GeoMx® DSP
Data Analysis
User Manual

MAN-10154-01 for v2.5 software | March 2022

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

www.nanostring.com

T: 206.378.6266
T: 888.358.6266
E: geomxsupport@nanostring.com

Sales Contacts

United States: us.sales@nanostring.com
EMEA: europe.sales@nanostring.com
Asia Pacific & Japan: apac.sales@nanostring.com
Other Regions: info@nanostring.com

EU Authorized Representative

NanoString Technologies
Germany Gmbh
Birketweg 31
80639 Munich
Germany

UK Authorized Representative

NanoString Technologies
Europe Limited
11th Floor Whitefriars
Lewins Mead
Bristol BS1 2NT
United Kingdom

Rights, License, & Trademarks

Use

For research use only. Not for use in diagnostic procedures.

Intellectual Property Rights

This GeoMx® Digital Spatial Profiler (DSP) User Manual and its contents are the property of NanoString Technologies, Inc. ("NanoString"), and are intended solely for use by NanoString customers, for the purpose of operating the GeoMx DSP System. The GeoMx DSP System (including both its software and hardware components) and this User Guide 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 the GeoMx DSP System contained in the product quotation, NanoString grants you a limited, non-exclusive, non-transferable, non-sublicensable, research use only license to use the proprietary GeoMx DSP System only in accordance with the manual 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 the proprietary GeoMx DSP System. Without limiting the foregoing, no right or license, whether express, implied or statutory, is granted by NanoString to use the GeoMx DSP 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, NanoString Technologies, the NanoString logo, GeoMx, and nCounter are trademarks or registered 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 document are the property of their respective owners.

Copyright

Table of Contents

Changes in this revision

This **GeoMx DSP Data Analysis User Manual** (MAN-10154-01) replaces the GeoMx-NGS Data Analysis User Manual (SEV-00090-05) and GeoMx-nCounter Data Analysis User Manual (SEV-00067-07).

NGS-specific information is separated from nCounter-specific information with colored text boxes, as described in **Conventions on page 7**.

Other changes in this manual revision include:

- Added instructions for Data Analysis Workflows **on page 61**
- Defined Queue View in **Figure 25**
- Clarified Cluster Analysis calculations **on page 44**
- Clarified Q3 Normalization calculation **on page 38**
- Clarified Q30 values in the Segment Summary tab **on page 18**

Conventions

The following conventions are used throughout this manual and are described for your reference.

Bold text is typically used to highlight a specific button, keystroke, or menu option. It may also be used to highlight important text or terms.

Blue underlined text is typically used to highlight links and/or references to other sections of the manual. It may also be used to highlight references to other manuals or instructional material.

A gray box indicates general information that may be useful for improving assay performance. These notes aim to clarify other instructions or provide guidance to improve the efficiency of the assay workflow.

IMPORTANT: This symbol indicates important information that is critical to ensuring a successful assay. Following these instructions may help improve the quality of your data.

WARNING: This symbol indicates the potential for bodily injury or damage to the instrument if the instructions are not followed correctly. Always carefully read and follow the instructions accompanied by this symbol to avoid potential hazards.

For NGS readout: Content in blue boxes denotes steps or information specific to NGS readout of GeoMx DSP. Follow these instructions if using Illumina® NGS to read out GeoMx DSP counts.

For nCounter readout: Content in green boxes denotes steps or information specific to nCounter readout of GeoMx DSP. Follow these instructions if using nCounter® MAX/FLEX, Pro, or SPRINT to read out GeoMx DSP counts.

GeoMx DSP Workflow

The GeoMx Digital Spatial Profiler (DSP) is a novel platform developed by NanoString. This product relies on antibody or nucleic acid probes coupled to photocleavable oligonucleotide tags. After probes hybridize to targets in slide-mounted tissue sections, the oligonucleotide tags are released from discrete regions of the tissue via UV exposure. Released tags are quantitated by nCounter technology or Illumina Next Generation Sequencing (NGS). Counts are mapped back to tissue location, yielding a spatially resolved digital profile of analyte abundance **(see Figure 1)**.

- **Day 1: Slide Staining.** Prepare slides and incubate biological targets with UV-cleavable probes. Prepare manually or using the BOND RX/RX^{m} Fully Automated IHC/ISH Stainer from Leica Biosystems®.
- **Day 2: Process Slides on GeoMx DSP.** Load prepared slides into the GeoMx DSP instrument and enter slide/study information. Slides are scanned to capture fluorescent images used to select regions of interest (ROIs). The instrument collects UV-cleaved oligos from the ROIs into the wells of a collection plate.

For NGS readout:

Day 3: Transfer the collected aspirates to a PCR plate and perform **Library Prep** with Seq Code primers. Pool and purify the products, then **Sequence** on an Illumina NGS instrument.

Day 4: Process FASTQ sequencing files into digital count conversion (DCC) files using **NanoString's GeoMx NGS Pipeline** on Illumina DRAGEN™ accessed via BaseSpace™ Sequence Hub, or using GeoMx NGS Pipeline standalone software. Upload DCC files on to the GeoMx DSP.

For nCounter readout:

Day 2, continued: Transfer the collected aspirates to a hybridization plate along with GeoMx Hyb Code reagents. Hybridization occurs overnight.

Day 3: Pool wells and **Process on an nCounter MAX/FLEX or Pro Analysis System or SPRINT Profiler.** Upload reporter count conversion (RCC) files to the GeoMx DSP.

- **Day 4 or 5**: Create a **Data Analysis** study in the Data Analysis suite and perform quality-control checks and data analysis, and generate analysis plots.

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

GeoMx DSP Workflow

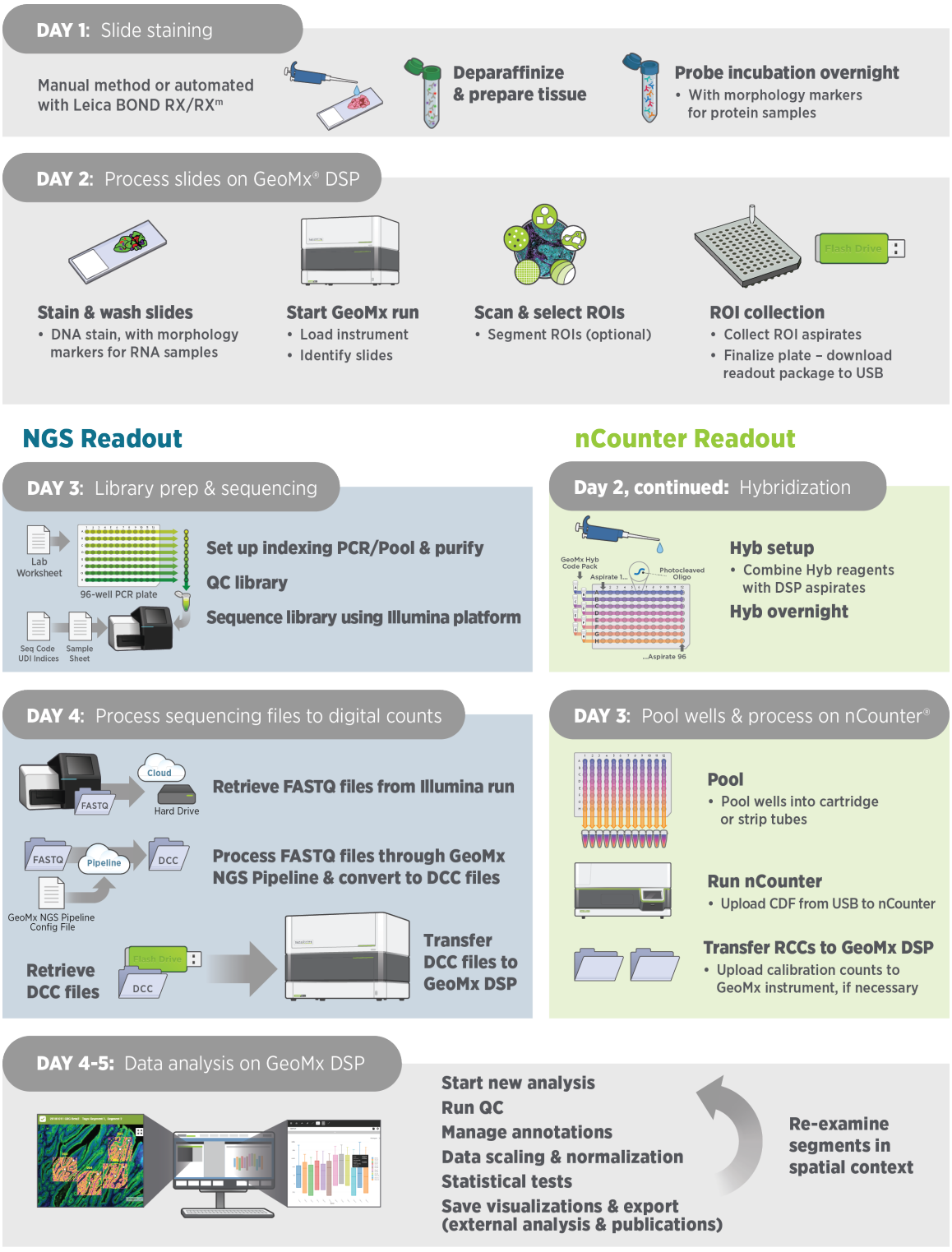


Figure 1: GeoMx DSP workflow summary

User Manuals and Resources

The GeoMx DSP workflow is divided into the following user manuals:

Workflow Step 1	GeoMx DSP Manual Slide Preparation User Manual MAN-10150 GeoMx DSP Automated Slide Preparation User Manual MAN-10151	
Workflow Step 2	GeoMx DSP Instrument User Manual MAN-10152	
Workflow Step 3	**For NGS readout:** GeoMx DSP NGS Readout User Manual MAN-10153	**For nCounter readout:** GeoMx DSP nCounter Readout User Manual MAN-10089
Workflow Step 4	GeoMx DSP Data Analysis User Manual MAN-10154	

User manuals and other documents can be found online in the NanoString University Document Library at https://university.nanostring.com.

Instrument and workflow training courses are available in NanoString University.

For NGS readout:
For documentation specific to the Illumina platform, see https://support.illumina.com.

For nCounter readout:
For documentation specific to the nCounter Pro, MAX/FLEX, and SPRINT instruments, see https://www.nanostring.com/support/support-documentation/ or the NanoString University Document Library at https://university.nanostring.com.

Data Analysis Introduction

The Data Analysis module can be accessed from the **Data Analysis** button at the top of the GeoMx DSP Control Center. From this, a menu appears where you have the option to select **New study from queue** or **Open Study**. You may also launch this module from the **Records** section of the GeoMx DSP Control Center. See Creating/Opening a Study on page 13.

Data Analysis Screen Orientation

The Data Analysis screen is divided into three panes and a toolbar (see Figure 2). Below, we describe the functions of each pane and the interactions between all three.

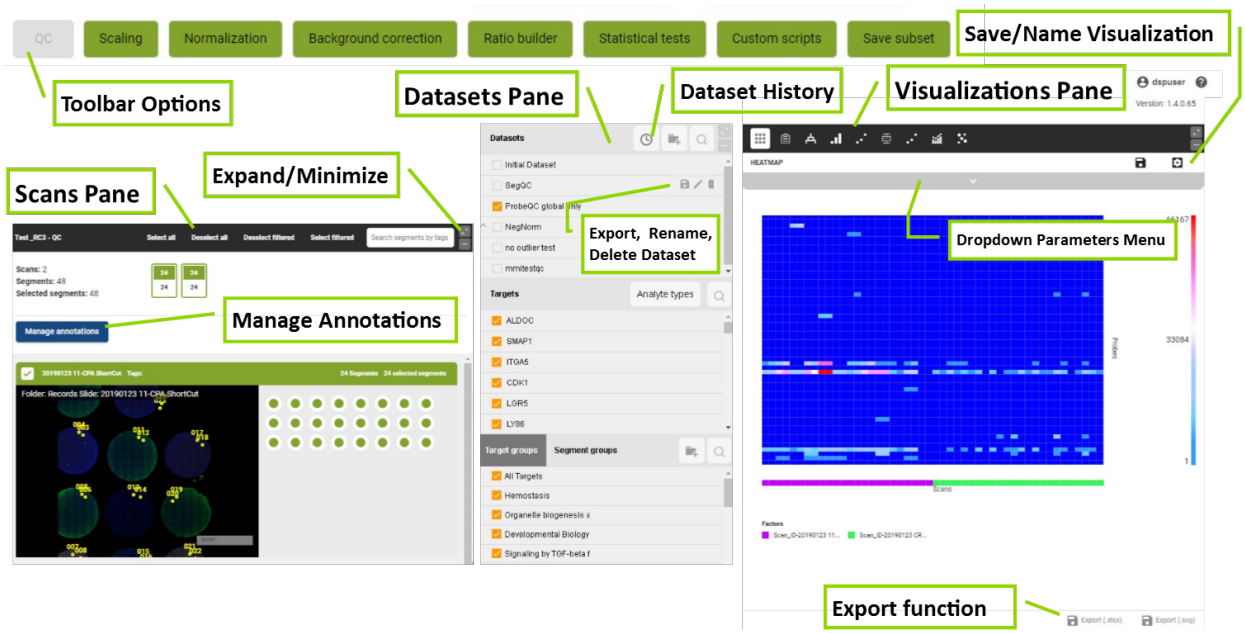


Figure 2: Data Analysis Suite diagram

The **Scans** pane is shown on the left side and contains individual image views of the scans. Scan images can be displayed as thumbnails or as interactive images. Scan icons, depicted as rectangles, represent each scan and are located at the top of this pane. Green circles (picker buttons) represent each segment (or ROI if ROIs are not segmented) and are located to the right of each image viewer. These icons and buttons are also toggle controls - a single click will switch the state of a scan or segment from selected to unselected and vice versa. Hovering over an icon or picker button displays additional information such as name, tags, etc. See Scans Pane on page 69.

Numerous **Toolbar Options** are available, however, when starting an analysis you will only be able to select **Data QC.** Once Data Quality Control (QC) has been run, other toolbar options will

become active. Toolbar options are active or inactive based on the current selected dataset and the tasks allowed for that dataset. After each data analysis step performed using the toolbar options, you will be prompted to save the new dataset. See Toolbar Options **on page 32**.

The **Datasets** pane is in the middle and lists all saved datasets, targets, target groups, and segment groups associated with the current study. At a minimum, you will see the **Initial Dataset** and the **All Targets** group. See **Datasets Pane on page 71**.

The **Visualizations** pane is on the right. Visualizations are visual representations of the dataset selected, the targets selected, and the adjustments applied to the data from those targets.

Select an area of interest in any visualization to see the respective spatial location of the segments highlighted in the Scans pane. Right-click on your selection in the visualization to choose an action. These can include:

- Creating a target group or segment group from the selection.
- Excluding the selected set of targets or segments from the current study.
- Creating a tag for the selected set of segments.

See **Visualizations Pane on page 78**.

Creating/Opening a Study

A new analysis can be launched from the Data Analysis button on the GeoMx DSP Control Center (see Figure 3); the new analysis will be built using the scans in the Data Analysis Queue.

For a scan to be eligible for inclusion in a study, count data must have been uploaded for the scan. For instructions to upload count data, see the GeoMx DSP Instrument User Manual (MAN-10152).

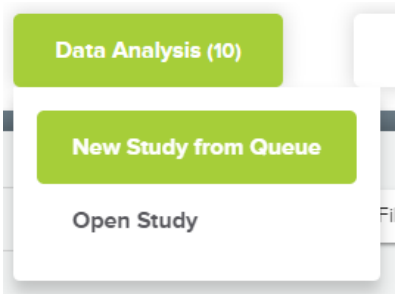


Figure 3: Data Analysis button

Assembling the Data Analysis Queue

1. Select the **Records** button from the DSP Control Center.
2. Select the folder containing your scans from the Navigation window on the left (see Figure 4).

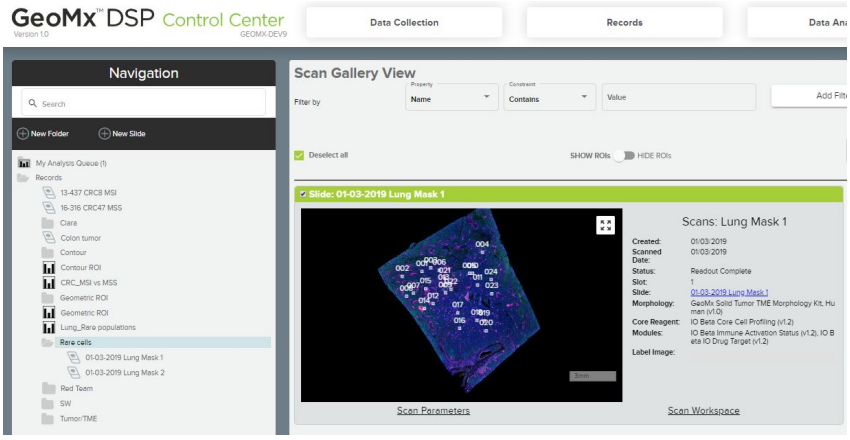


Figure 4: Scan Gallery View under Records; adding scans to Data Analysis queue

3. Review the status of each scan of interest. Only scans that are **Readout complete** can be added to the queue.

> **IMPORTANT:** If you previously uploaded counts and are now re-uploading counts, note that the new counts will replace the old counts for the slide records and for any future data analysis studies BUT any old data analysis studies will remain as they were with the old data.

4. Select each scan of interest by clicking the check box in the upper-left corner. This will turn the header green. You can view slides in the gallery format or individually.
5. Click the **Add Selected Scans to Queue** button in the upper right corner of the window to queue the scan(s) together for an analysis. The number of scans currently in the Data Analysis Queue will appear in parentheses on the Data Analysis button; this number will change dynamically as scans are added to or removed from the queue.
6. To view or edit the Queue, select the **My Analysis Queue** folder from the top row of the Navigation window (see Figure 5).

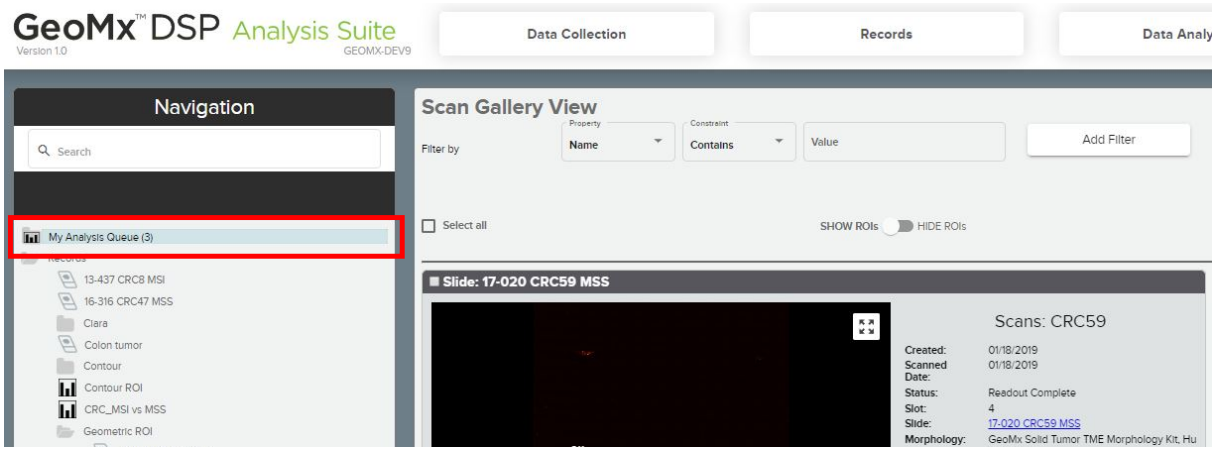


Figure 5: My Analysis Queue view

Filtering and Selecting Scans

Use slide and scan metadata to filter scans. Select search parameters from the dropdowns in the upper left of the window (see Figure 5): Property (name), Constraint (contains), and Value (customizable field). Select **Add Filter** to apply a filter rule. Applied filter rules will appear as gray boxes under the filter field. Use the **X** on each box to remove that filter, if desired. Select **Clear Filters** to clear all applied filter rules.

Select all scans using the check box "Select all" in the header (see Figure 5). Alternatively, select one scan at a time using the check box in each scan's upper left corner.

The **Scan Order** default is by scan date (from most to least recent); select another option from the drop-down (Scan Name or Slide Name), if desired.

