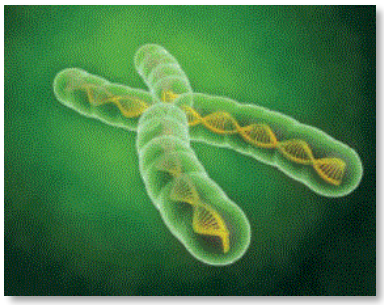
All About Fusion Analysis in nSolver & Advanced Analysis



This quick-start guide is designed to direct you through the steps of importing, processing, exporting, and analyzing your nCounter Fusion data in nSolver software, v4.0 and later and Advanced Analysis v2.0 and later. More detailed information may be needed and is available in the nSolver and Advanced Analysis user manuals.

Note: Advanced Analysis **must** be run to derive fusion detection calls for your nCounter Fusion assay data. Import and create experiments in nSolver and then run the Advanced Analysis module v2.0 or later.

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 RCC Files**. Navigate to your unzipped data folder and select your RCC files. Select the **Data Contains Fusion Probes** button to designate fusion probes (Figure 1). Fusion probes can be identified by consulting your CodeSet info.

Highlight fusion probes and select **Fusion** from the **Analyte Type** drop-down menu (Figure 2).

When complete, select **Next** (see QC section).

Once RCC files have been imported, select **Import RLF** and browse to select and import the RLF for your dataset.

