experiment** button on the main dashboard (Figure 25) and enter a unique experiment name. Select the study with which your experiment should be associated from the drop-down menu. The Owner, Protocol, and Description fields are optional. Select **Next**.

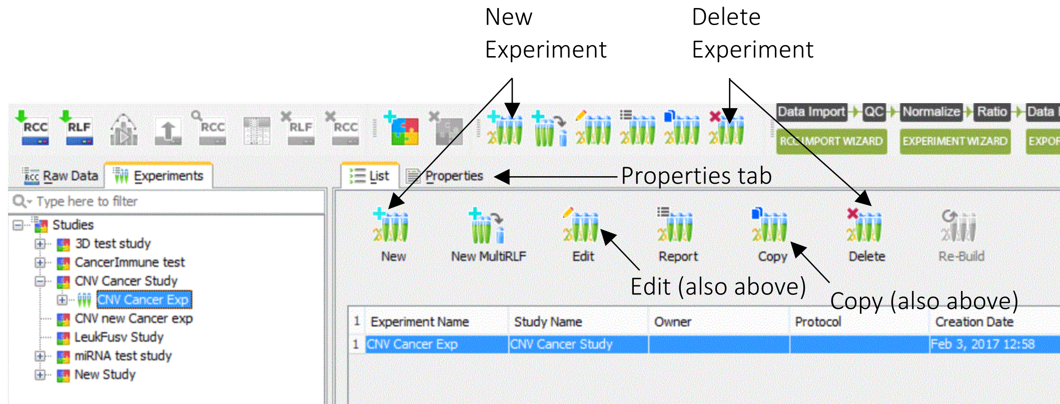



Figure 25: nSolver dashboard with buttons to create, edit, and delete experiments

In the *Add Sample/Lanes* window, choose the raw data that you want to add to your experiment. There are several ways to search for and select the data you are looking for. You can:

- **Select the analyte type** you want to study from the drop-down menu on the left; only data of the chosen analyte type will be displayed.
- **Select the CodeSet of interest** from the navigation tree on the left. All RCC files in the selected CodeSet *that contain your chosen analyte-type* will display in the central table.
- **Use the filters** above the central table window and click the **Go** button to display only files of interest.
- **Select rows** and use the **Keep Selected** or **Exclude Selected** buttons to filter out any unwanted samples. The **Show Excluded** button displays all files once again.

Once your desired samples are displayed, select **Next**.

Go to the *Annotations* section for next steps.



nanoString

If you later wish to edit the name of the experiment or other details, select your experiment and then select the Properties tab or the **Edit Experiment** button (see Figure 25; this has more options than the Properties tab). Edit your experiment and select **Save**.

You can also select **Copy Experiment** (see Figure 25) if you want to create a new experiment with most of the same settings as an existing experiment.

To delete an experiment, select your experiment on the Experiments tab and select the **Delete Experiment** button (see Figure 25).

## Creating Experiments Example

The following is an excerpt from the **3D Bio Data Example**, which can be found in *Appendix A*.

- Select a **New Experiment** button on the main dashboard and enter a unique experiment name. The study you had highlighted should be listed as the default, but if not, you can select it from the dropdown menu. Select **Next**.
- Select the **3D_SOLIDTUMOR_SIG** CodeSet from the list on the left, then select the samples to include in or exclude from the experiment. For this example, we will use all samples in the dataset, but if you needed to filter samples in a larger dataset you could:
  - Use the filter tool for a **File Name** that **contains 8h**, then select **Go** (see figure below).
  - Highlight the desired samples and select **Keep Selected**.
  - Highlight the samples you don't want to keep and select **Exclude Selected**.
- Select **Next.**

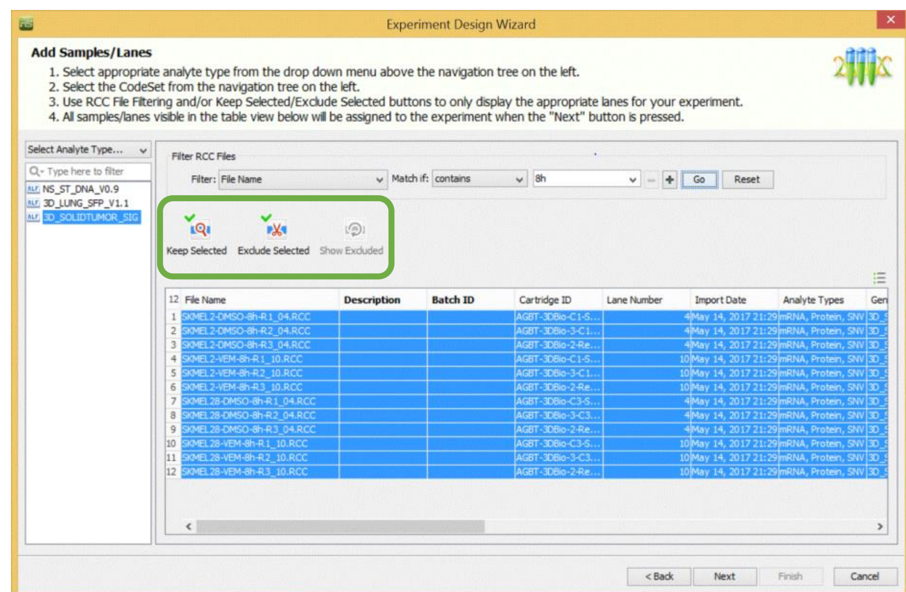



Figure 26: 3D Bio Data Example - creating an experiment

# Annotations

The next step is to **Add Sample Annotations**. This adds columns with which you can annotate each file and group for downstream analysis.

Click on the **Add Annotation** button to add a new column (Figure 27), then assign it a name. Under **Column Type**, you can choose **Text**, for an editable field, **True/False**, for a checkbox, or **Numeric**, for an editable field and neighboring **Unit Name** field.

Add your annotation to the columns (for examples or ideas, see *Annotations Example*, below). You can type directly into each of the cells or copy and paste from another source. If adding an annotation for one sample, *all rows must have data entered.* If the information you are adding is a sample-specific identifier (rather than a grouping), consider adding that information in the Description column instead of Annotations. See the *Viewing & Customizing Samples* section. Proceed to the *Background* section for next steps.

The **Remove Annotation** button is available, if needed.

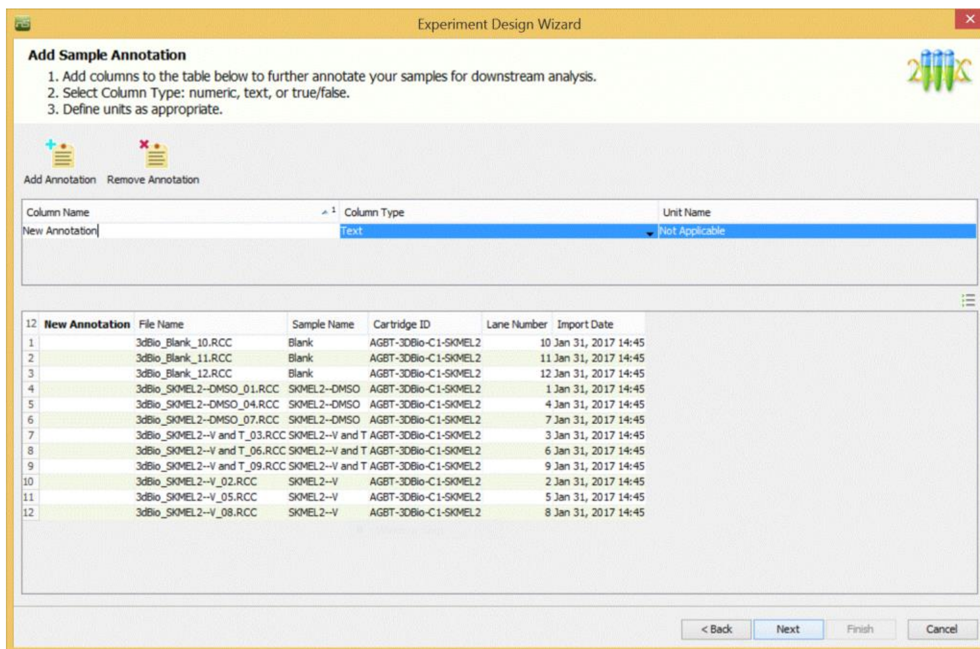



Figure 27: Add Sample Annotations window

## Annotations Example

The following is an excerpt from the **3D Bio Data Example**, which can be found in *Appendix A.*

The treatment type, Sample #, and time for each sample has been incorporated in the sample name. These are good categories to use for sorting and analyzing data, but it is only by adding this information as a separate annotation that it can be utilized as a variable for differential analysis. When we created an experiment using this data, we included samples from two cell types, SKMEL2 (BRAF WT) and SKMEL28 (BRAF mutant). Now, we will add annotations to separate these groups.

- Create two annotation categories by selecting **Add Annotation** twice.
- Click in the fields below **Column Name** and change **New Annotation** to **Treatment**, and **New Annotation 2** to **BRAF Genotype.**
- Under **Column Type**, use **Text.** See the *Annotations* section for information on other column types.
- The column heading will change dynamically to reflect the new column name. Add the specific annotations (or copy and paste from another source) under the new column, according to what is documented in the sample names:
  - **DMSO** or **VEM** for **Treatment**
  - **WT/WT** (SKMEL2) or **Mut/Mut** (SKMEL28) for **BRAF Genotype**

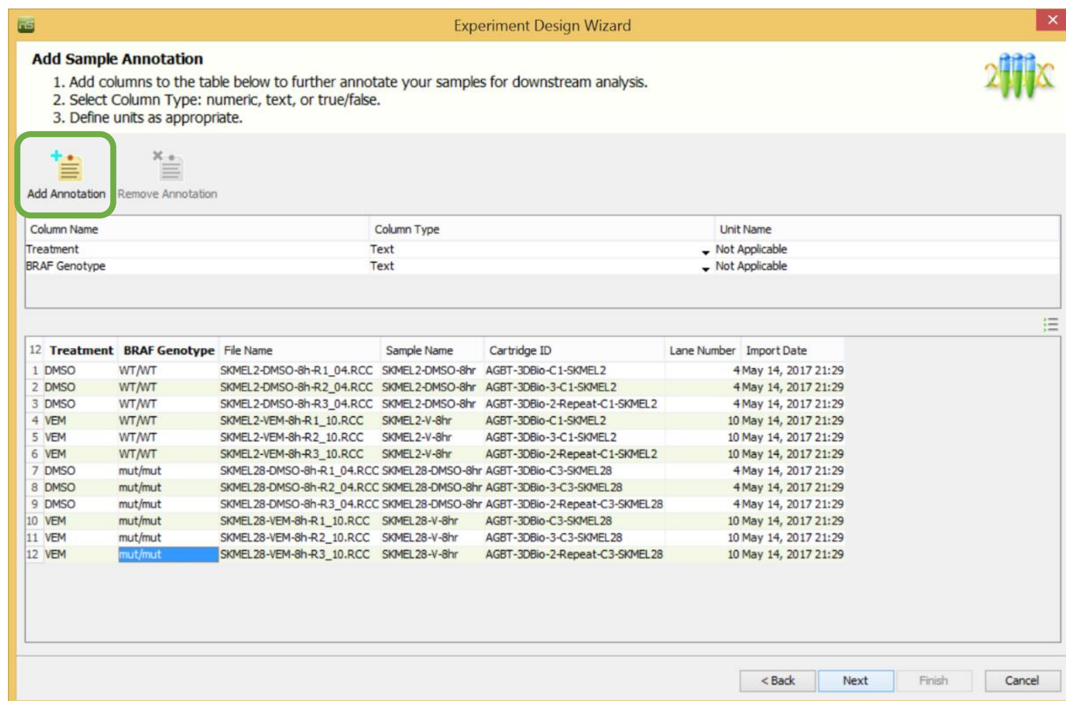



Figure 28: 3D Bio Data Annotations window

# Background Subtraction & Thresholding

Every assay has a certain level of background signal or “noise” and this can vary assay-to-assay, probe-to-probe, and lane-to-lane. For this reason, nSolver 4.0 prompts you to select a method to calculate and reduce background signal before normalizing and analyzing your samples. Background calculations utilize the counts from negative controls, several of which are included in each CodeSet. These negative controls are probes for which no target is present.

For most datasets, background correction will be unchecked initially, indicating that no background calculations will be performed. You can either maintain this setting, or choose one of the two general methods of calculation available: **Background Thresholding** or **Background Subtraction**.
Of these methods, NanoString recommends Background Thresholding for most analyte types and assays. See additional details in the respective assay-specific sections for *Plex$^2$* and *PlexSet* assays.

## Background Thresholding

There are two options to determine a value for Background Thresholding (see Figure 29):

- **Negative control thresholding** uses the average of the negative controls. You may choose which negative controls to use and which type of mean (arithmetic/geometric) to calculate on each analyte tab.
- **Threshold count value** allows you to choose a value as the threshold.

Once the background level has been determined, all raw counts at or below this value will be adjusted to it. Select **Next** to proceed to *Normalization.*

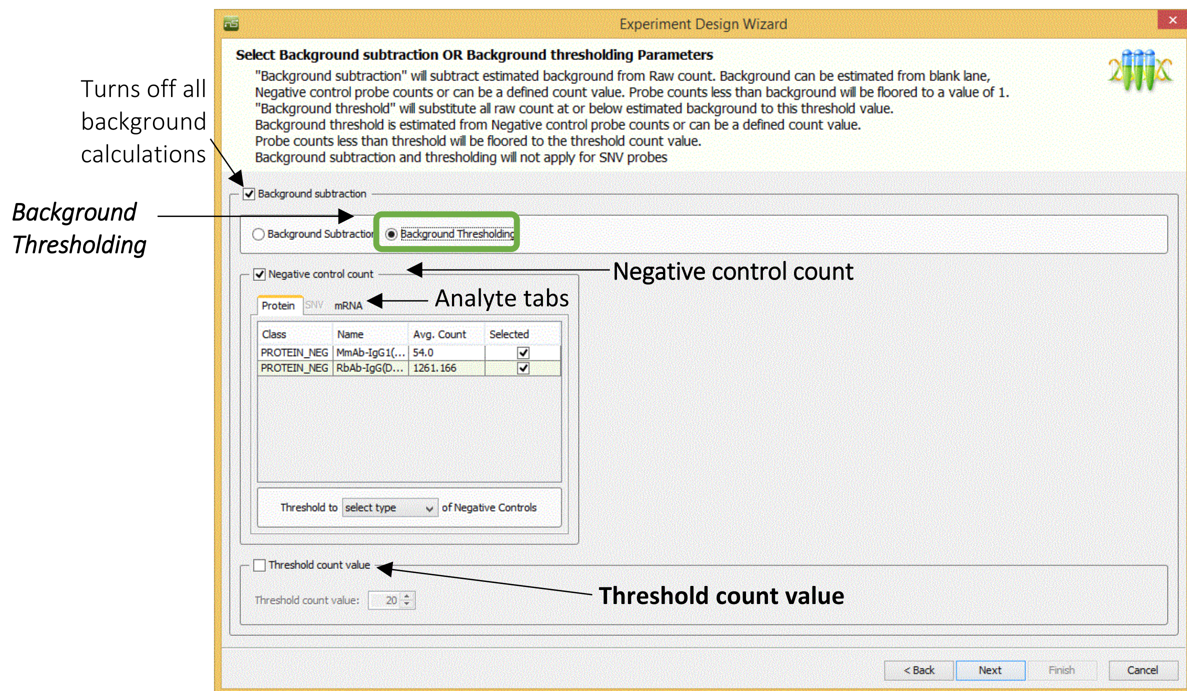



Figure 29: Background options window

## Background Subtraction

Though not recommended for most applications, Background Subtraction is available as an option. There are several ways to customize Background Subtraction (see Figure 30).

- **Negative control subtraction** uses the average of the negative controls for a particular lane. You may choose which negative controls to use and which type of mean (arithmetic/geometric) to calculate on each analyte tab.
- **User-definable value** allows you to choose a value to subtract.
- **Blank lane background subtraction** uses the value from a blank lane (if loaded). A blank lane is one in which nuclease-free water is added as input instead of RNA; this will generate a background measurement that will estimate probe-specific background levels instead of general background levels. Use green add/remove arrows to move any blank lanes from the *Experiment Lanes* field (on the left) to the *Blank Lanes* field (on the right).

Once the background level has been determined, it will be subtracted from each of the raw counts to determine the true counts. Select **Next** to proceed to *Normalization.*

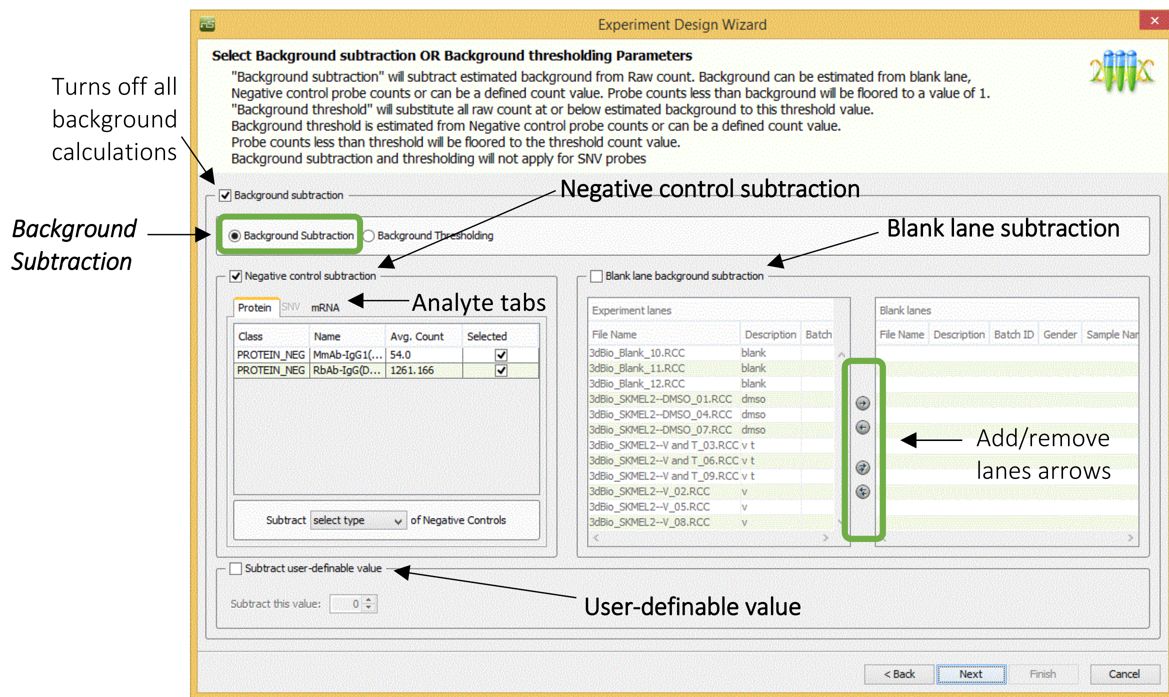



Figure 30: Background subtraction parameters orientation

## Background Example

The following is an excerpt from the **3D Bio Data Example**, which can be found in *Appendix A.*

**Background** noise in data can be filtered out using subtraction or thresholding; this is optional. By default, background correction will be checked off and in most cases, you will not need to deviate from this. If performing background correction on your data, thresholding is recommended for most analyte types.

For the example dataset, **3D Bio Data**, leave the **Background Subtraction/Thresholding** box unselected, which leaves background correction off.

Background Subtraction/ Thresholding

Background Subtraction  Background Thresholding

Negative control count

Protein SNV mRNA

| Class | Name | Avg. Count | Selected |
|---|---|---|---|
| PROTEIN_NEG | MmAb-IgG1(... | 154.666 | ✓ |
| PROTEIN_NEG | RbAb-IgG(D... | 326.0 | ✓ |

Threshold to select type of Negative Controls

Threshold count value

Threshold count value: 20

Figure 31: Background options window

# Normalization

## Steps of Normalization

Normalization is a two-step data transformation that balances counts between lanes, allowing you to make meaningful biological comparisons. Specific normalization options may differ among analyte types (see Figure 32). Generally, raw gene expression data is normalized in this two-step process:

1. A **Positive Control Normalization** factor is calculated using the positive controls that are spiked in to every sample. This normalization (left side of window) adjusts for variations that exist across samples, lanes, cartridges, and days and include differences in user technique, hybridization, magnetic bead purification, complex-to-slide binding, and imaging. The default range of acceptable values is **0.3-3.0**, as seen in the lower left portion of the Normalization Parameters window.
2. A **CodeSet Content Normalization** factor (also called Reference or Housekeeping Normalization factor) is calculated using reference genes to adjust for differences in analyte abundance and/or analyte quality across samples. This normalization (right side of window) removes input variance and accounts for different degradation states. The default range of acceptable values for this is **0.1-10**, as seen in the lower right portion of the Normalization Parameters window.

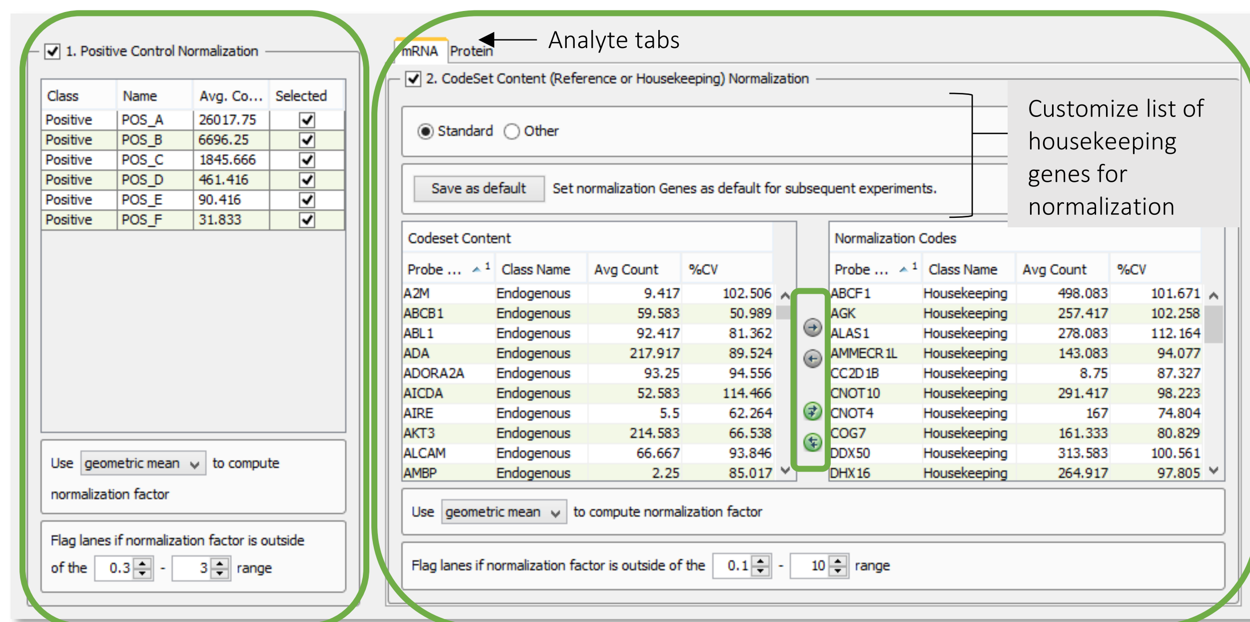



Figure 32: Normalization parameters orientation

*By default, nSolver 4.0 displays the recommended normalization method for each assay type.* It is not usually recommended or necessary to change these default settings. See additional details in the respective assay-specific sections for *Plex*[2] and *PlexSet* assays. De-select either the Positive Control Normalization checkbox or the CodeSet Content Normalization checkbox to turn off Normalization calculations.

## Normalization Settings

The Normalization Parameters window allows you to customize the data normalization process. Use checkboxes next to **Positive Control Normalization** and **CodeSet Content Normalization** to activate or deactivate normalization based on those types of control sequences.

### Positive Control Normalization

Use checkboxes to select which Positive Control Probes to include in Positive Control Normalization. You may consider unchecking POS_F, as its concentration is close to background levels. 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.

### CodeSet Content Normalization

All CodeSet Content probes listed in the *Normalization Codes* window 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.

**mRNA: Standard** is the recommended setting which takes the geometric mean of housekeeping genes. **Other** is also available but should only be used in a large CodeSet in which most genes are expected to be unaffected by experimental conditions; this setting scales to the total count of the lane. If working with a CodeSet with a low number of housekeeping genes, you may consider filtering the list of housekeeping genes to the most robust ones. To do this, click on the column heading, **Avg Count,** in the **Normalization Codes** window. This will sort the genes by count. You may consider removing genes with counts <100 from the list using the arrows. NanoString recommends removing housekeeping genes with counts below background (usually below 20).

**Protein:** Use radio buttons to choose either **Protein_Cell_Norm**, which automatically selects control probes (similar to Housekeeping genes for mRNA), or **All Proteins** for normalization.

Select the **Save as Default** button to designate the current settings as default in future experiments.

The drop-down menus below each of the **Normalization Probe** boxes are defaulted to use the **geometric mean** for calculations, since it weighs the low-concentration controls equally with the high-concentration controls. **Mean** should only be selected if normalization probes are of similar expression levels.

Once you have chosen the Normalization parameters, select **Next** to proceed to *Ratios*. Note: Plex$^2$, PlexSet, SNV, and multi-RLF expermints require calibration steps before ratio-building (see the *Analyte- & Assay-Specific Notes* section).

Possible error message
***No Normalization Codes Selected***
indicates that you need to identify housekeeping genes for CodeSet Content Normalization. Click on each of the analyte tabs to select these reference genes for each analyte type.

## Normalization Example

The following is an excerpt from the **3D Bio Data Example**, which can be found in *Appendix A.*

Use the tabs to review the settings for the different analytes, and adjust them, as noted below. Due to the nature of the calling algorithm and the complexity of **SNV** data, its normalization is hard-coded.

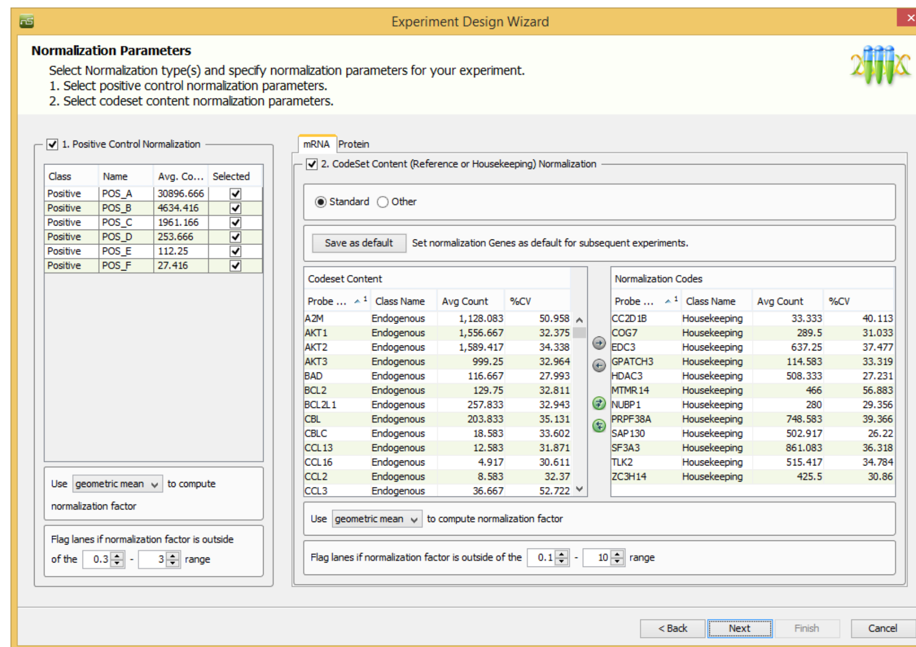



Figure 33: 3D Bio Data Example - normalization parameters

In the **Positive Control Normalization** field, note the following:

- There are no analyte tabs in this field since mRNA and Protein share a single set of Positive Control Normalization settings. Maintain these default settings, using the **geometric mean** with the default flagging range.

In the **CodeSet Content Normalization** field, note the following:

- The default **mRNA CodeSet Normalization** settings include all Housekeeping Genes labeled in the CodeSet.
- The default **Protein CodeSet Normalization settings** include the **histones** labeled in the CodeSet**.** Alternatively, you may select **All** or click on the %CV column heading and remove the highest %CV samples, but it is not necessary for the purposes of this example.

Select **Next**.

## What to Do If You Have a Content Normalization QC Flag

A QC flag for content normalization indicates that the flagged sample had a content (or housekeeping gene) normalization factor more than 10-fold different from the average sample in the same experiment. Review your Normalized Data Table (see the *Normalized Data Table* section) to see any Content Normalization flags your data may have.

Content Normalization flags can be caused by:

- A **significant reduction in overall assay** efficiency for that sample.
- An **effective reduction in quantity or quality (fragmentation)** of the input analyte. The likelihood of a reduction in assay efficiency can be assessed by the presence of any other QC flags for that sample.
- **Insufficient RNA targets** to count. If the sample has no other QC flags except that for Content Normalization, this indicates that the assay is working well, but contains low RNA concentrations or highly fragmented RNA (such as from an archival FFPE sample).