Creating a New Study

1. From the GeoMx DSP Control Center, hover over the **Data Analysis** button (see Figure 6).

2. **New Study from Queue** and **Open Study** buttons will appear. Select **New Study from Queue**. The number in parentheses on the Data Analysis button indicates the number of scans currently in the Queue. In the example here (see Figure 6), there are 10 scans presently in the queue.

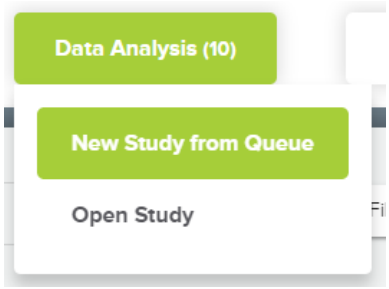


Figure 6: New Study from Queue button

> You can also create a study by opening the **Data Analysis Queue** under **Navigation** and then select **New Study from Queue**.

3. Enter a **Name**, **Description** (optional), and the folder where you would like this study saved (see Figure 7).

4. Select **Create** to build an analysis from the scans in your Data Analysis Queue. Folder permissions can be set so that only defined users can view a study (see the GeoMx DSP Instrument User Manual (MAN-10152)).
   - Only the scans that you have added to your **Data Analysis Queue** (represented by the number in parentheses on the Data Analysis button) will be used to build your study.

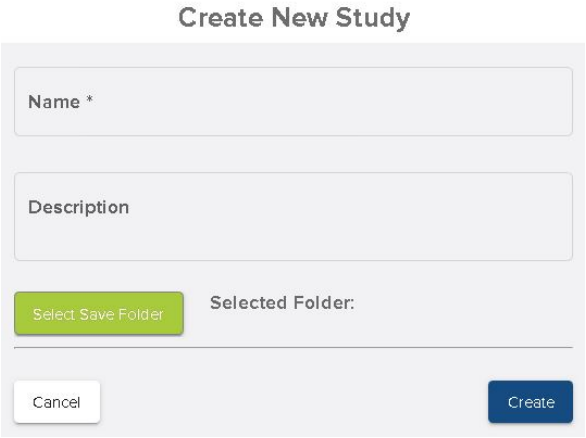


Figure 7: Create New Study window

5. A window will appear letting you know that a study is being created. Click the notifications bell to see the progress. Once ready, click the **Open Study** button in that window to open the study.

Opening a Study

Open a previously-saved study by double-clicking on its name in the Navigation window (see Figure 8).

Alternatively, hover over the **Data Analysis** button, then click **Open Study** (see Figure 9).

- In the pop-up window, browse previously created studies.
- Select your study and select **Open Study** .

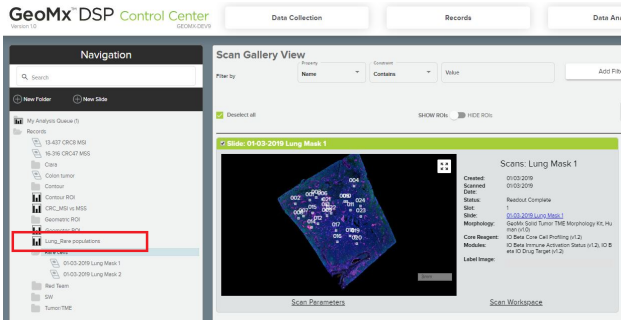


Figure 8: Locate study in Navigation window

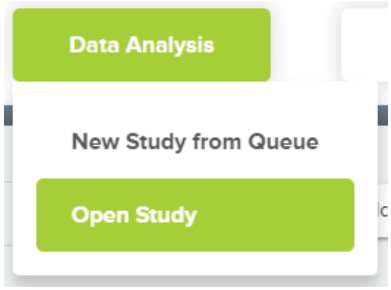


Figure 9: Open Study

Note that a study can only be opened by one user at a time. If encountering the message, “analysis is locked by another user”, check to see if another user is accessing this study on instrument or by remote access. If not, close the study, wait one minute, refresh the page, then re-open the study.

Once you have a study open, proceed to Data QC:

For NGS readout:
Refer to Data QC for NGS Readout on page 17.

For nCounter readout:
Refer to Data QC for nCounter Readout on page 26.

Data QC for NGS Readout

> **For NGS readout:** This section applies only to studies with NGS readout.

Assess data

The sequence run should be assessed prior to data QC using the line plot and study summary tools in the visualization pane:

Line plot

The line plot (see Figure 10) shows the **Raw reads**, **Trimmed reads**, **Stitched reads**, **Aligned reads**, and **Deduplicated reads** in Log_{10} space across all segments. For more information, see the GeoMx DSP NGS Readout User Manual (MAN-10153).

- Examine this graph for differences between the plotted lines. In general, the number of reads in all categories should be essentially the same, except for deduplicated reads. Deduplicated reads are expected to plot lower than the rest; this indicates tha PCR duplicates were sequenced and that you have representative sequencing. Look for dips in the plotted lines, as this may indicate low reads and ineffective sequencing. In this example (see Figure 10), the drop in reads on the left may be due to small area ROIs.
- Click the arrow under the line plot header to access the drop-down field. Choose **Linear**, **ratios**, or **Log_{10}** transformation. Select **tags** or **factors** to stratify by color.
- Click and drag to zoom in on any area of interest in the line plot.

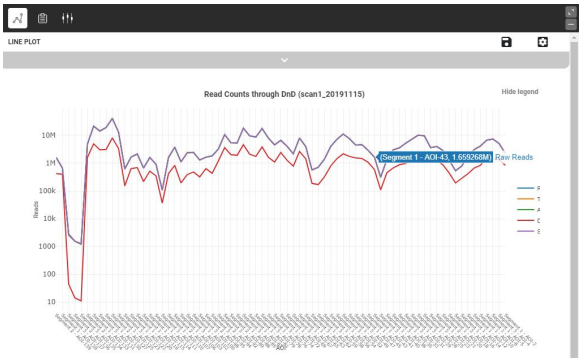


Figure 10: Line plot for NGS data

Study Summary

The NGS study summary (see Figure 11) plots the total number of reads for each segment by processing step outcome (trimmed, deduplicated, etc.). It also contains drop-down fields with **Segment Summary** (see Figure 12), **Probe Summary**, and **Target Summary**.

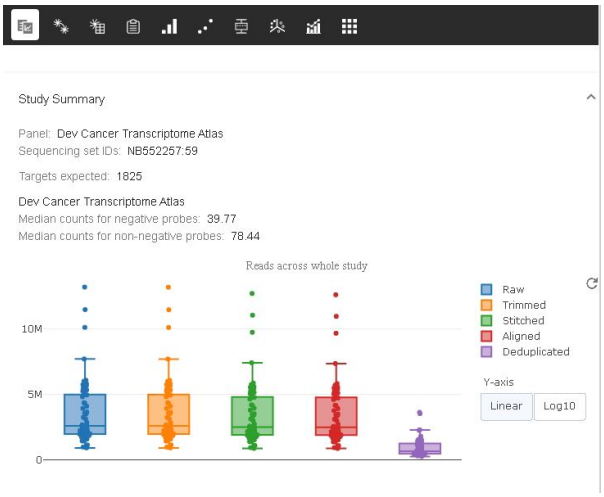


Figure 11: NGS Study Summary

Segment summary includes Q30 values. Q30 refers to the percentage of bases that get a Q score of at least 30 (99.9% accuracy). Q score is the most common metric to assess the accuracy of a sequencing platform (refer to https://www.illumina.com/documents/products/technotes/technote_Q-Scores.pdf for more details).

The terms **umiQ30** and **rtsQ30** refer to the percentage of bases with Q≥30 specifically in the UMI and RTS_ID portions of the reads, respectively.

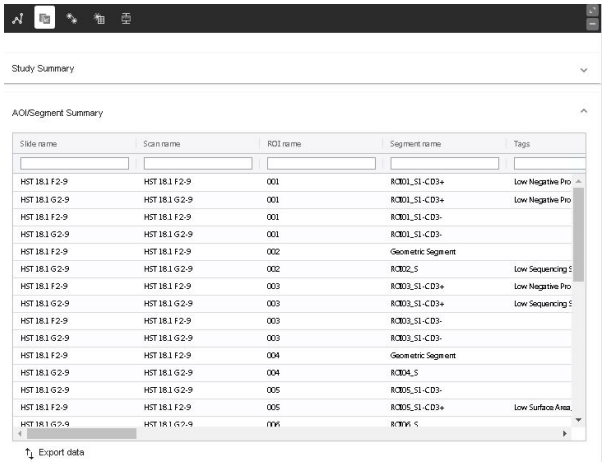


Figure 12: Segment Summary

Data QC

Data QC allows you to specify parameters for evaluating the quality of your data and to qualify segments by number of cells or by area. There are two steps in the QC process: **Segment QC** and **Biological probe QC**. Segments outside of desirable quality criteria are tagged with QC failure tags. Find more details on QC calculations in **Algorithm Details on page 101**.

Segment QC

In the QC process, some basic metrics collected from the run are compared to an expected range of values to determine the overall quality of the data.

Select the **QC** button. Review the QC parameters and default values selected for your study **(see Figure 13)**. You may change the selections and default values as appropriate based on the parameters of your study. After running QC, the software flags segments not meeting these QC parameters.

Flagged segments are not removed automatically, but are tagged for reference in subsequent data analysis steps.

Figure 13: QC parameters window

After each task (QC, Normalization, etc.), the resulting data will be saved as a new dataset and appear in the dataset list (see **Datasets**). Enter a unique name for each dataset, a description (optional), and any tags (optional).

Technical Signal QC

Technical signal QC assesses the quality of sequencing for each segment.

The **raw read threshold** is the lowest number of raw reads allowed for a segment.

The **percent aligned reads** establishes the minimum percent of raw reads that align to a target sequences that is allowed.

Sequencing saturation sets the minimum percent of sequencing saturation allowed. The percent of sequencing saturation is calculated as: 1 - (deduplicated reads/aligned reads) x 100. 100% sequencing saturation indicates a representative sample, while 0% sequencing saturation indicates that all reads were unique. Values below 50% may need to be resequenced.

Technical background QC

The technical background QC is a GeoMx DSP run control.

The **Negative probe count geomean** establishes the level of technical noise expected and flags segments with signal below that level. The negative probe count geomean will be calculated separately for each probe kit.

> The Negative Probe count geomean is not relevant for Protein NGS and should be un-checked.

The **No Template Control Count** establishes the level at which counts in the NTC will be flagged. The No Template Control (NTC) is used to detect contamination in the library prep.

Minimum Nuclei and Surface Area Count

The GeoMx chemistry is designed for an optimal range of nuclei and surface areas. The default value of each of these QC parameters will automatically fill with the minimum recommended value for the GeoMx product used in the present run. Note, however, that not all slides will have nuclei count, particularly if no DNA channel was selected. These may be safely deselected if ranges or consideration are not appropriate for the parameters of your study.

Select the **Run QC** button.

Viewing Segment QC Results

When you click **Run QC**, a pop-up window will appear with preliminary QC results **(see Figure 14)**. Scroll down the list.

Any samples not meeting the set QC criteria will appear with an orange WARNING in the QC status column. Check the **Show only segments with warning** box at the upper left to see this complete list.

Click the **Back** button to adjust QC parameters, if appropriate, or the **Run QC** button to accept these results. Segments designated with **Warning** will be retained in the analysis and tagged with a QC flag. You can select or deselect segments associated with this tag from the search bar in the Scans pane, for example to exclude flagged segments from a subsequent analysis step. See **Tags on page 66**.

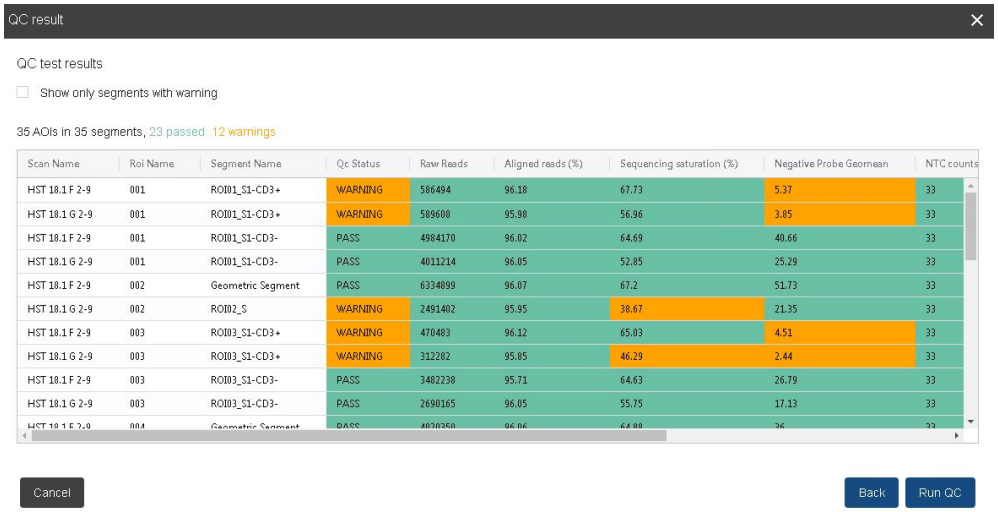


Figure 14: QC test results window

From the **Visualizations pane**, select the **Line plot** or **Summary** button to assess segment QC.

Biological Probe QC for RNA Assays

Once you have run initial QC on your segments, the **Biological probe QC** button will become active. Here, you determine the thresholds for excluding probes from your data set that appear to be outliers. The Protein and RNA Biological probe QCs differ and are documented separately here.

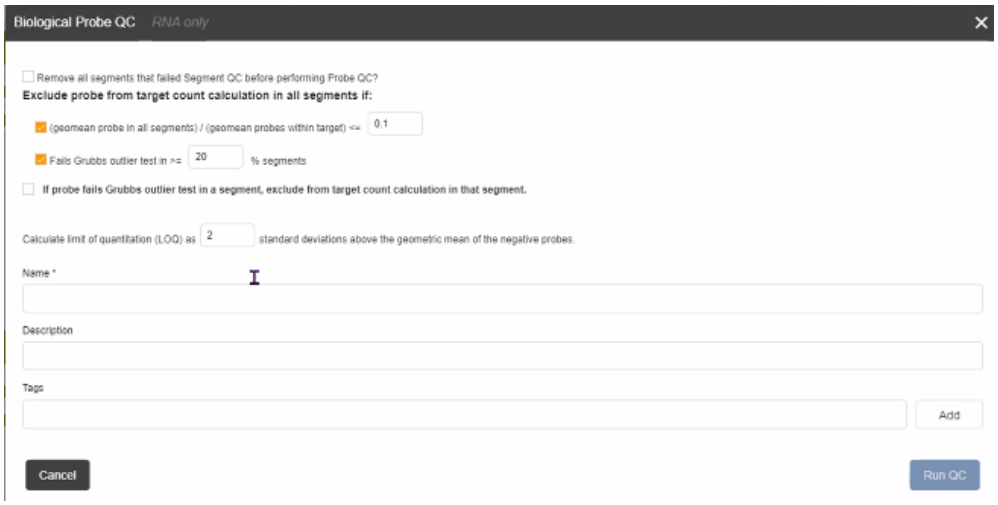


Figure 15: Biological Probe QC window

Select the **Biological probe QC** button. Review the QC parameters and default values selected for your study **(see Figure 15)**.

> **IMPORTANT:** Checking the first box in this window will remove all segments that had a **warning** in the first QC step. The remaining parameters outlined here determine when to exclude outlier probes from the data. You may change the selections and default values as appropriate based on the parameters of your study.

Geomean probe in all segments / Geomean probes within target

This is a parameter intended to catch dropout probes (low outliers). It excludes probes for which the average counts across all segments are <=10% (or the value you enter) of the counts for all probes to that gene. This test catches probes that performed poorly relative to other probes to the same target. If a single probe is used to quantify a target, the Grubbs test is not used.

Example: if four of five probes to GeneA yield 100 counts per segment on average and the fifth probe yields <10 counts on average, then the fifth probe will be called a global outlier (removed from ALL segments).

Fails Grubbs outlier test

This is a parameter intended to catch probes that are consistent outliers (always higher or lower than the other probes to the same gene) when multiple probes for a target are analyzed. A Grubbs outlier test is performed on the probe counts to a given target in a given segment (5 values if there are 5 probes to the target). This is done for every target in every segment. If a particular probe is designated an outlier in >=20% of segments, then it will be designated as a global outlier (removed from ALL segments). This test removes probes that are outliers in a user-defined proportion of segments from the entire dataset. If a single probe is used to quantify a target, the Grubbs test is not used.

If probe fails Grubbs outlier test in a segment, exclude from target count calculation in that segment

This is a toggle that determines whether or not local outliers are removed from the data. If this box is checked, probes that are outliers in a given segment will be removed from that segment, even if they are not global outliers.

Note that the box in front of the statement **If probe fails Grubbs outlier test in a segment, exclude from target count calculation in that segment** is checked by default.

Example: for all the Grubbs tests performed above, the outlier values will be removed from individual segments if this box is checked. If the box is not checked, the Grubbs tests are still performed to identify global outliers, but only global outlier probes are removed from the dataset.

It is only recommended to run the global outlier test on datasets with 24 or more ROIs/segments.

Calculate LOQ

The limit of quantitation (LOQ) is set to be the negative probe geomean multiplied by the geometric standard deviations of the negative probes.You may set LOQ to the value of your choice.

$$\mathbf{LOQ = GeoMean(Neg_{1\text{-}n}) * GeoStdev(Neg_{1\text{-}n})^{LOQ\ Threshold}}$$

This is a confidence threshold rather than a detection threshold. A value below LOQ does not necessarily mean that a target is not expressed, but if the value is >LOQ then we are confident that it is expressed. 2.0 is the default (and recommended) setting for high confidence detection. If an experiment included more than one probe kit, the LOQ will be calculated separately for each probe kit.

Click the **Run QC** button.

Viewing Biological QC Results

When you click **Run QC**, a pop-up window will appear with preliminary biological probe QC results **(see Figure 16)**. Scroll down the list.

Any samples not meeting the set Biological QC criteria will appear with an orange WARNING in the outlier status column. Check the **Show failed only** box at the upper left to see this complete list.

Click the **Back** button to adjust the Biological QC parameters, if appropriate, or the **Apply** button to accept these results. Probes designated with a **Warning** status will not be automatically removed from analysis; they will be tagged with a QC flag. Select or deselect segments associated with this tag from the search bar in the Scans pane, for example to exclude flagged segments from a subsequent analysis step. See **Tags on page 66**.

This QC step distinguishes local outliers from global outliers. A **local outlier** is a probe appearing as an outlier in a few segments whereas a **global outlier** probe appears as an outlier in >20 (or as defined) % of segments and will be removed from all segments.

Any QC parameter marked with a warning will materialize as a new tag. You can select or deselect segments associated with this tag from the search bar in the Scans pane. See **Tags on page 66**.

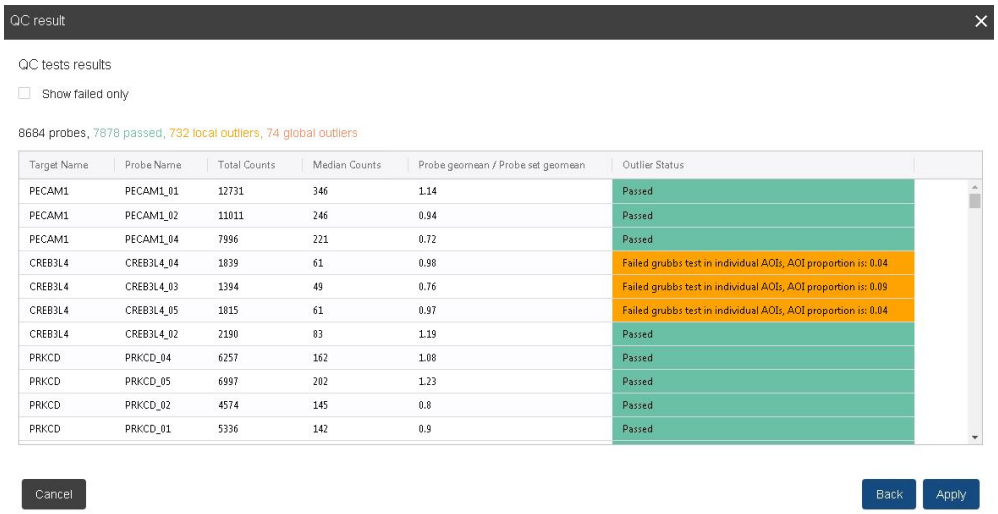


Figure 16: Biological QC results

Biological Probe QC for Protein Assays

Figure 17: Biological Probe QC for protein assays

Select the **Biological probe QC** button to open the Biological Probe QC window **(see Figure 17)**.