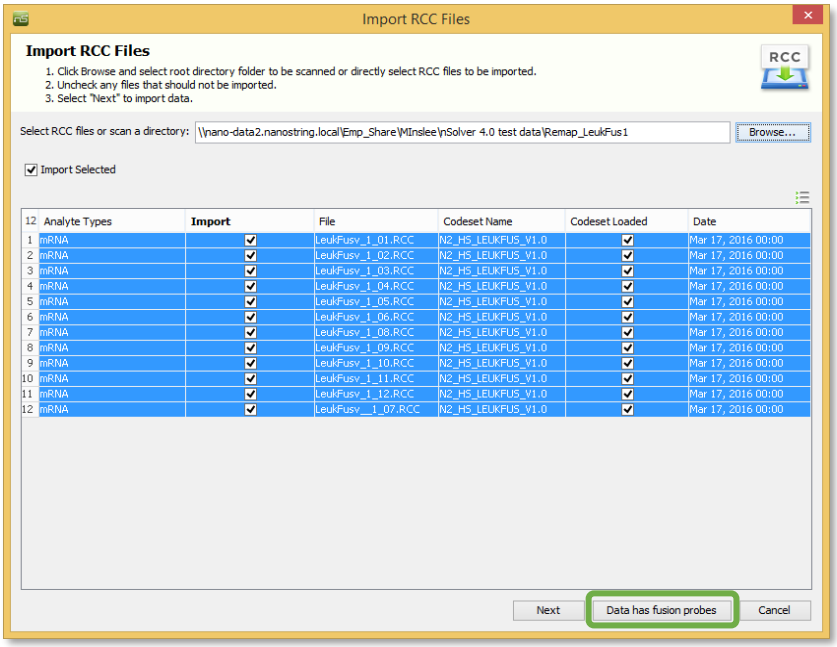


Figure 1: Importing RCC files

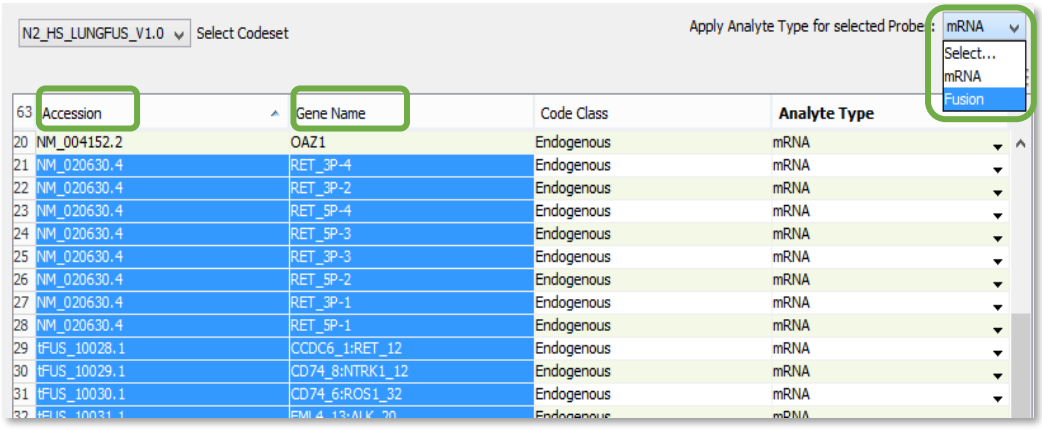


Figure 2: Designating Fusion probes

QC

Choose the RLF and then QC parameters (Figure 3). If hidden, select the double arrow at the right of the screen to reveal System QC parameters. Default settings are recommended.

- **Imaging QC** is a measure of the percentage of requested fields of view successfully scanned in each cartridge lane.
- **Binding Density QC** is a measure of reporter probe density on the cartridge surface within each sample lane.
- **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.
- **Positive Control Limit of Detection QC** indicates whether the counts for the POS_E control probe with target sequence spiked in at 0.5fM (assumed to be the system's limit of detection) is present above the counts of the Negative control probes with statistical significance.

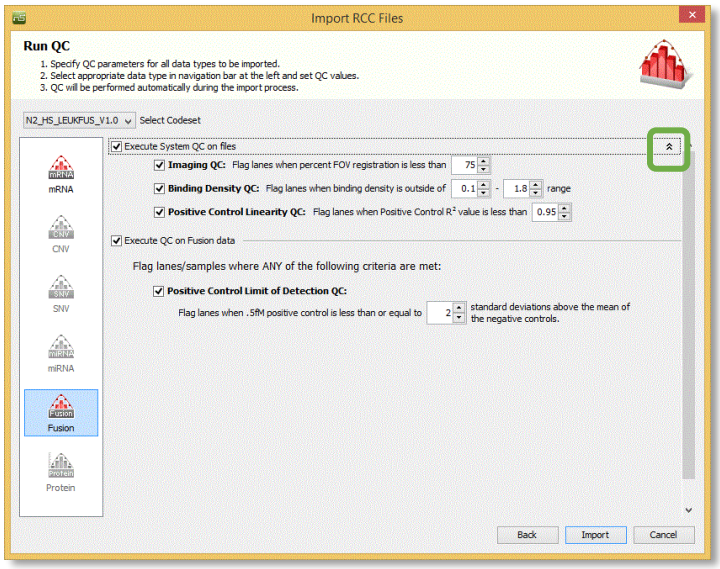


Figure 3: QC parameters

nSolver displays and applies the QC parameters recommended by NanoString. It is usually not necessary to adjust these default settings. Select **Import**.

NOTE: If your CodeSet contains probes for analytes other than Fusion, nSolver will offer those analyte-specific options in the experiment wizard. Consult the nSolver 4.0 Quick Start guide or the nSolver 4.0 User Manual for information on these options and their results.

Exploring Raw Data

Your RCC data files will now be stored under the corresponding RLF CodeSet on the **Raw Data** tab. Selecting the RLF name allows you to view all RCC files in a table format. Scroll to check for QC flags (Figure 4). 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.

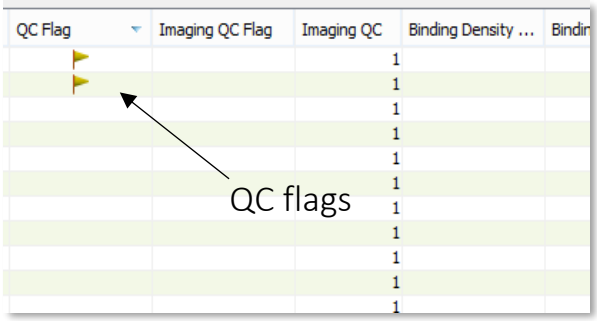


Figure 4: QC flags

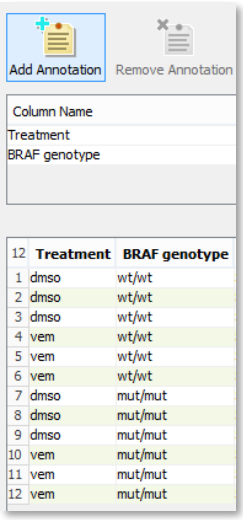
Creating Experiments

Within nSolver, a **Study** is an organizational folder used to store experiments, and an **Experiment** is a collection of samples that have been analyzed together to allow comparisons between samples or samples grouped in conditions. Any studies and experiments you create will be visible on the **Experiments tab**.