Considerations for what to do if you have a Normalization QC Flag:

- If you see a normalization flag for the POS controls, you may have had an assay-level problem.
- The biggest effect of any QC flag will be on low-expression targets.
- If the sample failed the QC by a wide margin, you may want to consider dropping the sample.
- If the normalization factor was only slightly outside the recommended range or if the genes being studied were moderately to highly expressed, you may consider overlooking the flag and keeping the sample for analysis.

## Ratios

A fold change (ratio) is computed by taking the mean of the normalized lanes associated with a common treatment group from the *Annotation Parameters* screen and then dividing it by the baseline specified on this screen. There are three ways to choose the baseline for computing ratios and fold change estimates (see Figure 34).

- The **all pairwise ratios** 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.*
- You may choose to calculate ratios by **partitioning by** one of your annotation values. *This is the most commonly-used method.*
- The **user-selected reference samples** option allows you to select any sample lane to be used as a baseline in building ratios. Move samples that you want to use in the baseline from the *All Samples* table on the left to *the Base* Samples table on the right by selecting them and using the green arrow buttons. The baseline will be 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).

nSolver 4.0 will automatically run a t-test if sample annotations are present and if more than one sample is added to the treatment groups (see the *T-test* section). Select the False Discovery Rate checkbox if you would like to calculate this value (see the *False Discovery Rate* section).

Once you have selected your preferences, select **Next**.

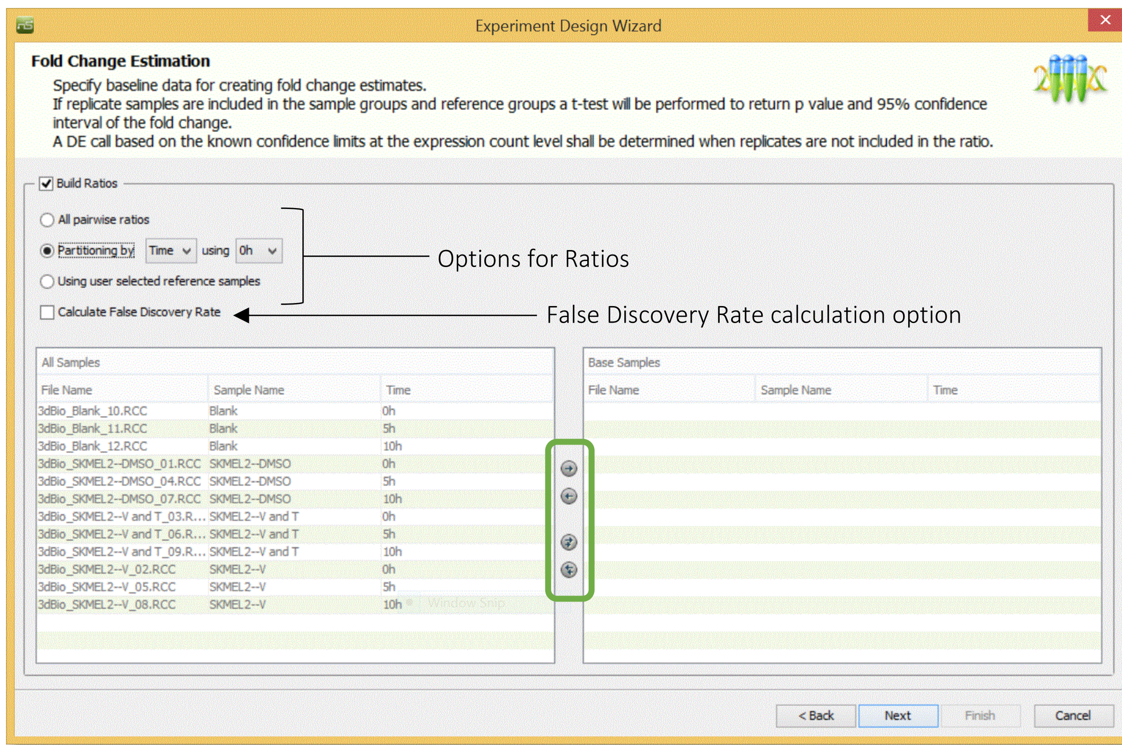



Figure 34: Fold change estimation (ratio) options in the Experiment Wizard

If working with SNV data, a **Ratio not created** window will appear, reminding you that ratios are not generated from SNV data.

The *Ratio Data Names* window prompts you to assign names and/or descriptions to your fold change data (Figure 35). You may unselect the check boxes to the right for any ratios you do not wish to calculate.

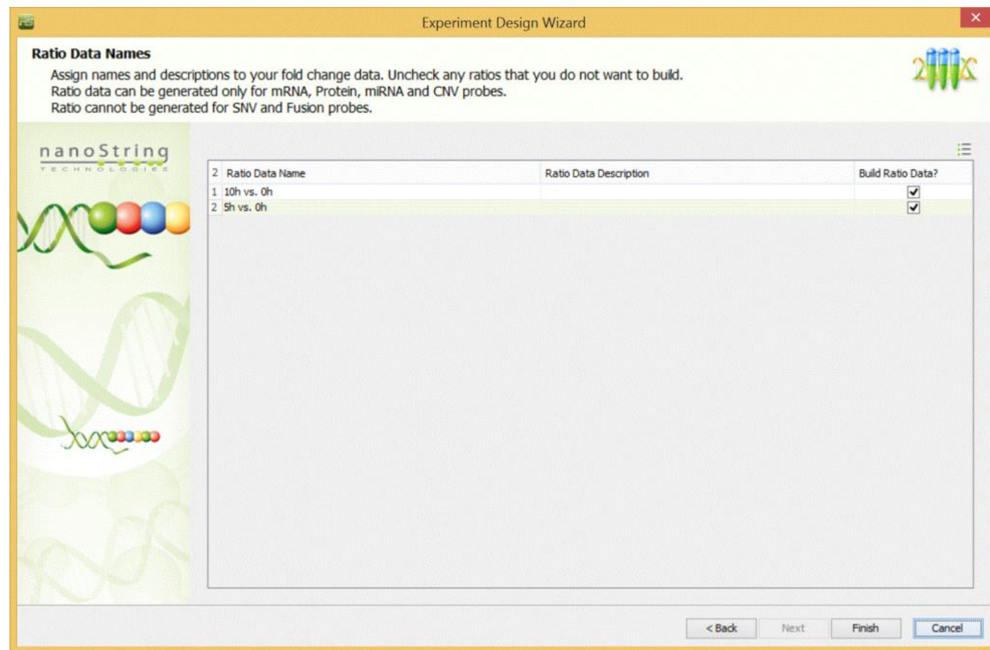



Figure 35: Ratio data names options in Experiment Wizard

Select the **Finish** button, which should now be activated.

Your experiment will be built and will then appear on the **Experiments** tab. Expand the experiment in the navigation tree to see the different levels of data and proceed to the *Data Output and Export* section.

Ratios Example

The following is an excerpt from the **3D Bio Data Example**, which can be found in *Appendix A*.

In creating ratios for the example dataset, **3D Bio Data**, select **Partitioning by.** It will default to one of the annotations we entered earlier, Treatment, and will choose a treatment type as the reference, in this case, DMSO. For this example, we will keep these defaults, but they can be changed using the dropdown menus, if desired.

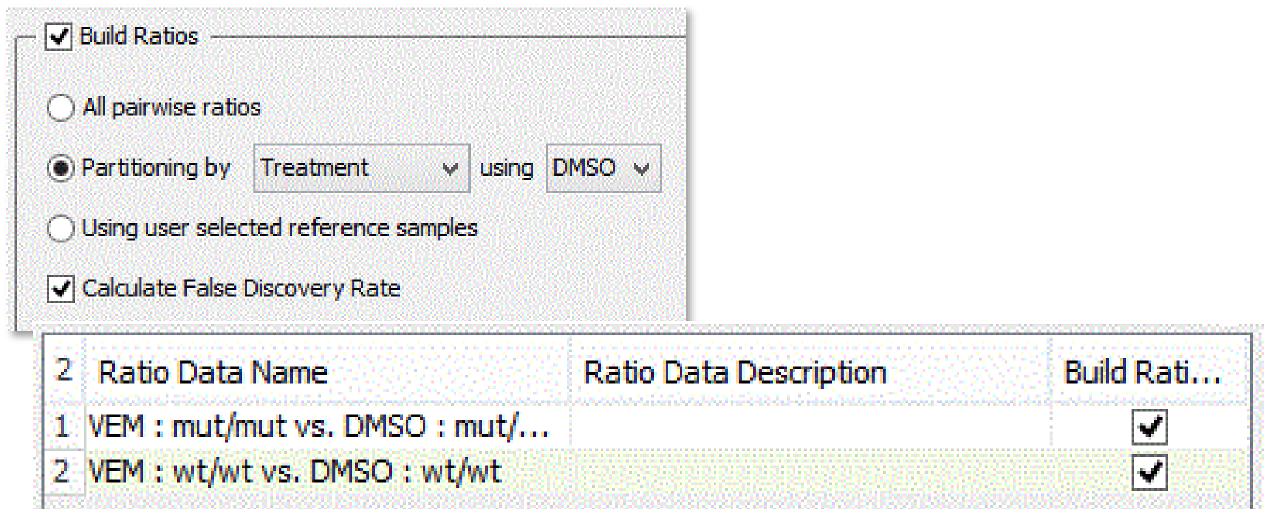



Figure 36: 3D Bio Data Example – ratios windows

This experimental design results in two ratios to confirm, partitioning by treatment within each cell type. You can use the checkboxes along the right side of the window to confirm or cancel building the ratio. You can change the name of this comparison, as well, if desired. Select **Finish**.

# Data Output & Export

nSolver offers several data export options, regardless of if you have only raw data or if you have normalized, grouped, ratio, and/or analysis data as well. Your experiment will be on the Experiments tab under the Study you designated for it. Expanding the navigation tree (selecting the **+** sign) will reveal the different types of analyzed data available to you (Figure 37).

## Types of Data

For most analyses, the following levels of data will be displayed:

- The **Raw Data** table contains unprocessed data for all samples in the experiment.
- The **Normalized Data** table contains the processed data for all samples.
- The **Grouped Data** table contains the geometric mean of expression levels within each group (as defined by the sample annotations, see the *Annotations* section). If no annotations were added in creating the experiment, the grouped data will appear the same as the normalized data.
- The **Ratio Data** table contains fold-change results and any statistical inferences surrounding these calculations. To view these additional values, right-click on an existing column header, choose **Select Columns** and select the columns you want to display. Alternatively, you can select the **Column Options Icon** and **Show All Hidden Columns**.
- The **Analysis Data** section contains results of any analyses carried out (see the *Analysis* section).

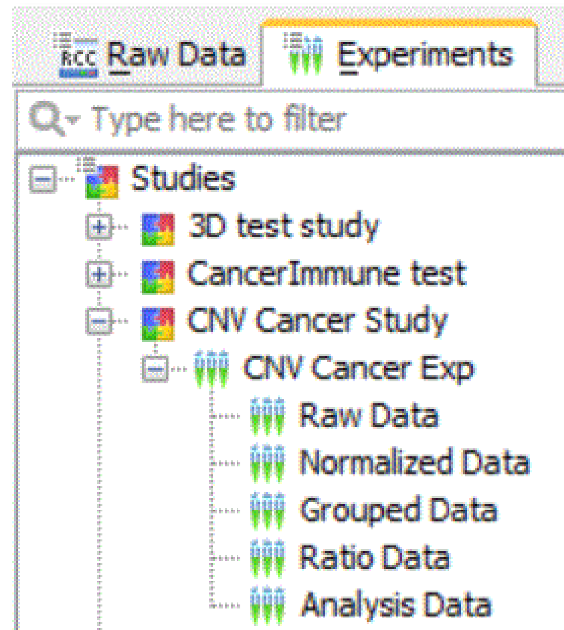



Figure 37: Experiment data navigation tree

To open any data table, highlight the data category of interest from the navigation tree and select one or more rows in the central table. Then, select **Table** (see Figure 38) to view samples in a more detailed table, **Export** (see the *Exporting Data* section) to export the table results without viewing first, or **Analysis** or **Advanced Analysis** (see the *Analysis & Advanced Analysis* section) to perform visual analyses using multiple graphing options. nSolver may offer additional or alternative button options, depending on the analyte types detected in your data.

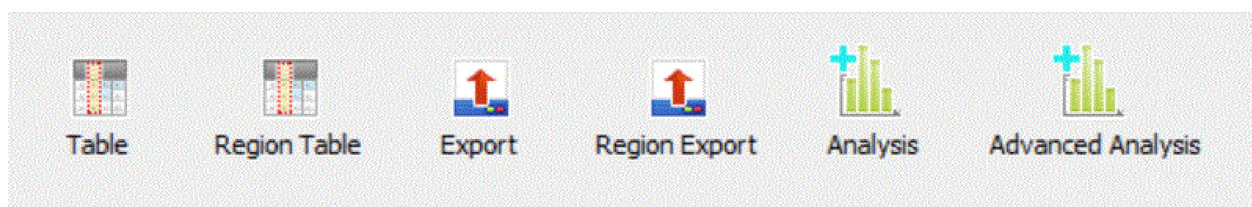



Figure 38: Experiment data options buttons – CNV data

## Viewing Tables

After highlighting the data category of interest (Raw/Normalized/Grouped/Ratio) in the **Experiments** tab, select the **Table** button to view the data table (Figure 39). You can **Filter** your data, select different **View Options** (such as counts, $log_{10}$, or $log_2$) and **Export** all or some of the table from that window, as well. Tables for each data category will contain detailed information on each probe, summarized information on each probe's overall performance (including min and max values and % samples above threshold), any QC flags, as well as the individual results for each sample. Throughout the tables, column headers in red indicate that a QC or Normalization Flag is associated with that lane.

### A note about $log_2$

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. $Log_2$ counts are commonly used when comparing counts to array data and are also appropriate for t-tests.

Filter data options

View output options

**Normalized Expression Data**

Counts are presented here in a matrix to allow for visual checking of data, filtering of probes of interest, and/or quick export.

Filter Expression Data

Filter: Probe Name Match if: is anything Go Reset

View Options

Show data as Counts

| 337 | Probe Name | Region Name | Annotation | Class Name | Analyte Type | Species Name | NS Probe ID | Target Sequence | Genome Strand |
|---|---|---|---|---|---|---|---|---|---|
| 1 | TP73-1 | TP73 | TP73 | Endogenous | CNV | | TP73-1 | | + |
| 2 | TP73-2 | TP73 | TP73 | Endogenous | CNV | | TP73-2 | | + |
| 3 | TP73-3 | TP73 | TP73 | Endogenous | CNV | | TP73-3 | | + |
| 4 | MYCL1-1 | MYCL1 | MYCL1 | Endogenous | CNV | | MYCL1-1 | | + |
| 5 | MYCL1-2 | MYCL1 | MYCL1 | Endogenous | CNV | | MYCL1-2 | | + |
| 6 | MYCL1-3 | MYCL1 | MYCL1 | Endogenous | CNV | | MYCL1-3 | | + |
| 7 | CDKN2C-1 | CDKN2C | CDKN2C | Endogenous | CNV | | CDKN2C-1 | | + |
| 8 | CDKN2C-2 | CDKN2C | CDKN2C | Endogenous | CNV | | CDKN2C-2 | | + |
| 9 | CDKN2C-3 | CDKN2C | CDKN2C | Endogenous | CNV | | CDKN2C-3 | | + |
| 10 | JUN-1 | JUN | | Endogenous | CNV | | JUN-1 | | + |
| 11 | JUN-2 | JUN | | Endogenous | CNV | | JUN-2 | | + |
| 12 | JUN-3 | JUN | | Endogenous | CNV | | JUN-3 | | + |
| 13 | MAGI3-1 | MAGI3 | | Endogenous | CNV | | MAGI3-1 | | + |
| 14 | MAGI3-2 | MAGI3 | | Endogenous | CNV | | MAGI3-2 | | + |
| 15 | MAGI3-3 | MAGI3 | | Endogenous | CNV | | MAGI3-3 | | + |
| 16 | REG4-1 | REG4 | | Endogenous | CNV | | REG4-1 | | + |
| 17 | REG4-2 | REG4 | | Endogenous | CNV | | REG4-2 | | + |
| 18 | REG4-3 | REG4 | | Endogenous | CNV | | REG4-3 | | + |
| 19 | MCL1-1 | MCL1 | | Endogenous | CNV | | MCL1-1 | | + |
| 20 | MCL1-2 | MCL1 | | Endogenous | CNV | | MCL1-2 | | + |

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

Export Export Close

Figure 39: Experiment data table options

## Raw Data Table

This table displays the raw counts for each sample. It is essentially the same as the one generated on sample import, however, it contains two additional columns. The **% Probes Above Threshold** column is useful for a second check on the overall quality of the data; a value that deviates a great deal from the values of the other samples may indicate a QC issue. The **Blank Lane Flag** column is helpful if you designated a blank lane when you ran your data.

## Normalized Data Table

This table contains normalized values for all samples. In addition to the general probe ID and performance info, it has the following columns: **Positive Normalization Flag, Content Normalization Flag, Positive Normalization Factor,** and **Content Normalization Factor** (see Figure 40). See the *What to Do If You Have a Content Normalization QC Flag* section.

If working with SNV data, an additional button, **Variant Table**, will be available to view variants. See the *SNV Data Analysis* section.

| CNV Positive Normalization Factor | CNV Content Normalization Factor |
|---|---|
| 0.91 | 1.48 |
| 0.87 | 0.96 |
| 1.77 | 0.59 |
| 1.76 | 1.14 |
| 1.03 | 0.69 |
| 1.27 | 1 |
| 0.81 | 0.99 |
| 0.83 | 0.84 |
| 0.99 | 1.11 |
| 0.92 | 1.29 |
| 0.78 | 1.47 |
| 1.02 | 1.27 |
| 0.95 | 1.08 |

Figure 40: Normalization table - norm flag columns

If working with CNV data, additional buttons, **Region Table** and **Region Export,** will be available to view and export specific regions of interest. See the *CNV Data Analysis* section.

### How to use the SNV variant table

The following is an excerpt from the **3D Bio Data Example**, which can be found in *Appendix A*.

When the **Normalized Data** level is selected, the **Variant Table** button becomes active. Select the normalized samples, then select this button. In the resulting table, scroll to the right so that you can see all the sample data columns. Then, scroll down. You should see the occasional **green flag** in some of the **Variant Call** columns, indicating a variant is present in your data at that gene. You can also click on each Variant Call column header to sort flagged probes on top.

| 104 | Probe Name | Class Name | Variant Call: SKMEL2-DMSO-8h-R1_04.RCC | Variant Call: SKMEL2-DMSO-8h-R2_04.RCC | Variant Call: SKMEL28-VEM-8h-R3_10.RCC |
|---|---|---|---|---|---|
| 1 | BRAF COSM476 (V600E) | SNV_VAR | | | |
| 2 | TP53 COSM44571 (L194R) | SNV_VAR | | | |
| 3 | FBXW7 COSM22932 (R465C) | SNV_VAR | | | |
| 4 | EGFR COSM6239 (G719A) | SNV_VAR | | | |
| 5 | JAK2 COSM333722 (G180A) | SNV_VAR | | | |
| 6 | APC COSM26697 (I1307K) | SNV_VAR | | | |
| 7 | ERBB2 COSM14060 (L755S) | SNV_VAR | | | |
| 8 | KRAS COSM520 (G12V) | SNV_VAR | | | |

Figure 41: 3D Bio Data SNV variant table

## Grouped Data Table

This table contains data grouped by annotations. In addition to the general probe ID and performance info, it has its own Normalization Flag column. See the *What to Do If You Have a Content Normalization QC Flag* section.

If working with CNV data, additional buttons, **Region Table** and **Region Export,** will be available to view and export specific regions of interest. See the *CNV Data Analysis* section.

## Ratio Data Table

This table contains ratio data. It has its own Normalization Flag column. In the event of a flag, see the *What to Do If You Have a Content Normalization QC Flag* section. It also has columns which relate to significance testing, which may be initially hidden. See the *Significance Testing* section, below.

**View Options** allows you view data as Fold Change, Ratio, or $\log_2$ Ratio.

A note about fold changes

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:

$$-1/(A/B) \quad or \quad -B/A$$

If working with CNV data, instead of the Table button, you may select the **Probe Data** button, which gives you options to sort/filter your data by probe name. **Region Table** and **Region Export** buttons are also available from this data level. See the *CNV Data Analysis* section.

### Significance Testing

nSolver provides three different methods to determine significance of ratios: **error model**, **t-test**, and **false discovery rate (FDR).** If these values are initially hidden on the **Ratio Data** table, right-click (or command-click) on any existing column header and choose **Select Columns**. Select the columns you want to display. Alternatively, you can select the **Column Options** icon and **Show All Hidden Columns**.

## Error Model (DE Call)

The DE Call (Differential Expression call) is an error model to assist in determining confidence of ratios when no replicates exist. The DE call can be viewed in the ratio data table. 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. It is *not* designed to provide conclusions about the populations from which samples were drawn.

- *Yes* indicates that the difference between two groups (numerator and denominator) is significant beyond technical noise.
- *No* indicates that the ratio can be explained by technical noise.

### A note about error model

At any given number of counts, the nCounter platform is subject to a certain degree of technical variability. The DE Call is a mapping of raw count values to an estimated level of 95% of technical variability. NanoString derived this error model from a large experiment in which thousands of gene/sample combinations were run on numerous CodeSets, allowing for measurement of the average expression and technical variability (on a log scale). The result is a conservative estimate of any gene's technical variability.

## T-test

When replicates have been assigned in Annotations (see *Experiments* section), a t-test runs between them automatically and a p-value is calculated and appears in the ratio data table.

nSolver calculates the 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 to assist in determining whether the fold change calculated is statistically significant. nSolver performs a two-tailed t-test on the log-transformed normalized data that assumes unequal variance.

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

### A note about t-tests

nSolver performs the t-test using seven significant figures for each data point, whereas it exports normalized data with two significant figures.

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

## False Discovery Rate

A gene's False Discovery Rate (FDR) is the proportion of genes with values at least as low as the gene in question that are expected to be false discoveries. FDR can be used as a more conservative and informative alternative to p-values.

When replicates have been assigned in Annotations (see *Experiments* section), the **Calculate FDR** option is enabled on the fold change estimations screen. Selecting this prompts nSolver to use the Benjamini-Yekutieli procedure to calculate the FDR from the p-values returned by the t-test.

### A note about false discovery rates

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.

By default, only fold-changes are shown in the Ratio Data Table, but clicking on the Column Options icon will allow viewing of additional hidden data fields. Note that the column titled **FDR** contains FDR adjusted p-values rather than FDR thresholds.

## Experiment Reports

Experiment Reports are available for all experiments. These reports contain information such as raw data, QC settings and results, background and normalization settings and results, and controls.

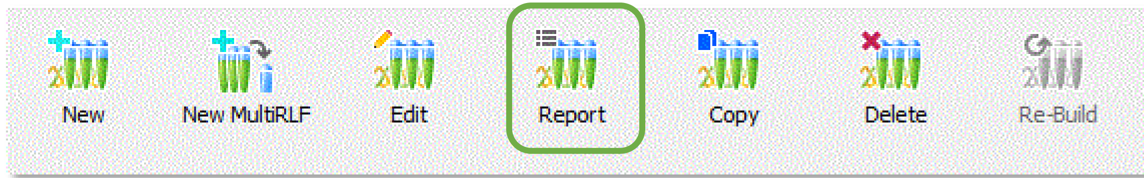



Figure 42: Experiment Report button

To open the report, select the experiment of interest in the navigation tree on the **Experiments tab** and select the **Report** button (Figure 42). The report will open in HTML format.

**To save**, use your browser tool bar to save as an HTML file.

If using Internet Explorer, you can **export directly to Excel** by right-clicking in the body of the report (see Figure 43) and selecting **Export to Microsoft Excel.** Using the yellow arrows, select what regions of the report to import (the yellow arrows will turn to green check marks when selected). Selecting the arrow in the upper left corner will indicate that you want to import the entire report. Select **Import.** The content will appear in an Excel sheet where you can view, edit, print and save.

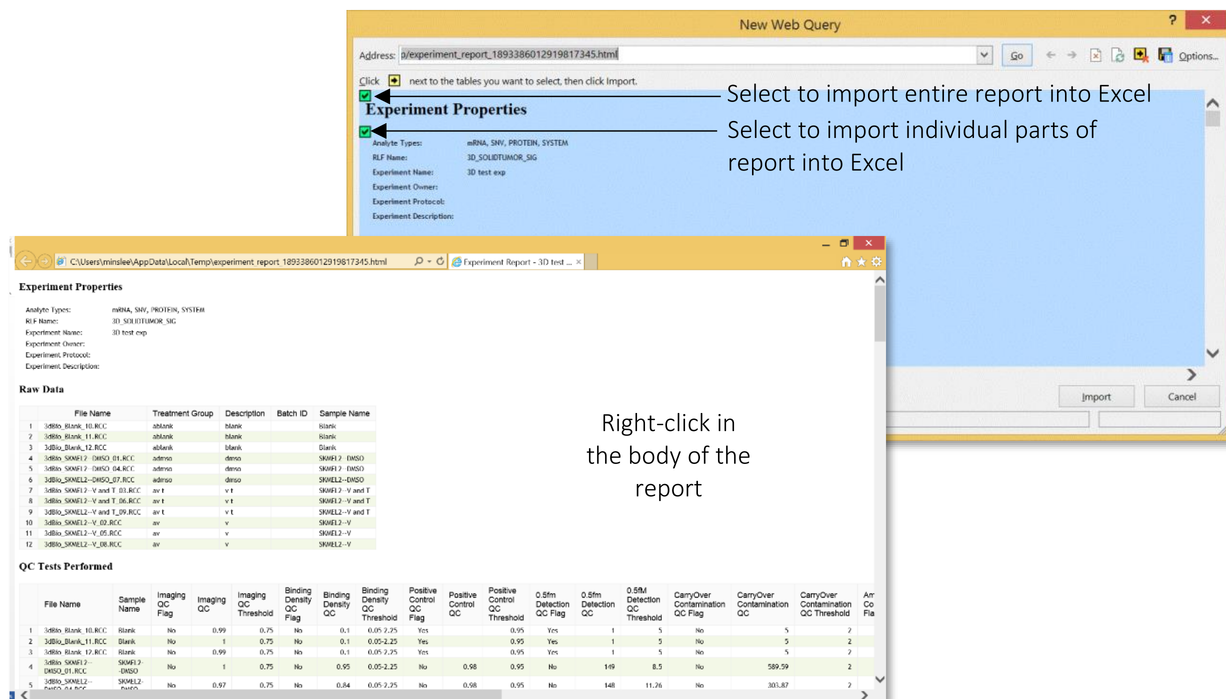



Figure 43: Experiment report windows

## Exporting Data

Before exporting data, check the export settings by selecting **Export** from the top toolbar (Figure 44). Select **Configure export** to select locale and delimiter settings (comma delimited, for example, is preferred for most U.S.–based applications). Select **Ok**.

To export, select samples and/or probes in your data table of interest and select the **Export** button. Specific options that follow will depend on the analyte types in your data and the type of data you are exporting. The ***Export Wizard*** will launch.

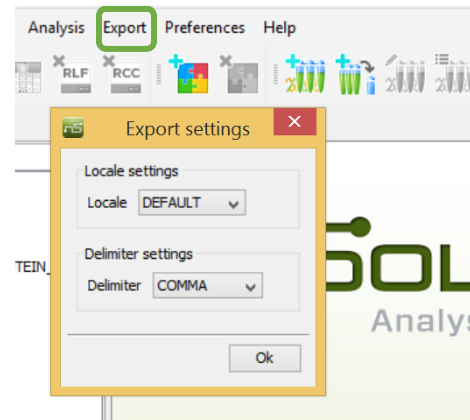



Figure 44: Configuring export settings

Select a format for export. Depending on the assay, you may see the following options:

- The **Custom Text Format** 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) and MeV (open source). You can export separate tables into separate files per lane/sample or combine all data into one file. You can also 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).
- The **RCC Collector Tool Format** is designed for customers who prefer a format which provides raw and normalized data outputs in exactly the same format. It is 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.
- The **Upload MultiSet to Ingenuity®** option is for those with an Ingenuity IPA® license. With this format, you can upload datasets with their associated calculations to the IPA system, which should launch automatically and prompt you to login before uploading.
- The **Partek Genomics Suite™** option is directly compatible for import into the Partek Genomics Suite system.
- The **BioDiscovery Nexus Copy Number™** option is available only for ratio CNV data. The nSolver system exports in a format directly compatible for import into BioDiscovery's Nexus Copy Number software. Nexus Copy Number requires log ratios, which is why it is only available at the ratio data level.
- The **iPathwayguide™ Support** option is for Advaita Bioinformatics iPathwayguide™ users. The format nSolver exports is directly compatible with the iPathwayguide™ software; gene expression data includes ratio data and associated p-values for the fold change estimates, as well.

Select your preference and select **Next**.

**Custom Text Format Export** allows you to select your export parameters (Figure 45) by selecting *Columns to Export*, *Group Names Options*, *Split Options*, *Format Options*, and the *Data Scale* (counts, $log_2$ or $log_{10}$). Options may vary with assay type.

Select **Finish**, name your file, and designate the location in which you would like it saved. Select **Save**. You will be given the option to select **OK** to open the file immediately or **Cancel** to export without opening.