Calculate LOQ

The limit of quantitation (LOQ) is set to be the negative probe geomean + some number of geometric standard deviations of the negative probes.

$$\mathbf{LOQ = GeoMean(Neg_{1\text{-}n}) * GeoStdev(Neg_{1\text{-}n})^{LOQ\ Threshold}}$$

This is a confidence threshold rather than a detection threshold. A value below LOQ does not necessarily mean that a target is not expressed, but if the value is >LOQ then we are confident that it is expressed. 2.0 is the default (and recommended) setting for high confidence detection. If an experiment included more than one probe kit, the LOQ will be calculated separately for each probe kit.

It is only recommended to run the global outlier test on datasets with 24 or more ROIs/segments.

Click the **Run QC** button.

Data QC for nCounter Readout

For nCounter readout: This section applies only to studies with nCounter readout.

Data QC allows you to specify parameters for evaluating the quality of your nCounter run and to qualify segments by number of cells or by area. The resulting dataset after QC is adjusted for inherent technical variability across the lanes of the nCounter cartridge. Segments outside of desirable quality criteria are tagged with QC failure tags. Find more details on QC calculations in **Algorithm Details on page 101**.

Once you have imported data, the **QC** button will be active. In the QC process, some basic metrics collected from the nCounter run are compared to an expected range of values to determine the overall quality of the data.

Select the **QC** button. Review the QC parameters and default values selected for your study **(see Figure 18)**. You may change the selections and default values as appropriate based on the parameters of your study.

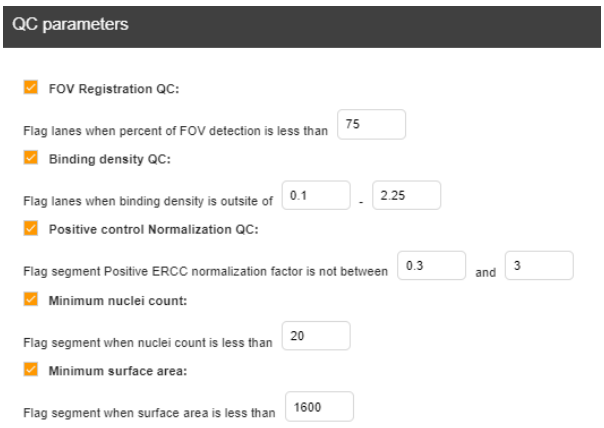


Figure 18: QC parameters window

After each task (QC, Normalization, etc.), the resulting data will be saved as a new dataset and appear in the dataset list (see **Datasets**). Enter a unique name for each dataset, a description (optional), and any tags (optional).

FOV Registration QC

This metric reports the percentage of fields of view (FOVs) the Digital Analyzer or SPRINT instrument was able to capture. At least 75% of FOVs should be successfully counted to obtain robust data. A failure in this area may indicate something as simple as a tilted or smeared cartridge, which can be remedied by rescanning (*MAX*/*FLEX* systems only; ideally within one week). However, consistently reduced percentages can be indicative of an issue with the instrumentation. Wells with consistently reduced percentages should not be included in further analysis. In troubleshooting FOV Registration QC results, you may:

- Check alignment of the cartridge in the nCounter instrument – if crooked, reload and rescan.
- Clean the bottom of the cartridge with **70% EtOH** and a lint-free wipe, reload, and rescan.
- Refer to the appropriate nCounter instrument user manual for additional information (see https://nanostring.com/support/support-documentation).

Binding Density QC

This metric is a measurement (in spots per square micron) of the concentration of barcodes detected by the nCounter instrument. The Digital Analyzer may not be able to distinguish each probe from the others if the density is too high. Generally, a segment's binding density is proportional to its quantity of hybridized probe tags. Very high binding density can result in decreased linearity in counts and loss of sensitivity. Low binding density can result in a loss of detectable signal from expressed targets. The default Binding Density ranges are:

- nCounter *MAX* or *FLEX* system: 0.1-2.25 spots per square micron.
- nCounter *SPRINT* system: 0.1-1.8 spots per square micron.

Measurements outside of these ideal ranges may result in a Binding Density QC failure, but should be checked to see how much they deviate from the ideal range. Large increased deviations can result in reduced FOV registration and may represent a saturated cartridge. Large decreased deviations may represent a failed lane. Wells exhibiting either type of large Binding Density deviation should be discarded from analysis. Some deviation may still result in good quality data, but this must be evaluated on a case-by-case basis. In troubleshooting Binding Density QC results you may:

- Check the input amount. More sample input will result in an increased Binding Density.
- Consider the expression level of the targets. If the targets are highly expressed, Binding Density will go up simply because more molecules are being targeted in your samples.

- Consider the size of the panel. If a panel contains probes for more targets, then Binding Density will usually be higher.

Positive Control Normalization QC

The GeoMx Hyb Code Pack contains one spike-in positive hybridization (pos-hyb) control probe per well for evaluating the quality of the nCounter readout step. A Positive Control Normalization factor is calculated using these pos-hyb controls. This normalization adjusts for variations that exist across lanes, cartridges, and days and account for differences in user technique, hybridization, magnetic bead purification, complex- to- slide binding, and imaging. Robust expression of the pos-hyb control probe is expected.

The default range of acceptable values is 0.3 - 3.0, but values outside this range may be determined to be acceptable based on empirical data. Very low counts for the pos-hyb probe in combination with a large Binding Density deviation may indicate issues with the hybridization step. In addition, large variations in ROI or segment size will contribute to an increased Positive Control Normalization range of values; in this case, care should be taken to ensure the positive hybridization controls are above 100 counts before positive control normalization.

As part of the QC process, data will be normalized to the positive controls in the run; the positive controls will be removed from the dataset after that point. Before running QC, it is recommended that you view the positive control results in the initial dataset to ensure that positive control counts are as expected.

Minimum Nuclei and Surface Area Count

The GeoMx chemistry is designed for an optimal range of nuclei and surface areas. The default value of each of these QC parameters will automatically fill with the minimum recommended value for the GeoMx product used in the present run. Note, however, that not all slides will have nuclei count, particularly if no DNA channel was selected. These parameters may be safely deselected if ranges or consideration are not appropriate for the parameters of your study.

Select the **Run QC** button.

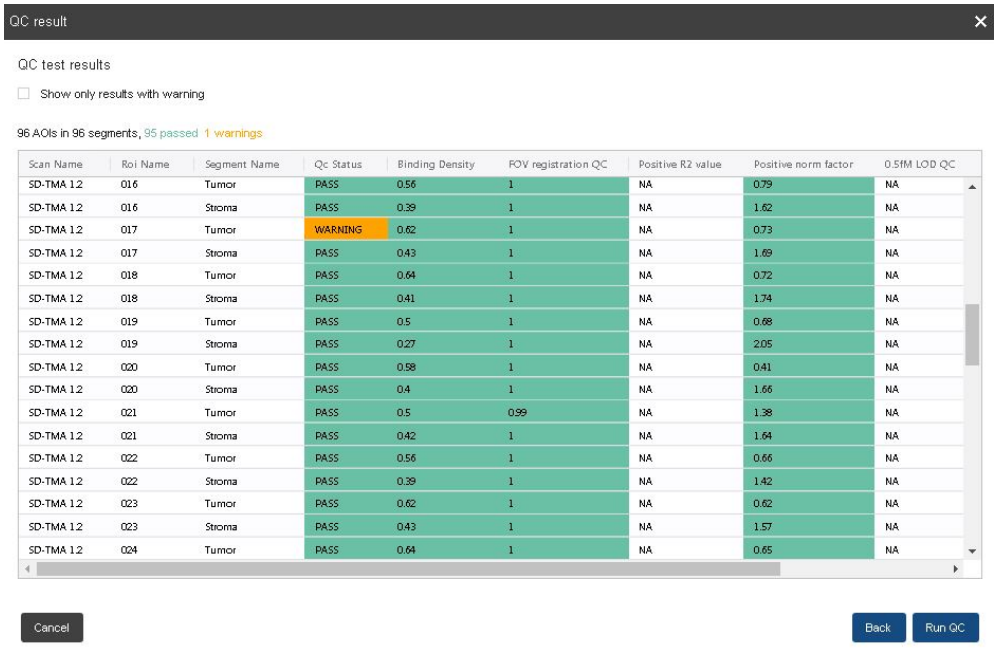
Viewing QC Results

When you click **Run QC**, a pop-up window will appear with preliminary QC results **(see Figure 19)**. Scroll down the list.

Any samples not meeting the set QC criteria will appear with an orange WARNING in the QC status column. Check the **Show only results with warning** box at the upper left to see this complete list.

Click the **Back** button to adjust QC parameters, if appropriate, or the **Run QC** button to accept these results. Segments designated with **Warning** will be retained in the analysis and tagged with a QC flag. You can select or deselect segments associated with this tag from the search bar in the Scans pane, for example to exclude flagged segments from a subsequent analysis step. See **Tags on page 66**.

Any QC parameters designated with a warning will materialize as a new tag. You can select or deselect segments associated with this tag from the search bar in the Scans pane. See **Tags on page 66**.



Scan Name	Roi Name	Segment Name	Qc Status	Binding Density	FOV registration QC	Positive R2 value	Positive norm factor	0.5fM LOD QC
SD-TMA 1.2	016	Tumor	PASS	0.56	1	NA	0.79	NA
SD-TMA 1.2	016	Stroma	PASS	0.39	1	NA	1.62	NA
SD-TMA 1.2	017	Tumor	WARNING	0.62	1	NA	0.73	NA
SD-TMA 1.2	017	Stroma	PASS	0.43	1	NA	1.69	NA
SD-TMA 1.2	018	Tumor	PASS	0.64	1	NA	0.72	NA
SD-TMA 1.2	018	Stroma	PASS	0.41	1	NA	1.74	NA
SD-TMA 1.2	019	Tumor	PASS	0.5	1	NA	0.68	NA
SD-TMA 1.2	019	Stroma	PASS	0.27	1	NA	2.05	NA
SD-TMA 1.2	020	Tumor	PASS	0.58	1	NA	0.41	NA
SD-TMA 1.2	020	Stroma	PASS	0.4	1	NA	1.66	NA
SD-TMA 1.2	021	Tumor	PASS	0.5	0.99	NA	1.38	NA
SD-TMA 1.2	021	Stroma	PASS	0.42	1	NA	1.64	NA
SD-TMA 1.2	022	Tumor	PASS	0.56	1	NA	0.66	NA
SD-TMA 1.2	022	Stroma	PASS	0.39	1	NA	1.42	NA
SD-TMA 1.2	023	Tumor	PASS	0.62	1	NA	0.62	NA
SD-TMA 1.2	023	Stroma	PASS	0.43	1	NA	1.57	NA
SD-TMA 1.2	024	Tumor	PASS	0.64	1	NA	0.65	NA

Figure 19: QC test results window

Manage Annotations

The GeoMx DSP system groups and filters data with annotations called tags and factors. For information on tagging data in visualizations, see **Tags on page 66**. Segment annotations can be imported by uploading a spreadsheet:

1. Click the **Manage Annotations** button in the Scans pane **(see Figure 20)**.

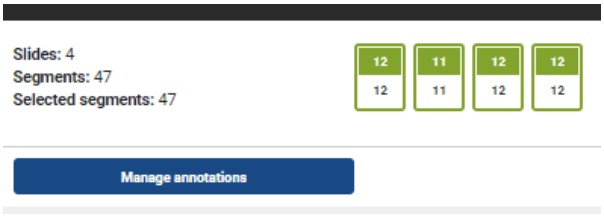


Figure 20: Manage Annotations window

2. Download the **Annotations template file** **(see Figure 21)** to a USB or, if accessing remotely, to a remote file location.

Figure 21: Manage Annotations window

3. On a separate computer (not the GeoMx), open the downloaded spreadsheet and review the pre-filled **Scan name**, **ROI (label)**, and **Segment (Name/Label)** columns. If you have any existing tags, they will be listed in the **Segment Tags** column **(see Figure 22)**. You can:
   - Modify the annotations pre-loaded in the **Segment Tags** column and/or
   - Create a new column(s) and add your tag names here. Column headers should be one word and cannot begin with "ROI" or the column will be mistaken for ROI_IDs by the software. Content in these columns can be imported as tags or factors. Factors are tags that have multiple sub-categories which collectively span all segments in a dataset. For example, the factor *Tissue Type* would have the sub-categories *tumor*, *stroma*, and *invasive margin*.
4. After you have added all desired tags, delete the instruction row at the top of the spreadsheet.

	Scan name	ROI (label)	Segment (Name/ Label)	Segment tags	Scan_ID	ROI_ID	Tags
5	DZ074 85C	1	roi-001-segment-001		DZ074 85C	1	
6	DZ074 85C	2	roi-002-segment-001		DZ074 85C	2	
7	DZ074 85C	3	roi-003-segment-001		DZ074 85C	3	
8	DZ074 85C	4	roi-004-segment-001		DZ074 85C	4	

Figure 22: Adding a tags column

5. Save this spreadsheet as an **.xlsx** file on a USB or remote file location and return to the **Manage Annotations** window.
6. From the **Import from file** tab, **Browse** to your saved spreadsheet.
7. Choose the **Update type** from the drop-down:
   - **Append** (default) adds new tags on to any existing tags in the dataset.
   - **Merge** adds new tags to any untagged segments and replaces the existing tags of any tagged segments.
   - **Replace** deletes any existing tags and replaces them with the tags specified in the spreadsheet.

   If importing factors only without modifying tags, check the **skip modifying tags** box.
8. Use the drop-downs to designate the columns that contain each piece of information requested **(see Figure 23)**. For **factors**, choose from **text**, **boolean** (yes/no), or **number** (numeric value) to describe the type of levels within the factor.

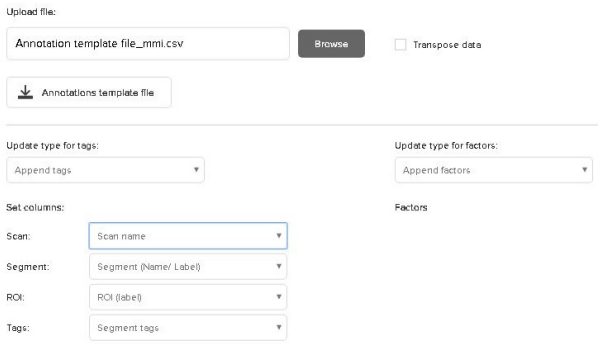


Figure 23: Importing annotations from file

9. Review the grid at the bottom of the window. Ensure that the desired columns of tags to be imported are selected; deselect any columns whose tags you do not want to be imported.
10. Select **Apply**.

If you encounter an error message when uploading an annotation file, check that you do not have leading or trailing spaces in the sample name or slide name. Ensure that you have not changed the Segment names, ROI names, or Scan Names in the Annotation file, or the software will not recognize the information.

To remove existing tags, create an annotation file from the template with blanks in the Segment_tags column. Choose **Replace** as "update type for tags" in the **Manage Annotations** window.

Toolbar Options

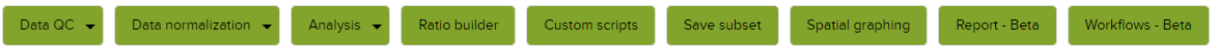


Figure 24: Data Analysis Toolbar

The toolbar (see Figure 24) contains many options to analyze your data. Active data analysis tasks are listed in the Queue View, accessible by clicking on the Queue View icon (see Figure 25).

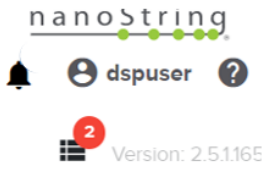


Figure 25: Queue View icon in red.

Data QC

Under the **Data QC** tab, find **QC** and **Filter** (described on the next page).

QC

You must run **Data QC** first (see Data QC for NGS Readout on page 17 or Data QC for nCounter Readout on page 26). After this, the other toolbar buttons will become activated (turn green). **Scaling** and **Background Correction** (within the Data Normalization tab) and **Ratio Builder** are designed to be run once each (after they have been run, they will be inactive).

Filter

The Filter Parameters allow you to establish an expression threshold and a frequency at which the targets or segments are allowed to be below that threshold **(see Figure 26)**.

Select an **expression threshold.** Targets or segments with expression at or below the threshold will be deselected in the present dataset.

For NGS readout: Choose from LOQ (limit of quantitation, see below), user defined value, or the greater of the two.	**For nCounter readout:** Select a user-defined value.

Note that a value below LOQ does not necessarily mean that a target is not expressed; if the value is > LOQ then we are confident that it is expressed above background levels and that we have accurately quantified the target.

Frequency defines the frequency at which a segment or target is allowed to be below the expression threshold and still be retained in the dataset. For example, "Keep targets with 90% segments above threshold" is more strict than "Keep targets with 10% segments above threshold."

Enter a **Dataset name** and any **Tags** you would like associated with the new dataset (optional). You may also check the **Create Segment/Target group from excluded** box.

Click the **Update** button to see your changes applied in the window.

Click **Save**.

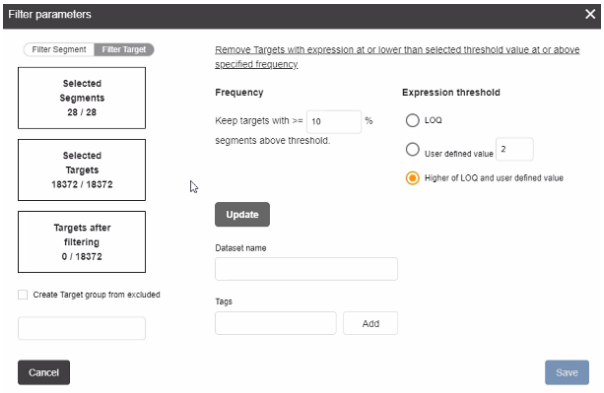


Figure 26: Filter parameters for NGS studies

Data Normalization

Under the **Data Normalization** tab, find **Scaling, Normalization**, and **Background Correction**.

Scaling

In the Scaling window, you can adjust counts to scale to the geometric mean, median, or mean of the detected area, or to the number of nuclei (see Figure 27). Choose the desired **Average type** from the drop-down. Find more details on scaling calculations in Algorithm Details on page 101.

After each task (QC, Normalization, etc.), the resulting data will be saved as a new dataset and appear in the dataset list (see **Datasets**). Enter a unique name for each dataset, a description (optional), and any tags (optional).

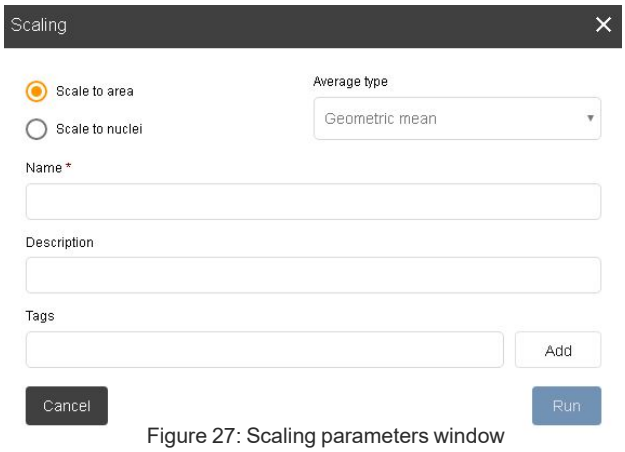


Figure 27: Scaling parameters window

Take care to avoid scaling to area or nuclei after normalizing to housekeepers or IgGs. Doing this effectively double normalizes your data to the same factor.

Scale to area

This is a type of normalization correcting for variable surface areas. The Data Analysis software calculates the ratio of the geometric mean, median, or mean surface area to the measured surface area of each segment. This ratio is then used to adjust the counts detected in that segment.

Scale to nuclei

This is a type of normalization based on counting nuclei, and may be preferable when studying expression per cell. The Data Analysis software calculates the ratio of the geometric mean, median, or mean number of nuclei to the measured number of nuclei of each segment. This ratio is then used to adjust the counts detected in that segment.

Select the **Run Scaling** button.

Viewing Scaling Results

From the **Visualizations pane**, select the **Dataset Summary** button and the **Scaling** tab (see Figure 28).