Select the **New Study** button to create a study.

Select the **New Experiment** button to create an experiment under that study. Follow the prompts to select the samples to include in your experiment.

Annotations to define sample groups should be assigned for experiments in which fold-change estimates and their statistical significance should be studied (Figure 5). These annotations can also be used in Advanced Analysis.



	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 5: Annotations

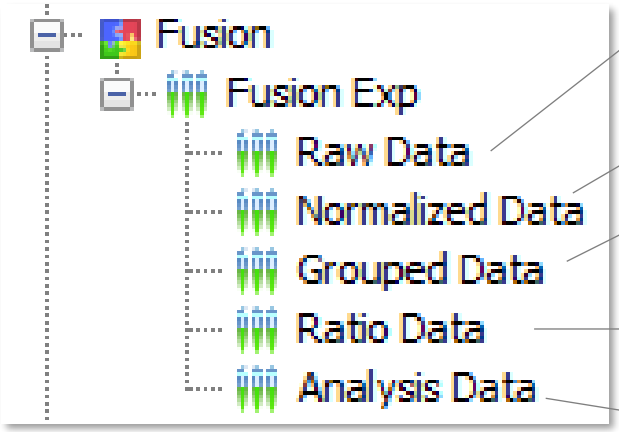
Background noise can be minimized in a few different ways. Confirm/select an option below (see wizard steps) and select **Next.**

- 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 a blank lane (if loaded).

Normalization can be accomplished by using the geometric mean of the **Positive Control** counts. Review the set defaults, set preferences, and select **Next**.

Data Export

Your experiment will now be visible under your study on the **Experiments tab**. Expand the navigation tree (Figure 6). Select the desired data table, highlight samples of interest in the main 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.



Raw Data table contains unprocessed data for all samples in this experiment. Use this for **Advanced Analysis** (next page).

Normalized Data table contains the processed data for all samples.

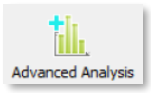
Grouped Data table contains the geometric mean of expression levels for all samples.

Ratio Data level contains any ratios created between the groups.

Analysis Data table contains any analysis you have run.

Figure 6: Experiment data file hierarchy

Advanced Analysis



Highlight your Raw or Normalized Data table and select **Advanced Analysis**. Assign a Name and select **Next**.

Select a unique identifier in the **Identifier** column such as File Name or Sample Name (shorter names are preferable). In the **Use in Analysis** column, select the boxes corresponding to Annotations you would like to study (see Figure 7); this will select them for analysis as **Covariates**.

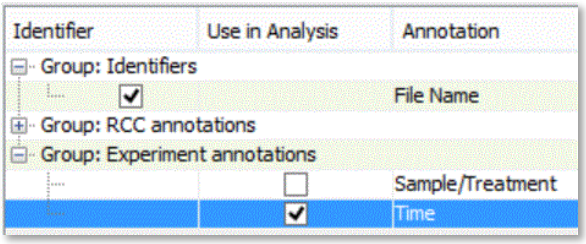


Figure 7: Choosing identifiers and covariates

You may select **Quick Analysis** and choose one covariate from the dropdown for analysis.

Alternatively, you may select **Custom Analysis** if you would like to run a multi-RLF analysis, choose multiple covariates, or customize your analysis in another way. The General Options tab will appear (see Figure 8). Select the **Fusion Analysis Parameters** button to change the p-value thresholds for Junction probe and End probe imbalance detection.