If exporting CNV data, the **Region Table** and **Region Export** buttons will be activated to view and/or export averaged values of probes from particular regions from the Normalized, Grouped, or Ratio data levels.

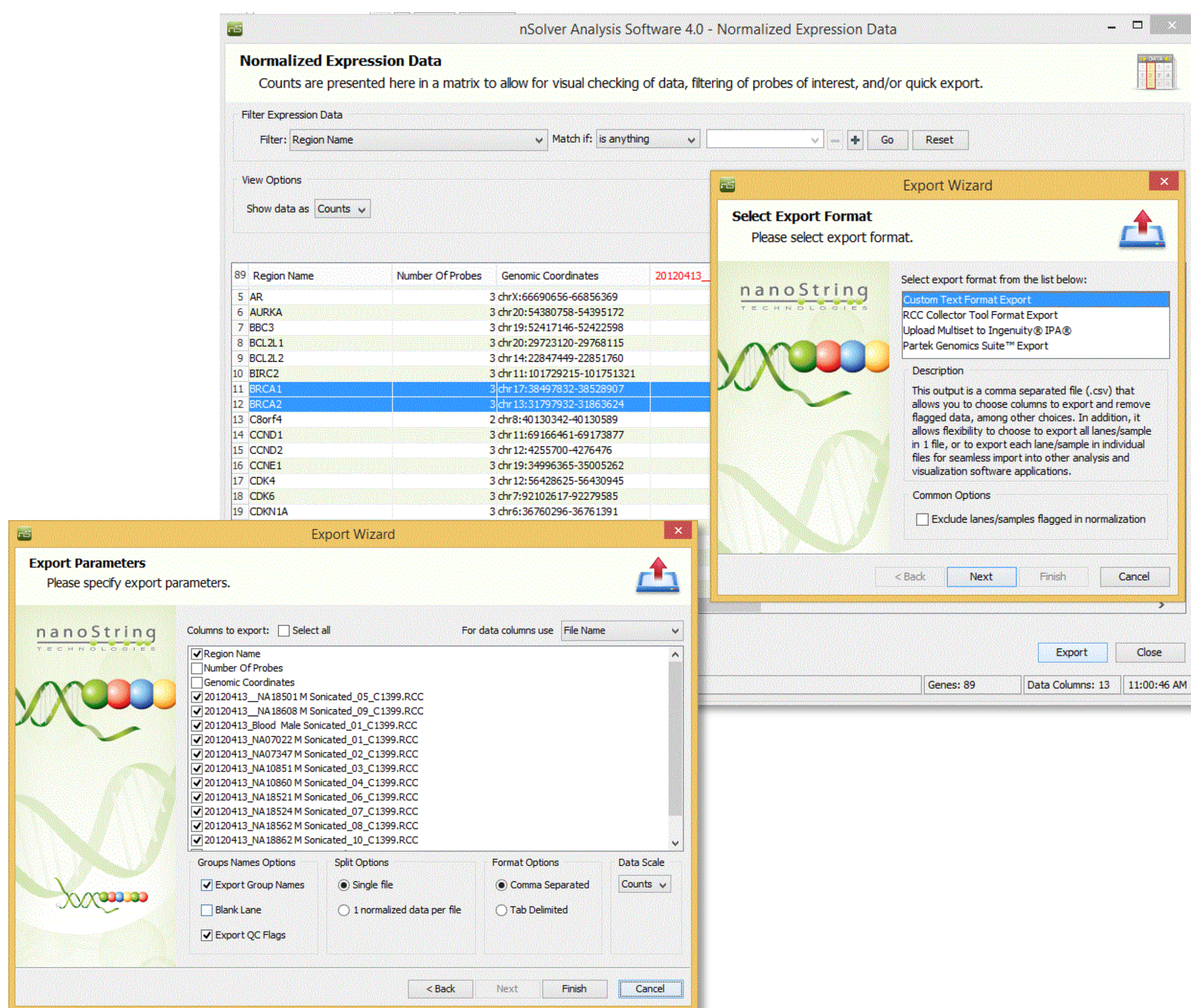



Figure 45: windows associated with data export

# Analysis & Advanced Analysis

Five basic visualizations are available using the Analysis option in nSolver 4.0 (Figure 46), which may be launched from any of the raw, normalized, grouped, or ratio data levels. Many more tools, particularly useful for pathway-based research, are made possible by Advanced Analysis, which makes use of the open source R statistical software. Advanced Analyses may only be launched when viewing the raw or normalized data levels (see the *Advanced Analysis* section).

## Analysis

Below are the five visualizations available for most analyte types in the basic Analysis option. This section will provide you with the basic steps to complete any analysis, then provide more details on each type of visualization, dedicating a section to each one.

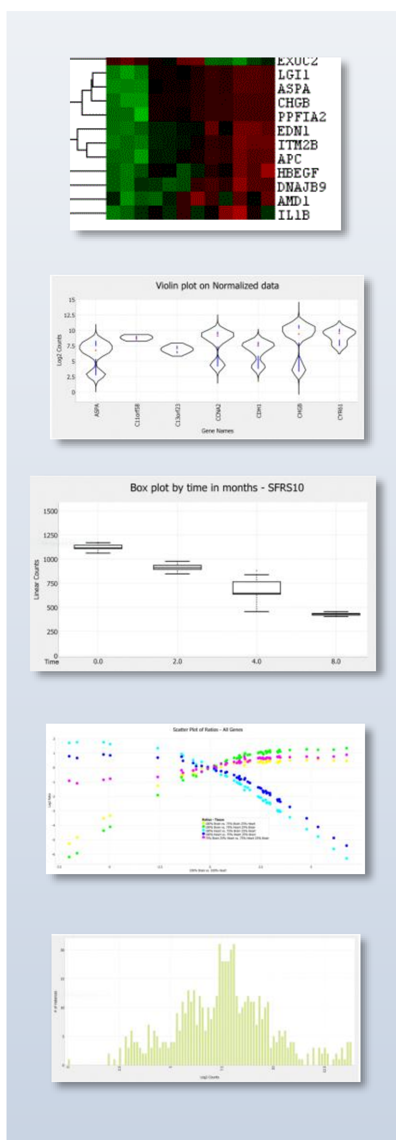



Figure 46: visualizations available in analysis step

**Heat maps** use agglomerative clustering, a bottom-up form of hierarchical clustering, which merges pairs of clusters as they move up the hierarchy. This makes inferences about the relationships of all datasets to each other and is often used to evaluate gene expression.

**Violin plots** are similar to box plots in that they display the range of data. However, they also show the density of values along an axis, much like a histogram. They can be used to illustrate relative gene expression in different cell populations over time.

**Box plots** are non-parametric analyses that display differences between subsets of an experiment without making any assumptions about the underlying statistical distribution. They display the range of data as well as the extents of each quartile.

**Scatter plots** compare raw, normalized, grouped, or ratio data by plotting individual data points using Cartesian coordinates and assigning one variable to each axis. They can be used to identify trends in the relationship between two variables with little to no manipulation.

**Histograms** display the distribution of data that have been binned into discrete intervals. They are used to estimate the probability distribution of a continuous variable.

## To Run an Analysis

On the Experiments tab, select the level of data you wish to work with (raw/normalized/grouped/ratio), then select the particular rows of interest in the central table. Select the **Analysis** button.

The ***Analysis Wizard*** will launch and will prompt you to select the **Analysis Properties** (Figure 47).

- Enter a **Name**.
- Select an **Analysis Type** (see following sections for details on each type).
- Check the box to **filter outlier samples** and genes from analysis using a **Histogram Filter** (see below for more details), if desired.
- Check additional boxes to exclude flagged samples, controls, and/or normalization probes from analysis, if desired.

Select **Next**.

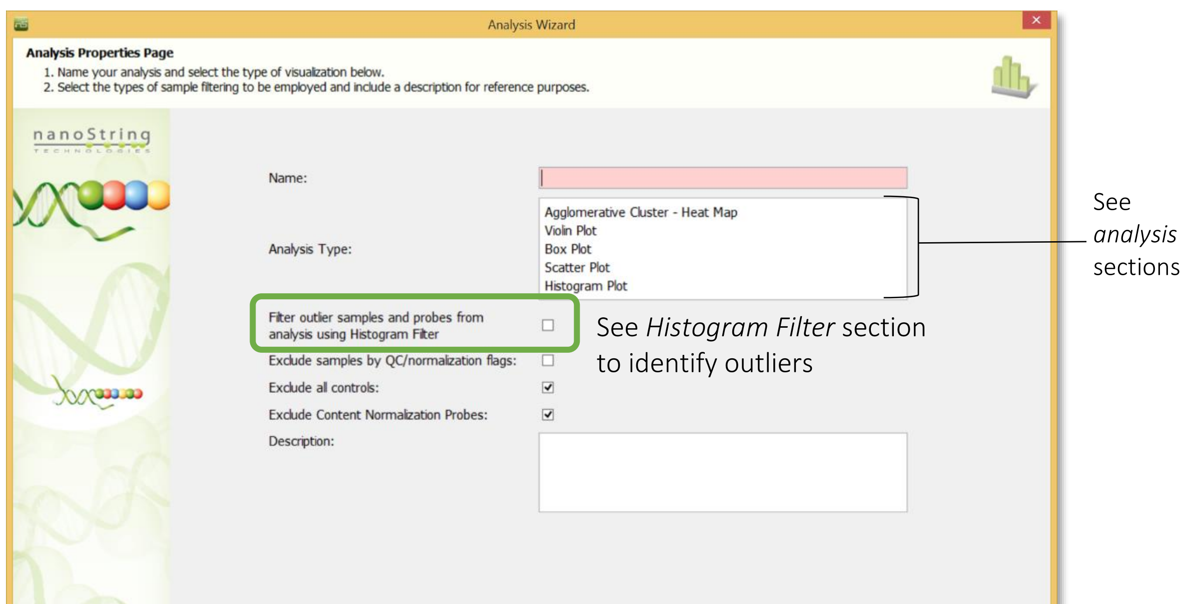



Figure 47: Analysis properties within the Analysis Wizard

## Histogram Filter

If you selected the checkbox, *Filter outlier samples and probes from analysis using Histogram Filter* on the *Analysis Properties* page, you will see two histograms, one representing the spread of your sample data and the other representing the spread of your probe data (Figure 48).

The Histogram Filter is generally designed to allow removal of poor quality samples or genes from downstream analysis and visualization. If you unchecked the *Filter outlier samples and probes from analysis using Histogram Filter* box on the Analysis Properties window, this step will be skipped.

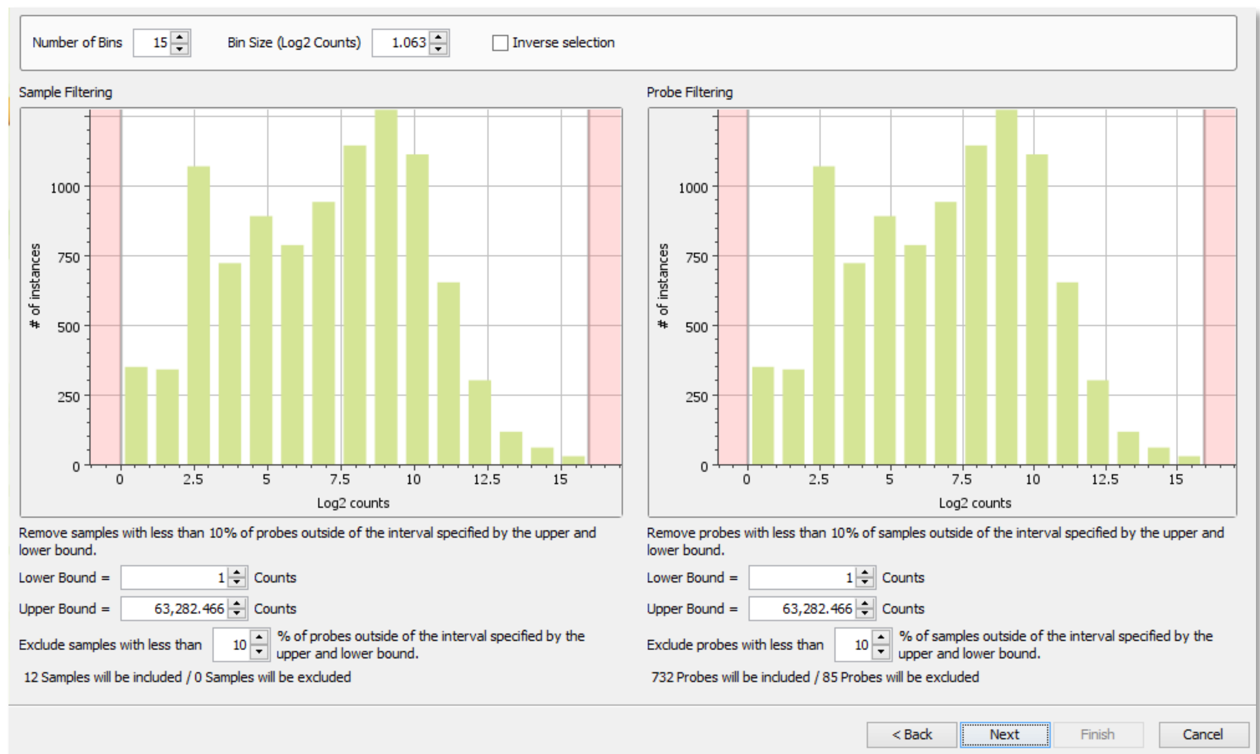



Figure 48: Histogram filtering option, performed before data visualization

You have three customizations you can make before adjusting the filter:

- **Number of Bins** allows you to change the number of bins (bars on the x-axis) in the histogram.
- **Bin Size** allows you to adjust the y-axis representation of number of instances.
- Selecting **Inverse Selection** allows you to filter out the center portion of your spread of data, rather than the outliers.

To adjust the filter, you can manually move the sliding pink bars to the right and left to select the spread of sample and probe data you want to include (pink-shaded data will be excluded). Alternatively, you can use the controls below the histogram to designate the specific counts at which the upper and lower bounds should be set. Once you have set your preferences, select **Next**.

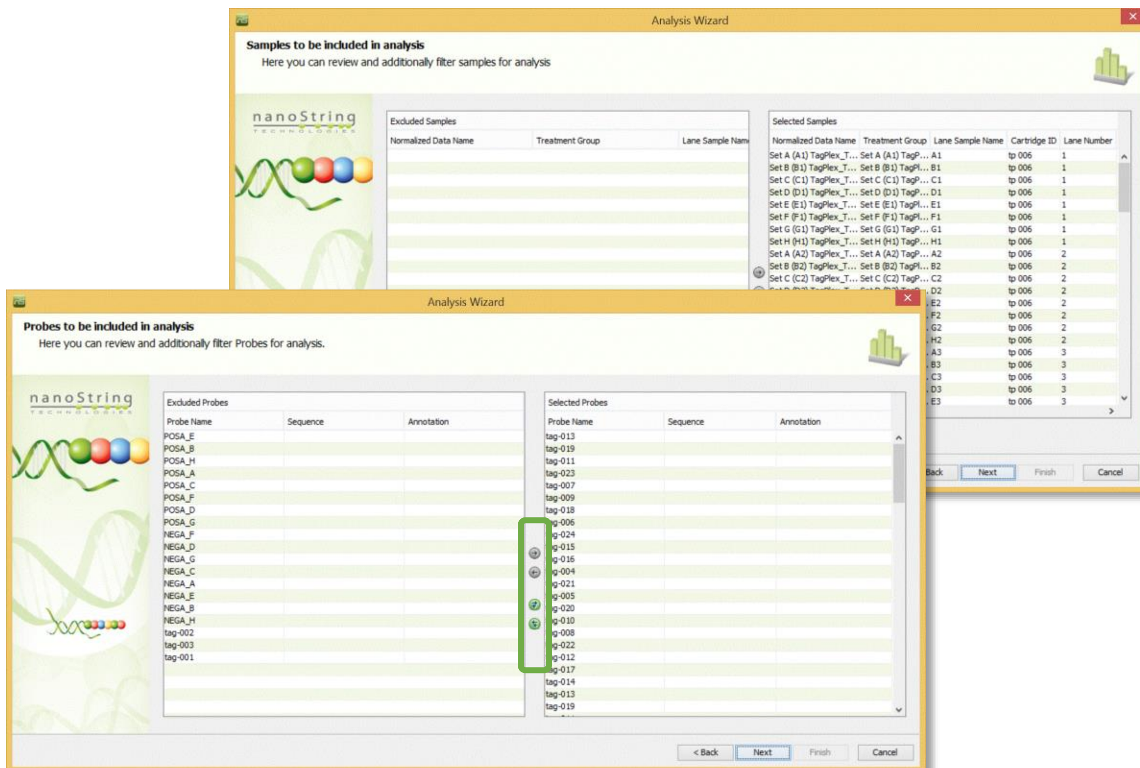



Figure 49: Windows associated with choosing samples and probes to be included in analysis

The next two windows display an adjustable list of samples excluded from analysis and an adjustable list of probes excluded from analysis (Figure 49). If you selected *Exclude QC and Normalization Flags* on the *Analysis Properties* page, any flagged samples will be listed as excluded on this page. For each window, use the green arrows to move any additional samples and/or probes from the **Selected** list to the **Excluded** list or vice versa.

If you are creating an *Agglomerative Cluster (Heat Map)* plot, select **Next** to enter the visualization-specific parameters.

For all other plots (*Violin*, *Box*, *Scatter,* or *Histogram*), select **Finish** to launch the visualization.

## Agglomerative Cluster (Heat Map)

Agglomerative clustering is a form of bottom-up hierarchical clustering, which makes inferences about datasets' relationships to each other. 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 an interactive heat map with a dendrogram tree (Figure 50). Datasets belonging to the same branch of a cluster are similar to each other at some level; datasets in separate branches are less similar at some level.

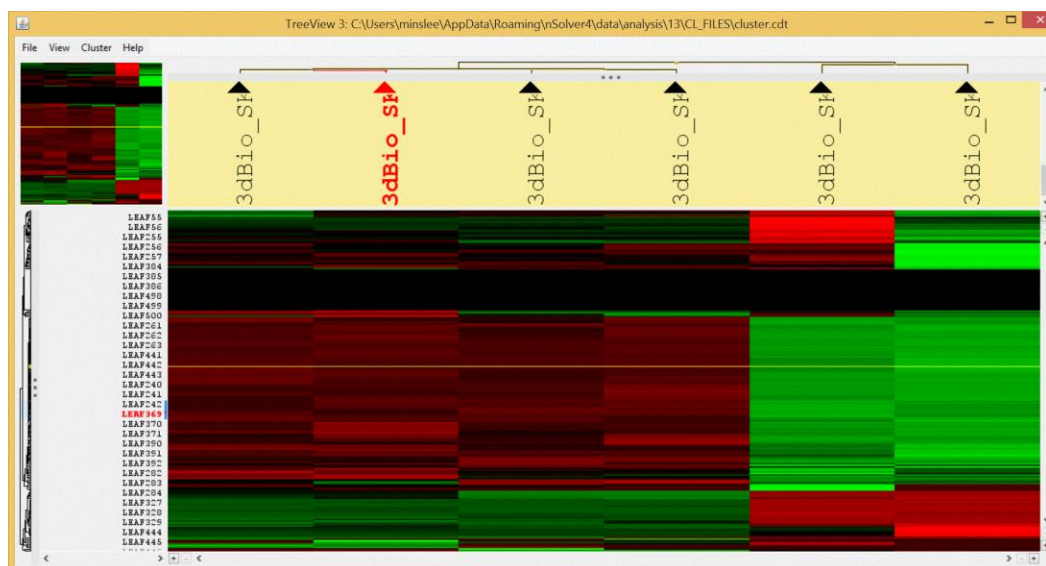



Figure 50: Heat map

To perform an analysis, start with the *To Run an Analysis* section. If you selected the Agglomerative Cluster analysis, you will be prompted to adjust the following functions (see Figure 51):

- If analyzing CNV data, a dropdown menu will allow you to choose **Regions** or **Probes** for analysis.
- You can choose to base the **Z-score transformation** on **Genes** and/or on **Samples**. This will dictate how the heat map is centered and scaled. Checking at least one of these boxes is recommended to produce the most easily-interpreted heat maps.
- **Distance metric**—Choose between Euclidean distance, Pearson correlation (this will run by default), and Spearman correlation. A short definition explaining each metric is dynamically displayed.
- **Linkage method**—Choose between Average, Median, Complete, Centroid, and Wards Minimum Distance linkage methods. A short definition explaining each method is dynamically displayed.
- **Sample data use**—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* window of the Analysis Wizard (use the **Back** button) in the order you want them to appear in the heat map. Then, choose the **Ordered Set** option, here. Otherwise, agglomerative clustering will be performed on the 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* window of the Analysis Wizard (use the **Back**

button) in the order you want them to appear in the heat map. Then, choose the **Ordered Set** option, here. Otherwise, agglomerative clustering will be performed on the 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.

When you have selected your parameters, select **Finish**. Your heat map will appear.

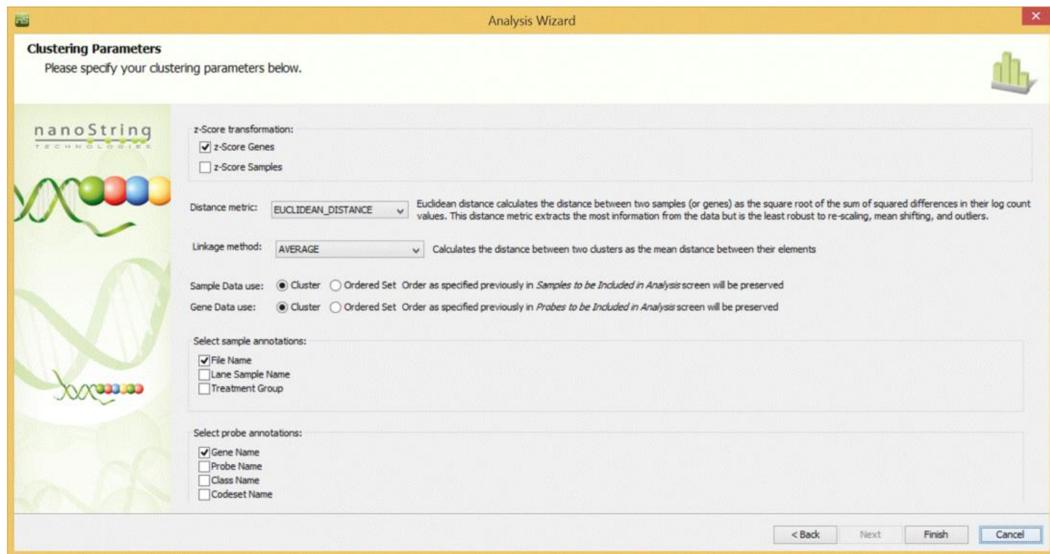



Figure 51: clustering parameters for heat maps

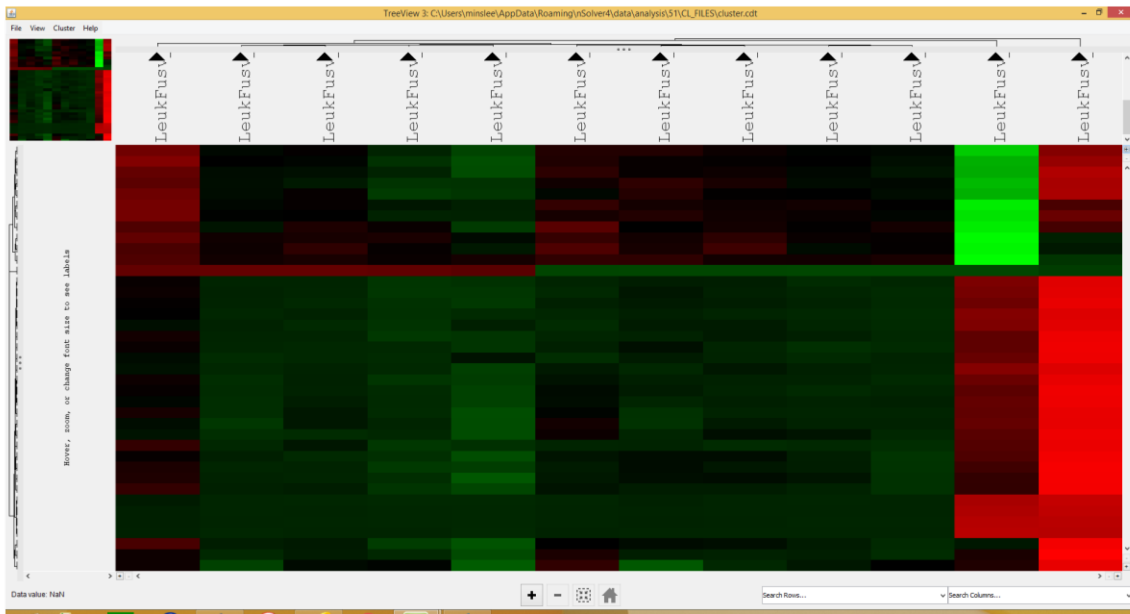


Figure 52: interactive heat map

The heat map (Figure 52) has the following features and options:

- **Sample names** are listed along the top (x-axis) and **gene names** (may initially aooear as **leaves**; adjust in view settings) make up the dendrogram tree on along the left side (y-axis). If labels are not visible, hover over the map and they will dynamically appear. You can also change the type of label or size of font by using the **View** menu on the toolbar (see below).
- Buttons and dropdown lists along the bottom of the map allow you to zoom and search for certain samples, probes, and regions.
- **File** on the top toolbar can be used to Open, Edit, or Export an analysis.
- **View** on the top toolbar can be used to adjust labels and colors, or to hide the dendrogram tree. Changing the font size and selecting the **Keep Fixed** box can make the image labels easier to read.
- **Cluster** on the top toolbar allows you to adjust how data is clustered.
- **Help** provides information on settings and keyboard shortcuts.

Your analysis will be automatically saved and can be accessed again by selecting the **Analysis Data** level under your experiment on the **Experiments** tab.

NSolver automatically launches the open source app, Java TreeView 3.0, and exports the clustered data to create a heat map. No changes to Java TreeView 3.0 code were made. All documentation, including technical documentation on Java TreeView, can be found at https://bitbucket.org/TreeView3Dev/treeview3/ and is based on the following publication: Keil C, Leach RW, Faizaan SM, Bezawada S, Parsons L, Baryshnikova A. (2016). TreeView 3.0 (alpha 3) - Visualization and analysis of large data matrices [Dataset]. Zenodo. http://doi.org/10.5281/zenodo.160573

## 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 median is depicted with a red dot (see Figure 53).

- The **first quartile** is the value at which 25% of data is below. It is also called the *lower quartile* or the *25th percentile*. It is depicted as the top value of the lower blue line in the violin plot.
- The **second quartile** separates the dataset in two halves and is also called the *median* or the 5*0th percentile*. It is depicted as the red dot in the violin plot.
- The **third quartile** is the value at which 75% of data is below. It is also called the *upper quartile* or the *75th percentile*. It is depicted as the bottom value of the upper blue line in the violin plot.

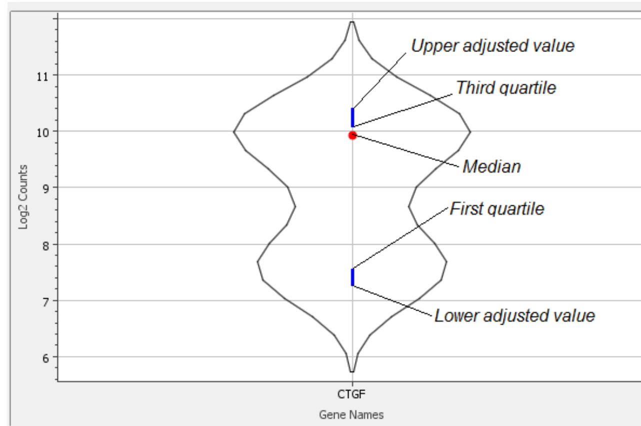



Figure 53: violin plot diagram

To perform an analysis, start with the *To Run an Analysis* section. If you chose a Violin Plot, you can select genes of interest from the left side of the Violin Plots page and watch the plot build dynamically on the right side.

The Violin Plots page (see Figure 54) contains several features you can use to quickly identify the genes you want to plot. You may use the **Data Filtering** function at the top of the page, or sort genes by clicking the column headers in the **Gene** field.

The **Show by condition** checkbox allows you to see expression with respect to a single gene or small group of genes plotted across the experimental annotation. This is available if annotations were added to create treatment groups during experiment building (see the *Annotations* section).

If multiple genes are selected, you can identify which gene belongs to which plot by hovering over each one with the cursor.