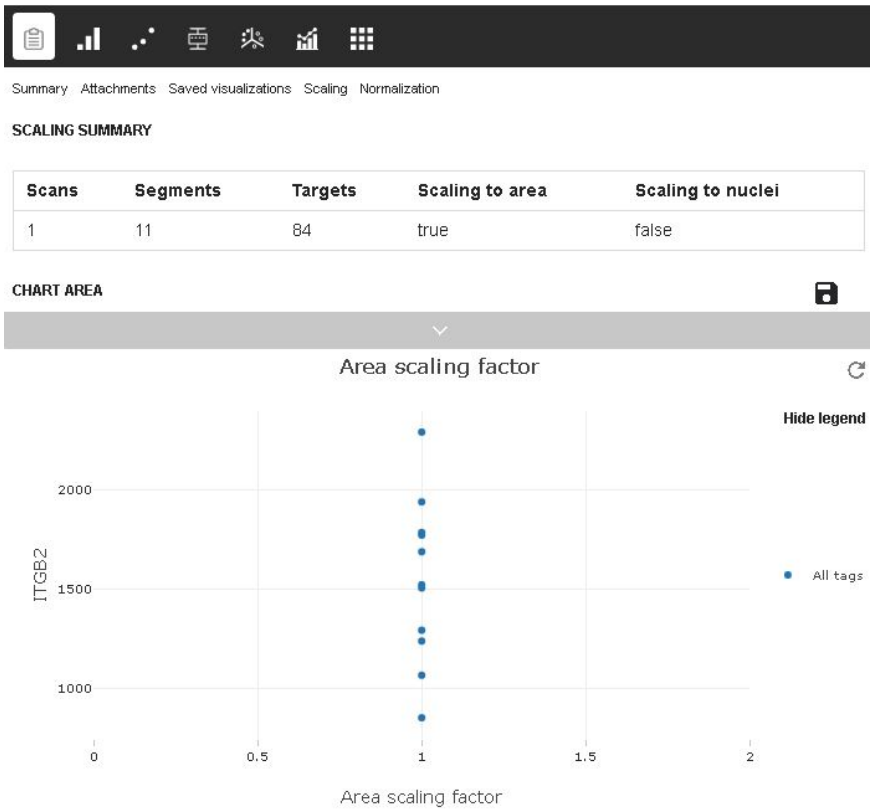


Figure 28: Scaling results plot

- The **Scaling Summary** section features a summary of the scans, segments, and probes that were included in the scaling, as well as any of the selections made.
- The **Chart area** provides visualizations plotting the scaling factors against adjusted counts for the selected probe; change this probe using the **Select probe** field.
- Click the arrow to access the drop down field (see Figure 29).

  • Select **Area scaling factor**, **Surface area**, **Nuclei scaling factor**, or **Nuclei counts** to see the respective chart. The **Summary Table** lists the scaling factors for each segment.

  • Select the probe of interest.

  • **Select tags** or **factors** to filter and color the results based on tag.

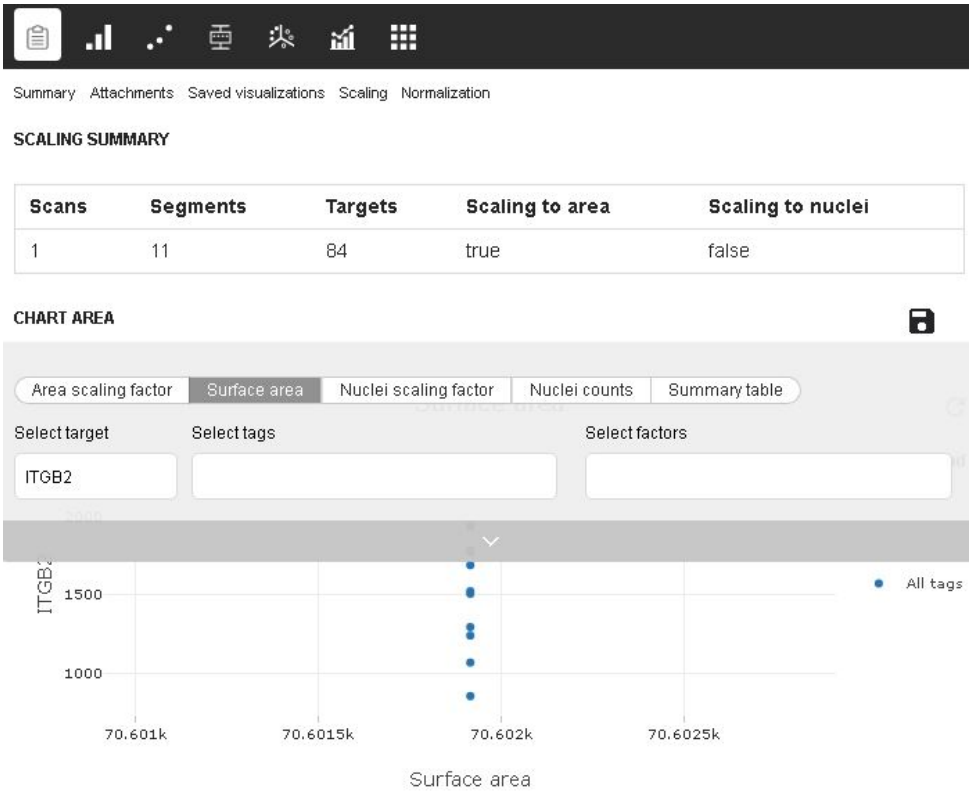


Figure 29: Scaling summary drop down menu

Normalization

Normalization is a data transformation that balances the results between segments within an analysis using the counts from a specific set of probes. Find more details on normalization calculations in **Algorithm Details on page 101**.

After each task (QC, Normalization, etc.), the resulting data will be saved as a new dataset and appear in the dataset list (see **Datasets**). Enter a unique name for each dataset, a description (optional), and any tags (optional).

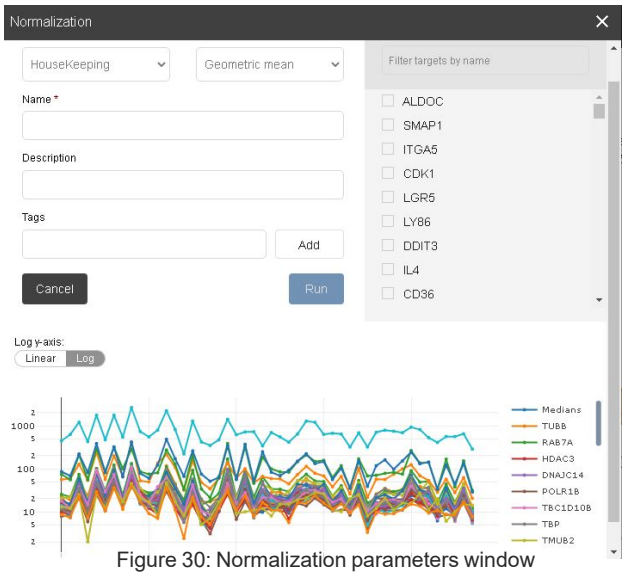


Figure 30: Normalization parameters window

A line plot displaying the expression of normalization probes will appear at the bottom of the Normalization window **(see Figure 30)**. If more than 40 targets are selected, this plot will show the median. You may choose to customize normalization using the following fields:

Normalization types:

Housekeeping: This option uses a CodeSet Content Normalization factor (also called Reference or Housekeeping Normalization factor), which is calculated using reference targets endogenous to the biological sample to adjust for differences in analyte abundance and/or analyte quality across samples. This method assumes that a reference target is consistently expressed across all samples. If the assumptions are met this normalization can account for differences in the amount of cells analyzed.

Predefined target genes classified as controls will be automatically selected. Use the check boxes to make desired alterations. The default range of acceptable scaling values based on this normalization type is 0.1 - 10.

Target Group: Target group normalization allows you to normalize using the target group of your choice. Use the check boxes in the list to select a target group. This normalization has no set defaults.

Custom: Custom normalization allows you to choose the probes to be applied as reference genes when predefined genes either don't exist or are not appropriate for the dataset. Ensure that the segments in your analysis are "like" segments. Use the check boxes to select probes for custom normalization.

Negative: Probes classified as Negative controls will be auto-selected. Negative normalization scales counts in a segment based on how the negative probe counts in that segment compare to those in other segments. If a segment has higher negative probes than average, its target counts will be reduced, and vice versa. This method is similar to calculating a signal-to-noise ratio, except counts are multiplied by a normalization factor rather than being divided by the negative probe geomean.

Average types:

Choose from Geometric mean, Median, Average, Sum, Minimum, Maximum, and (for NGS studies) Q3. The drop-down menu defaults to the geometric mean, since it weighs the low-concentration controls equally with the high-concentration controls.

For NGS readout:
RNA Assays: Upper quartile or **Q3 normalization** is the recommended normalization method for all targets that are above the limit of quantitation (LOQ). Q3 normalization divides the counts in one segment by the 3rd quartile value for that segment, then subsequently multiplies that value by the geometric mean of the 3rd quartile values of all segments. Q3 normalization should be performed after filtering the dataset to remove targets below the LOQ. To use this method, select **Normalization Type: Target Groups**, **AllTargets**, then select **AverageType: Q3**.

Protein Assays: the recommended normalization method is to use the geomean of **Housekeepers** (commonly S6, Histone H3, and GAPDH). To use this method, select **Normalization Type: Housekeeping** and **Average Type: Geometric mean**.

For nCounter readout:
RNA and Protein Assays: Housekeepers and Negative controls are commonly used to normalize nCounter RNA and Protein data. Determine the best normalization method by following **Introduction to GeoMx Normalization** and associated NanoString University tutorials. Certain **workflows** help evaluate the best normalization strategy. See **Workflows (Beta) on page 61**.

Q3 normalization rescales the gene expression data such that all segments have similar gene expression ranges. It reduces variance from segment size, segment cellularity, and other technical factors. Underlying the Q3 method is an assumption that, despite biological variation, segments should have similar gene expression count distributions. Thus, Q3 is only appropriate when a probe panel is large and diverse, such as one that targets the full transcriptome. When probe panels are small and targeted, or when segments are small and only a small fraction of a probe panel is detected in those segments, Q3 is typically not an appropriate normalization method. Additionally, Q3 is performed only with regard for assay signal, not assay background, and can become biased if signal is not adequately above background. This issue commonly arises when segments are small.

Viewing Normalization Results

From the **Visualizations pane**, select the **Dataset Summary** button and the **Normalization** tab.

If scaling (area or nuclei) or background correction was performed, you may want to evaluate the visualization depicting normalization factor vs. scaling factor or normalization factor vs. background factor. High values for normalization factor, scaling factor, or mean background should be carefully evaluated.

- The **Normalization Summary** section features a summary of the scans, segments, and probes that were included in the normalization, as well as any of the selections made.
- The **Chart area** features visualizations that compare the normalization trend to the scaling or background adjustment trends (if those tasks have been run).
- Click the arrow to access the drop down field. Select **Scaling** or **Background** from the slider to see the respective chart. **Select tags** or **factors** to filter and color the results based on tag.

Background Correction

Background Correction allows you to adjust for any non-target-specific adherence of probes to the tissue.

In the Background Correction window (see Figure 31), you can customize background adjustment. Calculations can be performed by the DSP Data Analysis software to minimize the impact of the background on the data counts. Choose probes that have no ligand in your datasets' segments (for protein assays, this may be IgG controls, and for RNA assays, this may be all negative probes).

After each task (QC, Normalization, etc.), the resulting data will be saved as a new dataset and appear in the dataset list (see **Datasets**). Enter a unique name for each dataset, a description (optional), and any tags (optional).

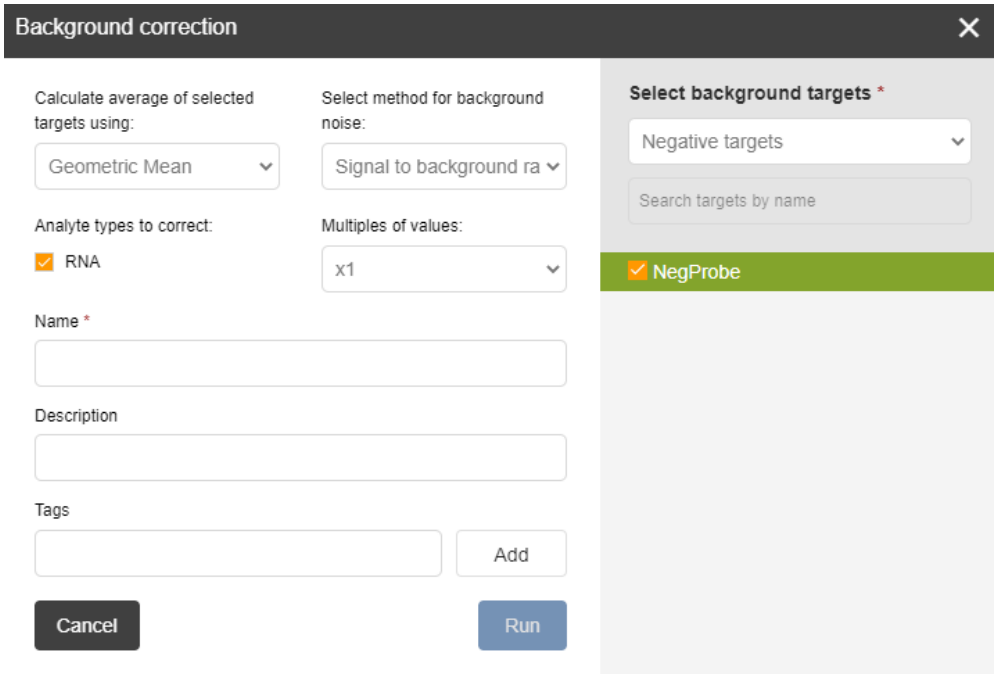


Figure 31: Background correction parameters window

Average methods

Choose from Geometric Mean, Maximum, Median, Mean, or Mean+n (n represents the number of standard deviations).

Background Method

Choose one of the following background correction methods.

Signal to Background Ratio

For every probe in your segment, the Data Analysis software will calculate a ratio of the probe counts to the average represented background, using the averaging method you selected and the counts from the probes you designated as negative.

Thresholding

Choose the negative controls from the probes field for a negative control threshold. The system will take the average (using the method specified in the **Average method** field) of the negative control probe counts to determine the background level. Once the background level has been determined, all raw counts at or below this value will be set to this treshold value.

Subtraction

Though not recommended for most applications, Background Subtraction is available as an option. Choose the negative controls from the probes field for a negative control subtraction. The system will take the average (using the method specified in the **Average method** field) of the negative control probe counts to determine the background level. Once the background level has been determined, it will be subtracted from the raw counts value of each probe.

Background probes

Select groups of background probes (using **negative probes**, **probe sets**, or **tags**) using the drop-down menu and search bar. You can also select **manually** and select probes from the list.

Select the **Background correction** button.

Viewing Background Results

From the **Visualizations pane**, select the **Dataset Summary** button and the **Background** tab.

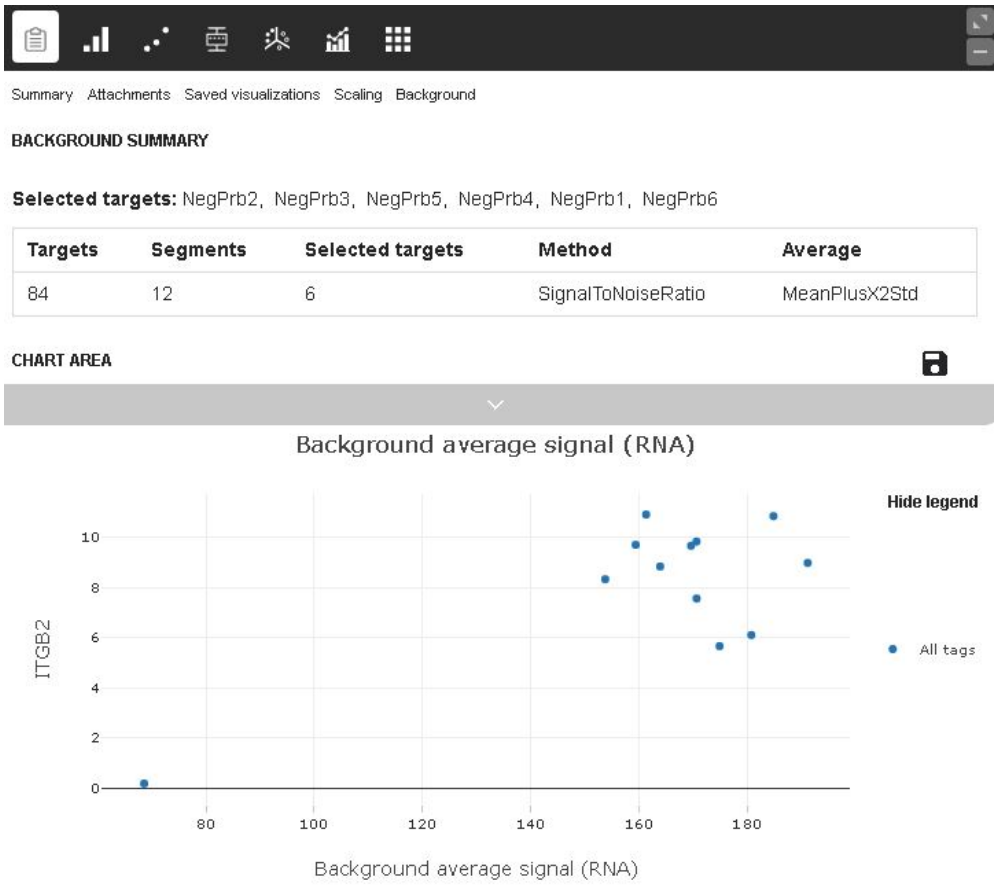


Figure 32: Background results plot

- The **Background Summary** section features a summary of the scans, segments, and probes that were included in the background correction, as well as any of the selections made.
- The **Chart area** (see Figure 32) provides visualizations plotting the background average signal against adjusted counts for the selected probe; change this probe using the **Select probe** field.
- Click the arrow to access the drop down field (see Figure 33). Select **Protein** or **RNA** from the slider to see the respective chart. **Select tags** or **factors** to filter and color the results based on tag.

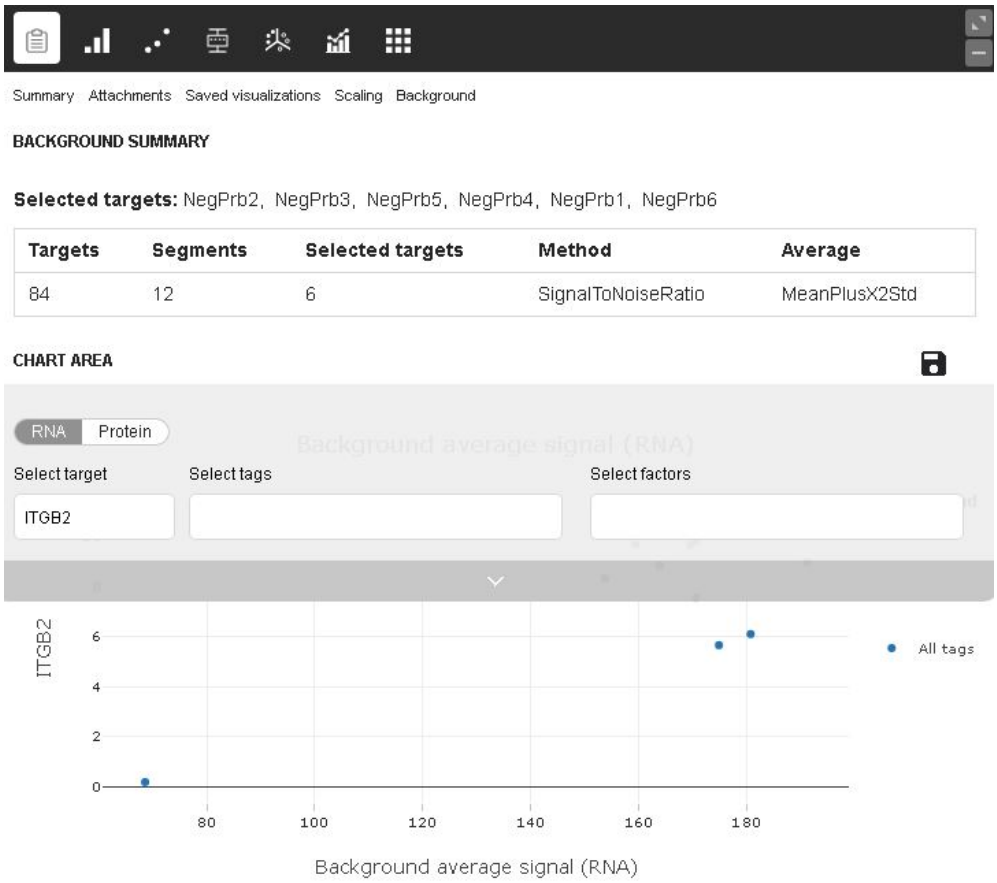


Figure 33: Background summary drop down field

Analysis

Under the **Analysis**, tab, find **Cluster**, **Statistical Analysis**, and **Pathway Analysis** (if eligible).