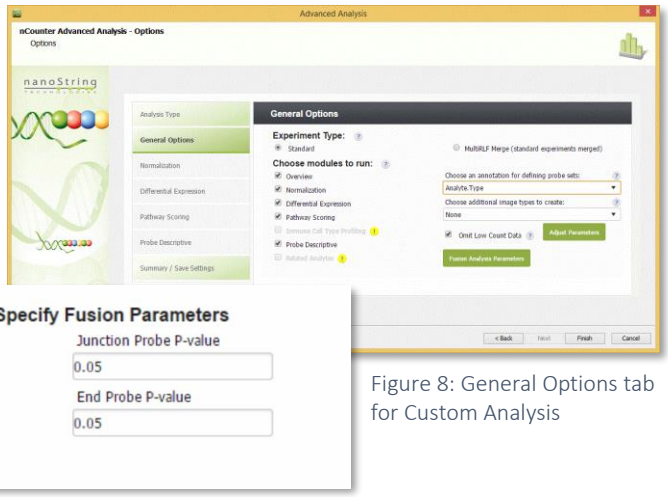


Figure 8: General Options tab for Custom Analysis

Select **Finish**. Return to the nSolver dashboard. Highlight your analysis' name from the Analysis level and select **Analysis Data**. Your analysis will appear in an HTML window. Select the Fusion tab to view results.

The **Fusion Detection Summary** (Figure 9) is a sample-gene matrix, color-coded according to whether a gene tested positive or negative for evidence of a junction (relying on a sequential outlier test for Junction Probe detection) and/or positive or negative for evidence of an End Probe imbalance (using a t-test for over-expression of the 3' probe). This figure summarizes the fusion calls by plotting each gene tested on the vertical axis and each sample tested horizontally. Color (see plot key) indicates whether a fusion was detected and the type of evidence used to make the call (Junction +/-, End Probe +/-, or both). The Junction and End probes provide different levels of evidence for fusion events (see Table 1). The **Fusion Summary Report** summarizes the results in a table format and provides specific details on the probe results used to make the calls.

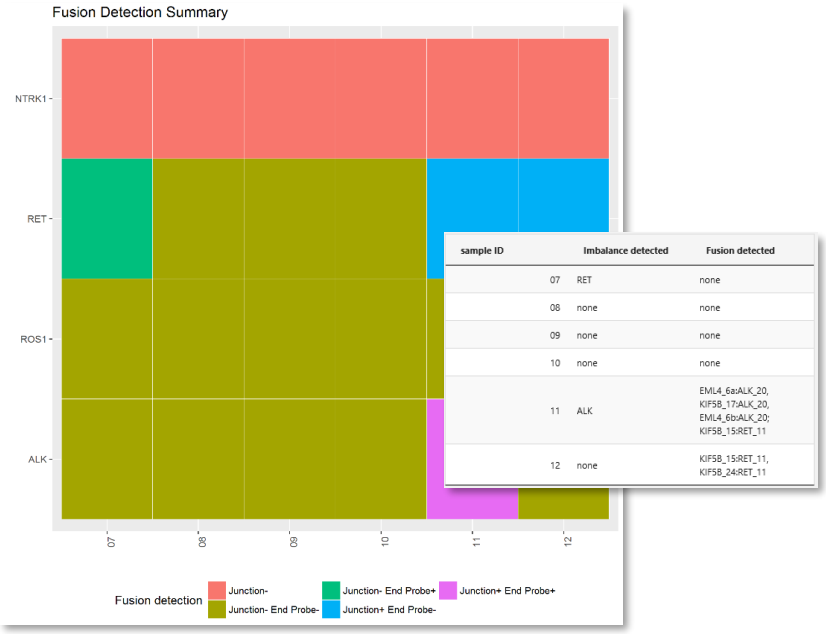


Figure 9: Fusion Detection Summary and Fusion Summary Report

The **Heatmap** displays the $\log_2$ raw counts for the different fusion probes and allows you to view **All** probes, just the **End probes**, or just the **Junction probes**. Many fusions will be immediately obvious in these heatmaps, either through high counts of a Junction probe or through strongly imbalanced 5'/3' probes for a gene within a sample. These heatmaps can also reveal technical artifacts.