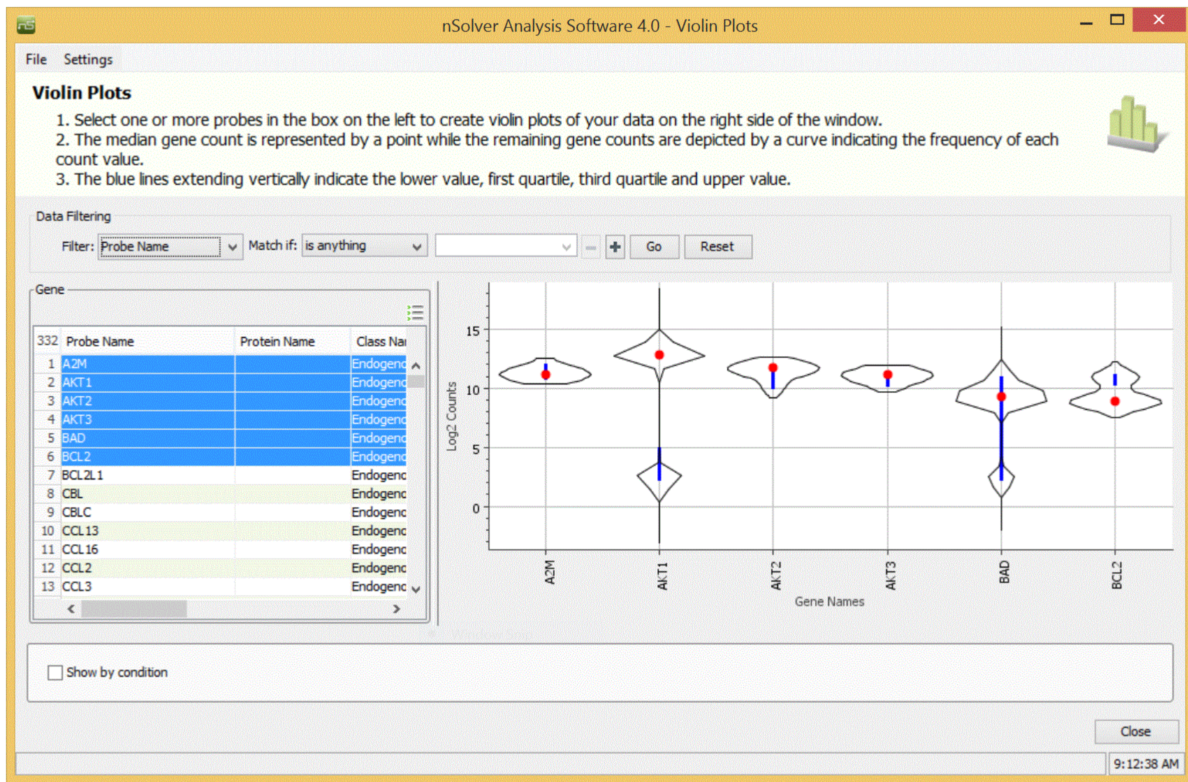



Figure 54: violin plot

Use the **File** menu on the toolbar to Save Image, Save Settings, or Print.

Use the **Settings** menu on the toolbar to specify plot settings, including chart title, axis font, and axis label alignments. You may also specify scale type, choosing to display violin plots using linear count values, $\log_2$ values, or $\log_{10}$ values.

Your analysis will be automatically saved and can be accessed again by selecting the **Analysis Data** level under your experiment on the **Experiments** tab.

## 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 (see Figure 55). Box plots display differences between subsets of an experiment without making any assumptions about the underlying statistical distribution; they are non-parametric.

- The **first quartile** is the value at which 25% of data is below. It is also called the *lower quartile* or the *25th percentile*. It is depicted as the bottom of the box in the box plot.
- The **second quartile** separates the dataset in two halves and is also called the median or the *50th percentile*. It is depicted as the center line of the box in the box plot.
- The **third quartile** is the value at which 75% of data is below. It is also called the *upper quartile* or the *75th percentile*. It is depicted as the top of the box in the box plot.

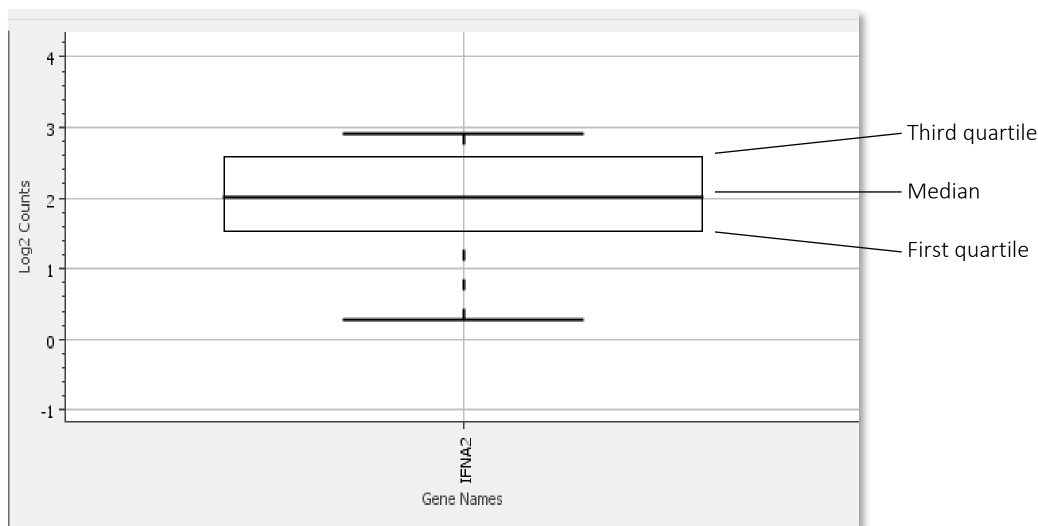



Figure 55: box plot diagram

To perform an analysis, start with the *To Run an Analysis* section. If you chose a Box Plot, you can select genes of interest from the left side of the Box Plots page and watch the plot build dynamically on the right side.

The Box Plots page contains several features you can use to quickly identify the genes you want to plot (Figure 56). You may use the **Data Filtering** function at the top of the page, or sort genes by clicking the column headers.

The **Show by condition** checkbox allows you to see expression with respect to a single gene or small group of genes plotted across the experimental annotation. This is available if annotations were added to create treatment groups during experiment building (see the *Annotations* section).

If multiple genes are selected, you can identify which gene belongs to which plot by hovering over each one with the cursor.


nanoString

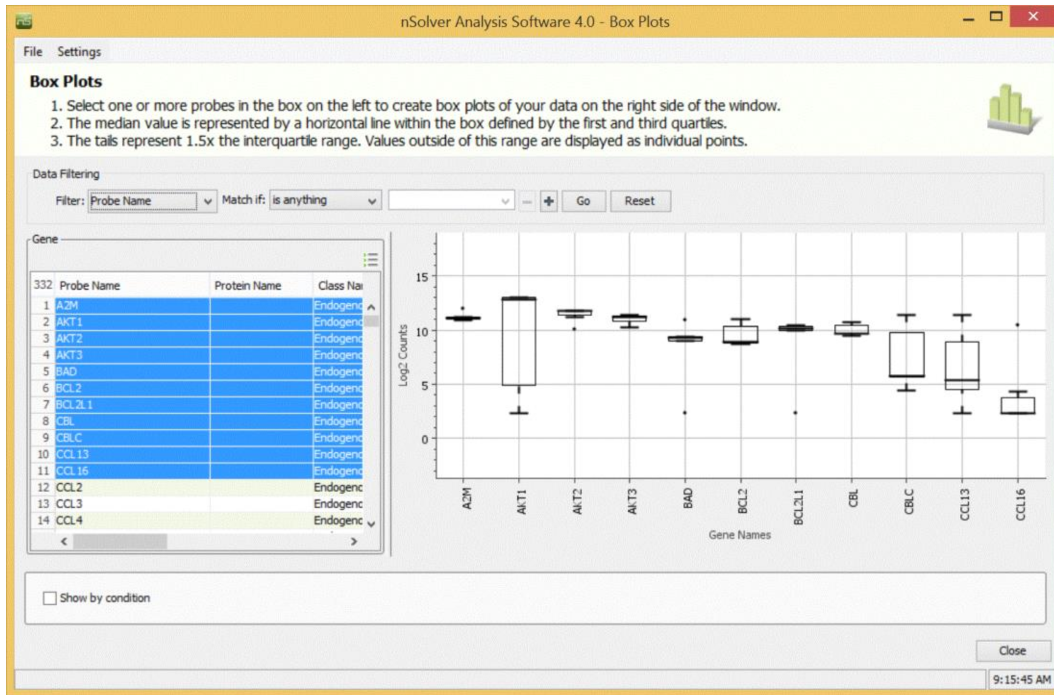



Figure 56: box plot

Use the **File** menu on the toolbar to Save Image, Save Settings, or Print.

Use the **Settings** menu on the toolbar to specify plot settings, including chart title, axis font, and axis label alignments. You may also specify scale type, choosing to display box plots using linear count values, $\log_2$ values, or $\log_{10}$ values.

Your analysis will be automatically saved and can be accessed again by selecting the **Analysis Data** level under your experiment on the **Experiments** tab.

## Scatter Plots

The scatter plot is a visualization that plots one of the samples on the x-axis and the remaining samples on the y-axis.

To perform an analysis, start with the *To Run an Analysis* section. If you selected a scatter plot, you will be asked to specify the samples to include in the scatter plot on the final dialog box of the Analysis Wizard.

- 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), the Analysis Wizard will prompt you to select at least one ratio to be used in the scatter plot. If no p-values exist, then you must select at least two ratios to use in the scatter plot.

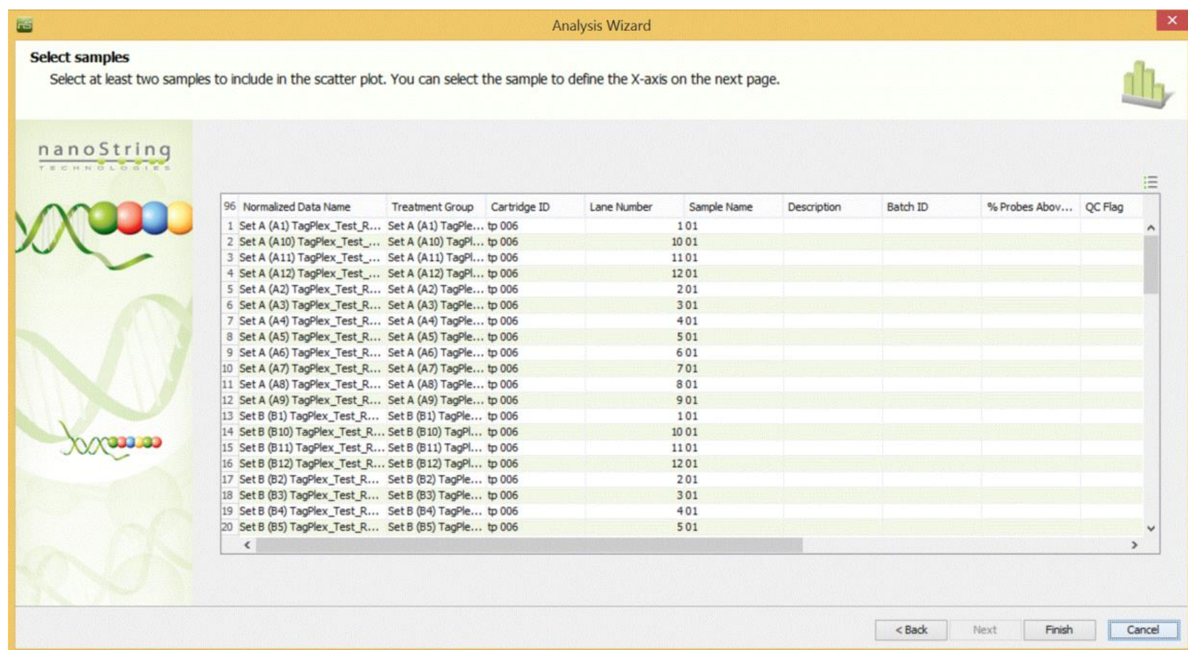



Figure 57: sample selection for scatter plot

Highlight your samples of interest (see Figure 57) and select **Finish.**

Your scatter plot will appear (Figure 58). Customize the plot:

- Adjust the plot dynamically by adding or removing samples by selecting the checkboxes on the y-axis.
- Change the sample depicted on the x-axis using the dropdown menu at the bottom left.
- Use the **Color by…** dropdown menu to color the scatter plot points 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.

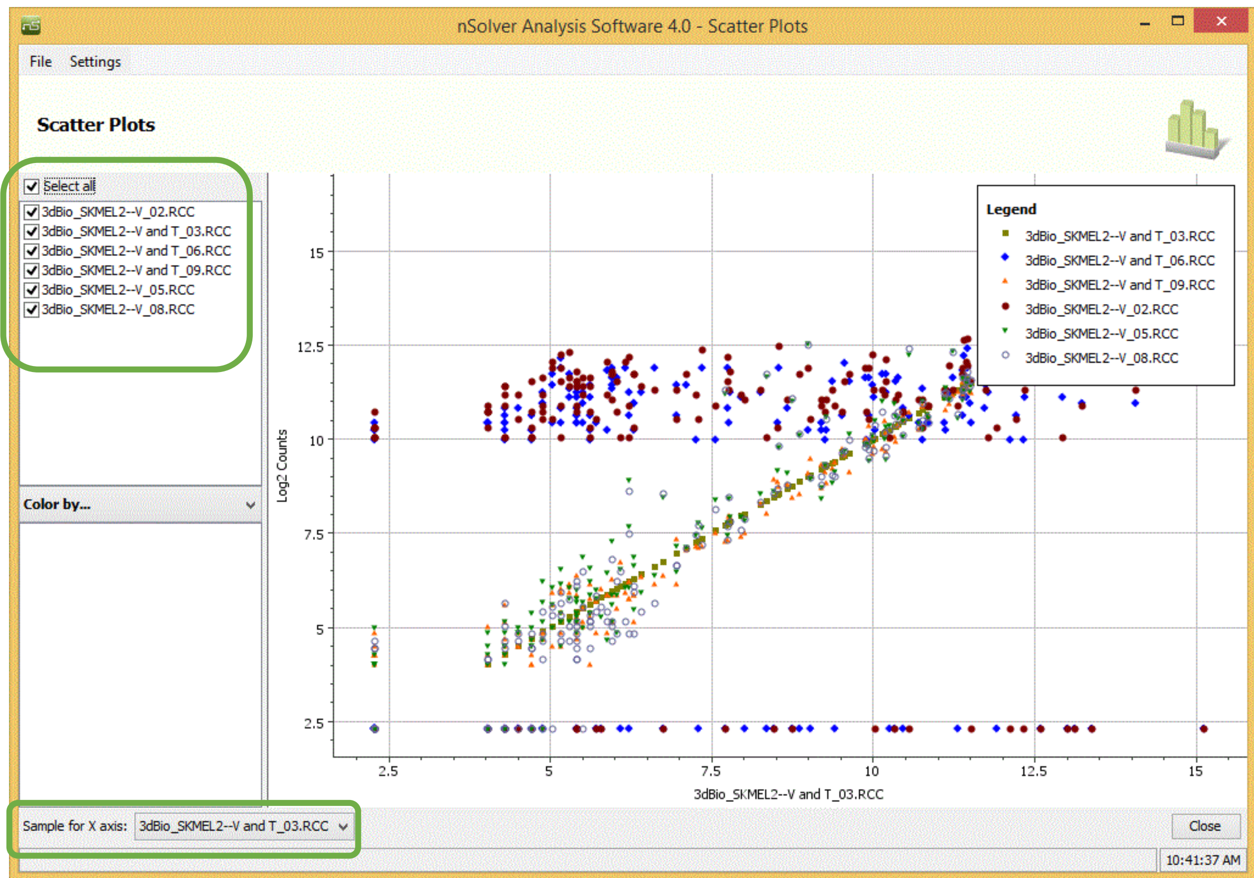



Figure 58: scatter plot

Use the **Settings** menu on the toolbar to access Common Settings, Legend and data point setting, and Scale type.

- Using **Common Settings**, you may customize settings such as the appearance and content of the chart's title, as well as the axes titles, labels, and ranges. You may also choose whether the scatter plot's regression should be shown or not, and with what settings.
- The **Legend settings** allow you to adjust the color of the data points and the content of the legend.
- 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.

Use the **File** menu on the toolbar to Save Image, Save Settings, or Print.

Your analysis will be automatically saved and can be accessed again by selecting the **Analysis Data** level under your experiment on the **Experiments** tab.

## 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.

To perform an analysis, start with the *To Run an Analysis* section. If you chose a Histogram Plot, you can select genes of interest from the left side of the Histogram Plots page (see Figure 59) and watch the plot build dynamically on the right side.

To quickly identify the genes you want to plot, you may use the **Data Filtering** function at the top of the page or sort genes by clicking the column headers.

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

Use the **File** menu on the toolbar to Save Image, Save Settings, or Print.

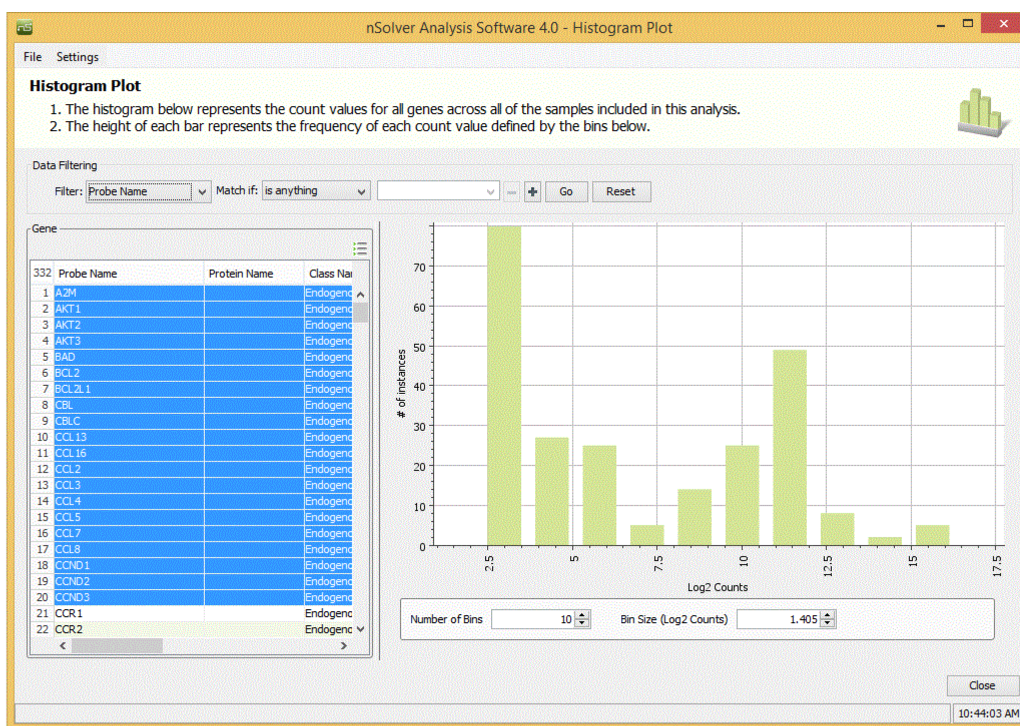



Figure 59: Histogram plot

Use the **Settings** menu to specify plot settings, including chart title, axis font, and axis label alignments.

Your analysis will be automatically saved and can be accessed again by selecting the **Analysis Data** level under your experiment on the **Experiments** tab.

## Analysis Table Options

Highlighting the analysis level of your experiment on the Experiments tab reveals the list of analyses associated with it. Clicking on an analysis in this list activates the analysis table options (Figure 60).

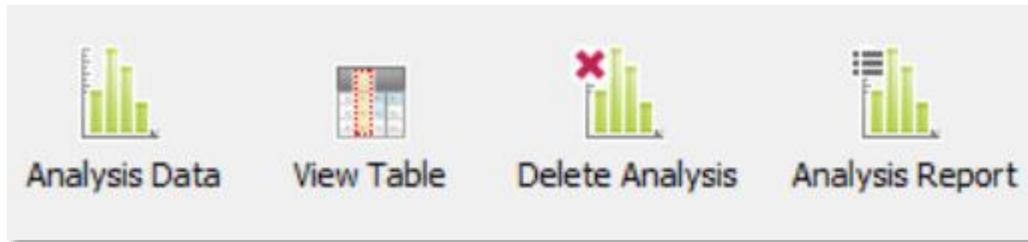



Figure 60: Analysis option buttons

Select **Analysis Data** to view a previously-created plot.

Select **View Table** to view the data table upon which the analysis was performed.

Select **Delete Analysis** to delete the analysis.

Select **Analysis Report** to view a report containing information such as included and excluded samples and genes, as well as plot data. The Analysis Report can be exported and saved in the same way as the Experiment Report (see the *Experiment Reports* section).

## Advanced Analysis

Advanced Analysis is conveniently provided as a link from the nSolver dashboard. It brings together powerful academic open-source analysis tools, provides a simple interface to guide you through analysis, and displays results in an interactive HTML document. Each Advanced Analysis is performed using R, a powerful statistical software program. Familiarity with R is not required as users only need to interact with a simple wizard within nSolver 4.0. See the *Installation* section for download instructions. The information here is intended to be a summary only; for more details on Advanced Analysis, see the ***Advanced Analysis User Manual*** (MAN-10030) for the version you are running.

### Running an Advanced Analysis

To run an Advanced Analysis, navigate to the **raw** or **normalized data** level (other data levels cannot be used) using the navigation tree on the **Experiments** tab. Select samples from the central table which contain **mRNA**, **protein**, **fusion** and/or **SNV** data (other analytes cannot be used) and click the **Advanced Analysis** button.

A warning will appear, prompting you to **import the RLF** for your dataset, if you haven't already.

Type a **Name**, confirm the **Analysis Type** to be used, and use **Browse** to select an output path for the files generated. Select **Next**. A warning will appear if nSolver detects a version of R which is incompatible with the program (R version 3.3.2 is required for Advanced Analysis 2.0).

You will be prompted to establish a unique **identifier** and **covariate(s)** for analysis (Figure 61). To view existing sample annotations, select the **View Annotations** button. To import sample annotations from another source, select the **Import** button.

#### Possible Error Message

***No RLF was loaded***: for an assay to be analyzed in Advanced Analysis, the corresponding RLF file must be loaded. See the *RLFs* section.

***Not compatible R version:*** the version of R you have installed on your computer is not compatible with the version of Advanced Analysis you are trying to run. Advanced Analysis 2.0 requires R version 3.3.2. See the *Downloading R 3.3.2* section.

- A checkbox will appear in the **Identifier** column if one of the existing **Group Identifiers** such as file name, sample name or lane number distinguishes each sample from the others. Select one.

- In the **Use in Analysis** column, choose the covariates you would like to use in analysis. Expand the navigation tree to choose one of the **RCC annotations** and/or select one of the **Experiment annotations** (annotations created during experiment-building).

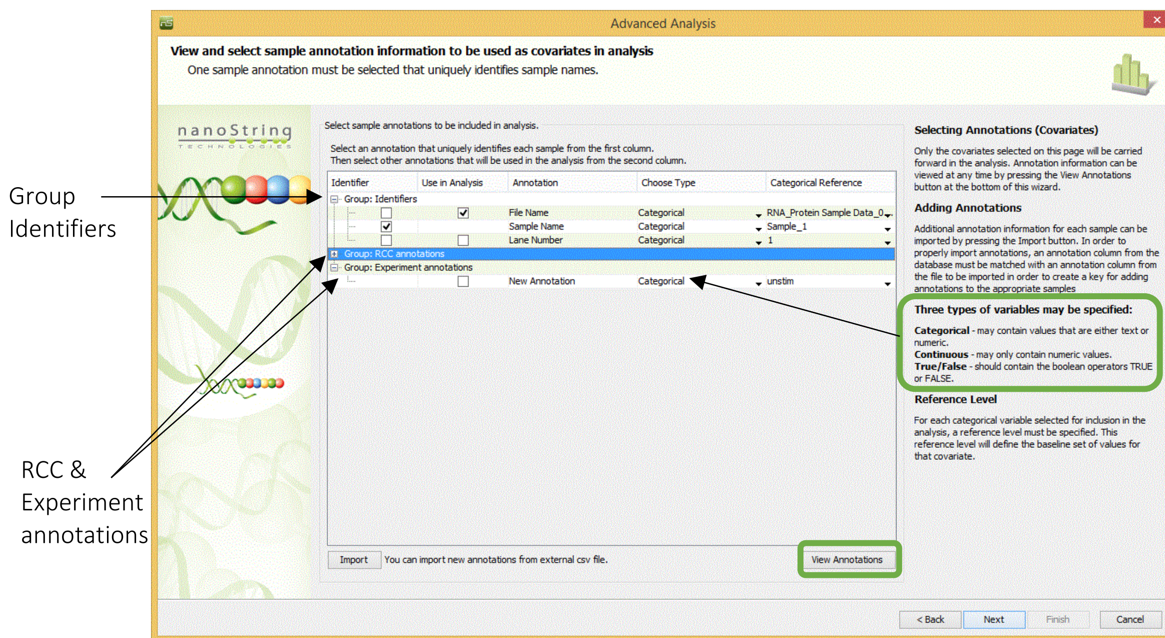



Figure 61: Advanced Analysis – assigning identifiers and covariates

- You can also import a list of annotations from another source. Select the **Import** button and **Browse** to navigate to the .csv file that contains the desired annotations. Use the drop-down menus to choose one annotation from the imported data and the corresponding 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.
- 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. Some analyses require that you establish a baseline reference for categorical variables.

Importing sample annotations

**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). Unselect the checkbox and try again. Alternatively, open the .csv file to verify the new annotations are correct and do not contain any omissions or typos.

Select **Next**.

Choose the type of analysis you would like to run (Figure 62).

- A **Quick Analysis** is one with default settings and is built on only one of your chosen covariates. Select the **annotation** upon which the quick analysis should be built.
- A **Custom Analysis** can be built on multiple covariates and requires that you visit each resulting tab to customize the analysis. You may select or deselect additional modules for analysis.
- The **Summary/Save Settings** tab allows you to review your settings and save them for a future analysis, as well.
- The message at the bottom of this *Analyte Type* window will tell you whether probe annotations were loaded for your dataset. You can use the **click here** link to view and/or edit probe annotations or to download a template in order to create your own.
- The **Load Settings** button retrieves saved settings so that you can apply them to a present analysis.

Select **Finish.**

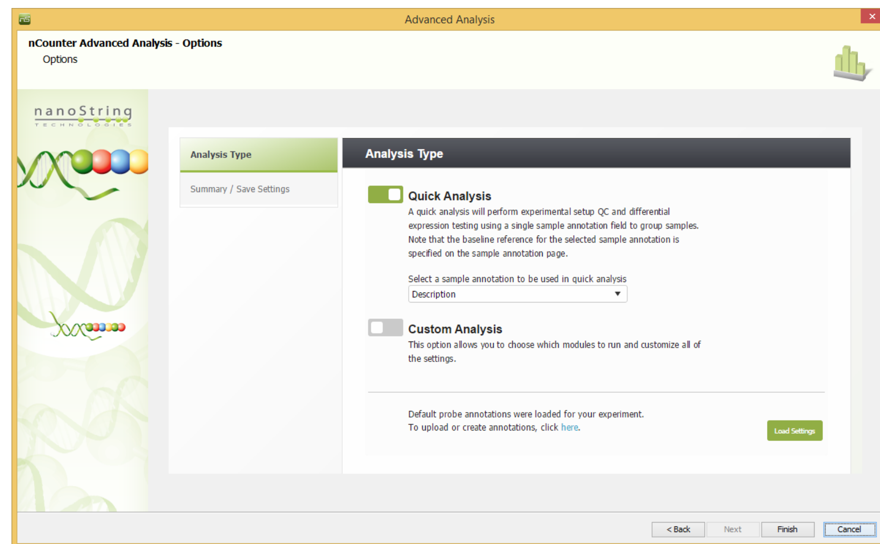



Figure 62: Advanced Analysis Quick or Custom Analysis option

You will be retuned to the nSolver dashboard. Highlight your analysis in the list and select **Analysis Data** to view your plots and options.

This will open an HTML window and dynamically display the progam's status (Figure 63). When complete, a summary screen will appear (Figure 64). Click through the different plots and options for viewing data. For more information, refer to the *Advanced Analysis User Manual* (MAN-10030).