Cluster

The cluster or dendrogram displays an unsupervised hierarchical clustering of the segments in the study. Clustering is performed based on correlation to determine the position in the dendrogram. The cluster heatmap plots each segment-probe cell according to determined position, with color representing the z- score. When exporting data from the Cluster visualization, you will be exporting only the values of the visualization, which are z-scores.

After each task (QC, Normalization, etc.), the resulting data will be saved as a new dataset and appear in the dataset list (see **Datasets**). Enter a unique name for each dataset, a description (optional), and any tags (optional).

Before running a Cluster analysis, set up your dataset with the desired segments and targets selected. Click the **Cluster** button on the toolbar. In the Cluster window, enter a dataset name and any tags desired. Click the **Calculate** button **(see Figure 34)**.

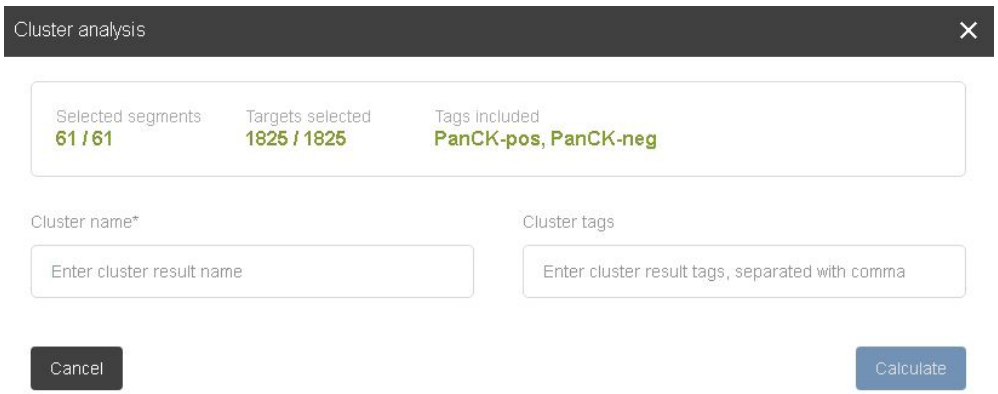


Figure 34: Cluster window

The Cluster analysis will populate under your current dataset. Click the arrow next to the dataset name in the Datasets pane to view the Cluster analysis in the Visualizations pane **(see Figure 35)**.

Data points belonging to the same branch of a cluster are similar to each other at some level; data points in separate branches are less similar.

Segments are aligned along the x-axis and targets on the y-axis.

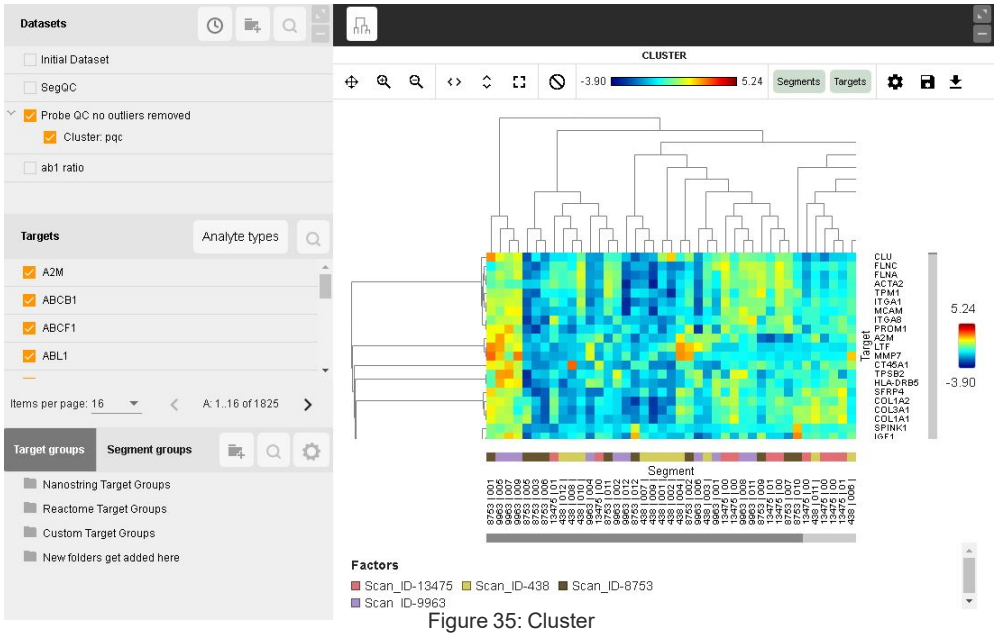


Figure 35: Cluster

- Hover over an area of the dendrogram to see a pop-up displaying the segment, target, count, and any tags.
- Select a branch of the dendrogram to zoom.
- To name the visualization, select the **gear icon** (see Figure 36), enter the desired name, and select **Save**. Use the check boxes to **Export Columns** and **Display Legend**.


Figure 36: Save and gear icon

- To save the visualization, select the **Save** icon. Access the saved image later under the **Summary** tab. For more information on accessing saved visualizations, exporting visualizations, or exporting data from visualizations, see Saved Visualizations on page 82.

Click and drag to select part or all of the dendrogram. Right clicking within this selected area summons a menu (see Figure 37) with which you can:

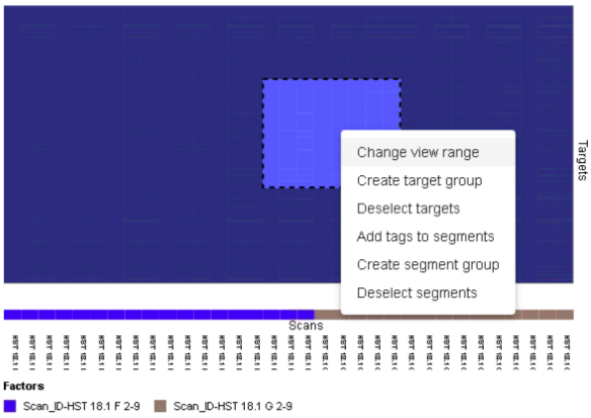


Figure 37: Right-click visualizations menu

- Create a probe group comprised of the selected probes.
- Deselect the selected probes from the current analysis.
- Add tags to the selected segments.
- Create a segment group comprised of the selected segments.
- Deselect the selected segments from the current analysis.

Click the arrow to access the drop-down field (see Figure 38).

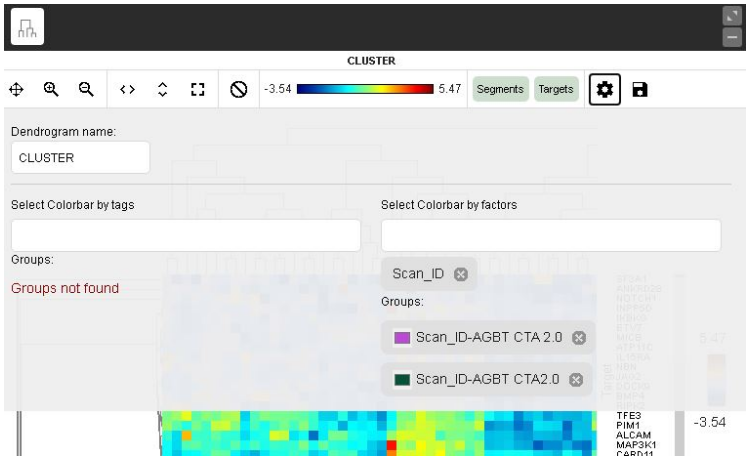


Figure 38: cluster drop down field

Here, you may customize the graph by deciding whether to:

- Use the toolbar at the top to zoom, fit to page (vertically, horizontally, or both), reset view, change the color scheme, or show/hide segments and targets.
- Use the scroll bar at right to scroll the cluster plot through long lists of targets.
- Click the gear icon to name the visualization, choose linear or log data, and establish coloring by tags or factors.

Statistical Tests

The statistical tests tool is under the Analysis drop-down. This function can be used to define groups within a selected dataset and test hypotheses.

When up to two groups are established by selecting tag combinations, the following tests can be run:

- Unpaired T-test
- Paired T-test
- Mann-Whitney U-test
- Linear mixed models

After each task (QC, Normalization, etc.), the resulting data will be saved as a new dataset and appear in the dataset list (see **Datasets**). Enter a unique name for each dataset, a description (optional), and any tags (optional).

When performing statistical analysis, if background subtraction has already been run, beware of statistical anomalies that can occur when values are thresholded to background or subtracted to 1.

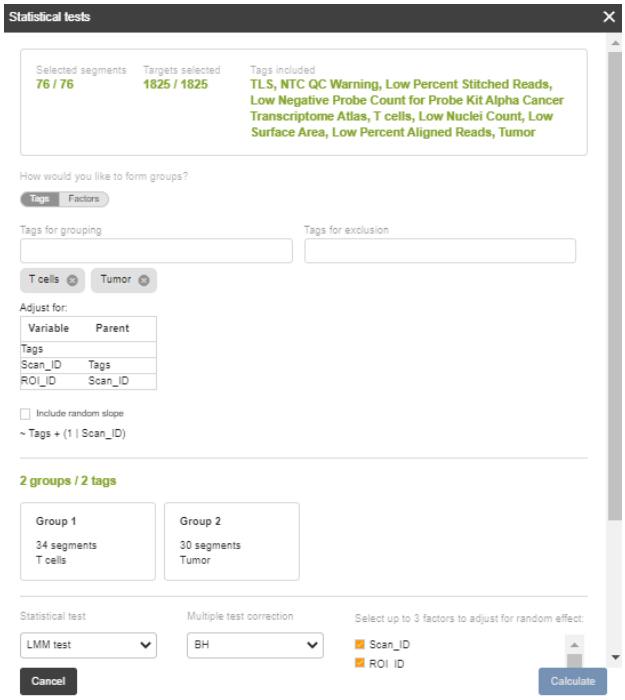


Figure 39: Statistical tests parameters window

1. In the box at the top of the Statistical tests window (see Figure 39), review the selected segments, targets, and tags to be included in this analysis. If these are not correct, exit this window and select or create another dataset.

2. Select the tag you'd like to use in your first group in the **Tags for grouping** field. Select **Add tag** to the right of this field.
3. Select tags you'd like to exclude from your groups in the **Tags for exclusion** field. Select **Add tag** to the right of this field.
4. Review the groups created. For each group, hover over the gear icon to review the segments included in that group; click off any you would like to exclude from analysis. Select **Update** if you made any changes.
5. Select the Statistical Test desired.

   A study with tissue microarray slides and 1 ROI per patient most often should use the T-test. A study with tissue sections and multiple ROIs per patient most often should use the Linear mixed model test.

   - **T-test (unpaired)**: The unpaired t-test should be selected when comparing the means of two groups from independent observations. For example, an unpaired t-test would be used to test whether the expression of a probe in the tumor microenvironment (TME) is significantly different between tissue specimens originating from tumors which responded to treatment vs. tissue specimens originating from tumors which did not respond to treatment. T-tests assume a normal distribution.

     Under the null hypothesis, there is not a significant difference in the mean target probe expression between the two groups.

   - **T-test (paired)**: The paired t-test should be selected to compare means between paired observations, such as between segments originating from the same ROI. For example, a paired t-test would be used in comparing the expression of a probe target in the TME before and after treatment for matched tissue samples. T-tests assume a normal distribution.

     Under the null hypothesis, there is not a significant difference in the mean target probe expression between the two groups.

   - **Mann-Whitney test**: The Mann-Whitney U-test is a non-parametric test of the null hypothesis. It does not require the assumption of normal distribution.

     Under the null hypothesis, the distributions of both populations are equal.

     Under the alternative hypothesis, the distributions of the populations are not equal; a randomly selected observation from one group will be greater or smaller than a randomly selected observation from the second group. The test returns the probability that the median expression of a target probe in one group is greater than the other.

- **Linear Mixed Model**: Linear Mixed Models (LMMs) are designed to handle data with repeated measurements from the same sampling unit. For example, in a study where 6 ROIs are collected from each sample, the ROIs from any one sample will all be influenced by the idiosyncratic biology of that sample, sharing a tendency for higher expression of some genes/proteins and lower expression of others. LMMs explicitly model this tendency.

  Specify at least one random effect to adjust for; the default is Scan_ID. LMM is only envisioned for random effect blocks and is not valid for cross designs. Many levels should be present for each random effect with several segments per group. Not enough replicates for some levels may cause an error in the underlying model code where the number of observations is not enough for the number of random effects. If only one fixed effect variable is tested, consider using subsets of data to explore LMMs separately. Choose random slope when your fixed effect factor has multiple levels within each of one of your random effect factors.

  In software v2.5.1, the LMM lowest level random effect is no longer a requirement, as it was in earlier software versions. This change does not affect the formula for LMM, since earlier software versions automatically dropped the lowest level random effect.

  In contrast to LMM, the t-test is often dramatically wrong in data with repeated measurements. By treating multiple ROIs from the same sample as independent observations, the t- test inappropriately inflates the sample size of the study. For example, given 100 ROIs from one treated and one control sample, the t-test will act as if there are 100 independent samples from the treated and control groups, while a LMM will correctly acknowledge that there is just one independent sample from each group, with many repeated measurements.

  Unlike random intercept only model, a random slope model allows each group line to have a different slope and that means that the random slope model allows the explanatory variable to have a different effect for each group. For example, when each group has different ROI types and their effect is of interest, the variance of group-specific ROI type effect needs to be estimated via the random slope model to properly evaluate the global ROI type effect.

6. Select the Multiple test correction desired. The method chosen will be used to adjust the p-values of individual genes. A gene's False Discovery Rate (FDR) reflects the proportion of significant genes that are expected to be false discoveries. FDR can be used as a more conservative and informative alternative to p-values.

- **None**: this selection is an option to forgo calculating FDR.
- **Permutation q-Value**: this method requires a large number of tests and may not be the most effective method for GeoMx DSP analysis.
- **BH**: the Benjamini-Hochberg procedure is a standard method in controlling FDR and is the **recommended method to be used for GeoMx DSP protein analysis**. It assumes that the tests are positively correlated for a subset of genes/proteins, which usually holds for gene/protein expression data.
- **BY**: the Benjamini-Yekutieli procedure is an adjustment to BH. It controls FDR in a more general setting, but poses a more strict threshold that leads to fewer significant calls. With the BY method, it's very unlikely to identify any significantly differentially expressed genes in DSP analysis.

7. Enter a unique Test result name and list the test result tags.
8. Select **Calculate.**

Viewing Statistical Test Results

The results from the statistical tests task can be accessed by clicking the arrow next to the dataset on which the test was run in the second pane. Select the test of interest from the drop-down list and the summary of results will appear in the 3rd Data Analysis pane.

Pathway Analysis

> **For NGS readout:** Pathway Analysis is only available for NGS datasets.

The Pathway Analysis button is only active with NGS datasets after a t-test or linear mixed model has been run in datasets with 1000 or more targets. Due to the number of required targets, Pathway Analysis is not available for most NGS Protein studies.

1. Select a dataset that fits the above criteria and click the **Pathway analysis** button.
2. In the Pathway analysis window (see Figure 40), set thresholds based on suitable parameters for your study and click the **Run** button. This enters the fold change values for each target from the statistical test results into a gene set enrichment analysis (GSEA).

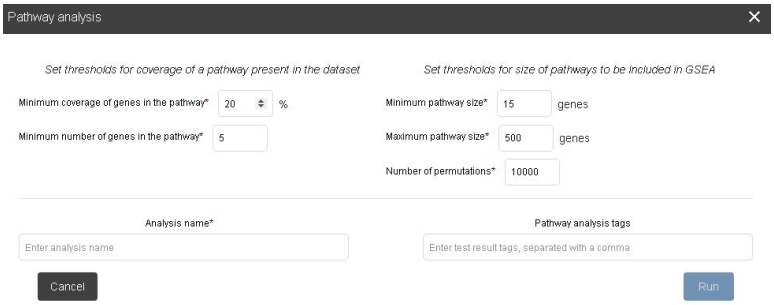


Figure 40: Pathway Analysis window

Resulting visualization options include the **Pathway Analysis Table**, **Bargraph**, **Volcano plot**, and **Pathways map**.

Select pathways of interest in the table (see Figure 41) to activate the **Create Target Group** button. Click on a point (pathway) of interest on the Volcano plot to reveal the heatmap of targets within that pathway below the volcano plot. Search for pathways using the search fields.

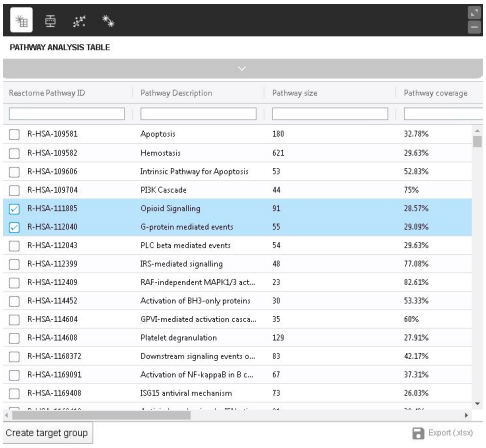


Figure 41: Pathway Analysis window

Ratio Builder

Ratio Builder allows you to create ratios and estimate the relative expression of target probes. This can be used to compare expression of targets in a segment to the average of a group of segments delineated by tags, compare expression of a target between segment pairs from within an ROI, or estimate relative expression of a target probe within a segment.

Ratios can only be built for the segments and probes that are actively selected in the Scans pane.

Choose one of the three tab options **(see Figure 42)** to build ratios:

After each task (QC, Normalization, etc.), the resulting data will be saved as a new dataset and appear in the dataset list (see **Datasets**). Enter a unique name for each dataset, a description (optional), and any tags (optional).

- The **Ratio to a segment/group of segments** option calculates the ratio of each probe in each segment to the average of one or more segments selected as the baseline (denominator).
- The **Binary ratios** option calculates the relative expression of selected targets between a pair of segments belonging to the same ROI. For example, a binary ratio could compare a particular target probe's abundance between a tumor segment and a TME segment in a common ROI.
- The **Target ratios** option calculates the relative expression of a probe pair within each segment. For example, one could use the probe ratio function to create a dataset comparing phospho-AKT counts to AKT counts for all selected segments in the dataset.

Ratio to a segment/group of segments

1. Before you open Ratio Builder, select the scans and segments you would like to include in the numerator in the Scans pane (see **Scans Pane on page 69**). To use tags to filter segments, see **Tags on page 66**.

   It is important to select the appropriate segments and probes from the dataset prior to specifying parameters for ratio building; it is also important to choose the appropriate tags in selecting denominators.

2. Open **Ratio Builder** and confirm you are on the **Ratio to a segment/group of segments** tab **(see Figure 42)**.

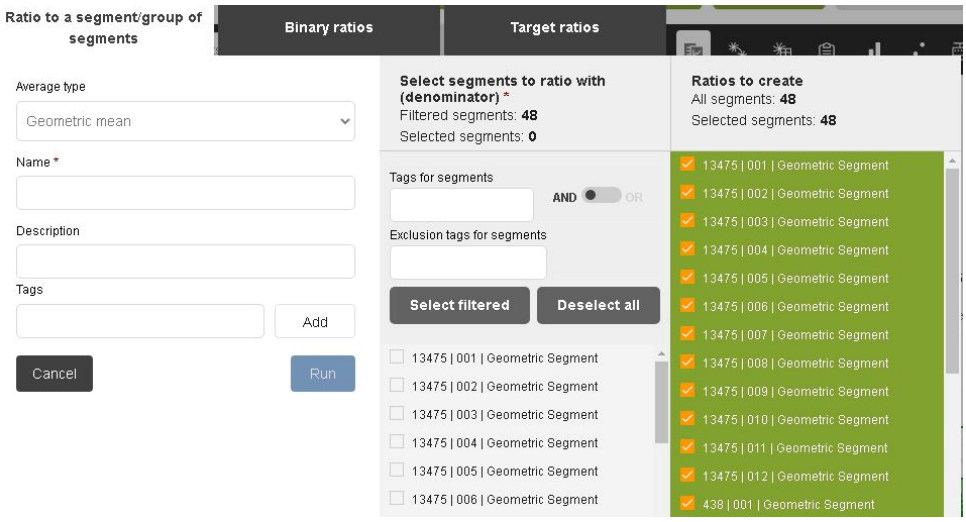


Figure 42: Ratio to average parameters window

3. Choose the **average type** you'd like to use: Geometric mean (default and recommended), Median, Average, Sum, Minimum or Maximum.
4. In the central field, select the segment(s) you would like to include in the **denominator** of your ratio. Use the following steps to find segments of interest:
   - Use the search bar here to filter segments by tags. Select a tag from this drop-down.
   - The number of filtered segments will be reflected at the top of the central field and the names of the filtered segments will be listed in the central field. Review the list.
   - Click **Select filtered** to include all the segments in this list in the denominator or individually select the ones you'd like included.
   - All segments that are selected for the denominator will be averaged without further stratifying the data.
5. In the right-most field, all segments chosen for the numerator will be listed, indicating that a ratio for each of these segments will be created. Click off any that you do not wish to create.
6. In the left-most field, enter a unique name for the resulting dataset, a description (optional), and any tags (optional).
7. Click the **Run** button. You will be returned to the main Data Analysis screen; your new dataset will be selected and the live visualization will reflect this, as well.