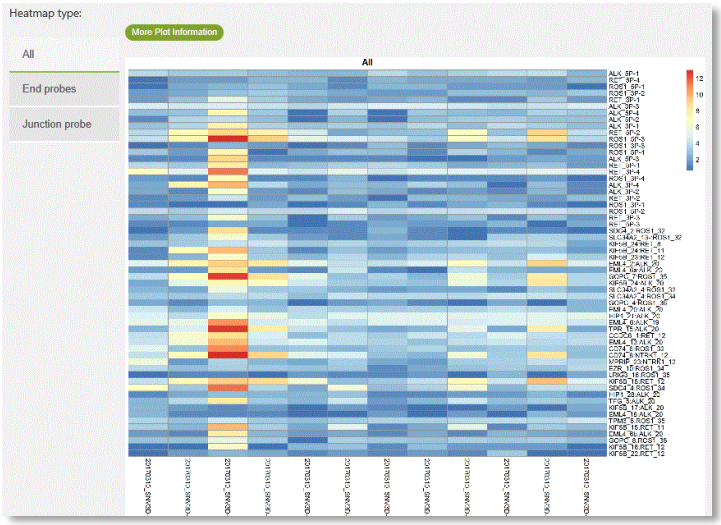


Figure 10: Heatmap of fusion results

On the **By Samples** tab, the **End probes** plot shows the $\log_2$ raw counts of the 5' and 3' probes. If an End probe detection call was made in the Fusion Detection Summary plot, confirm in this plot that the 3' probes' counts are visibly higher than the 5' probes' counts. A **p-value** for this is provided in the upper left of the plot. The **Junction probes** plot shows the $\log_2$ raw counts of the Junction probes. If a Junction probe detection call was made in the Fusion Detection Summary plot, this plot can provide a double check for the calling algorithm's results. The numbers above the bars show the ranking of detection where **1** conveys the highest confidence in detection and subsequent higher rankings convey decreasing confidence. A ranking of **0** means undetected. In the absence of a fusion, all probes will fall in the background of the system (with a 0 rank).

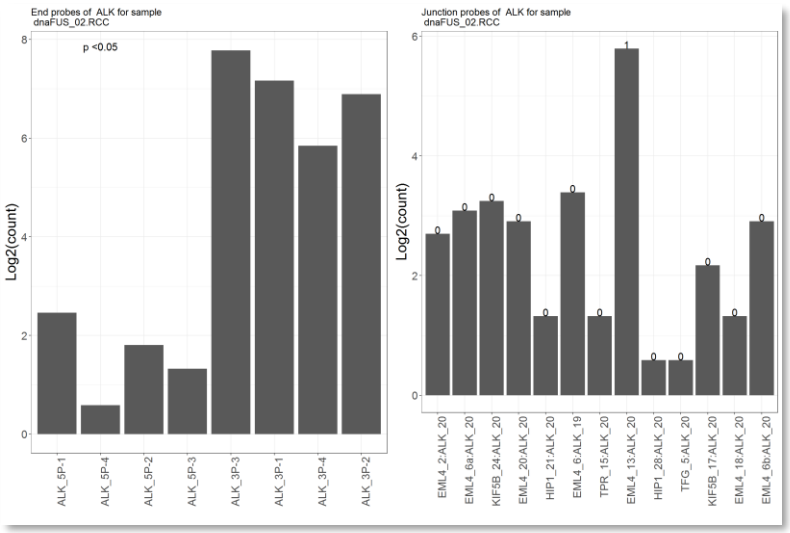


Figure 11: Barplots of Junction and End probe fusion results

Table 1: Categories for Fusion Calls

Result	Category	Summary	Example conclusion
End probe detection call; Junction probe detection call	Detected Gene Fusion, Variant Conclusive	There is a high probability that the sample is positive for a specific gene fusion variant	Positive ALK gene fusion event at EML4_13:ALK_20
End probe detection call; no Junction probe detection call	Detected Gene Fusion, Variant Inconclusive	There is a high probability that the sample is positive for a fusion event but the variant is inconclusive. May indicate the variant is not currently included in the fusion-specific probes (potentially a new variant)	Positive ALK gene fusion event, location unknown
End probe undetected call; Junction probe undetected call	Non-Detected Gene Fusion	There is a high probability that the sample is negative for a fusion event	No gene fusion variants detected
End probe undetected call; Junction probe detection call	Inconclusively Detected Gene Fusion	It is possible that the junction probe hit is a false positive, or that a fusion is truly present but has insufficient expression to be detected with the End probe test.	Possible low-level expression of GOPC_4:ROS1_36 atop high wild type ROS1 expression.