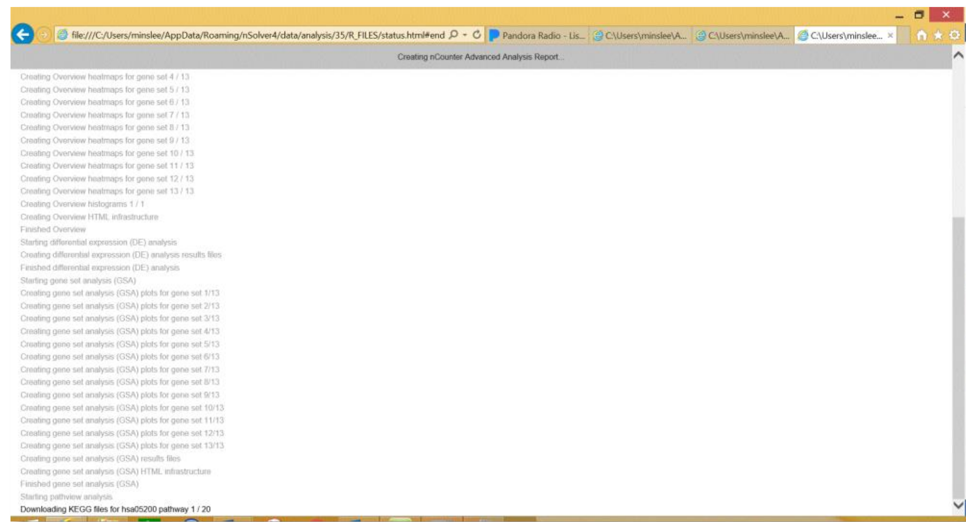


Figure 63: HTML window with Advanced Analysis status

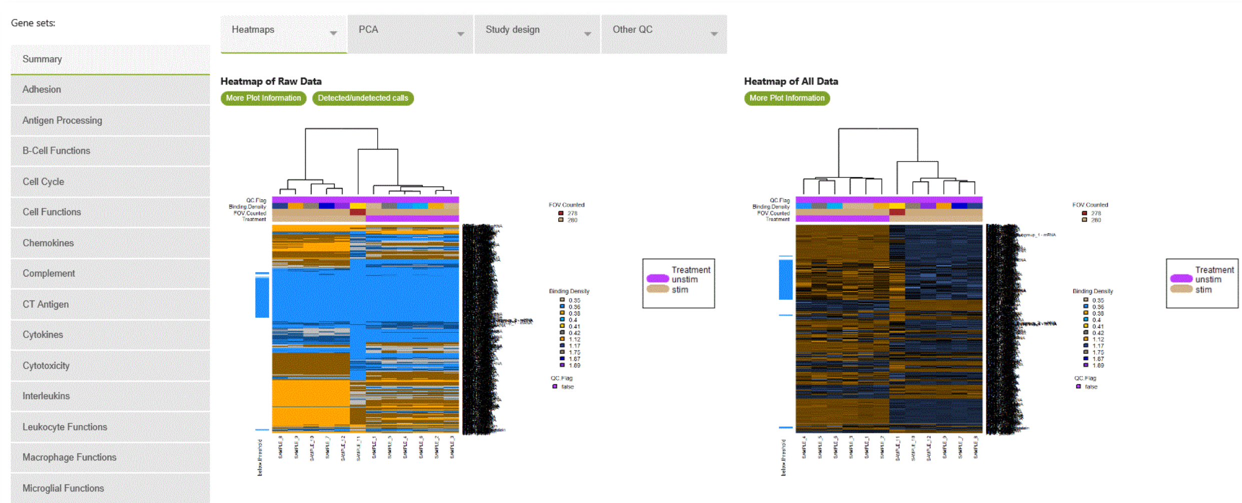



Figure 64: Overview window of Advanced Analysis

# Analyte- & Assay-Specific Notes

nSolver 4.0 resolves data from a diverse range of assays and analyte types. As a result, customization of the workflow is necessary for some processes. Use the *Quick Start* section or main body of the nSolver 4.0 User Manual for the general workflow of data analysis; this section is supplementary and provides additional details encountered when working with *Plex$^2$*, *PlexSet*, *SNV*, *CNV*, and *multi-RLF* data.

## Plex$^2$ Data Analysis

Due to the multi-plexed nature of the Plex$^2$ assay, its nSolver workflow is slightly different from that of other assays.

The Plex$^2$ assay uses 48 wells of a plate for hybridization and allows you to pool up to 4 samples into a single cartridge lane for analysis on the Digital Analyzer (see Figure 65). The barcodes in the Plex$^2$ kit are divided into 4 sub-CodeSets and are labeled such that the data can be sorted by sub-CodeSet (1-4), Pool (1-12), or Sample (1-48) in nSolver.

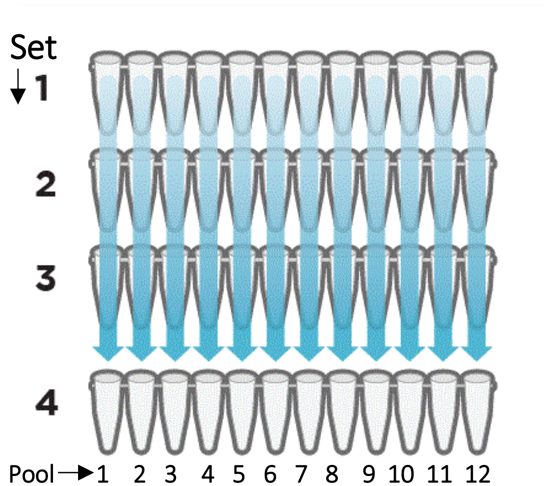


Figure 65: Plex$^2$ sample set up

**Importing RCC files** is done in the same manner as the standard workflow.

**Quality Control** uses only the Binding Density and Image Quality in Plex$^2$ assays.

When **exploring raw data**, note that the main raw data table columns are labeled as *Set 1*, *2*, *3,* and *4* (see Figure 66). Rows are labeled with RCC file names, which correspond to the different cartridge lanes or *Pools 1-12*. You may choose to enter individual sample names here.

| 12 | File Name | Description | Batch ID | Set 1 Sample Name | Set 2 Sample Name | Set 3 Sample Name | Set 4 Sample Name | Sample Name | Cartridge ID |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 20111220_Plex2_Pool-10_10.RCC | | | | | | | Pool-10 | Plex2_Cartridge |
| 2 | 20111220_Plex2_Pool-11_11.RCC | | | | | | | Pool-11 | Plex2_Cartridge |
| 3 | 20111220_Plex2_Pool-12_12.RCC | | | | | | | Pool-12 | Plex2_Cartridge |
| 4 | 20111220_Plex2_Pool-1_01.RCC | | | | | | | Pool-1 | Plex2_Cartridge |
| 5 | 20111220_Plex2_Pool-2_02.RCC | | | | | | | Pool-2 | Plex2_Cartridge |
| 6 | 20111220_Plex2_Pool-3_03.RCC | | | | | | | Pool-3 | Plex2_Cartridge |
| 7 | 20111220_Plex2_Pool-4_04.RCC | | | | | | | Pool-4 | Plex2_Cartridge |
| 8 | 20111220_Plex2_Pool-5_05.RCC | | | | | | | Pool-5 | Plex2_Cartridge |
| 9 | 20111220_Plex2_Pool-6_06.RCC | | | | | | | Pool-6 | Plex2_Cartridge |
| 10 | 20111220_Plex2_Pool-7_07.RCC | | | | | | | Pool-7 | Plex2_Cartridge |
| 11 | 20111220_Plex2_Pool-8_08.RCC | | | | | | | Pool-8 | Plex2_Cartridge |
| 12 | 20111220_Plex2_Pool-9_09.RCC | | | | | | | Pool-9 | Plex2_Cartridge |

Figure 66: Plex$^2$ raw data table layout

Highlighting your samples of interest and selecting the **Table** button allows you to view the individual counts of each sample (see Figure 67); column headers are sorted by well number (*Set1-Pool1*, *Set2-Pool1*, etc.).

| Set 1 () 20111220_Plex2_Pool-1_01.RCC | Set 2 () 201112... | Set 3 () 201112... | Set 4 () 201112... | Set 1 () 201112... | Set 2 () 20111220_Plex2_Pool-2_02.RCC |
|---|---|---|---|---|---|
| 209 | 193 | 207 | 144 | 106 | 168 |
| 847 | 869 | 588 | 238 | 416 | 1,278 |
| 3 | 7 | 13 | 9 | 7 | 17 |
| 14 | 14 | 21 | 9 | 23 | 18 |
| 13 | 6 | 9 | 1 | 10 | 14 |
| 2 | 2 | 8 | 5 | 4 | 8 |
| 8 | 3 | 7 | 14 | 21 | 12 |
| 14 | 16 | 4 | 11 | 13 | 18 |
| 145 | 146 | 94 | 86 | 236 | 184 |
| 7 | 7 | 7 | 7 | 8 | 4 |
| 16 | 12 | 13 | 14 | 28 | 14 |
| 5 | 8 | 9 | 3 | 6 | 12 |

Figure 67: Plex$^2$ raw data table columns

Creating an **Experiment** takes you through annotations, background options and normalization, just as in the standard workflow. You can maintain the default settings:

- **Background** correction will be deselected. If you choose to correct for background, choose **Thresholding**, and set to a count value of **20**.
- **Positive Normalization** should be set to the **geo mean** of all **POS** counts (A-F).
- **CodeSet Content Normalization** can be selected if you have designated Housekeeping Genes; move these from the *CodeSet Content* window to the *Normalization Codes* window using the arrows. If you don't have designated housekeeping genes in the present CodeSet, you may skip CodeSet Content Normalization.

After normalization, you have the option to designate your reference lane in the *CodeSet Calibration* window. For **calibration**, a reference sample should have been loaded in all 4 wells (one corresponding with each CodeSet, or "Set") of at least one of the pools so that at least one known sample is run across the entire Plex$^2$ CodeSet. Designate this pool in the *CodeSet Calibration* window (Figure 68).

Select the **Sample Reference Normalization** checkbox in the upper left to activate the options in the window. Select the lane in which you loaded your reference sample in the *Subcode Samples* window (on the left). Use the arrows to move the desired lane to the *Selected Samples* window (on the right).

The different Sets of data are available on the different tabs. Select one Set as your reference by checking the **Use as Reference** checkbox at the bottom of the appropriate tab. Checking the box **Warn if count of genes is less: …** and selecting a value from the dropdown will ellicit a popup, warning you of the genes whose counts fall below that value (see below).

Select **Next**.

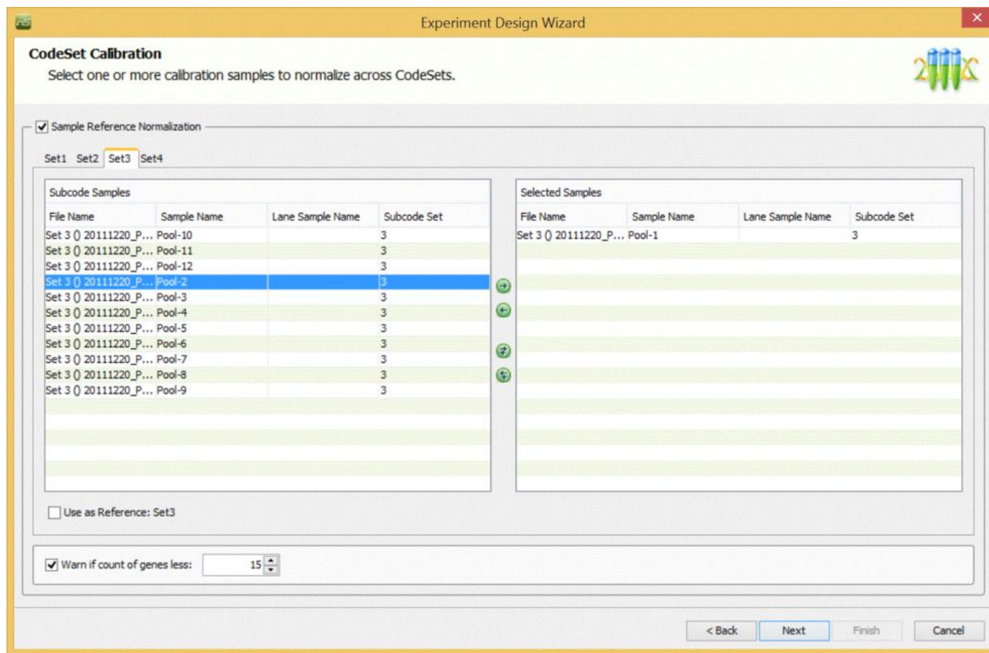



Figure 68: PlexSet CodeSet Calibration

If the box titled **Warn if count of genes is less: …** was checked in the CodeSet Calibration window, you will be alerted to any low counts. Select **Continue**.

You will be prompted to establish a baseline data for ratios, then asked to assign ratio data names per the standard workflow. Select **Next**, then **Finish**.

Data viewing and analysis follow the same guidelines as standard sample assays.



nanoString™

## PlexSet Data Analysis

Due to the unique nature of the PlexSet assay, its nSolver workflow is slightly different from that of other assays.

The PlexSet assay uses a 96-well plate for hybridization and allows you to pool all 8 samples in a column (A-H) into a single cartridge lane for analysis on the Digital Analyzer (see Figure 69). The barcodes in the PlexSet sets (A-H) are labeled such that the data can be sorted by well number (A1, B1, etc.) or by cartridge lane number (1-12) in nSolver.

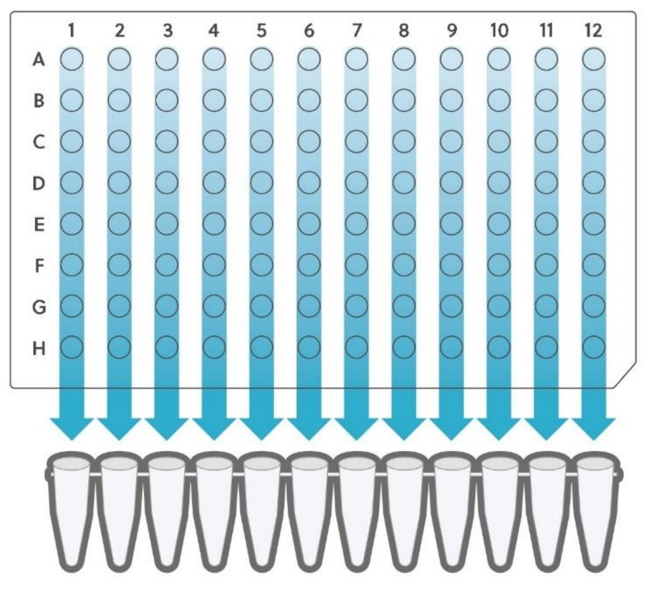



Figure 69: PlexSet multiplexing

### Sample Setups

You can use a few different configurations for setup while being mindful of important downstream QC considerations (see below for examples). For more information, see the *PlexSet Reagents Manual* (MAN-10040) and the *All About PlexSet Technology Data Analysis in nSolver Software guide* (MAN-10044).

**Example setup 1: Running a single probe set across multiple plates**. Reserve lane 1 of the first plate for your Calibration Sample. Subsequent PlexSet cartridges using the same probe sets will not need an additional reference sample lane. With our current nSolver data analysis software, your experiments should be organized down columns. One RLF is used for these scenarios.

**Example setup 2: Running less than 96 samples per cartridge.** PlexSets A-D can be run on one cartridge, and PlexSets E-H can be run on another. When subsequent PlexSet kits are run with different PlexSet configurations (e.g. all PlexSets), a reference sample should be re-run across all PlexSets for calibration. One RLF is used for these scenarios.

**Example setup 3: Running PlexSet kits with different probe sets.** If multiple probe sets are run on the same plate, probe sets should be organized down columns to allow downstream analysis with nSolver software. Two RLFs are used for these scenarios (one for each probe set).

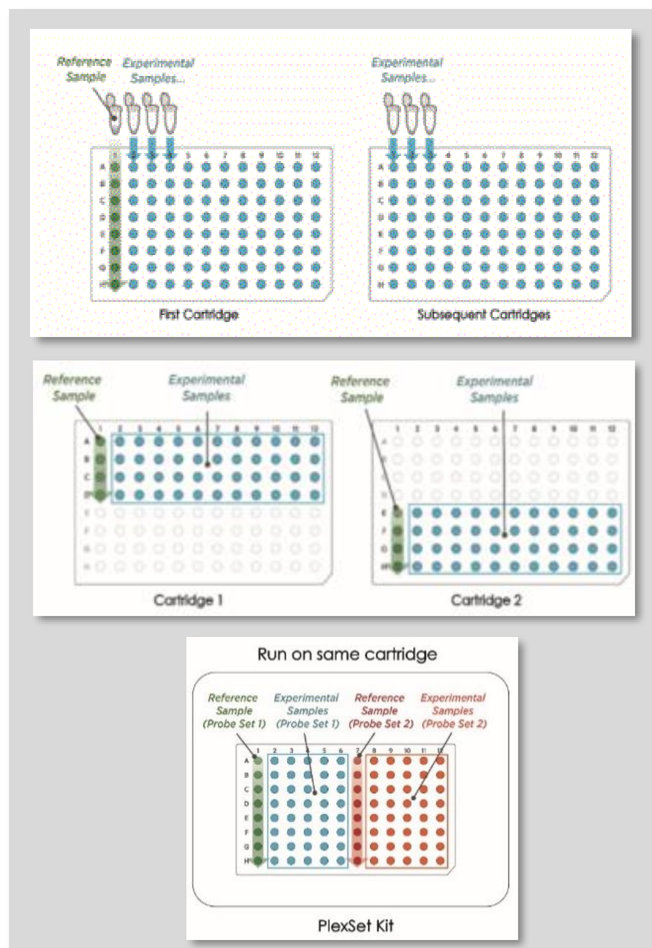



Figure 70: PlexSet example setups

## Titration

Prepare a **Titration Run** at least once per study, as running a Titration Kit provides information on the optimal sample input amounts for the PlexSet assay. The Titration Kit contains nCounter XT TagSet reagents that correspond to the PlexSet assay (e.g. PlexSet-48 will use nCounter XT TagSet-48). See the *PlexSet Reagents Manual* (MAN-10040) for more information. It contains the same types of POS and NEG controls as most other PlexSet assays.

**Import your Titration data** into nSolver (see the *Quick Start* section and below). Check your raw data, paying particular attention to any lanes with QC flags. Create an Experiment using the following settings:

- **Background** correction will be de-selected. Leave this default.
- **Positive Normalization** should be set to the **geo mean** of the **Top 3** POS counts. De-select CodeSet Content Normalization.
- De-select **Reference Calibration**, as you should skip this step.
- De-select **Build Ratios**, as you should skip this step. Select **Finish** to build the experiment.

Highlight relevant lanes of **Normalized Data**, select **Export**, then use default settings of the **RCC Collector Tool Format Export**. Save the resulting .csv file and open it in the spreadsheet program of your choice.

Copy the columns containing **probe name**, and those containing **the counts for each titration category** in the cell type of interest. Paste them into another spreadsheet tab and label them appropriately (see Figure 71).

| | | | Lung | | |
|---|---|---|---|---|---|
| **Tag** | **Barcode** | **Gene Name** | **50** | **100** | **200** |
| tag-001 | GRBRGB | GAPDH | 46738 | 89410 | 173993 |
| tag-005 | BRBYRY | B2M | 22574 | 43690 | 84695 |
| tag-086 | RBRBYG | HLA-DRA | 6268 | 12115 | 23433 |
| tag-077 | GYGBGB | CD9 | 6029 | 11148 | 22135 |
| tag-054 | YBRYGB | FN1 | 4542 | 9137 | 17875 |
| tag-068 | BGRBYB | ITGB1 | 3833 | 7141 | 14027 |
| tag-006 | GBGYRG | TUBB | 2900 | 5650 | 11083 |
| tag-029 | YGBYGR | TFRC | 2843 | 5625 | 11368 |
| tag-053 | GBRBRY | XBP1 | 2322 | 4393 | 8439 |
| tag-027 | GRBRGR | CTNNB1 | 2085 | 3979 | 8231 |

Figure 71: Exported and formatted titration data

Average any duplicate counts.

Add the counts, excluding POS and NEG controls, in each column (use the **SUM** function in Excel). See Figure 72.

Highlight the summed counts and the titration levels and **Insert** a **Line Graph**. The titration categories (sample input) should be set as the x-axis, and Total Normalized Counts should be set as the y-axis. See Figure 73.

| | | Lung | | |
|---|---|---|---|---|
| | **Input (ng)** | **50** | **100** | **200** |
| | **SUM** | 134,906 | 258,459 | 505,265 |
| | **Equation** | **y = 2468.9x + 11503** | | |
| | **y intercept** | **11503.0** | | |
| | **slope** | **2468.9** | | |
| | **Correlation (R2)** | **1.000** | | |
| **MAX/FLEX** | **Input (ng) for 150,000 Counts** | 56 | | |
| **SPRINT** | **Input (ng) for 400,000 Counts** | 157 | | |

Figure 72: titration data sums and line graph equations

View the equation for the line in the format:

$y = mx + b$

Copy this into a new cell (see Figure 72) and set y to the applicable value:

y= 150,000 for MAX/FLEX platforms

y= 400,000 for SPRINT platforms

Solve for **x**. This is your optimum input amount for the sample.

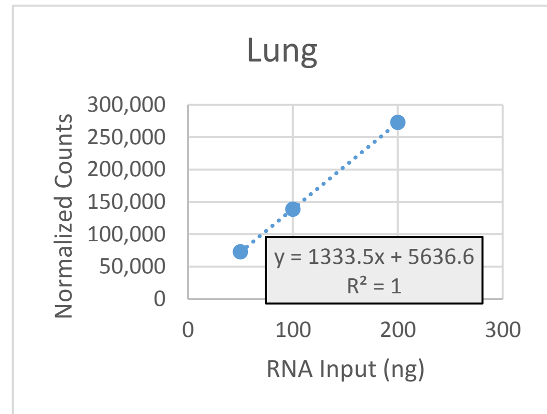



Figure 73: titration data line graph

## PlexSet Sample Data Analysis

**Importing RCC files** is done in the same manner as the standard workflow. If running multiple cartridges, import them all; you'll separate out your samples of interest later, when you create an experiment (see below).

**Quality Control** uses only the Binding Density and Image Quality in PlexSet assays. Neither Positive QC parameter is measured. Instead, identical positive controls are included in each PlexSet set (one for each row A-H). These eight POS controls are listed with each sample's counts, acting collectively as lane controls (see below). One negative control is included in each PlexSet set (one for each row A-H), as well.

When **exploring raw data**, note that the main raw data table columns are labeled as *Set A*, *Set B*, etc. through *Set H* (see Figure 74). Rows are labeled with RCC file names, which correspond to the different cartridge *lanes 1-12*.

| 36 | File Name | Description | Batch ID | Set A Sample Name |
|---|---|---|---|---|
| 1 | PlexSet_OneProbeSet_FullPlate_01.RCC | | | A01 |
| 2 | PlexSet_OneProbeSet_FullPlate_02.RCC | | | A02 |
| 3 | PlexSet_OneProbeSet_FullPlate_03.RCC | | | A03 |
| 4 | PlexSet_OneProbeSet_FullPlate_04.RCC | | | A04 |
| 5 | PlexSet_OneProbeSet_FullPlate_05.RCC | | | A05 |
| 6 | PlexSet_OneProbeSet_FullPlate_06.RCC | | | A06 |
| 7 | PlexSet_OneProbeSet_FullPlate_07.RCC | | | A07 |
| 8 | PlexSet_OneProbeSet_FullPlate_08.RCC | | | A08 |
| 9 | PlexSet_OneProbeSet_FullPlate_09.RCC | | | A09 |
| 10 | PlexSet_OneProbeSet_FullPlate_10.RCC | | | A10 |
| 11 | PlexSet_OneProbeSet_FullPlate_11.RCC | | | A11 |

Figure 74: PlexSet data table layout

Highlighting your samples of interest and selecting the **Table** button allows you to view the individual counts of each sample (see Figure 75); column headers are sorted by well number (A1, B1, etc.). Note that all eight POS and NEG controls are listed for each sample. Collectively, these act as lane controls.

| 112 | Probe Name | Accession # | NS Probe ID | Class Name | Analyte Type | Annotation | Set A (A01) PlexSet_OneProbeSet_FullPlate_01.RCC | Set B |
|---|---|---|---|---|---|---|---|---|
| 96 | ZEB1 | NM_001128128.1 | ZEB1 | Endogenous | mRNA | | 984 | |
| 97 | NEG_1 | ERCC_00019.1 | NEG_1 | Negative | SYSTEM | | 12 | |
| 98 | NEG_2 | ERCC_00019.1 | NEG_2 | Negative | SYSTEM | | 4 | |
| 99 | NEG_3 | ERCC_00019.1 | NEG_3 | Negative | SYSTEM | | 6 | |
| 100 | NEG_4 | ERCC_00019.1 | NEG_4 | Negative | SYSTEM | | 11 | |
| 101 | NEG_5 | ERCC_00019.1 | NEG_5 | Negative | SYSTEM | | 3 | |
| 102 | NEG_6 | ERCC_00019.1 | NEG_6 | Negative | SYSTEM | | 7 | |
| 103 | NEG_7 | ERCC_00019.1 | NEG_7 | Negative | SYSTEM | | 5 | |
| 104 | NEG_8 | ERCC_00019.1 | NEG_8 | Negative | SYSTEM | | 7 | |
| 105 | POS_1 | ERCC_00002.1 | POS_1 | Positive | SYSTEM | | 3,586 | |
| 106 | POS_2 | ERCC_00002.1 | POS_2 | Positive | SYSTEM | | 2,302 | |
| 107 | POS_3 | ERCC_00002.1 | POS_3 | Positive | SYSTEM | | 2,186 | |
| 108 | POS_4 | ERCC_00002.1 | POS_4 | Positive | SYSTEM | | 2,900 | |
| 109 | POS_5 | ERCC_00002.1 | POS_5 | Positive | SYSTEM | | 2,705 | |
| 110 | POS_6 | ERCC_00002.1 | POS_6 | Positive | SYSTEM | | 2,607 | |
| 111 | POS_7 | ERCC_00002.1 | POS_7 | Positive | SYSTEM | | 2,793 | |
| 112 | POS_8 | ERCC_00002.1 | POS_8 | Positive | SYSTEM | | 1,973 | |

Figure 75: PlexSet data table columns

Creating an **Experiment** takes you through annotations, background options and normalization, just as in the standard workflow. Select your samples of interest. You can maintain the default settings:

- **Background** correction will be deselected. If you choose to correct for background, choose **Thresholding**, set to a count value of **20**.
- **Positive Normalization** should be set to the **geo mean** of the **Top 3** POS counts.
- **CodeSet Content Normalization** can be selected if you have designated Housekeeping Genes; move these from the *CodeSet Content* window to the *Normalization Codes* window using the arrows. If you don't have designated housekeeping genes in the present CodeSet, you may skip CodeSet Content Normalization.

After normalization, you have the option to designate your reference lane in the *CodeSet Calibration* window (see Figure 76). For **calibration**, a reference sample should have been loaded in all corresponding wells of one column of the 96 well hybridization plate so that at least one known sample is run across the entire PlexSet (see example setups, above).

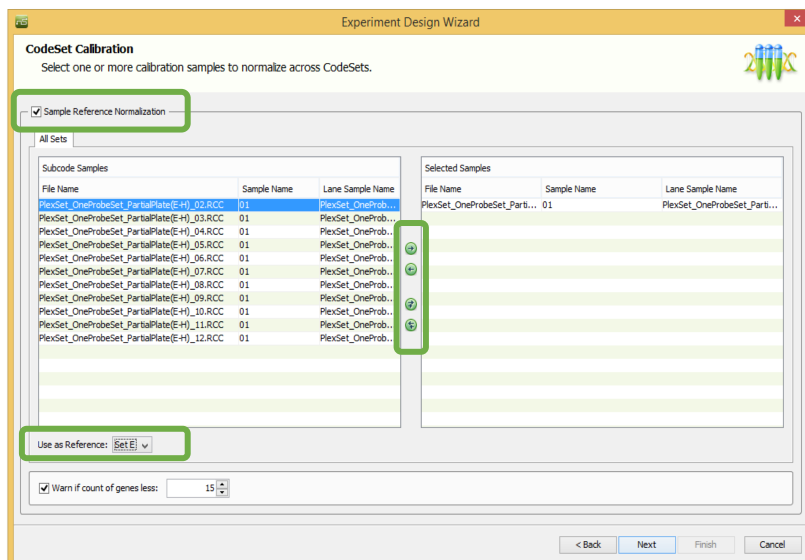



Figure 76: CodeSet Calibration for PlexSet

Select the **Sample Reference Normalization** checkbox in the upper left to activate the options in the window. Select the lane in which you loaded your reference sample in the *Subcode Samples* window (on the left). In most setup configurations, this is lane 1. Use the arrows to move the desired lane to the *Selected Samples* window (on the right).

Select one of the rows in which your PlexSet was loaded in the **Use as Reference** dropdown at the bottom of the window. The default is *Set A*; if running a partial plate in rows E-H (as in Cartridge 2 of Example 3, above), you must change this default to *Set E, F, G*, or *H*. Checking the box **Warn if count of genes is less: …** and selecting a value from the dropdown will ellicit a popup, warning you of the genes whose counts fall below that value (see below). Select **Next**.

If the box titled **Warn if count of genes is less: …** was checked in the CodeSet Calibration window, you will be alerted to any low counts. If you ran a partial plate, some of these low counts may be due to readings from empty wells on the plate and can be disregarded (see Figure 77). Select **Continue**.

**Low Count Genes**

The following genes have low count levels for at least one reference sample(s). This cross RLF experiment may not support accurate comparisons. Press cancel to adjust reference sample(s), or proceed to continue with the selected reference samples.

| 96 | Gene Name | Set A | Set B | Set C | Set D | Set E | Set F | Set G | Set H |
|---|---|---|---|---|---|---|---|---|---|
| 1 | ABCF1 | 4749.0 | 5762.0 | 1248.0 | 3322. | 1.0 | 1.0 | 1.0 | 1.0 |
| 2 | ABL1 | 741.0 | 3208.0 | 8392.0 | 7820. | 1.0 | 3.0 | 1.0 | 2.0 |
| 3 | ALAS1 | 6613.0 | 6475.0 | 11464.0 | 6142. | 1.0 | 1.0 | 1.0 | 1.0 |
| 4 | B2M | 6462.0 | 13960.0 | 11836.0 | 14584. | 1.0 | 2.0 | 1.0 | 1.0 |
| 5 | BCL2 | 14116.0 | 14388.0 | 11814.0 | 492. | 1.0 | 4.0 | 1.0 | 1.0 |
| 6 | BCL6 | 5701.0 | 14467.0 | 9393.0 | 11605. | 1.0 | 1.0 | 1.0 | 1.0 |
| 7 | C3 | 14037.0 | 10041.0 | 13752.0 | 14324. | 1.0 | 1.0 | 1.0 | 3.0 |
| 8 | CASP3 | 10961.0 | 1215.0 | 2998.0 | 2755. | 2.0 | 2.0 | 1.0 | 2.0 |
| 9 | CASP8 | 11294.0 | 12542.0 | 5153.0 | 13637. | 1.0 | 1.0 | 1.0 | 1.0 |
| 10 | CCL2 | 8649.0 | 10982.0 | 8235.0 | 11832. | 1.0 | 1.0 | 1.0 | 1.0 |
| 11 | CCL20 | 9210.0 | 7232.0 | 5935.0 | 3990.0 | 1.0 | 1.0 | 1.0 | 1.0 |

Figure 77: Low Genes alert window from a partial PlexSet plate. Sets E-H were not loaded.