Binary ratios

The binary ratios feature calculates the relative expression of a target between a pair of segments belonging to the same ROI. Ensure that the input data includes segments from a common ROI and that each segment type has a distinct tag.

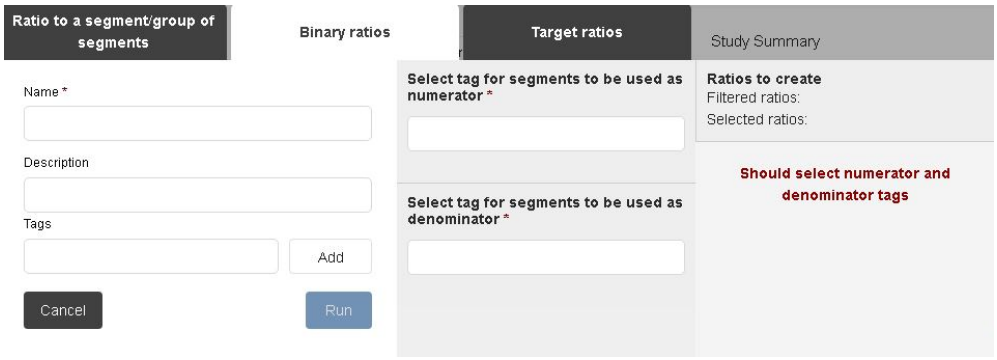


Figure 43: Binary ratios parameters window

1. Before you open Ratio Builder, select the scans and segments you would like to include in the numerator in the Scans pane (see Scans Pane on page 69). To use tags to filter segments, see Tags on page 66.
2. Open **Ratio Builder** and click on the **Binary ratios** tab (see Figure 43).
3. In the central field, select the tag(s) you would like to choose for the **numerator**, then the **denominator** of your ratio. See note above. Use the search bars to select tags.
4. In the right-most field, review the ratios to be built. Click off any you do not wish to create.
5. In the left-most field, enter a unique name for the resulting dataset, a description (optional), and any tags (optional).
6. Click the **Run** button. You will be returned to the Data Analysis screen; your new dataset will be selected and the live visualization will reflect this, as well.

Target ratios

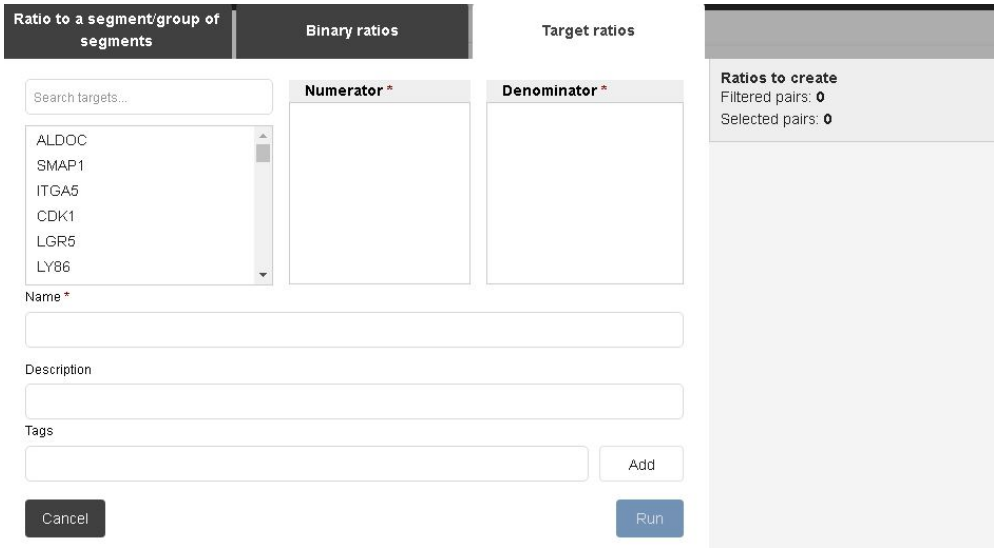


Figure 44: Probe ratios parameters window

1. Before you open Ratio Builder, select the scans and segments you would like to include in the numerator in the Scans pane (see Scans Pane on page 69). To use tags to filter segments, see Tags on page 66.

   The Probe Ratios feature calculates the relative expression of a probe within a segment. Ensure you have selected probes from a common segment.

2. Open **Ratio Builder** and click on the **Probe ratios** tab (see Figure 44).
3. In the left-most field, click and drag probes from the list into the **numerator** and **denominator** boxes. Use the search bar to find probes, if necessary.
4. In the right-most field, review the ratios to be built.
5. In the left-most field, enter a unique name for the resulting dataset, a description (optional), and any tags (optional).
6. Click the **Run** button. You will be returned to the Data Analysis screen; your new dataset will be selected and the live visualization will reflect this, as well.

Once you run **Ratio Builder**, certain toolbar buttons will be inactive. Returning to the Initial Dataset (checking the Initial Dataset box in the central pane) activates the buttons again.

Custom Scripts

Custom Scripts allows you to run custom R scripts on your data. Find custom scripts developed by NanoString in the NanoString GeoScript™ Hub or create your own.

Adding a Custom Script

1. Select a dataset that has been through **QC**.
2. Select the **Custom Scripts** button from the **Toolbar Options**.
3. On the **Manage** tab, click the **Add** button **(see Figure 45)**.
4. Enter a name, description (optional) and any tags associated (optional).
5. In the **Main Scripts** field, click the **+** button.
6. Browse for the desired R file on a USB or, if accessing remotely, in remote file location. Select it, and click **Open**.

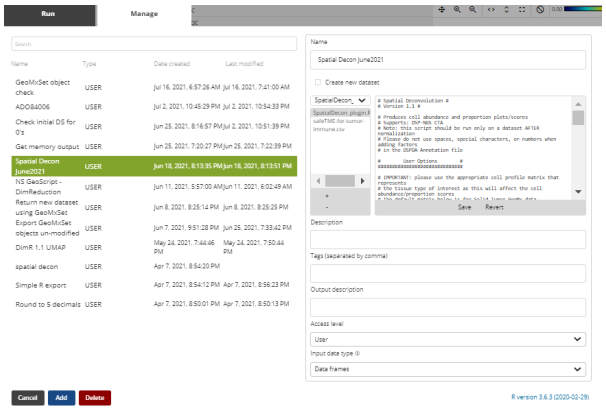


Figure 45: Adding a custom script

7. Click the **Save** button. The script will appear in the **Main Script** box. You can manually make changes here.

   Text at the top of the script may include information on the environment recommended to run the script. **Data frames** is typically used for existing scripts and **GeoMxSet** for new scripts. Select the environment from the **Input data type** field at the bottom of the window.

8. Click **Save** in the **Main Script** box to save any changes made, or **Revert** to go back to the originally uploaded version. This script will now be available to run.

Running a Custom Script

1. Select the **Custom Scripts** button from the **Toolbar Options**.

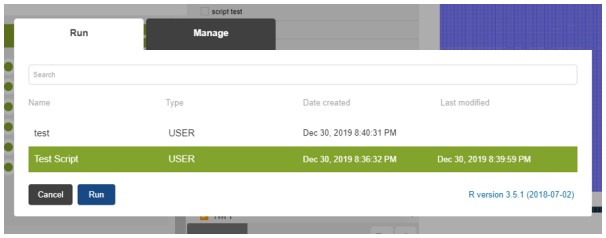


Figure 46: Running a custom script

2. On the **Run** tab, select a saved script from the list and select **Run** **(see Figure 46)**.
3. Select the **Manage** tab to Add, Edit, or Delete a script **(see Figure 47)**.
4. Check the **Create new dataset** box to save as a distinct dataset.
5. A **Successful!** message will appear, notifying you that the custom R script successfully executed. Select **OK**.
6. Navigate to the Visualizations pane of the Data Analysis Suite. Click on the **Dataset Summary** icon and select the **Attachments** tab. The Custom Scripts you just ran should appear.
7. Select the **Save** icon to download this visualization or the **Trash** icon to discard it.
8. Upon running a Custom Script and closing the Custom Scripts window, the original dataset may still be selected in the Datasets pane. Be sure to select the dataset generated by the Custom Script to view the output of the script.

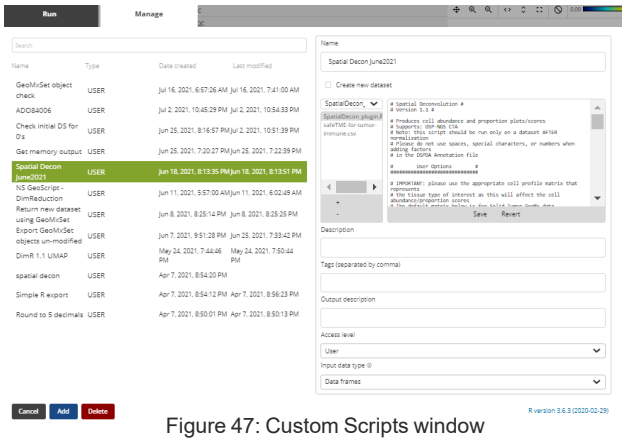


Figure 47: Custom Scripts window

Save Subset

The **Save Subset** function is available to save your current dataset settings and selections as a distinct dataset. For example, you may use the search bar to select and remove a subset of undesirable segments from a dataset (such as flagged QC) or to separate segments of interest to analyze separately (such as only tumor segments). Each of these selections can be saved as a subset.

To save your current dataset settings and selections, select the **Save subset** button.

Define: **Name**, **Description** (optional), and **Tags** (optional). Select **Save.**

Spatial Graphing

Spatial Graphing allows you to overlay a bubble plot of target expression over your scan image (see Figure 48).

1. Click the **Spatial Graphing** button from the toolbar.
2. Click the drop down arrow to view the parameters.
3. Select a **Target**.
4. Adjust parameters to affect the values displayed and size range of the bubbles. Use tag or factor to change color of displayed bubbles.
5. Click **Update** to see changes.
6. Drag and drop scans (if desired) to change order of appearance in window.

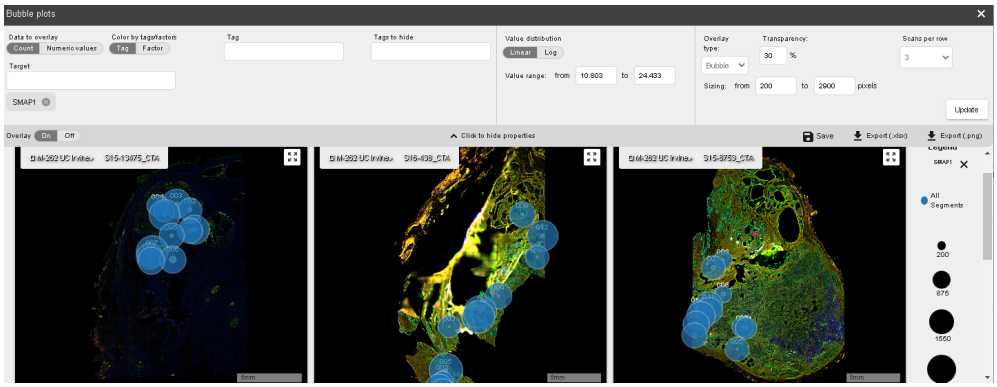


Figure 48: Spatial Graphing

Data Summary Report (Beta)

The Report tool in the GeoMx DSP Data Analysis Suite provides you with scan images and a subset of saved visualizations from your study. It allows you to share your data with collaborators as an HTML file. Note that once a report is generated, it replaces the previous report on the instrument.

Creating a Report

1. First, save any visualizations you want to share in your report. Note that Spatial Graphing and Pathway analysis plots are not available to share in the report.
2. Click the **Report** button on the Toolbar.
3. Fill out **Report Information**, including Project Description (information about your study) and Study Executive Summary (notes on your analysis). Names should not exceed 255 characters; Description should not exceed 1000 characters; and Executive Summary should not exceed 5000 characters. Do not use special characters in the Report Title.
4. Make selections on the **Include in Report** field.

- **Scans** will autopopulate.
- Under **Datasets**, click the drop down arrow to choose the dataset of interest (the default will be the most recently generated). Note that all datasets that directly preceded that dataset will be included automatically (if any dataset has multiple analysis paths, only the parent of the selected datasets will be included).
- Below, the available visualizations for this dataset **and the visualizations for each preceding dataset** will appear. Click the **x** on any visualization you would like to exclude from the Report.
- Choose **Targets** to use as markers. These are targets whose expression you might like to use for validating staining, etc. Choose up to four. They will generate a bar chart displaying expression data for each ROI/segment.

5. Click **Generate Report**. The **% Complete** will appear in the upper left of the header.

Downloading and Viewing a Report

1. A **Download** icon will appear in the upper right of the Data Analysis header when a report has finished generating and is ready to view. Click the icon to download.
2. Locate the downloaded file and unzip.
3. Open the **Open Me** HTML file in a Chrome browser.

4. Information about the report will appear at the top left: the name of study, the time of report generation, and the version number (related to how many reports have been generated from a single study.
5. The Report has tabs along the top.

- The **Summary** tab has study information.
- The **GeoMx** tab has an overview of the technology.
- The **Scan List** tab includes the scan images and a thumbnail of each ROI.
- One tab for each dataset included in your Report will be present in the header. The **Dataset** tabs include visualizations for the chosen dataset and all its preceding datasets, a pipeline view of preceding datasets, QC, and normalization metrics applied to each dataset.

6. In the **Scan List**:

- Use the **segment** button to view/hide each segment.
- Click on each scan for scan information or on each ROI for information on that ROI.
- Click the **bar graph** button to see a bar graph of the chosen markers below the scan.

7. In each Dataset tab:

- Click the triangle to choose whether to display scans or a scan with visualization.
- Click **Save** to save a .svg file.

Additional Notes

- Some images in the report are interactive with zoom, pan, and tooltips, while others are static.
- JPG images of scans are available in the report folder: <report folder name>\data\assets\data\scans
- Generating multiple reports from a single study results in multiple versions.
- Prior to building a report, save your visualizations with a name of your choosing. The plot title shown in the report pulls from the saved visualization name.
- If a scan name has a leading or trailing space, it will not appear in the report.
- Email geomxsupport@nanostring.com with user feedback on the Data Summary Report Tool while it is in Beta phase.Your feedback will help us improve this functionality so that it best fits your needs.

Workflows (Beta)

Workflows are a new feature of the GeoMx software v2.5. Workflows combine multiple data analysis steps into a single streamlined operation. Ten workflows are pre-loaded into the GeoMx DSP Data Analysis suite. Their parameters can be customized to your specifications. At this time, it is not possible to create a new workflow from scratch.

Available workflows:

- **RNA NGS Workflow 1**: First-pass quality control and normalization. This workflow performs Q3 normalization and cluster analysis on all remaining targets after QC and filtering. This workflow also produces a negative normalized dataset for comparison.
- **RNA NGS Workflow 2**: Statistical test and pathway analysis. This workflow requires annotations and modification of the statistical test parameters to complete fully. This workflow performs a statistical test and pathway analysis on a Q3 normalized dataset after QC and filtering steps.
- **RNA NGS Workflow 3**: Normalization after modified QC and filtering. This workflow performs segment QC with modified parameters and removes failed segments during biological probe QC. Then a modified filtering step keeps targets as long as they are above LOQ in more than 1% of segments. The resulting data are finally Q3 normalized and a cluster analysis is performed using all targets remaining after the modified filtering.
- **RNA NGS Workflow 4**: Compare normalization options. This workflow requires the **Evaluate Normalization for Protein script** from the NanoString GeoScript Hub to be set up in the **Custom Scripts** menu under name **Evaluate normalization** (the script works for protein and RNA studies). Once configured, this workflow will provide the output from this script as an attachment to the filtered dataset. This workflow also performs cluster analysis on separate Q3 and Negative normalized datasets. This workflow completes segment QC and biological probe QC with an additional step to filter out segments based on QC flags.
- **Protein NGS Workflow 1**: Compare normalization options. This workflow requires the **Evaluate Normalization for Protein script** from the NanoString GeoScript Hub to be set up in the **Custom Scripts** menu under the name **Evaluate normalization**. Once configured, this workflow will provide the output from this script as an attachment to the filtered dataset. This workflow also generates four datasets to compare normalization methods: scaling to area, scaling to nuclei, normalization to housekeeper genes, and a background correction using signal-to-noise ratio.

- **Protein NGS Workflow 2**: Compare normalization options without custom script. This workflow generates four datasets to compare normalization methods: scaling to area, scaling to nuclei, normalization to housekeeper genes, and a background correction using signal-to-noise ratio. A final dataset is filtered from the signal-to-noise dataset.
- **Protein NGS Workflow 3**: Compare normalization options after filter. This workflow requires the **Evaluate Normalization for Protein script** from the NanoString GeoScript Hub to be set up in the **Custom Scripts** menu under name **Evaluate normalization**. Once configured, this workflow will provide the output from this script as an attachment to the filtered dataset. This workflow also performs cluster analysis on separate Q3 and Negative normalized datasets. This workflow completes segment QC and biological probe QC with an additional step to filter out segments based on QC flags.
- **RNA nCounter Workflow 1**: Compare normalization option after segment filter. This workflow requires the **Evaluate Normalization for Protein script** from the NanoString GeoScript Hub to be set up in the **Custom Scripts** menu under name **Evaluate normalization** (despite the name, the script works for protein and RNA studies). Once configured, this workflow will provide the output from this script as an attachment to the filtered dataset. This workflow also generates four datasets to compare normalization methods: scaling to area, scaling to nuclei, normalization to housekeeper genes, and a background correction using signal-to-noise ratio.
- **Protein nCounter Workflow 1**: Compare normalization options after segment filter. This workflow requires the **Evaluate Normalization for Protein script** from the NanoString GeoScript Hub to be set up in the **Custom Scripts** menu under name **Evaluate normalization**. Once configured, this workflow will provide the output from this script as an attachment to the filtered dataset. This workflow also generates four datasets to compare normalization methods: scaling to area, scaling to nuclei, normalization to housekeeper genes, and a background correction using signal-to-noise ratio.
- **Protein nCounter Workflow 2**: QC, normalization and statistical test. This workflow requires annotations and selection of factors for the statistical test step. Once configured, this workflow will perform a statistical test and cluster analysis after stepping through QC, filtering out segments flagged during QC, and normalizing the remaining data to the housekeeper genes.

To run a data analysis workflow:

1. Click on **Workflows - Beta** to open the Run Workflow window.
2. See the compatible workflows for this dataset in the left **Workflows** pane. Hovering on the workflow name brings up a tooltip with the workflow description. Clicking on the workflow name brings up a schematic of the workflow steps in the **Workflow Structure** pane (see Figure 49). Use the **Filter** field to locate a particular workflow.
3. When a workflow is displayed in the **Workflow Structure** pane, hover or click on a step to see the **Selected Step Settings** in the right pane. Pan around the Workflow Structure pane by clicking and dragging the hand icon in the white space.