You will be prompted to establish a baseline data for ratios, then asked to assign ratio data names per the standard workflow. Select **Next**, then **Finish**.

When viewing and exporting data, ***select only the lanes in which your PlexSet was loaded***. Data in rows that were not run may appear over-normalized, meaning you may see high counts in these fields; these fields should be disregarded. Data analysis follows the same guidelines as standard sample assays.

## SNV Data Analysis

The nature of the SNV assay and resulting data analysis differs slightly from that of other analytes.

**Importing RCC files** is done in the same manner as the standard workflow. Select **SNV** from the analyte drop-down box to filter datasets and display only those containing SNV data (see Figure 78).

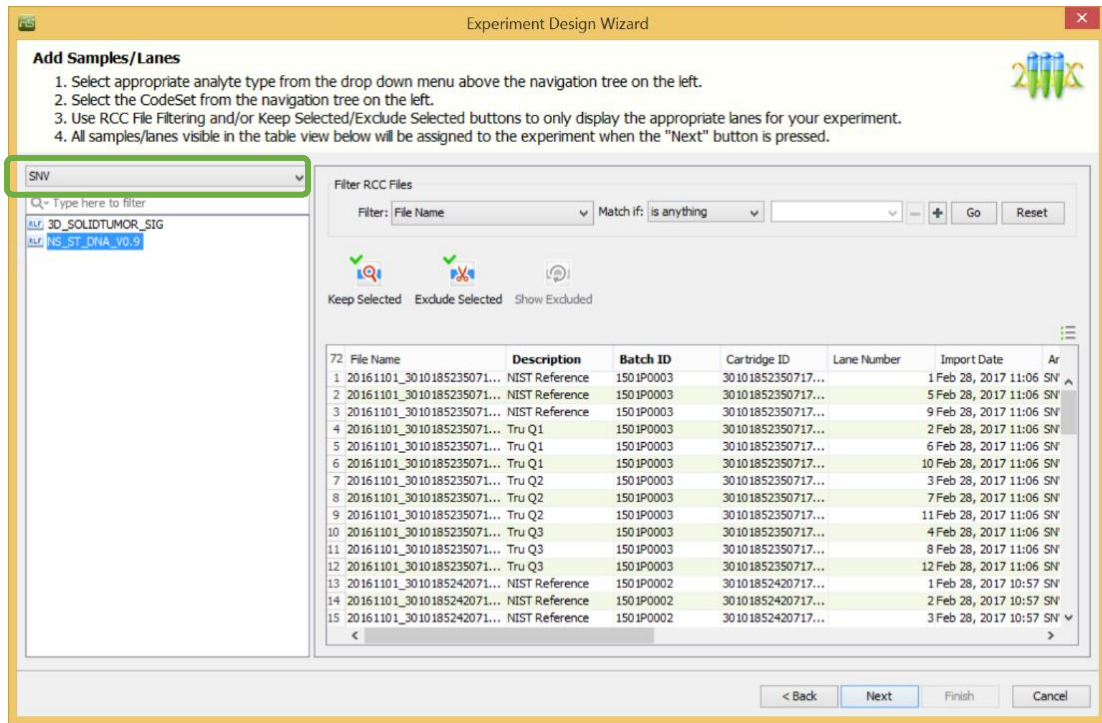



Figure 78: Importing SNV samples

There are several **QC parameters** unique to the SNV assay. See the *Analyte-Specific QC* section.

**Exploring raw data** is done in the same manner as the standard workflow.

Creating an **Experiment** takes you through creating annotations and normalization, similar to the standard workflow.

- No **Background** subtraction/thresholding will be initiated for SNV analysis.
- **Normalization** for SNV data is hardcoded; no customization options are available.
- **Probe Calibration** must be performed on at least 10 reference samples. These reference samples are used as a benchmark to which probe count levels are compared to detect variants. Datasets from different RLFs are represented on different tabs (see Figure 79). Move selected reference samples from the *Subcode Samples* window to the *Selected Samples* window using arrows.
- No **Ratio** calculations will be performed on SNV data.

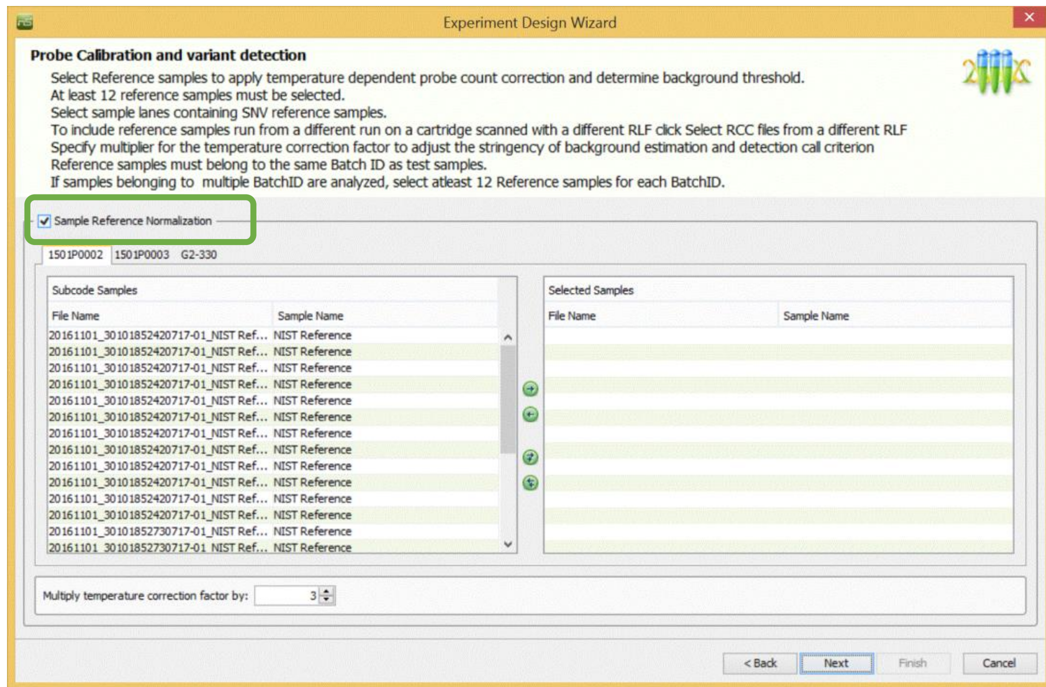



Figure 79: Probe calibration window for SNV data

Once built, find your experiment on the **Experiments** tab and expand the navigation tree (by clicking the **+** sign) to see the different levels of data tables. Select either **raw** or **normalized** data, highlight samples of interest, and select the **Table** button to create a table.

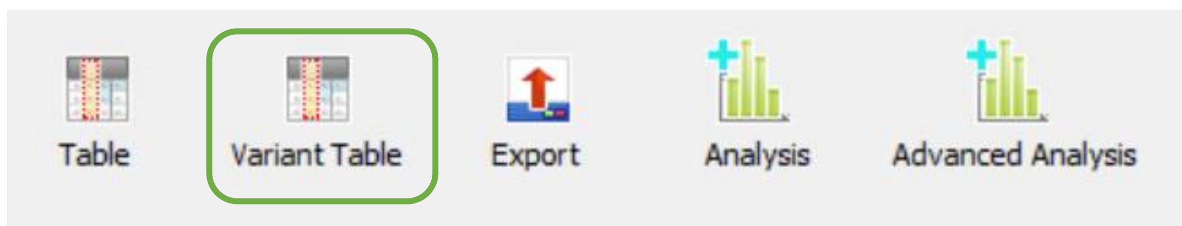



Figure 80: Table options buttons for SNV data

When the **Normalized Data** level is selected, the **Variant Table** button becomes active (see Figure 80). Select the normalized samples of interest, then select this button. In the resulting table, scroll to the right so that you can see all the sample data columns. Each sample will have its own **Variant Call** column. Scroll down to view each of these columns' results.

There are three categories of calls possible in the **Variant Call** columns:

- Variants detected in the data will be signified with a **green variant flag** (see Figure 81).
- A possible variant with inconclusive data may be signified with a **yellow "no call" flag**.
- A reference allele will exhibit no flag.

For visualizations, select raw or normalized data and select the **Advanced Analysis** button. Refer to the *Advanced Analysis* section of this manual as well as the *Advanced Analysis User Manual* (MAN-10030). There are limited tools in the basic Analysis option which will visualize SNV data.

| 104 | Probe Name | Class Name | Gen... | Genome Build | Chrom... | Start Position | End Position | SNV_VAR/SNV_... | Preliminary Vari... | SNV_VAR/SNV_... | Preliminary Va |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 72 | NRAS COSM564 (G12D) | SNV_VAR | - | hg19 | chr01 | 115,258,707 | 115,258,775 | 0.00125762 | | 0.00158081 | |
| 73 | TP53 COSM10733 (Q192*) | SNV_VAR | + | hg19 | chr17 | 7,578,240 | 7,578,312 | 0.00138191 | | 0.00134279 | |
| 74 | APC COSM13127 (R1450*) | SNV_VAR | + | hg19 | chr05 | 112,175,607 | 112,175,686 | 0.00058019 | | 0.00065962 | |
| 75 | TP53 COSM11513 (E68*) | SNV_VAR | - | hg19 | chr17 | 7,579,448 | 7,579,515 | 0.00275895 | | 0.0031568 | |
| 76 | ALK COSM144250 (G1202R) | SNV_VAR | - | hg19 | chr02 | 29,443,574 | 29,443,640 | 0.00054623 | | 0.00047338 | |
| 77 | ALK COSM99137 (L1196M) | SNV_VAR | - | hg19 | chr02 | 29,443,574 | 29,443,641 | 0.00054623 | | 0.00047338 | |
| 78 | NRAS COSM584 (Q61R) | SNV_VAR | - | hg19 | chr01 | 115,256,484 | 115,256,557 | 35.9107132 | | 36.81834412 | |
| 79 | ROS1 NOCOSM12 (G2032R) | SNV_VAR | - | hg19 | chr06 | 117,638,310 | 117,638,386 | 0.00185875 | | 0.00176139 | |
| 80 | GNA11 COSM52969 (Q209L) | SNV_VAR | + | hg19 | chr19 | 3,118,915 | 3,118,979 | 0.00131865 | | 0.00088145 | |
| 81 | APC COSM13121 (Q1367*) | SNV_VAR | + | hg19 | chr05 | 112,175,387 | 112,175,462 | 0.00119234 | | 0.00208594 | |

Figure 81: SNV Variant Table

## What To Expect From Your SNV Positive Controls

Unlike the majority of NanoString's assays, SNV POS Controls are not linear. The actual POS control counts can vary widely (see Figure 82a). It is less important to have a specific count or be within the range, than it is to see the controls show the relationship seen in the graph in Figure 82b.

| | Typical Range |
|---|---|
| A | 17000 - 48000 |
| B | 1700 - 6600 |
| C | 550 - 1800 |
| D | 16000 - 48000 |
| E | 3000 - 16000 |
| F | 6000 - 20000 |

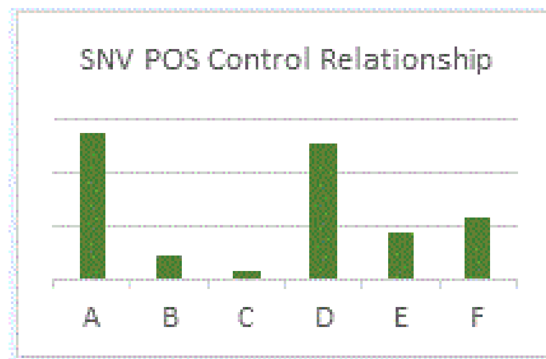



Figure 82a: SNV POS control typical range and 82b: Relationship and distribution.

## Potential SNV Cross-Hybridization Interactions

Due to the complex, competitive hybridizations that form the foundation of SNV chemistry, there are certain assays that, in order to ensure sensitivity down to 5% allele frequency, may also have affinity for other variant sequences in the assay. These interactions can result in false-positive calls among related probes assaying the same hotspot regions in the genome. Known potential variant cross-hybs are listed in the tables below. Exercise caution when analyzing data that shows positive results in these pairs of assays. The strongest call will likely be the assay listed in the "…When True Positive Present" column, and a weaker, secondary call may appear for the assay listed in the column "Putative False Positive…"

For example, in the Heme panel, when CSF1R COSM947 (Y969C) is present, you have a low chance that CSF1R COSM948 (Y969F) calls will be falsely elevated.

Table 1: Heme Panel Potential Hybridization Pairs

| Putative False Positive… | …When True Positive Present | Probability |
|---|---|---|
| CSF1R COSM948 (Y969F) | CSF1R COSM947 (Y969C) | Low |
| DNMT3A COSM52944 (R882H) | DNMT3A COSM99740 (R882P) | Low |
| FLT3 COSM27650 (D835A) | FLT3 COSM784 (D835V) | Low |
| IDH1 COSM28748 (R132S) | IDH1 COSM28749 (R132G) | Medium |
| IDH2 COSM41875 (R140L) | IDH2 COSM41590 (R140Q) | Medium |
| KIT COSM1310 (D816Y) | KIT COSM1311 (D816H) | High |
| KIT COSM1311 (D816H) | KIT COSM1310 (D816Y) | Medium |
| KRAS COSM512 (G12F) | KRAS COSM516 (G12C) | Medium |
| KRAS COSM512 (G12F) | KRAS COSM520 (G12V) | Medium |

Table 2: Solid Tumor Panel Potential Hybridization Pairs

| Putative False Positive… | …When True Positive Present | Probability |
|---|---|---|
| BRAF COSM473 (V600K) | BRAF COSM476 (V600E) | Low |
| BRAF COSM475 (V600E) | BRAF COSM476 (V600E) | Low |
| EGFR COSM12370 (L747_P753>S) | EGFR COSM12369 (L747_T751delLREAT) | High |
| EGFR COSM12370 (L747_P753>S) | EGFR COSM6255 (L747_S752delLREATS) | High |
| EGFR COSM12384 (E746_S752>V) | EGFR COSM12416 (E746_T751>VA) | High |
| EGFR COSM6223 (E746_A750delELREA) | EGFR COSM6225 (E746_A750delELREA) | High |
| EGFR COSM6255 (L747_S752delLREATS) | EGFR COSM12382 (L747_A750>P) | High |
| KRAS COSM549 (Q61K) | KRAS COSM550 (Q61E) | Low |
| KRAS COSM555 (Q61H) | KRAS COSM554 (Q61H) | Low |
| NRAS COSM585 (Q61H) | NRAS COSM586 (Q61H) | Low |

## CNV Data Analysis

The nature of the CNV assay and resulting data analysis differs slightly from that of other analytes.

**Importing RCC files** is done in the same manner as the standard workflow. Select **CNV** from the analyte drop-down box to filter datasets and display only those containing CNV data.

There are some **QC parameters** unique to the CNV assay. See the *Analyte-Specific QC* section.

**Exploring raw data** is done in the same manner as the standard workflow. The **Gender** column is a customizable drop-down menu (see implications of gender designation, below).

Creating an **Experiment** takes you through annotations, background, normalization, and ratios, similar to the standard workflow.

Find your experiment on the **Experiment tab** and expand the navigation tree (by clicking the **+** sign) to see the different levels of data tables. Select the data on which you would like to do your analysis. In addition to the standard workflow options, the following exist for CNV data:

- The **Normalized** and **Grouped** data levels have **Region Table** and **Region Export** buttons, which allow you to view and/or export data from specified regions.
- The **Ratio** data level has the Region Table and Region Export buttons and also has the **Probe Data** button, which gives you options to sort/filter your data by probe name.

For **autosomal probes**, the copy number estimates will be 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 will not be displayed for probes in the X and Y chromosome.

## Multi-RLF Experiments & Batch Calibration

In creating a New Multi-RLF Experiment, you can select a **CrossRLF/Batch Calibration** or a **MultiRLF Merge** experiment. You will need to have **RLF files imported** for any data you will be including in a Multi-RLF Experiment. You also must have run at least **one reference sample under each RLF** for calibration in a CrossRLF/ Batch calibration experiment

- If you ran a set of *identical samples* across *multiple CodeSets* and wish to consolidate the data, you will want to create a **MultiRLF Merge** experiment. nSolver will use the *Sample ID* to identify matched samples and will scale the normalized data of these identical samples based on the expression of the CodeSets' overlapping probes.
- To consolidate datasets of *non-identical samples* run across *multiple CodeSets,* you will want to create a **CrossRLF** experiment. At least one identical sample must be run across multiple CodeSet RLFs which is used to calibrate counts of overlapping probes between CodeSets.
- A **Batch Calibration** experiment is similar to a CrossRLF experiment and can be useful in calibrating across multiple lot numbers or instruments for samples run with the same RLF/CodeSet.

To create a new Multi-RLF experiment for any of the above purposes, select a **New MultiRLF** Experiment button on the main dashboard. Select the type of experiment you want to create (see above) and select **OK** (see Figure 83). Enter a unique experiment name and select the study with which your experiment should be associated from the dropdown menu. The Owner, Protocol, and Description fields are optional. Select **Next**. For next steps, see the *CrossRLF and Batch Calibration* section or the *MutiRLF Merge* section.

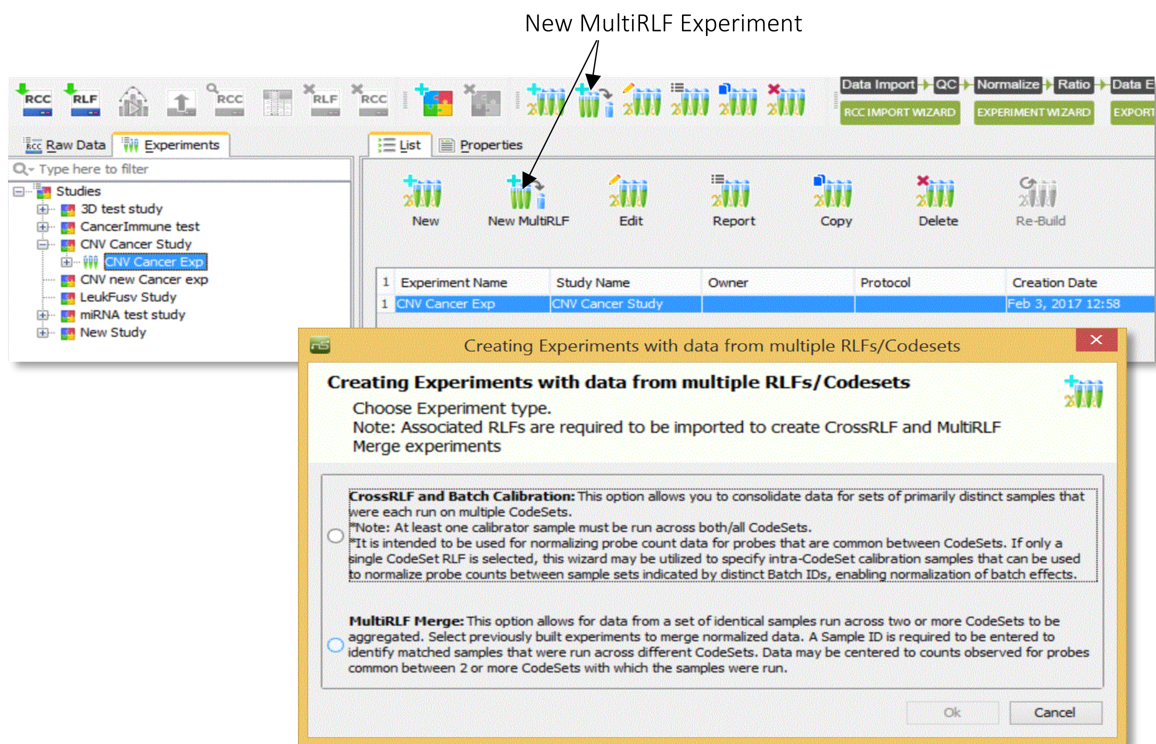



Figure 83: Creating a Multi-RLF Experiment

## CrossRLF and Batch Calibration

Once you have chosen a **CrossRLF** or **Batch Calibration**, you will be asked to add your samples to the experiment. On the *Add Sample/Lanes* screen, choose the raw data that you want to add, using control-click (or command-click) to select multiple files. There are several ways to search for the data you are looking for and select it for use in your experiment (see Figure 84).

- **Choose the analyte** you want to study from the drop-down menu on the left; only data of the chosen analyte type will be displayed.
- **Select the CodeSet** of interest from the navigation tree on the left. Use ctrl-click (or command-click) to select multiple CodeSets.
- **Use the filters** at the top of the main window to display only files of interest.
- **Select rows** in the main table viewer. To select all, click on the number in the upper left corner. You can also use the *Keep Selected* or *Exclude Selected* buttons to filter out any unwanted samples/lanes. The *Show Excluded* button displays all files once again.
- For a **Batch Calibration,** used the **Batch ID column** to document the lot number or instrument number you want to calibrate across.

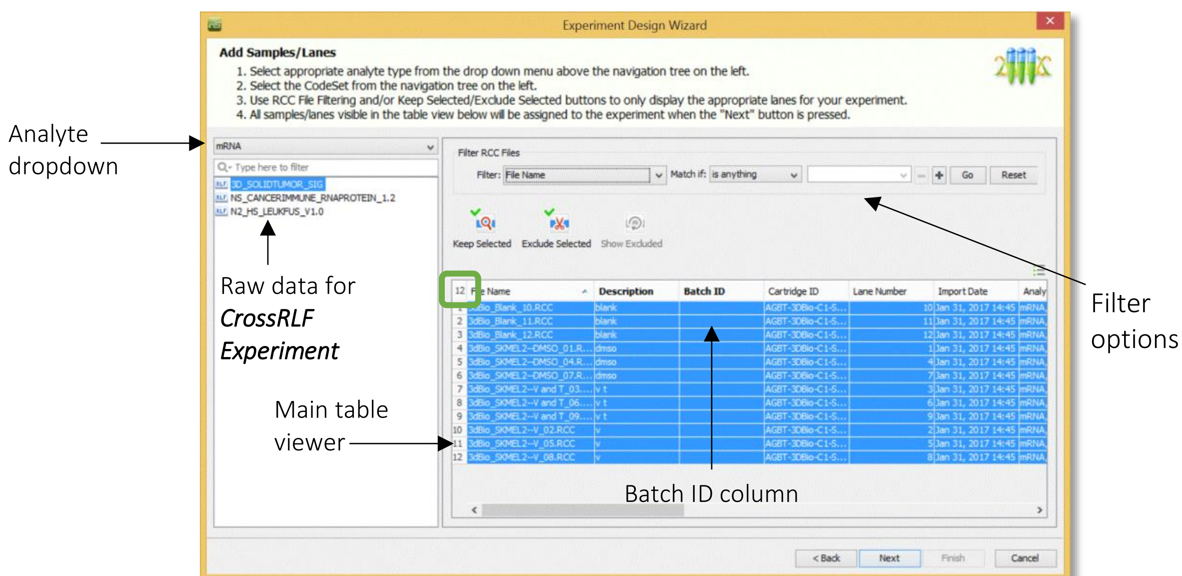



Figure 84: CrossRLF experiment creation

Once your samples of interest are displayed in the main table viewer, select **Next.** All samples displayed in the table will be included in the experiment.

The **Overlapping Probes** window (Figure 85) displays the probes which your selected RLFs have in common. CrossRLF experiments can be created even if a single probe overlaps, but it is advisable that most of the probes in all the CodeSets overlap. Select **Next.**

Figure 85: Overlapping probes window

Adding **annotations** can be done in the same fashion as the general workflow (see the *Annotations* section).

**Background** adjustment follows the general workflow (see the *Background Subtraction & Thresholding* section) with one caveat: if using Background Subtraction with a blank lane, you will need to have run a blank lane under each CodeSet; select each lane under the tabs provided in the window.

**Normalization** follows the general workflow (see the *Normalization* section).

Possible error messages

*Unsuitable Analyte Types*: CodeSets for cross-RLF experiment must belong to the same type and cannot be Sample Plex data.

*Selected less than two Experiments*: Select two or more experiments for merge.

*No RLF loaded*: MultiRLF Experiment requires RLF to be loaded. There are experiments with no RLF loaded.

For **calibration**, a reference sample should have been loaded and run under each CodeSet. Select the **Sample Reference Normalization** checkbox to activate the options in the window (see Figure 86). Select the lanes in which you loaded your reference sample in the *Subcode Samples* window (on the left). Use the arrows to move the desired lanes to the *Selected Samples* window (on the right). Select the tab for the next RLF, and repeat this process under this tab. Check the box for Use as Reference CodeSet to designate as the baseline CodeSet / Batch. Select **Next**.

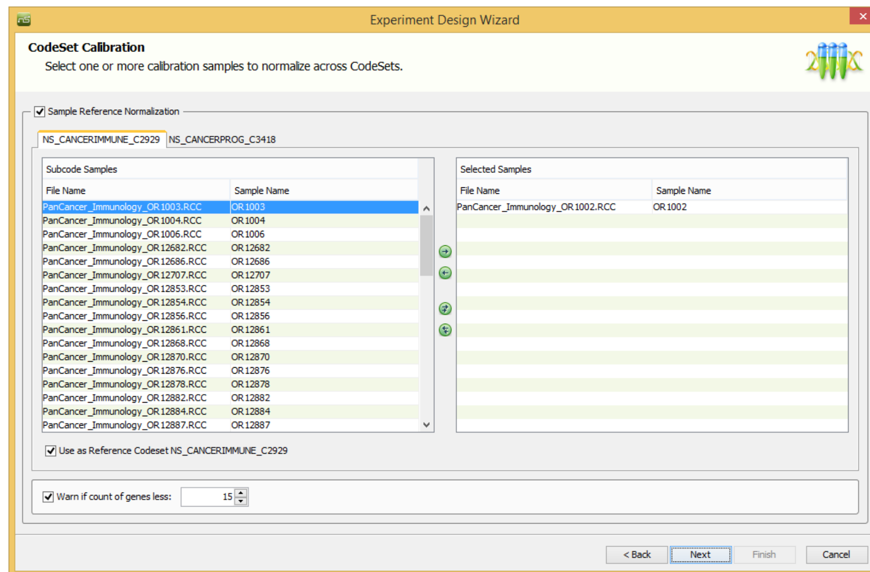
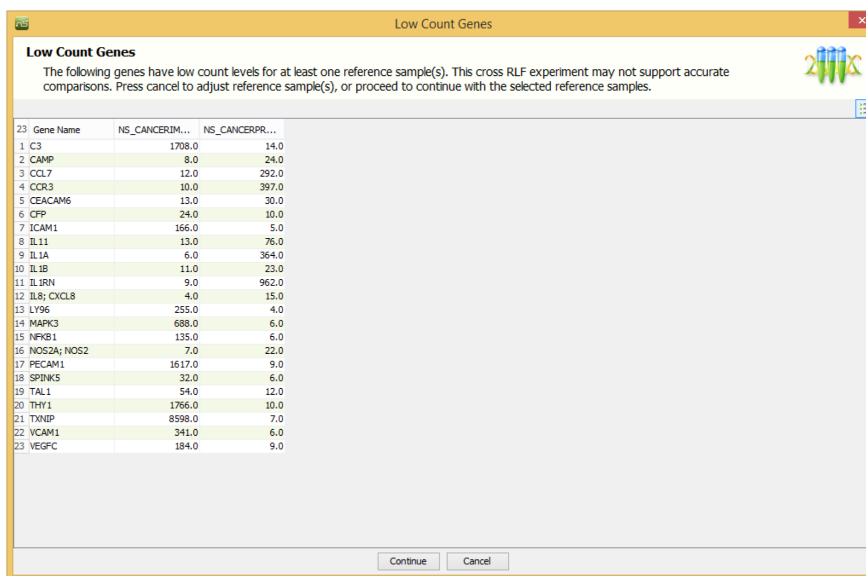



Figure 86: CodeSet Calibration for CrossRLF experiment

You will be alerted to any low gene counts (see Figure 87). Select **Continue**.




| 23 | Gene Name | NS_CANCERIM... | NS_CANCERPR... |
|---|---|---|---|
| 1 | C3 | 1708.0 | 14.0 |
| 2 | CAMP | 8.0 | 24.0 |
| 3 | CCL7 | 12.0 | 292.0 |
| 4 | CCR3 | 10.0 | 397.0 |
| 5 | CEACAM6 | 13.0 | 30.0 |
| 6 | CFP | 24.0 | 10.0 |
| 7 | ICAM1 | 166.0 | 5.0 |
| 8 | IL11 | 13.0 | 76.0 |
| 9 | IL1A | 6.0 | 364.0 |
| 10 | IL1B | 11.0 | 23.0 |
| 11 | IL1RN | 9.0 | 962.0 |
| 12 | IL8; CXCL8 | 4.0 | 15.0 |
| 13 | LY96 | 255.0 | 4.0 |
| 14 | MAPK3 | 688.0 | 6.0 |
| 15 | NFKB1 | 135.0 | 6.0 |
| 16 | NOS2A; NOS2 | 7.0 | 22.0 |
| 17 | PECAM1 | 1617.0 | 9.0 |
| 18 | SPINK5 | 32.0 | 6.0 |
| 19 | TAL1 | 54.0 | 12.0 |
| 20 | THY1 | 1766.0 | 10.0 |
| 21 | TXNIP | 8598.0 | 7.0 |
| 22 | VCAM1 | 341.0 | 6.0 |
| 23 | VEGFC | 184.0 | 9.0 |

Figure 87: Low count gene window

You will be prompted to establish baseline data for ratios, then asked to assign **ratio** data names as in the standard workflow. Select **Next**, then **Finish**.

Data viewing and analysis follow the same guidelines as standard sample analysis.

## MultiRLF Merge

Once you have chosen a **MultiRLF Merge**, you will be asked to add your experiments to the merge. There are several ways to search for the data you are looking for and select it for use in your experiment (see Figure 88).

- You may filter by analyte type using the **Select Analyte Type** drop down list
- Choose the **existing experiments** that you want to add, using control-click (or command-click) to select multiple files.
- **Use the filters** at the top of the main window to display onlyfiles of interest.
- **Select rows** in the main table viewer. To select all, click on the number in the upper left corner. You can also use the *Keep Selected* or *Exclude Selected* buttons to filter out any unwanted samples/lanes. The *Show Excluded* button displays all files once again.

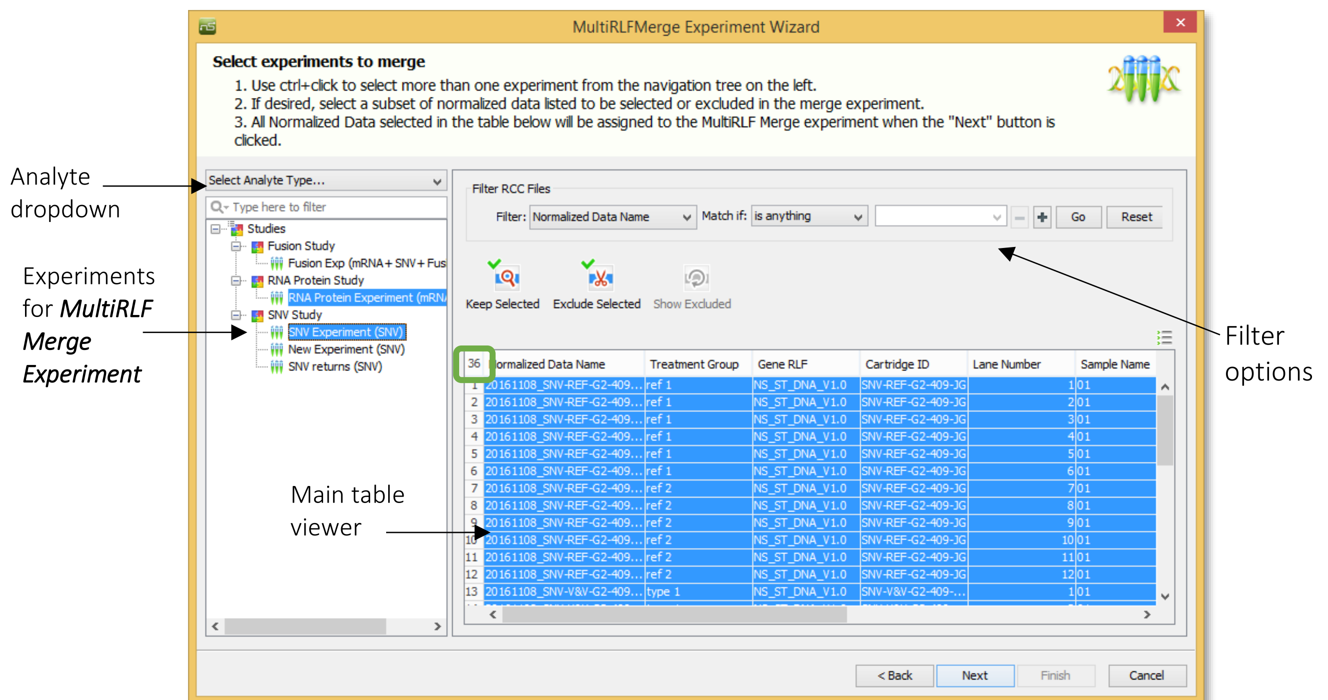



Figure 88: MultiRLF Merge experiment setup

Once your samples of interest are displayed in the main table viewer, select **Next.** All samples displayed in the table will be included in the experiment.

The **Overlapping Probes** window (Figure 89) displays the probes which your selected experiments have in common. MultiRLF Merge experiments can be created even if no probes overlap between the selected experiments. A minimum of 5 overlapping probes, however, is required for merged data to be on a comparable scale in the merged experiment. Select **Next**.

MultiRLFMerge Experiment Wizard

**Overlapping Probes**

Information for probes overlapping all experiments

| 178 | Probe Name | Gene Name | Annotation of N... | Annotation of N... | Avg. Count Of ... | Avg. Count Of ... |
|---|---|---|---|---|---|---|
| 1 | SOX17 | SOX17 | | | 50.13 | 28.28 |
| 2 | MTOR | MTOR | | | 376.72 | 254.04 |
| 3 | FIGF | FIGF | | | 42.47 | 45.18 |
| 4 | TP53 | TP53 | | | 417.65 | 600.11 |
| 5 | HSP90B1 | HSP90B1 | | | 1,573.95 | 994.46 |
| 6 | NR4A3 | NR4A3 | | | 107.18 | 120.11 |
| 7 | COL4A6 | COL4A6 | | | 36.78 | 55.77 |
| 8 | HGF | HGF | | | 132.17 | 169.04 |
| 9 | TNFSF10 | TNFSF10 | | | 1,866.27 | 2,097.35 |
| 10 | CDKN2A | CDKN2A | | | 38.15 | 54.4 |
| 11 | VEGFC | VEGFC | | | 50.9 | 124.75 |
| 12 | TGFBR2 | TGFBR2 | | | 557.4 | 659.37 |
| 13 | SMAD4 | SMAD4 | | | 598.03 | 642.04 |
| 14 | ITGB7 | ITGB7 | | | 75.65 | 78.32 |
| 15 | FGFR2 | FGFR2 | | | 753.53 | 335.75 |
| 16 | WNT5B | WNT5B | | | 59 | 106.91 |
| 17 | IFNG | IFNG | | | 19.08 | 18.09 |
| 18 | AKT1 | AKT1 | | | 2,547.45 | 2,453.81 |
| 19 | IL15 | IL15 | | | 41.12 | 36.28 |
| 20 | EGFR | EGFR | | | 291.28 | 313.74 |
| 21 | LAMA1 | LAMA1 | | | 27.3 | 318.75 |
| 22 | RB1 | RB1 | | | 382.67 | 355.51 |
| 23 | RBX1 | RBX1 | | | 1,186.73 | 1,088.33 |
| 24 | IL13RA2 | IL13RA2 | | | 17.03 | 17.96 |
| 25 | TFDP1 | TFDP1 | | | 588.57 | 886.46 |
| 26 | THBS1 | THBS1 | | | 2,390.88 | 3,592.32 |
| 27 | LAMA3 | LAMA3 | | | 76.77 | 49.58 |
| 28 | SOX9 | SOX9 | | | 1,042.68 | 1,084.93 |

< Back Next Finish Cancel

Figure 89: Overlapping probes window

Follow the prompts in the **Merge Normalized Data** window (see Figure 90). Note that data rows with identical Sample IDs will be merged. **The Fill Sample ID** button can be used if you would like to use the Sample Name for ID. The **View Result** button is available to check names and merged combos. Select the checkbox for **Use Scaling** to scale all merged data to the geometric mean of overlapping probes between the CodeSets. Select **Next**.

MultiRLFMerge Experiment Wizard

**Merge Normalized Data**

1. Assign Sample ID to each normalized sample.
2. Data rows with identical Sample ID will be merged. Data will be centered to probes that are common between two CodeSets if a minimum of 5 common probes exists and counts for these probes will be averaged. Scaling shall not be performed when CodeSets do not have at least 5 overlapping probes and caution should be used in interpreting merged data.
3. Click "View Result" to inspect Normalized Data associated with each Sample ID.
4. Click "Next" to continue.

☑ **Use scaling** Fill in Sample ID with Sample Name [Fill Sample ID]

| 117 | Normalized Dat... | Treatme... | Gene RLF | Cartridge... | Lane Nu... | Sample N... | Description | Batch ID | % Probe... | QC Flag | **Sample ...** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | PanCancer_Path... | Normal | NS_CANC... | 20140320... | 10 | OR1002 | Normal | | 89.351 | NO | |
| 2 | PanCancer_Path... | Normal | NS_CANC... | 20140320... | 11 | OR1003 | Normal | | 87.273 | NO | |
| 3 | PanCancer_Path... | Normal | NS_CANC... | 20140320... | 12 | OR1004 | Normal | | 85.065 | NO | |
| 4 | PanCancer_Path... | Normal | NS_CANC... | 20140326... | 10 | OR1006 | Normal | | 86.364 | NO | |
| 5 | PanCancer_Path... | B | NS_CANC... | 20140319... | 1 | OR12682 | B | | 83.766 | NO | |
| 6 | PanCancer_Path... | D | NS_CANC... | 20140319... | 12 | OR12686 | D | | 81.039 | NO | |
| 7 | PanCancer_Path... | D | NS_CANC... | 20140319... | 5 | OR12707 | D | | 81.948 | NO | |
| 8 | PanCancer_Path... | A | NS_CANC... | 20140320... | 8 | OR12853 | A | | 85.455 | NO | |
| 9 | PanCancer_Path... | C | NS_CANC... | 20140325... | 10 | OR12854 | C | | 84.545 | NO | |
| 10 | PanCancer_Path... | D | NS_CANC... | 20140319... | 11 | OR12856 | D | | 82.208 | NO | |
| 11 | PanCancer_Path... | B | NS_CANC... | 20140319... | 10 | OR12861 | B | | 85.195 | NO | |
| 12 | PanCancer_Path... | B | NS_CANC... | 20140320... | 2 | OR12868 | B | | 83.247 | NO | |
| 13 | PanCancer_Path... | B | NS_CANC... | 20140320... | 9 | OR12870 | B | | 83.506 | NO | |
| 14 | PanCancer_Path... | B | NS_CANC... | 20140320... | 5 | OR12876 | B | | 82.727 | NO | |
| 15 | PanCancer_Path... | B | NS_CANC... | 20140320... | 7 | OR12878 | B | | 84.026 | NO | |
| 16 | PanCancer_Path... | B | NS_CANC... | 20140320... | 1 | OR12882 | B | | 79.74 | NO | |
| 17 | PanCancer_Path... | A | NS_CANC... | 20140320... | 11 | OR12884 | A | | 84.675 | NO | |
| 18 | PanCancer_Path... | B | NS_CANC... | 20140320... | 6 | OR12887 | B | | 80.39 | NO | |

View result of merging [View Result]

< Back Next Finish Cancel

Figure 90: Merge normalized data

Select the columns to be displayed in the merged data table (see Figure 91) and select **Next.**

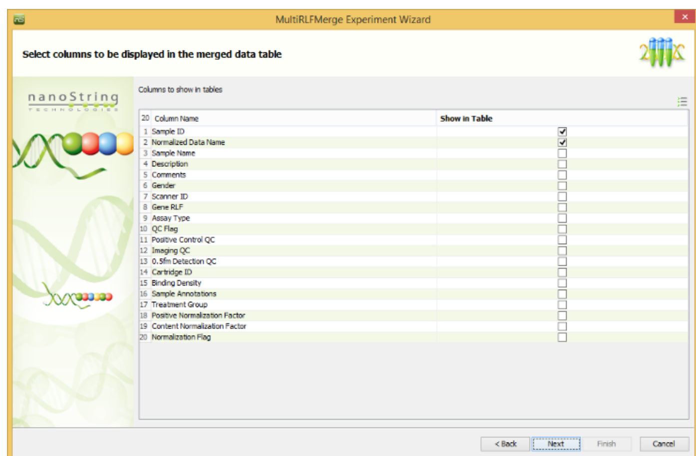



Figure 91: Columns to be displayed

Possible error messages

*Unsuitable Analyte Types*: CodeSets for cross-RLF experiment must belong to the same type and cannot be Sample Plex data.

*Selected less than two Experiments*: Select two or more experiments for merge.

*No RLF loaded*: MultiRLF Experiment requires RLF to be loaded. There are experiments with no RLF loaded.

Adding **annotations** can be done in the same fashion as the general workflow (see the *Annotations* section).

Because **Background** adjustment and **Normalization** were performed when the original experiments were created, those steps are skipped in the MultiRLF Merge. If 5 or more overlapping probes are present and **Use Scaling** was selected, each sample's ratio of the geometric mean of the overlapping probes between the first and second Codesets will be multiplied to the normalized data to get scaled values.

You will be prompted to establish baseline data for ratios, then asked to assign **ratio** data names per the standard workflow. Select **Next**, then **Finish**.

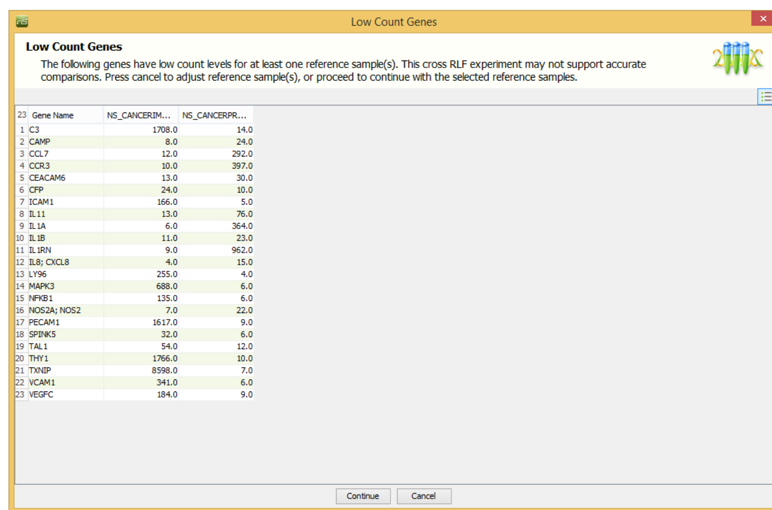



Figure 92: Low count genes window

You will be alerted to any low gene counts (Figure 92). Select **Continue**.

Data viewing and analysis follow the same guidelines as standard sample analysis.

# Appendix A: 3D Bio Data Example

The dataset, **3D Bio Data**, is included when you download the nSolver 4.0 Analysis Software. This data is a result of three biological replicates of two different melanoma cell lines, SKMEL28, which has a known mutation (c.1799T>A; p.V600E) in both copies of the BRAF gene, and SKMEL2, which has two normal copies of the BRAF gene. Both cell lines were treated with either DMSO (vehicle)or vemurafenib (a specific inhibitor of the V600E mutant BRAF protein) dissolved in DMSO for 8 hours.

Throughout the nSolver 4.0 User Manual, you will find excerpts of this dataset's analysis.

## Import

To import the files contained in the data folder, **3D Bio Data:**

1. Unzip the file by right-clicking or command-clicking and selecting **Extract All** (or by running your preferred file extraction program). Once the extraction is complete, you should be able to see one RCC file for every cartridge lane.

   *There should be two RLFs and 24 RCC files (12 samples and 12 SNV references).*

2. Open **nSolver 4.0** and select one of the **Import RLF** buttons. **Browse** to navigate to the folder with the unzipped data, highlight the **3D_SolidTumor_Sig.rlf** file, and select **Open**, then **Import.** Repeat this step to import the other RLF in the file, **NS_ST_DNA_v1.1.rlf**.

3. Select one of the **Import RCC Files** buttons. **Browse** to navigate to the folder with the unzipped data, highlight all RCC samples, and select **Open**.

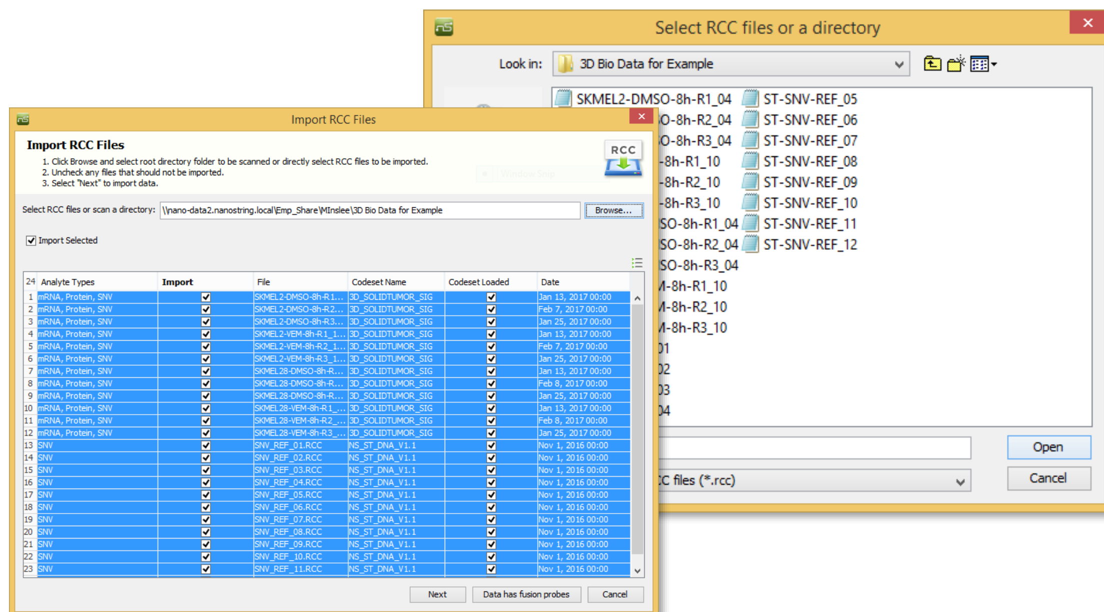



Figure 93: 3D Bio Data import screens

4. Note the following columns, then select **Next**.
   - The **Analyte Types** column lists mRNA, Protein, and SNV for the samples (and only SNV for the SNV reference samples).
   - All boxes in the **Import** column are checked, indicating that you want to import each sample.
   - The **CodeSet Name** column; your data will ultimately be stored under this name on the nSolver raw data tab.
   - The boxes in the **CodeSet Loaded** column are checked, indicating that you have imported the RLF.

## QC

In choosing the QC parameters, note the following:

- You must use the **double arrow** icon in the right corner of the screen to reveal the top three (System QC) parameters.
- The activated buttons in the panel of **analytes** along the left side of the window represent the analytes detected in your data (mRNA, SNV, and Protein). Selecting an analyte reveals the default QC parameters associated with it.
- You may change the QC parameters, but this is not usually recommended nor necessary.

Select **Import**.

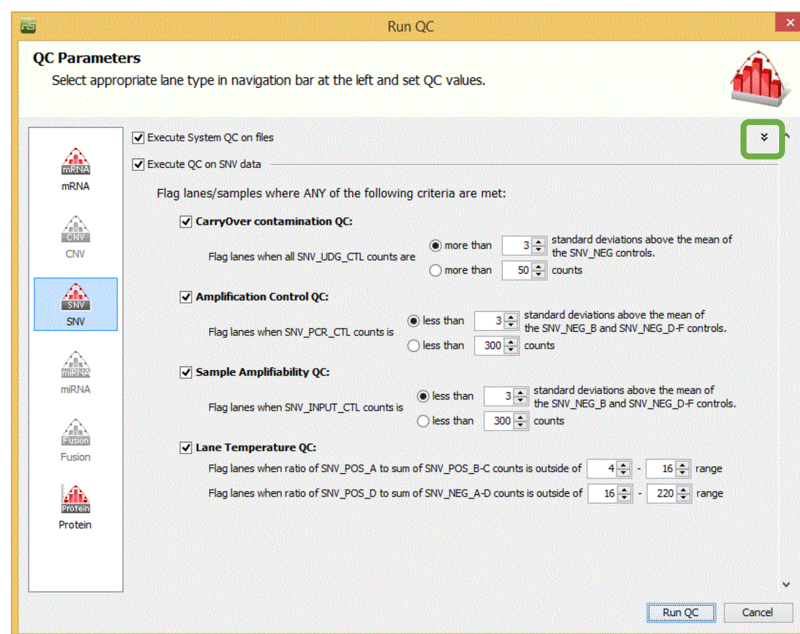



Figure 94: 3D Bio Data Example – QC

## Explore Raw Data

Once you have completed the import process, your RCC data files will be visible under the corresponding RLF CodeSets on the **Raw Data tab**: samples under **3D_SOLIDTUMOR_SIG** and SNV references under **NS_ST_DNA_V1.1**. Here, we explore the raw data and assess any QC flags.

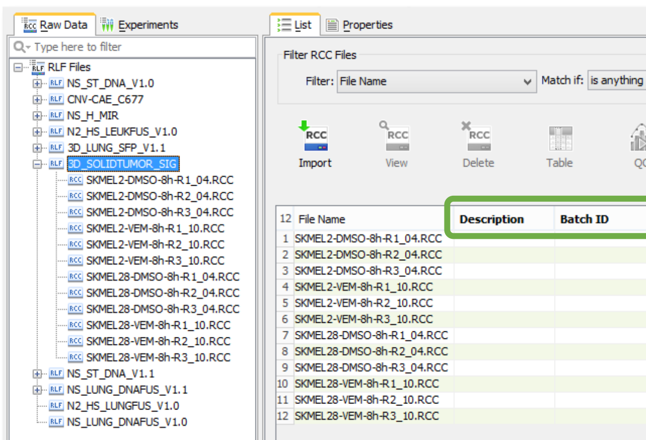



Figure 95: 3D Bio Data Example - exploring raw data

1. Selecting the CodeSet name, **3D_SOLIDTUMOR_SIG,** allows you to view all RCC files under it in a table format. You may edit fields in the Description and Batch ID columns, if desired.

2. As you scroll right in this table, notice that the first QC column is labeled simply **QC Flag**. A flag in this column indicates that there is a QC flag of some sort associated with this sample. Since there are no flags in the QC Flag column, we know that there are no QC flags associated with this data, but we will still scroll farther to the right and note the individual QC columns.

3. Select the **Column Options** icon in the upper right corner and select **Show All Hidden Columns** (this option will not be available if all columns are already being shown). This will reveal the QC columns with numerical data. Note that the values of the QC metrics vary only slightly from sample to sample. This is one indication of good data.

| QC Flag | Imaging QC Flag | Imaging QC | Binding Density ... | Binding De... | Positive Control ... | Positive Co... | Limit of Detectio... | Limit of Detectio... | CarryOver Cont... | CarryOver... | Amplification Co... | Amplificati... |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0.96 | | 0.8 | | 0.98 | | 12.07 | | 252.45 | | 34.84 |
| | | 1 | | 1.03 | | 0.98 | | 11.19 | | 366.74 | | 25.07 |
| | | 1 | | 0.66 | | 0.98 | | 7 | | 299.41 | | 21.16 |
| | | 0.99 | | 0.8 | | 0.98 | | 9.5 | | 273.07 | | 24.96 |
| | | 0.99 | | 0.77 | | 0.98 | | 10.26 | | 221.98 | | 22.22 |
| | | 1 | | 0.56 | | 0.98 | | 6.83 | | 162.9 | | 24.51 |
| | | 0.99 | | 1 | | 0.98 | | 16.1 | | 376.76 | | 38.38 |
| | | 1 | | 0.98 | | 0.98 | | 16.36 | | 276.84 | | 19 |
| | | 0.98 | | 0.72 | | 0.98 | | 9.11 | | 352.96 | | 17.25 |
| | | 1 | | 0.96 | | 0.99 | | 11.83 | | 373.14 | | 48.35 |
| | | 1 | | 0.87 | | 0.98 | | 10.98 | | 304.11 | | 27.68 |
| | | 1 | | 0.64 | | 0.98 | | 11.56 | | 291.76 | | 22.51 |

Figure 96: 3D Bio Data Example - examining QC flags

4. Regardless of whether you have QC flags in your data, you should review the results from your positive and negative controls. **Highlight all samples** and select the **Table** button. In the **Probe Name** column, look for **positive and negative controls.** You can use the **Filter Expression Data** tool above the table to search, as well. Note that the expression levels for the negative and positive control probes is relatively stable across samples. If one of your samples had a QC flag associated with it, its column heading would be red. For details and troubleshooting tips for specific flags, see the QC section of the nSolver 4.0 User Manual.

## Creating Studies & Experiments

To create an experiment with the example dataset, **3D Bio Data**, follow these steps.

1. Select the **Experiment** tab.

2. Select a **New Study** button on the main dashboard and enter a unique study name. The Owner, Protocol, and Description fields are optional. Select **Save**.

3. Find your new study on the **Experiments tab** and highlight it.

4. Select a **New Experiment** button on the main dashboard and enter a unique experiment name. The study you had highlighted should be listed as the default, but if not, you can select it from the dropdown menu. Select **Next**.

5. Select the **3D_SOLIDTUMOR_SIG** CodeSet from the list on the left, then select the samples to include in or exclude from the experiment. For this example, we will use all samples in the dataset, but if you needed to filter samples in a larger dataset you could:
    - Use the filter tool for a **File Name** that **contains 8h**, then select **Go** (see figure below).
    - Highlight the desired samples and select **Keep Selected**.
    - Highlight the samples you don't want to keep and select **Exclude Selected**.

6. Select **Next.**

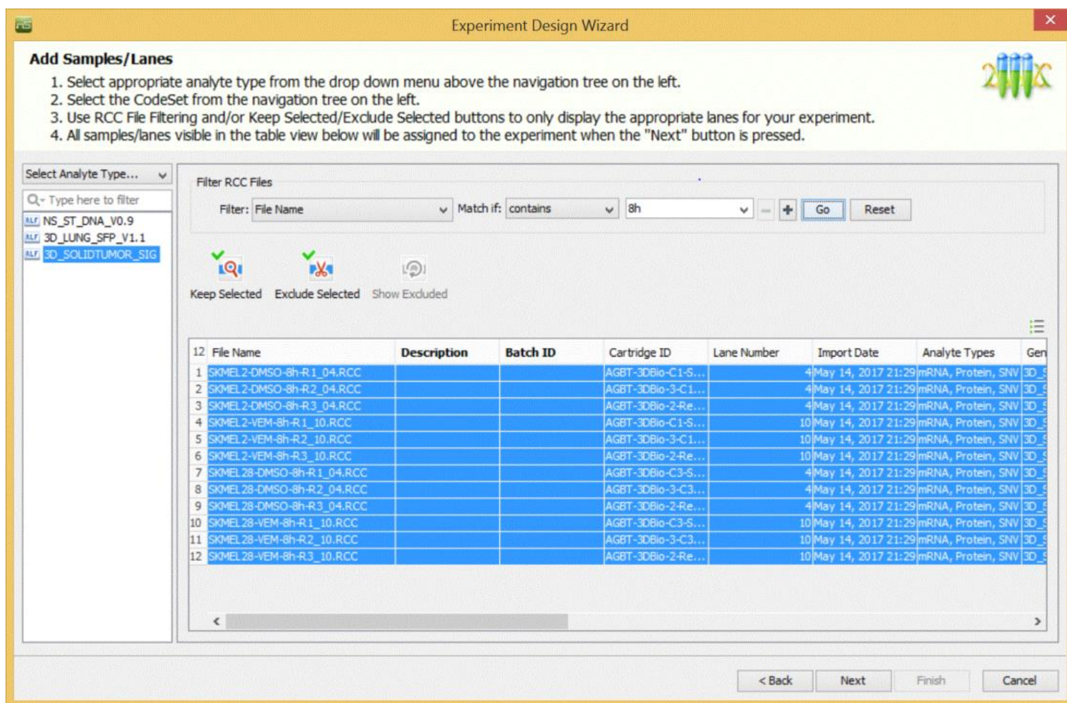



Figure 97: 3D Bio Data Example - creating an experiment

## Annotations

In the example dataset, **3D Bio Data**, the treatment type, Sample #, and time for each sample has been incorporated in the sample name. These are good categories to use for sorting and analyzing data, but it is only by adding this information as a separate annotation that it can be utilized as a variable for differential analysis.

When we created an experiment using this data, we included samples from two cell types, SKMEL2 (BRAF WT) and SKMEL28 (BRAF mutant). Now, we will add annotations to separate these groups.

1. Create two annotation categories by selecting **Add Annotation** twice.

2. Click in the fields below **Column Name** and change **New Annotation** to **Treatment**, and **New Annotation 2** to **BRAF Genotype.**

3. Under **Column Type**, use **Text.** See the *Annotations* section for information on other column types.

4. The column heading will change dynamically to reflect the new column name. Add the specific annotations (or copy and paste from another source under the new column (**DMSO** or **VEM** for **Treatment**, and **WT/WT** (SKMEL2) or **Mut/Mut** (SKMEL28) for **BRAF Genotype**, according to what is documented in the sample names).