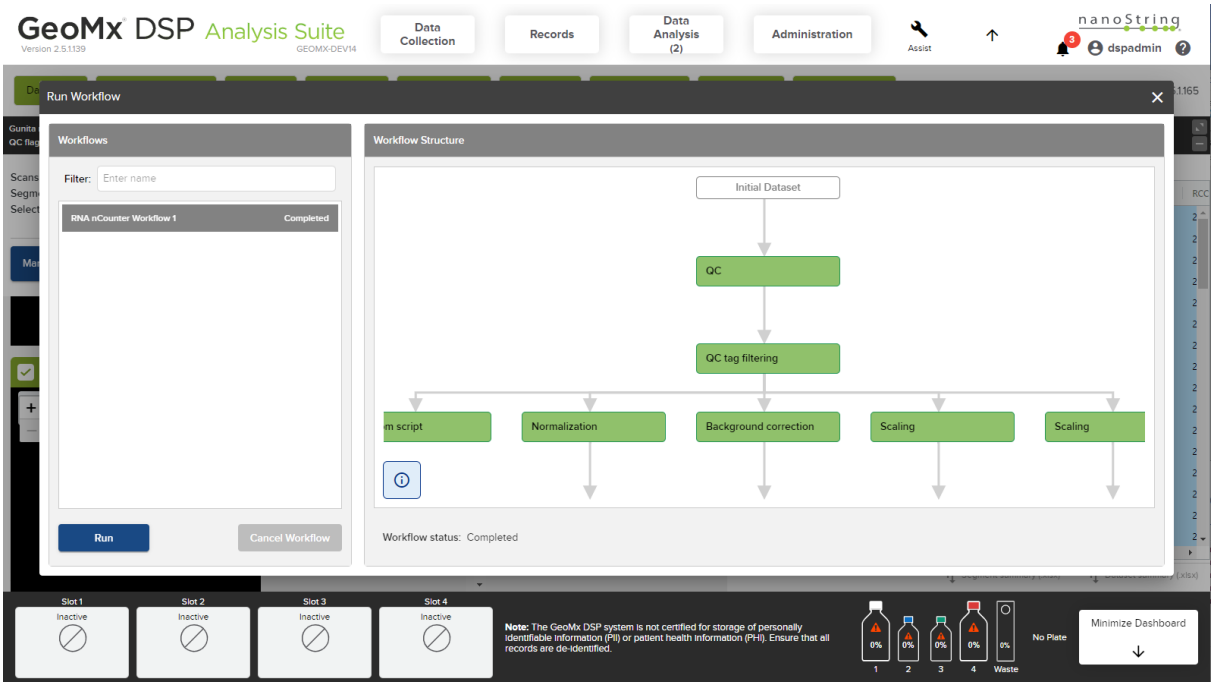


Figure 49: Run Workflow window

4. **Review workflow parameters** in the Selected Step Settings pane to check that the QC parameters match what you want to run. The default settings may be too restrictive, filtering out most data points and leading to downstream errors.
5. To edit the parameters for a particular step, click the **pencil** icon next to the step name in the Workflow Structure pane, or **Edit Parameters** in the Selected Step Settings pane (see Figure 50).
6. If you make changes to a workflow, you'll be prompted to save it with a new name. NanoString recommends using a naming scheme for your workflows, and relying on the Filter function to

locate workflows, since it is currently not possible to delete workflows. Saved workflows are only accessible within the study they were saved in.

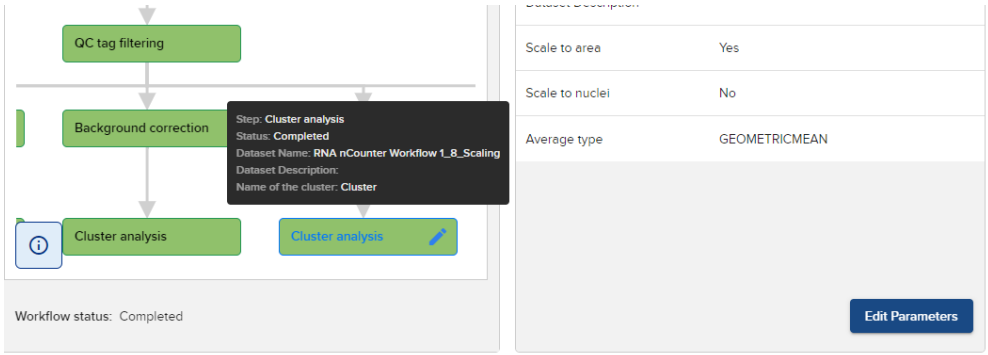


Figure 50: Edit parameters and pencil icon.

7. Click **Run**. The **Run Workflow** window opens. Name the run and enter any tags you wish to apply to the output dataset. Click **Run**.
8. **Workflow status** is shown along the bottom of the Run Workflow window. The steps of the workflow change from gray to yellow to green as they are completed. Click the **X** in the top right to close the Run Workflow window. The workflow will continue to run in the background.
9. The **dataset** produced by the workflow appears as a folder in the **Datasets pane**. Click on the dataset folder name to expand the datasets for each step within the workflow **(see Figure 51)**.
10. If you don't see the expected dataset output (or target group/segment group output), refresh the page (Ctrl+R or F5 key) to update the Datasets pane.

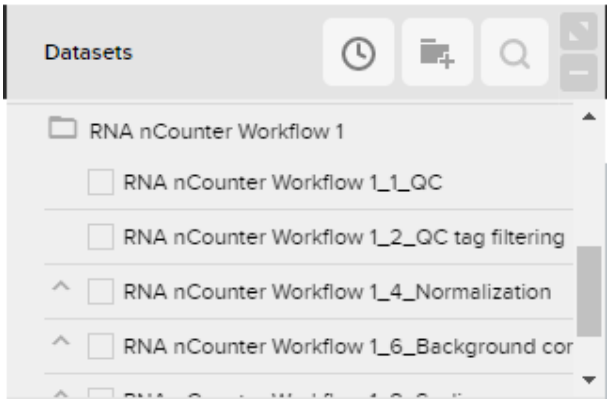


Figure 51: Dataset output from a workflow.

Email geomxsupport@nanostring.com with user feedback on the Workflows Tool while it is in Beta phase. Your feedback will help us improve this functionality so that it best fits your needs.

Analyze in Spatial Context

Use the inter-connectivity of the three panes **(see Figure 52)** to re-examine segments in spatial context. See **Scans Pane on page 69**, **Datasets Pane on page 71**, and **Visualizations Pane on page 78** for more information on each pane's features and how they interact.

On most visualizations, you can select a set of data points representing segments and add tags to them. In the heatmap and cluster heatmap, you can click-and-drag a box across one or more columns, selecting those segments of interest. You can then right-click and add the tags you'd like to be associated with the selected segments.

Turning to the Scans pane, you'll see that the selected segments in the plots are indicated by pink picker buttons. When you zoom into the scan image, you will see that the selected segments are highlighted with a pink border to easily identify them and to evaluate their spatial context.

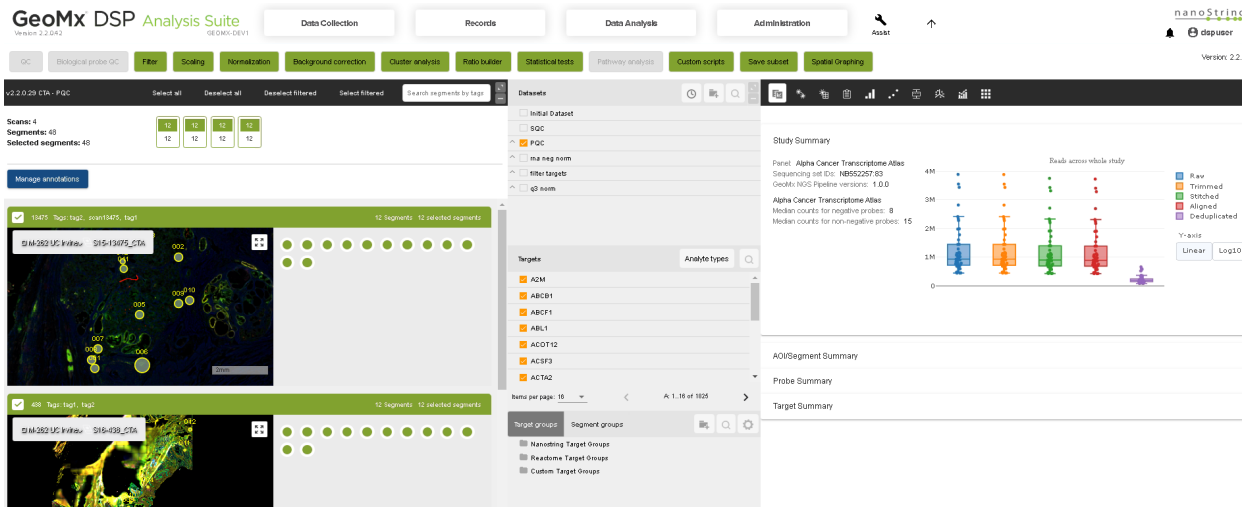


Figure 52: Data Analysis Suite

Tags

Tags allow you to group segments by type and can then be used to categorize and filter data for analysis.

Creating a new tag for a specific group of segments

1. Select only those segments of interest in the first Data Analysis pane (see Figure 53).

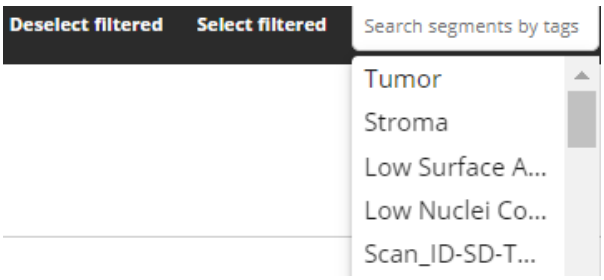


Figure 53: Selecting segments using segment picker buttons

2. Click and drag a box across the whole resulting heatmap in the third (visualizations) pane, such that the box includes all segments (it does not matter what probes are included in this box).
3. Right-click on this selection box. From the resulting menu (see Figure 54), select **Add tags to segments**.

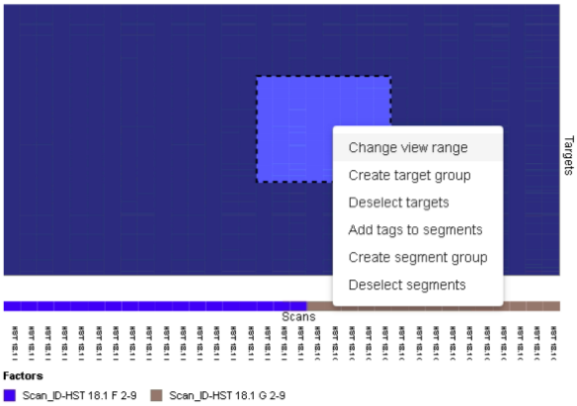


Figure 54: Adding tags to selected segments

4. In the **Add tags to segments** window (see Figure 55), review the list of **Selected segments**, then type the name of the first tag you would like to add to these segments and select the gray **Add tag** button.
5. Type the name of any additional tags you would like to add to these segments and select the gray **Add tag** button for each.
6. Select the blue **Add tags** button when you are done tagging this group of segments.

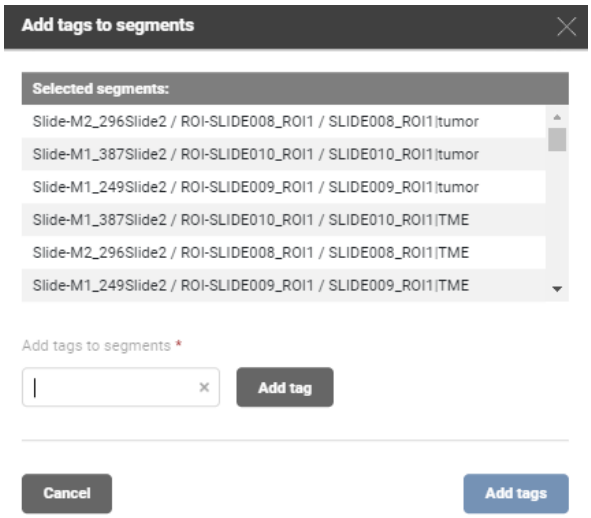


Figure 55: Segments selected for tag

Filtering data based on tags

1. In most cases, you'll want to start with all scans and segments selected; each slide icon should be green (see Figure 56); if a slide icon is orange or gray, click on it until it toggles to the green status (meaning all segments are selected).
2. Click in the search bar at the top of the Data Analysis window to select an available tag(s). To create a new tag, see Tags on page 66.

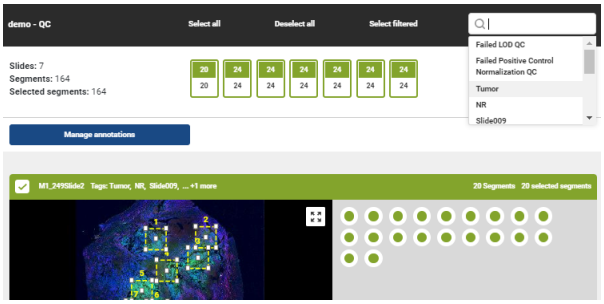


Figure 56: All scans selected

3. Any available tags will appear in the drop-down. Select the first tag of interest. Typing into the search box will limit the list of tags displayed to those that have at least a partial match to the typed term.

4. The chosen tag will appear as a gray button and the scan icons will show a preview of the segments included in this selection **(see Figure 57)**. Note that the filter *has not yet been applied*; the segment picker buttons next to each image viewer indicate the current selection.

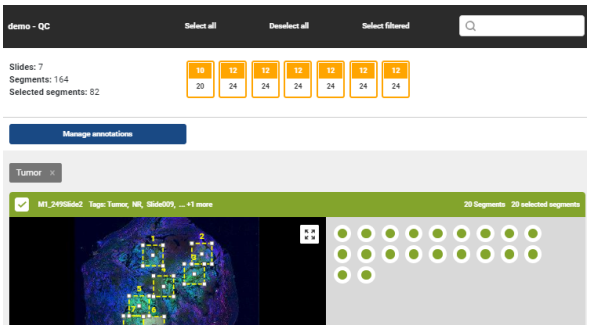


Figure 57: Tag selected for filtering, but not yet applied

5. To apply this filter to the analysis, click **Select Filtered**. Alternatively, you may click **Deselect Filtered** to remove the segments associated with this tag from analysis. The selected scans and segments should change dynamically to reflect this tag choice **(see Figure 58)**. In addition, the plot in the visualization pane will be redrawn to reflect the current selection.

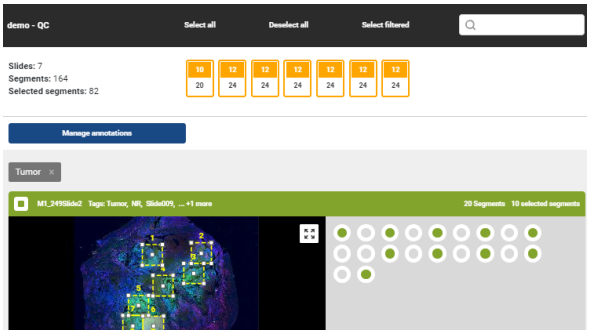


Figure 58: Segment selection, filtered by tag

6. Repeat this process to filter by additional tags, if desired. Select the **x** on a button to stop filtering using that tag.

Selections can be changed in a number of ways including by clicking off a gray tag box (see step 4), by clicking Select all/Deselect all, or by adding additional filters. To save your current selection, select the **Save Subset** button from the Toolbar.

Scans Pane

In the first pane of the DSP Data Analysis interface (see Figure 59), you may select the scans and segments to include in your study. All scans in the study will be represented as scan icons at the top of this pane and as scan images listed below. Scan images may be viewed in **Thumbnail** or **Interactive** mode.

Use the Maximize button to make this pane fill the screen or the Minimize button to hide it as a tab.

Select All/Deselect All

Select all and **Deselect all** buttons at the top of the window select and deselect (respectively) all scans and all segments within them.

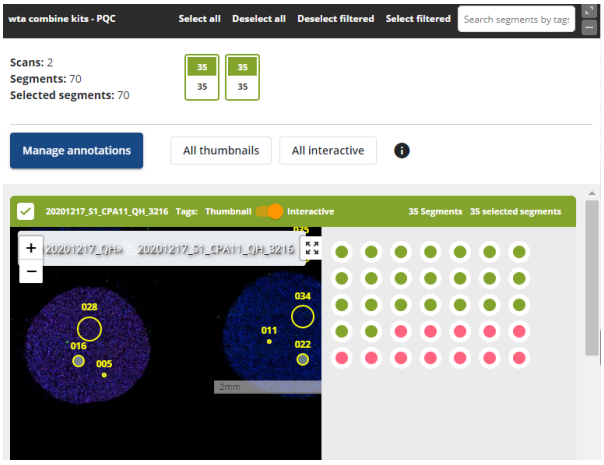


Figure 59: Scans pane

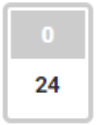
Scan icons

Scan icons provide a visual preview of:

- The number of segments selected for analysis (top half)
- The total number of segments on the scan (bottom half)
- The general proportion of segments selected for analysis:

All segments selected for analysis is indicated by a green outline and top half.

Some segments selected for analysis is indicated by an orange outline and top half.



No segments selected for analysis is indicated by a gray outline and top half.

Click on a slide icon to toggle between selecting (green or orange) or deselecting (gray) for analysis. Specify which segments will be selected using the slide images (see **Image Viewer on page 70**).

Manage Annotations

Segment annotations can be uploaded to a Data Analysis study from a spreadsheet using the Manage Annotations button in the Scans Pane. See **Manage Annotations on page 30**.

Image Viewer

Each image viewer **(see Figure 60)** portrays the scan and the spatial placement of the ROIs and segments. The check box in the upper left corner indicates whether that scan is selected for analysis.

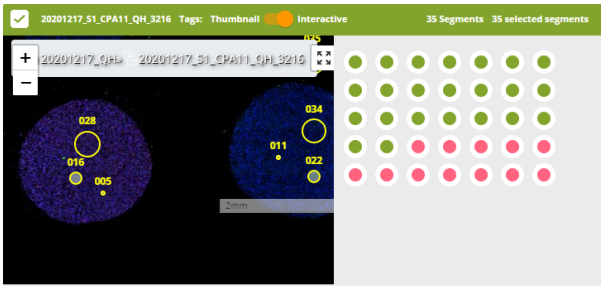


Figure 60: Image viewer and segment picker buttons

- Selected scans have a green header
- Deselected scans have a white header.

You may adjust the scan image to assist in viewing, selecting, and deselecting segments.

- Toggle between **Thumbnail** and **Interactive** modes using the slider button in the header. Use the **All Thumbnails** button or **All Interactive** button at the top of the Scans pane to toggle all image viewers at once.
- **Zoom in** and **out** using the scroll wheel on your mouse or zoom feature on your touch screen or touch pad.
- Click the **full-screen button** to expand the image to full-screen.

Each ROI in the image corresponds with one or more segment picker buttons to the right of the image. Hovering over an ROI on the image viewer will initiate a pop-up providing more information.

Datasets Pane

Use the Maximize button to make this pane fill the screen or the Minimize button to hide it as a tab.

The central pane of the DSP Data Analysis screen lists all datasets and target groups associated with the current study. At a minimum, you will see the **Initial Dataset** (the raw set of imported data; this will appear at the top of the Dataset list) in the Data sets field. The **All Targets** group, as well as any additional target groups defined in your core and module kit configuration files (see the GeoMx DSP Instrument User Manual (MAN-10152) for details regarding Scan Parameters) will populate the Target groups field.

Datasets

Anytime you run a task (QC, Normalization, etc.), you will be prompted to save the updated dataset. This new dataset will be listed in the dataset list (see Figure 61).

Select the **new folder icon** on the **Datasets** header to create a new folder to organize your datasets. Click and drag datasets to place them in the folder of your choice.

Select the search button on the **Datasets** header to search by tag or text.

To save a selected subset of scans and/or segments as a distinct dataset, use the **Save Subset** button on the Toolbar.

Manage Datasets

- At any time, you may select the **History** icon at the top of the datasets pane (see Figure 61) to see the adjustments that have been made to the selected dataset.

Figure 61: History bar in Datasets pane

- An **arrow** next to a dataset indicates that statistical tests have been run on that dataset; click on this **arrow** and select a test result box (see Figure 62) to see visualizations with those results.

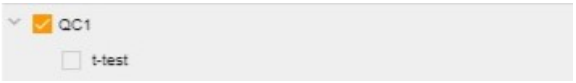


Figure 62: Statistical tests datasets

- Hovering over a dataset name reveals **export**, **edit**, and **trash** icons (see Figure 63).
- Export the whole dataset in .xlsx format by selecting the **export** icon that appears when you hover over it. The downloaded file will be a multi-tab spreadsheet.
- Rename the dataset by selecting the **edit** icon that appears when you hover over it.

Figure 63: Export, edit, and trash icons in the datasets pane

- Delete the dataset by selecting the **trash** icon that appears when you hover over it.

Merge Datasets

You may merge two datasets into one if the two datasets have either:

- **Different segments** from each other but the *same set of probes* or
- **Different probes** from each other but the *same set of segments*.