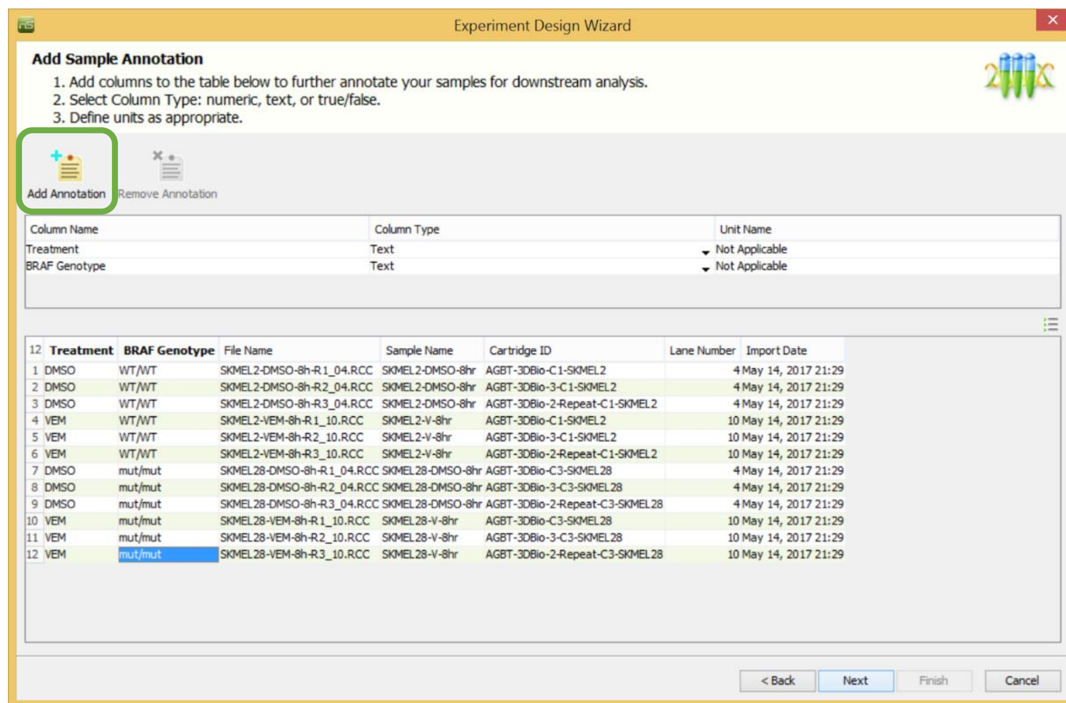



Figure 98: 3D Bio Data Example - creating annotations

## Background

**Background** noise in data can be filtered out using subtraction or thresholding; this is optional. By default, background correction will be checked off and in most cases, you will not need to deviate from this. If performing background correction on your data, thresholding is recommended for most analyte types.

For the example dataset, **3D Bio Data**, leave the **Background Subtraction/Thresholding** box unselected, which leaves background correction off.

Figure 99: 3D Bio Data – Background window

## Normalization

The recommended normalization settings for most analyte types will appear by default. **Normalization** of the example dataset, **3D Bio Data**, can be accomplished by using the Positive Control counts and by using selected normalization gene counts. Use the tabs to review the settings for the different analytes, and adjust them, as noted below. Due to the nature of the calling algorithm and the complexity of **SNV** data, its normalization is hard-coded.

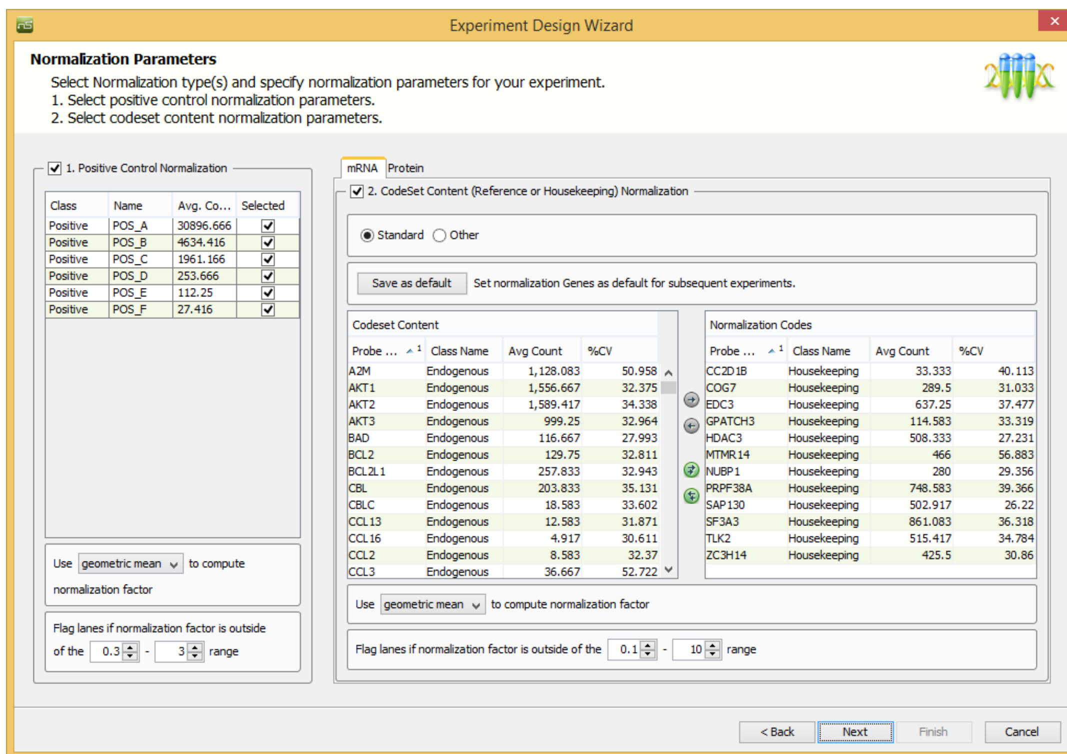



Figure 100: 3D Bio Data Example - normalization parameters

In the **Positive Control Normalization** field, note the following:

- There are no analyte tabs in this field since mRNA and Protein share a single set of Positive Control Normalization settings. Maintain these default settings, using the **geometric mean** with the default flagging range.

In the **CodeSet Content Normalization** field, note the following:

- The default **mRNA CodeSet Normalization** settings include all Housekeeping Genes labeled in the CodeSet.
- The default **Protein CodeSet Normalization settings** include the **histones** labeled in the CodeSet. Alternatively, you may select **All** or click on the %CV column heading and remove the highest %CV samples, but it is not necessary for the purposes of this example.

Select **Next**.

## SNV Probe Calibration

The next step in the analysis of the example dataset, **3D Bio Data**, is SNV Probe Calibration. This step is specific to the SNV assay and requires at least 10 reference SNV samples.

For the purposes of this example, we have included 12 SNV reference samples from the NS_ST_DNA_v1.1 CodeSet. If two of those samples were low-quality and we decided to drop them, we would still be able to run the analysis.

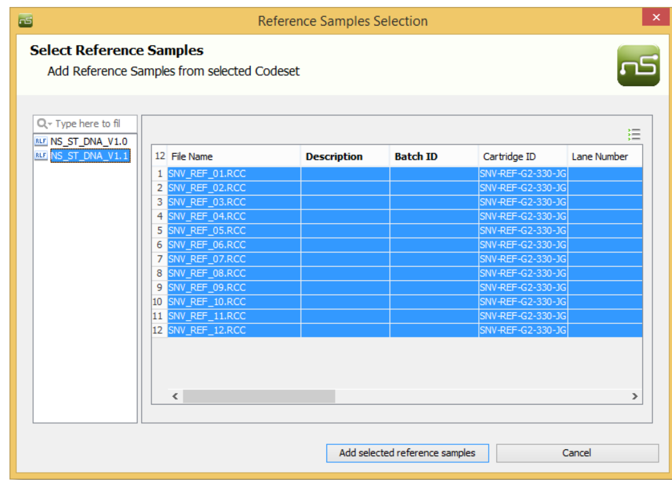



Figure 101: 3D Bio Data Example - SNV reference selection

1. Use the **Select RCC files from a different RLF** button since there are no SNV reference samples under the present RLF.

2. Highlight the **NS_ST_DNA_v1.1** CodeSet. Highlight all SNV reference samples and select the **Add selected reference samples** button.

3. You should see your SNV reference samples in the Selected SNV Reference Samples window. Select **Next**.

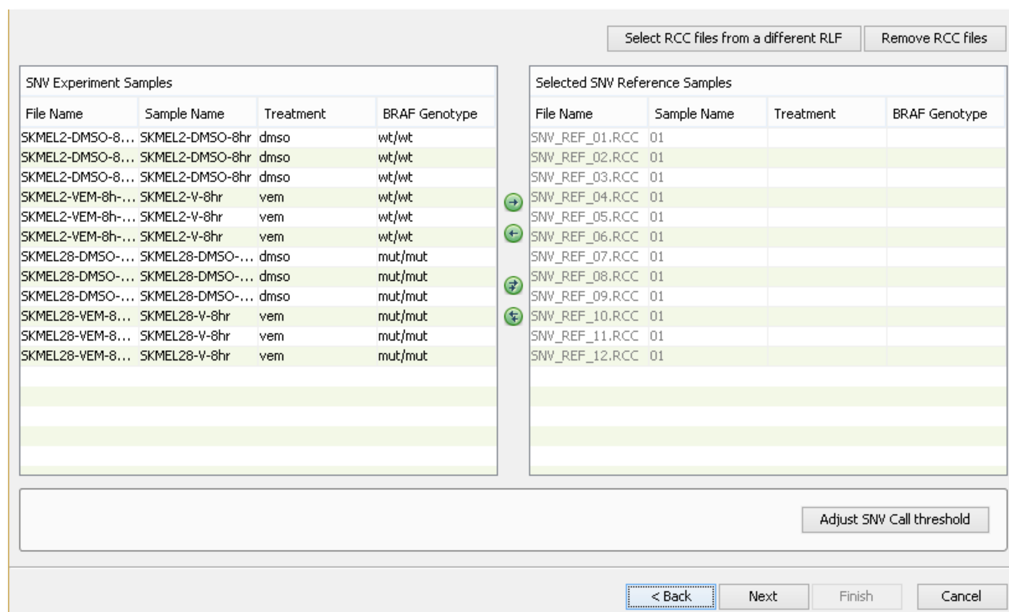



Figure 102: 3D Bio Data Example - SNV reference selection

## Fold Changes (Ratios)

In creating ratios for the example dataset, **3D Bio Data**, select **Partitioning by.** It will default to one of the annotations we entered earlier, Treatment, and will choose a treatment type as the reference, in this case, DMSO. For this example, we will keep these defaults, but they can be changed using the dropdown menus, if desired.

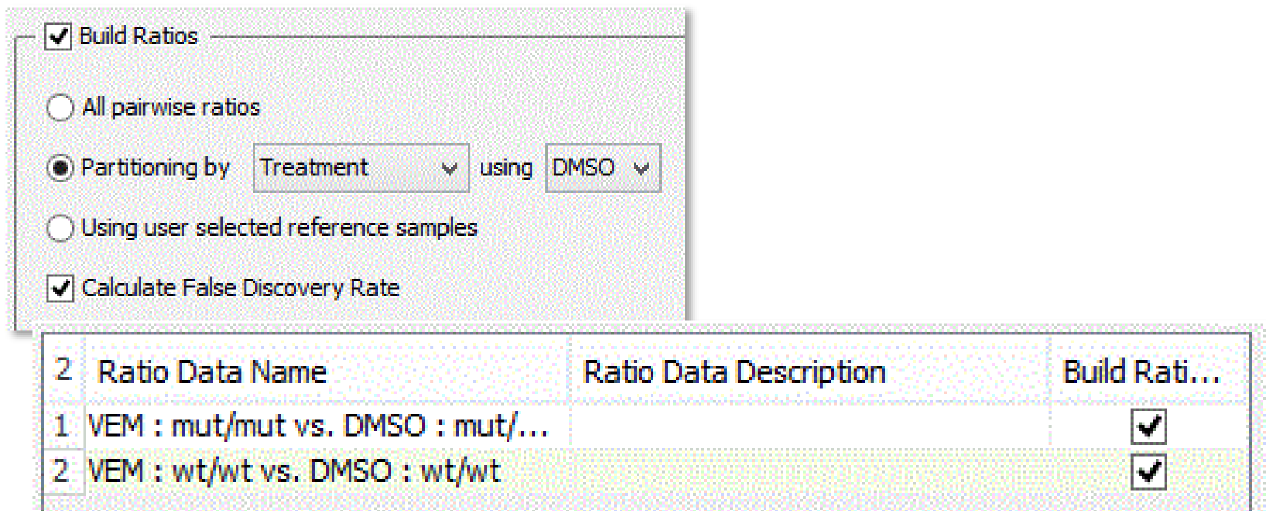



Figure 103: 3D Bio Data Example - ratios windows

This experimental design results in two ratios to confirm, partitioning by treatment within each cell type. You can use the checkboxes along the right side of the window to confirm or cancel building the ratio. You can change the name of this comparison, as well, if desired. Select **Finish**.

## Analysis

Your experiment for the example dataset, **3D Bio Data**, will now be stored under the corresponding study on the **Experiments** tab. Expanding the navigation tree next to it (clicking on the **+** signs) allows you to view all levels of data in the experiment. Here, we will discuss what level of data to use for different analyses, what analysis tool to use for each analyte type, and how to look at your 3D data holistically.

### Levels of Data

- The **Raw Data** is useful for checking the overall quality of the data and in determining how the data is clustering and what experimental variables impact this. We will use this level of data to create a visualization, below.
- The **Normalized Data** can also be used for QC purposes, and should be used for multi-RLF experiments.
- The **Grouped Data** can further assess clustering within the annotations you chose to create earlier.
- The **Ratio Data** has several columns which refer to the significance of the data within the ratios you chose to create earlier.

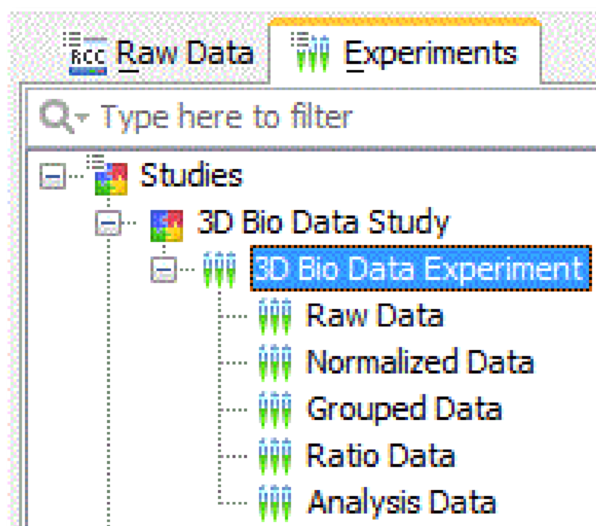



Figure 104: 3D Bio Data Example – data levels

In addition to visualizing the data in nSolver, you may export it as a .csv file. This file will open in Excel and can be imported into any common data analysis software.

### How to use the SNV variant table

When the **Normalized Data** level is selected, the **Variant Table** button becomes active. Select the normalized samples, then select this button. In the resulting table, scroll to the right so that you can see all the sample data columns. Then, scroll down. You should see the occasional **green flag** in some of the **Variant Call** columns, indicating a variant is present in your data at that gene. You can also click on each Variant Call column header to sort flagged probes on top.

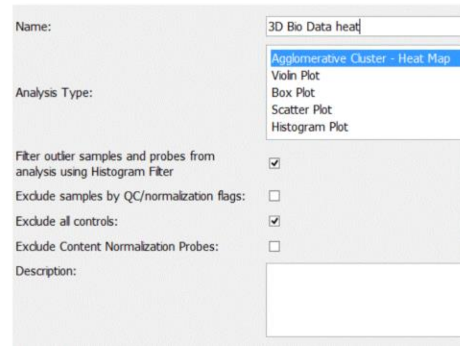
| 104 | Probe Name | Class Name | Variant Call: SKMEL2-DMSO-8h-R1_04.RCC | Variant Call: SKMEL2-DMSO-8h-R2_04.RCC | Variant Call: SKMEL28-VEM-8h-R3_10.RCC |
|---|---|---|---|---|---|
| 1 | BRAF COSM476 (V600E) | SNV_VAR | | | |
| 2 | TP53 COSM44571 (L194R) | SNV_VAR | | | |
| 3 | FBXW7 COSM22932 (R465C) | SNV_VAR | | | |
| 4 | EGFR COSM6239 (G719A) | SNV_VAR | | | |
| 5 | JAK2 COSM333722 (G180A) | SNV_VAR | | | |
| 6 | APC COSM26697 (I1307K) | SNV_VAR | | | |
| 7 | ERBB2 COSM14060 (L755S) | SNV_VAR | | | |
| 8 | KRAS COSM520 (G12V) | SNV_VAR | | | |

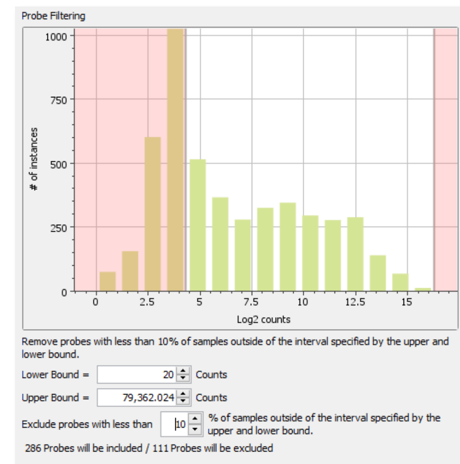
Figure 105: 3D Bio Data Example - SNV variant table

## How to look at mRNA and protein data

There are many ways to view mRNA and protein data, including violin plots and box plots. Here, we will create a heat map to see how the data is clustering.

1. After selecting the **Normalized data** under the experiment name, select all samples.
2. Select the **Analysis** button, which will launch the Analysis Wizard. Specify the **Name** for the analysis.
3. Under **Analysis Type**, select **Agglomerative Clustering**, and the **Histogram Filter** and **Exclude all controls** checkboxes. Check off the others. Select **Next**.
4. In histogram filter window, we will filter out all probes with counts 20 and lower in the **Probe Filtering** plot (we will leave the Sample Filter plot alone). Either drag the pink filter from the left or type 20 in the **lower bounds** box under the Probe Filtering histogram. Select **Next**.
5. All 12 samples should already be in the Selected Samples window. Select **Next**.
6. Most **Endogenous** genes and **Proteins** should already be in the **Selected** window (unless they fell below 20 counts), and all control genes and SNV probes should be in the **Excluded** window. Select **Next.**
7. Most heatmap settings can be left in the default mode. Under **Select sample annotation**, each annotation (**Treatment** and **BRAF Genotype**) box should be checked. Under **Select probe annotations**, **Gene Name** and **Protein Name** should be checked. Select **Finish.**



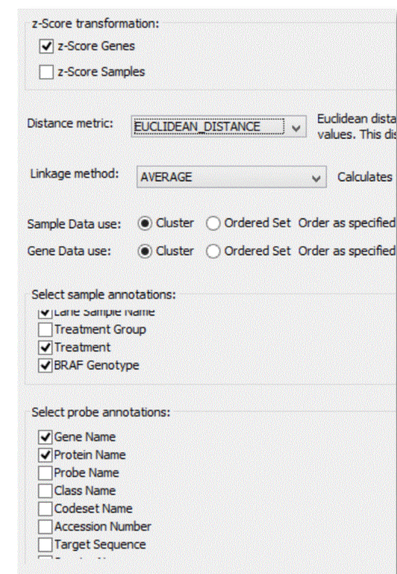




Figure 106: 3D Bio Data Example - windows associated with analysis



nanoString™

Figure 107: 3D Bio Data Example - heatmap

8. A java window will appear with the heatmap. Adjust the label settings, if needed. Use the slide bar to view the names of samples (columns, top). Select **View**, then **Labels**, and change **Rows** to **NAME** to view gene names (rows, left side). Adjust label font and size and use the **Keep fixed** checkbox to fix the column names in these settings.

   Note that:

   - The samples have clustered by cell type (SKMEL2 on the left, SKMEL28 on the right).
   - The SKMEL28 samples have clustered by treatment (DMSO on the left, VEM on the right).
   - The area with increased expression (green bars) in SKMEL28 samples and decreased expression (red bars) of SKMEL2 samples (see yellow selected box in the figure above) corresponds with one of the genes in the MAP-K pathway (MAPK1). This is what we would expect from a BRAF-mutated sample.

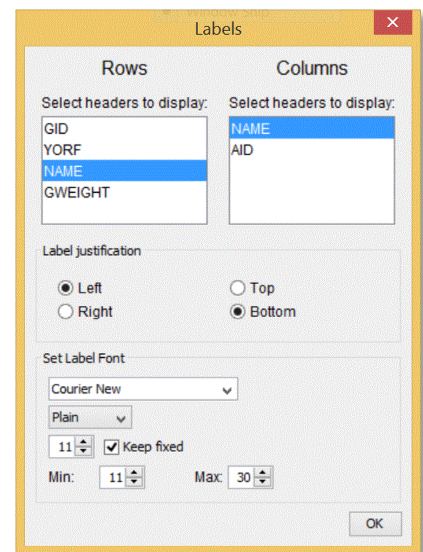



Figure 108: 3D Bio Data Example – heatmap label settings

# Glossary

This section defines terminology associated with the nSolver 4.0 Analysis Software System.

**Accession:** A unique identifier assigned to a particular gene or protein sequence. A text field associating the gene accession information with each nCounter Reporter Probe pair is included in some of the table outputs.

**Analyte Type:** A categorization of molecular variants that can be detected on nCounter platforms and analyzed with nSolver 4.0 Analysis Software System. The following analyte types are currently supported by NanoString platforms and software (as of this document's release): mRNA, miRNA, CNV, SNV, Fusion and Protein. A sample can contain multiple analyte types.

**Annotation**: A type of notation that can be used to establish groups of samples or probes.

**Assay Type**: A categorization of the nSolver Analysis procedures, each one designed to detect a specific type of molecule or perform a specific type of analysis.

**Assay:** A pre-designed laboratory test, usually focused on accomplishing a specific task. Each nCounter assay is capable of simultaneous, single-tube, multiplexed detection of up to 800 targets. Except for Plex$^2$ and PlexSet, there is typically one assay per lane.

**Background**: The inherent "noise" produced during testing. Background subtraction or thresholding calculations can be performed by nSolver to minimize the impact of the background on the data counts arising from actual variants.

**Batch ID**: A text field that can be used to document the lot number of a kit or reagent. This is useful when an experiment or CodeSet spans multiple lot numbers and the user must calibrate between them.

**Binding Density:** Number of fluorescent spots per $\mu m^2$ of the lane surface.

**Calibration Sample:** A reference sample that is used to correct for variability due to a technical variable such as CodeSet lot or instrument. This sample is assayed across all variations of that variable.

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

**Cartridge:** A 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.

**Copy Number Variation (CNV):** An assay type designed to detect copy number variations in user-defined regions of a genome.

**Code Class:** A categorization of reporter probe pairs. A text field with this designation is included in some of the table outputs.

**CodeSet:** The collection of probes that comprise a physical assay used to tag specific RNA/DNA sequences.

**CodeSet Plus:** An additional set of Reporter Probes customers can buy and add onto any custom CodeSet or Panel.

**Comments:** A text annotation field customized by the user in the CDF (see Cartridge Definition File, above).

**Differential Expression (DE) Call**: see Error Model.

**Digital Analyzer:** The nCounter instrument which collects data by taking images of the immobilized fluorescent reporters in the sample cartridge with a CCD camera through a microscopic 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. It tallies the barcodes in the images and the results are exported as a zipped file that can be downloaded via memory stick.

**Endogenous Probes:** Probes designed to target biological content defined by the customer. For Plex$^2$ and PlexSet 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.

**Error Model**: A calculation which determines whether the difference between two groups (numerator and denominator) is significant. This is usually used when no replicates exist in a dataset.

**Experiment:** A method that allows you to normalize, group, and create fold change estimates with your data. This then allows you to perform detailed analyses. Experiments can be performed on a single CodeSet (single-RLF) or on multiple CodeSets (multi-RLF). The Experiment Wizard facilitates this process. Upon completion of the wizard, raw, normalized, grouped, ratio data tables and analyses are available for viewing and/or export.

**False Discovery Rate (FDR)**: The portion of genes with values at least as low as the gene in question that are expected to be false discoveries. FDR can be used as a more conservative and informative alternative to p-values.

**Field of View (FOV):** An area of the cartridge surface discretely imaged by the Digital Analyzer. The Digital Analyzer uses multiple FOVs in the process of counting barcodes.

**File Attributes:** Data fields imported from the CDF (see Cartridge Definition File, above).

**File Name:** The unique name assigned to a lane by the digital analyzer. This is also known as the RCC file name.

**FOV Count:** Total number of FOVs (see Field of View, above) successfully counted per lane.

**Fusion**: An assay which detects areas in which a chromosomal fusion has taken place.

**Gene:** A genetic element that is targeted by a NanoString probe; for the purposes of this document, this may be a transcript (mRNA) or a microRNA (miRNA) molecule.

**Gene RLF:** The RLF used to associate gene names with counts specified by the user in the CDF (see Cartridge Definition File, above).

**Genomic Region**: A particular region of interest of the genome, identified by a single accession number. A set of probes may be made that target a particular genomic region; there may be multiple probes per region.

**GX:** A shorthand notation for the custom gene expression assay. This assay detects user-defined mRNA transcripts.

**Housekeeping Probes:** Probes designed to target transcripts that vary minimally in expression level from sample to sample. These probes are used in the normalization process during mRNA analysis.

**Invariant Probes:** Probes designed to target regions of the genome that have only two copies in most 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:** The physical flow cell of the cartridge surface over which nCounter probes are immobilized. Each lane is 2600 µm x 8000 µm.

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

**Lane ID:** The numerical identifier of the imaged lane (always a value from 1 to 12).

**Messages:** A text output field for Digital Analyzer QCs. A binding density greater than the platform's set threshold will generate a warning indication.

**miRGE:** An assay type which includes a custom combination of GX and miRNA targets for the detection of user-defined miRNA and mRNA transcripts.

**miRNA (micro RNA):** A small RNA of ~21-23 bases, typically associated with multiple proteins, which represses transcription of a specific target mRNA. NanoString offers panel assays targeting pre-defined sets of miRNAs in multiple different species (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.

**mRNA (messenger RNA):** A nucleic acid of 400-10,000 bases which serves as a template for protein synthesis (translation). MRNA panels are offered stand-alone, in Gene Expression, and with miRNA panels in the miRGE kits.

**Multiplexing:** A data field in nSolver that is checked if the cartridge lane has been utilized for a Plex$^2$ or PlexSet assay, and is unchecked if not.

**Name:** The 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.

**Normalization**: Normalization is a two-step data transformation that uses results from positive controls to balance counts between lanes, allowing the user to make meaningful biological comparisons.

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

**P-value**: The output of the t-test (see t-test). The lower the p-value, the stronger the evidence that the two groups have different expression levels.

**Plex$^2$:** An nCounter assay type which allows for up to 4 samples to be profiled in a single lane, trading increased samples per tube for decreased probes per CodeSet.

**PlexSet:** An nCounter assay type which allows for up to 8 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, the Prep Station automates immobilization of the sample onto an imaging surface for subsequent data collection.

**Probe:** A molecule NanoString designs to measure the level of a specific gene; customers can choose up to 800 probes in a single reaction to test. In many cases, there is a 1:1 relationship between a probe ID and a RefSeq ID and Gene Name, however, in some cases there may be a many to one relationship between probe designs and a single RefSeq ID or Gene Name.

**Quality Control flag**: An integer value specifying if the lane is flagged as “Failed”. Output only.
Radio Button – A GUI button control operating in a group of two or more, such that when the user selects a radio button, any previously selected radio button in the same group becomes deselected.

**Quality Control (QC) Parameters**: A variety of metrics, each measuring a different factor in the assay. Some, such as FOV and Binding Density, relate to the quality of the scanning process. Others, such as positive controls, measure the detection of known fragments included in each assay.

**Ratio:** An expression which compares two values, usually the result from an experimental sample to the corresponding result from an established baseline sample.

**Reporter Code Counts (RCC) File:** A comma-separated value data file 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 each target name and annotation with its respective unique fluorescent barcode in the CodeSet. While optional for some applications, 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:** The date information specified by the user in the CDF.

**Sample ID:** The sample identifier specified by the user in the CDF.

**Scanner ID:** The numerical identifier of the Digital Analyzer.

**Single Nucleotide Variant (SNV)**: An assay type which tests for variants which differ from the standards by only one nucleotide.

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

**Sub-CodeSet:** One of the multiple CodeSets that are assayed in a single lane of a cartridge when using the Plex$^2$ or PlexSet assays. Each sub-CodeSet has a designated 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 (up to 8 if loading a PlexSet assay).

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

**T-test**: A calculation which determines whether the difference between two groups (numerator and denominator) is significant. This is only used when replicates exist. The output is a p-value.

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