To merge:

1. Select the first dataset you'd like to use in the merge.
2. Hold down the Control (**Ctrl**) button and select any other datasets you would like to merge with the first.
3. Right-click and select **Merge datasets**.
4. Confirm or edit your selections using the check boxes. If a message appears, prohibiting this merge, revisit the restrictions listed above.
5. Define the **Name**, **Description** (optional), and **Tags**.
6. Select **Merge**.

Targets

All targets detected in the configuration file (see the GeoMx DSP Instrument User Manual (MAN-10152) for details regarding Scan Parameters) of the dataset will be listed in the probe list in the Datasets pane; long lists of targets will load in "pages". Each target with a checked box will be included in the study (see Figure 64).

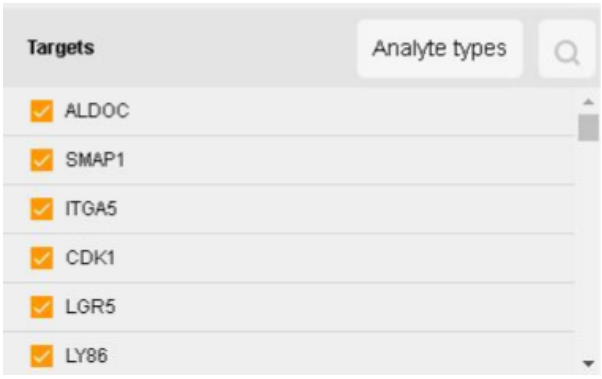


Figure 64: Probes list section of datasets pane

Selecting/deselecting targets

Uncheck boxes to remove targets from analysis.

Filter targets using one of the following buttons:

- Analyte Type
- Search: use the **Search** button to search for probes by **text** and/or by **tag.**

Target Groups & Segment Groups

Target groups and segment groups are listed in the Datasets pane (see Figure 65) . The folder **NanoString Target Groups** contains pre-defined target groups defined in the core or module kit configuration file (see the GeoMx DSP Instrument User Manual (MAN- 10152). Any new target groups you create will be saved in the folder **Custom Target Groups** . Each group is listed with a checkbox that indicates selected or deselected.

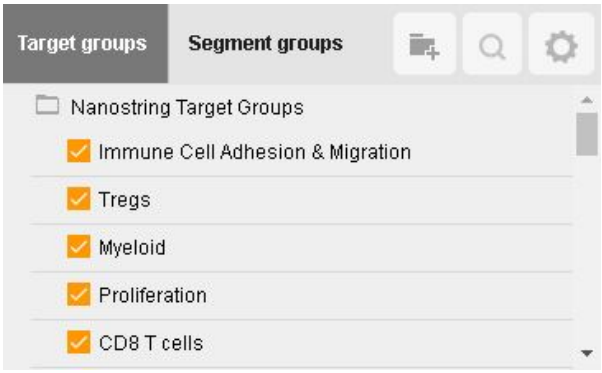


Figure 65: Probe and segments group section of Datasets pane

- Orange checkbox: all targets selected
- Grey checkbox: some targets selected
- Unchecked checkbox: no targets selected

Select the **new folder icon** on the **Target/Segment groups** header to create a new folder to organize your target groups or segment groups. Click and drag target/segment groups to place them in the folder of your choice.

Select the **search button** on the **Target/Segment groups** header to search by tag or text.

Select the **gear icon** for the Groups management window (see Figure 66). Use this window to edit target groups.

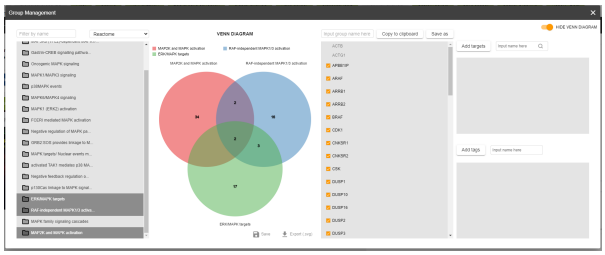


Figure 66: Groups management window

- To see the overlap of targets in each group, use the slider to **Show Venn Diagram**, then control-click target groups to view the Venn Diagram.
- Click on a region of the Venn Diagram and use **Save As** to save the targets in that region as a custom target group.
- Use the **Copy to Clipboard** button to copy selected groups of targets to the clipboard for pasting in a spreadsheet or other application.
- Use **Save** and **Export** buttons to save/export the .svg image of the Venn Diagram.

Creating New Target Groups

Create new target group from existing group

1. Start with an existing target group (such as **All Targets**).
2. Alter this group by deselecting targets you'd like to remove and/or clicking on targets you'd like to add. You can also "highlight" targets by selecting an area on a visualization in the third pane.
3. Right-click in the targets list and select **Create Group (selected)** (if you clicked on/off probes in the Targets list) or **Create Group (highlighted)** (if you highlighted targets using a visualization) to save this as a new target group **(see Figure 67)**.

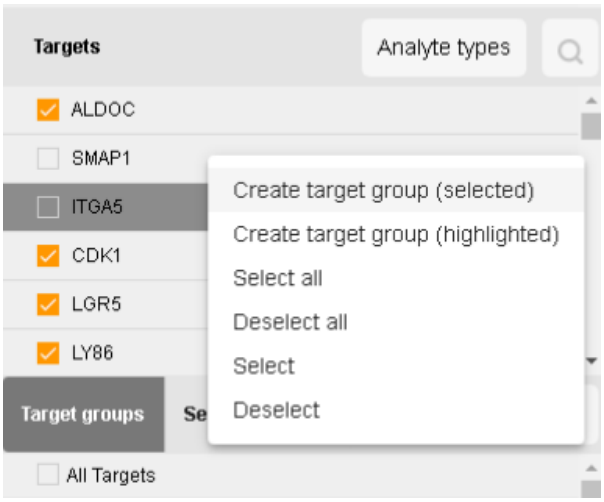


Figure 67: Creating a probe group from the probes section of the Datasets pane

4. Enter the Target Group name and add any tags desired (optional).
5. Click the **Create Target Group** button. The new target group will appear under the **Custom Target Groups** folder.

Create new target group from visualization

You may also select a region of the live visualization in the third pane and right-click on it to create a new target group from those targets in the selected region or to deselect them **(see Figure 68)**.

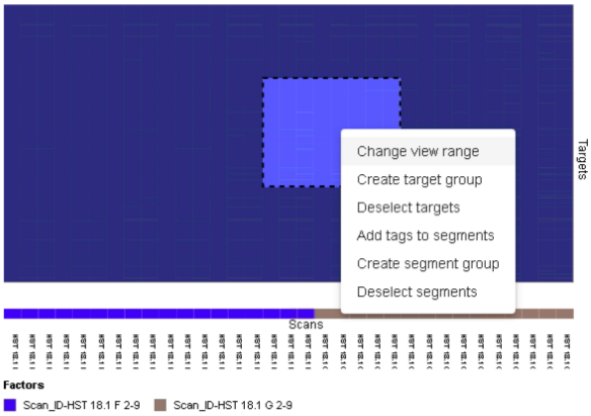


Figure 68: Creating a probe group from a visualization

Create new target group manually

1. Navigate to the **Manage Groups** window (via the gear icon) in the Target Groups window **(see Figure 69)**.
2. Click the **Add new group** button.
3. Enter a group name and add tags, if desired.
4. Add the appropriate delimiter for the list of targets you plan to copy and paste (new line, tab, comma).
5. Paste your list of targets into the field.
6. Click **Create**.

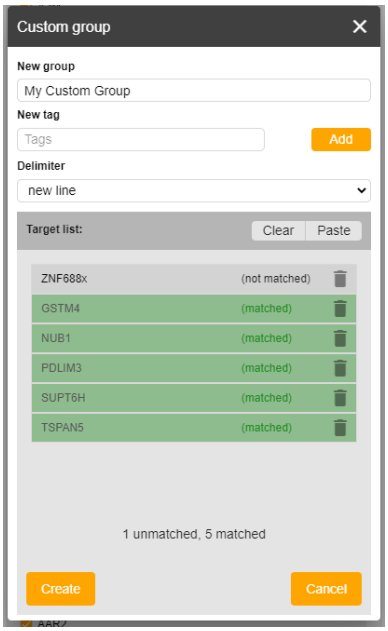


Figure 69: Add new group window

Create New Segment Group

Select a region of the live visualization in the third pane, right-click, and select **Create segment group** from the segments in the selected region **(see Figure 70)**. You can also deselect them from this menu.

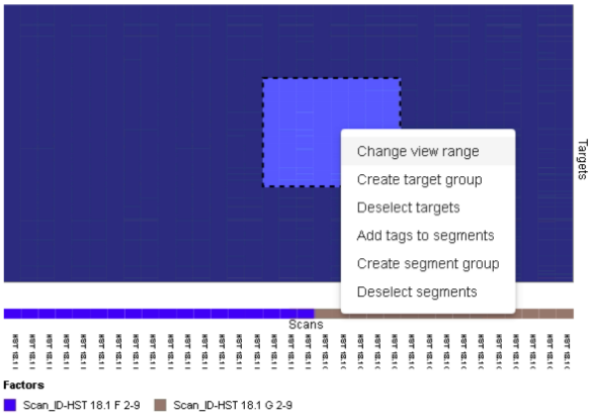


Figure 70: Creating a segment group from a visualization

Manage Target Groups & Segment Groups

Hovering over a target group or segment group name reveals **edit** and **trash** icons.

- Rename the group by selecting the **edit** icon that appears when you hover over it.
- Delete the group by selecting the **trash** icon that appears when you hover over it. You cannot delete the **All Targets** group.

Visualizations Pane

The third pane of the DSP Data Analysis window is the **Visualizations pane**. Visualizations are visual representations of the selected dataset and probes. They will change dynamically in response to the dataset, segments, and/or probes selected for analysis. Find more details on some visualizations in **Algorithm Details on page 101**.

Use the Maximize button to make this pane fill the screen or the Minimize button to hide it as a tab.

Visualization option icons line the header of this pane **(see Figure 71)**.

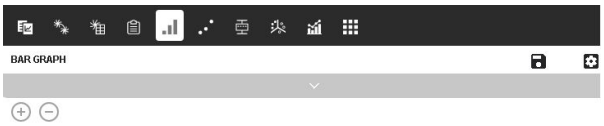


Figure 71: Visualization option icons

The **Save** and **Gear** icons are available to save and name visualizations, respectively. The **Download** icon is available for downloading the visualization.

Click and drag to select any area of interest on a visualization. Any segments in this area of interest will be shown as highlighted in the **Scans pane**. Right-click on your selection to view options:

- Creating a probe group or segment group for the selection.
- Excluding the selected set of probes or segments from the present study.
- Creating a tag for the selected segments.

Study Summary (NGS only)

> **For NGS readout:** Study Summary is only available for studies with NGS readout.

The study summary (see Figure 72) plots the total number of reads for each segment by each processing step outcome (trimmed, deduplicated, etc.). It also contains drop-down fields with **Segment Summary** (see Figure 73), **Probe Summary**, and **Target Summary**.

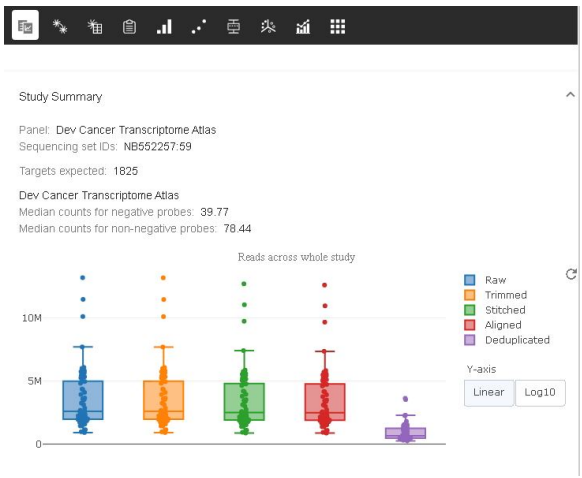


Figure 72: NGS Study Summary

Segment summary includes Q30 values. Q30 refers to the percentage of bases that get a Q score of at least 30 (99.9% accuracy). Q score is the most common metric to assess the accuracy of a sequencing platform.

The terms **umiQ30** and **rtsQ30** refer to the % of bases with Q≥30 specifically in the UMI and RTS_ID portions of the reads, respectively.

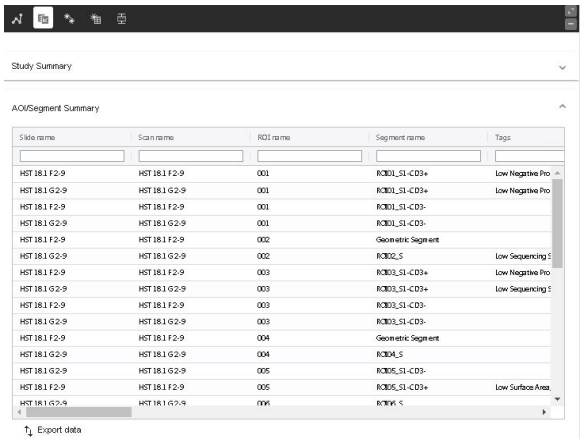


Figure 73: Segment Summary

Pathway Map (NGS only)

> **For NGS readout:** Pathway Map is only available for studies with NGS readout.

The pathway map **(see Figure 74)** shows the coverage of pathways for the entire dataset.

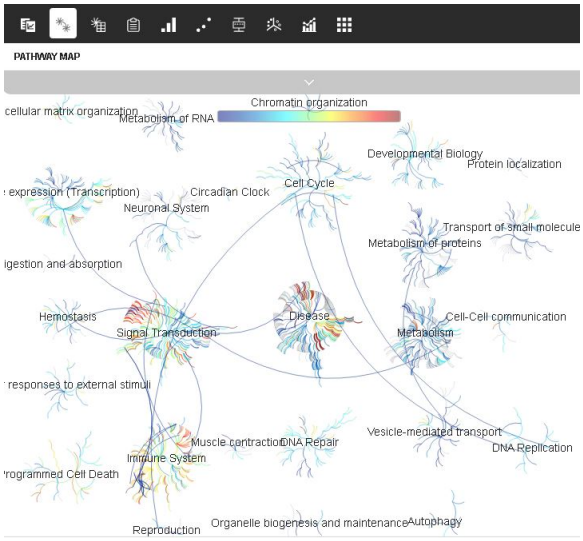


Figure 74: Pathway Map

Click the arrow to access the drop-down field **(see Figure 75)**. Here, you can change the color scheme of the Pathway Map.

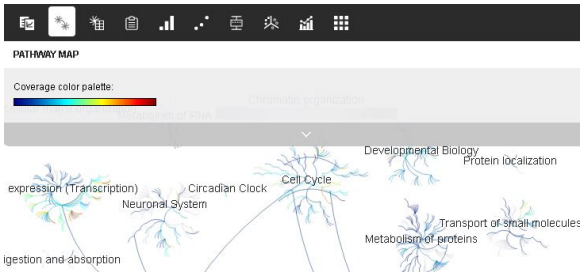
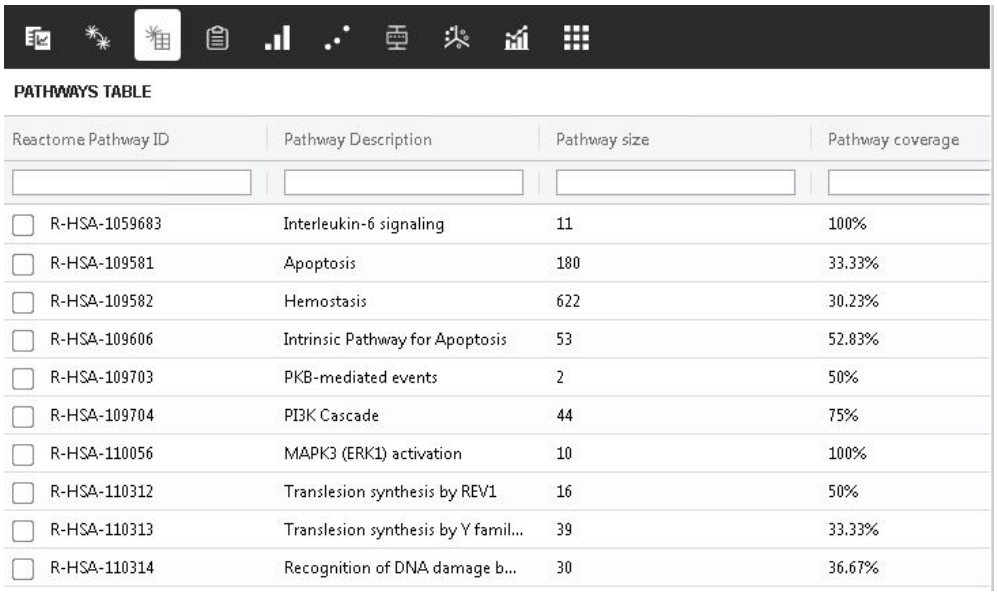


Figure 75: Pathway Map drop-down

Pathway Table (NGS only)

> **For NGS readout:** Pathway Table is only available for studies with NGS readout.

The Pathway Table summarizes the Pathway Map coverage **(see Figure 76)**. This value is calculated by dividing the **number of genes present in the dataset** by the **number of genes in present in the pathway**. The genes in the dataset are determined by the panel used for testing and the current selection in your dataset. The pathway information is obtained from reactome.org.



PATHWAYS TABLE

Reactome Pathway ID	Pathway Description	Pathway size	Pathway coverage
R-HSA-1059683	Interleukin-6 signaling	11	100%
R-HSA-109581	Apoptosis	180	33.33%
R-HSA-109582	Hemostasis	622	30.23%
R-HSA-109606	Intrinsic Pathway for Apoptosis	53	52.83%
R-HSA-109703	PKB-mediated events	2	50%
R-HSA-109704	PI3K Cascade	44	75%
R-HSA-110056	MAPK3 (ERK1) activation	10	100%
R-HSA-110312	Translesion synthesis by REV1	16	50%
R-HSA-110313	Translesion synthesis by Y famil...	39	33.33%
R-HSA-110314	Recognition of DNA damage b...	30	36.67%

Figure 76: Pathway Table

Dataset Summary

Selecting the summary button in the Visualizations pane of the Data Analysis window may reveal multiple tabs, depending on what tasks have been run: **Segment Summary**, **Attachments**, **Saved visualizations**, **Background**, **Normalization**, **Scaling**, and **Statistics**.

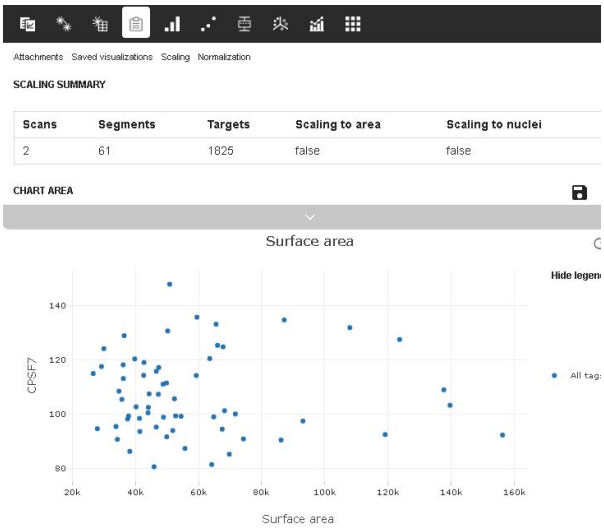


Figure 77: Summary in visualizations pane

Segment Summary

This table lists the scan name, ROI name, segment name, surface area, nuclei count, and any tags associated with each segment. Select or deselect listed segments to add or remove them from the current analysis.

Depending on the analysis performed, additional data may appear here to summarize information for each segment. This includes QC status, binding density, FoV registration QC, positive normalization factor, QC flags, and average background.

Attachments

This tab contains any files generated from custom scripts for the current dataset. See **Pathway Analysis on page 51**.

Saved Visualizations

This tab contains any saved visualizations for the current dataset. See **Saving a visualization on page 99** and **Accessing saved visualizations on page 100**.

Background Summary - see **Viewing Background Results on page 42**.

Normalization Summary - see **Viewing Normalization Results on page 39**.

Scaling Summary - see **Viewing Scaling Results on page 35**.

Statistics Summary - see **Viewing Statistical Test Results on page 50**.

Bar graph

The bar graph (see Figure 78) represents the count values of all probes across all segments included in the study. Segments are listed along the x-axis and counts along the y-axis. The height of each bar represents the frequency of each count defined by the bins.

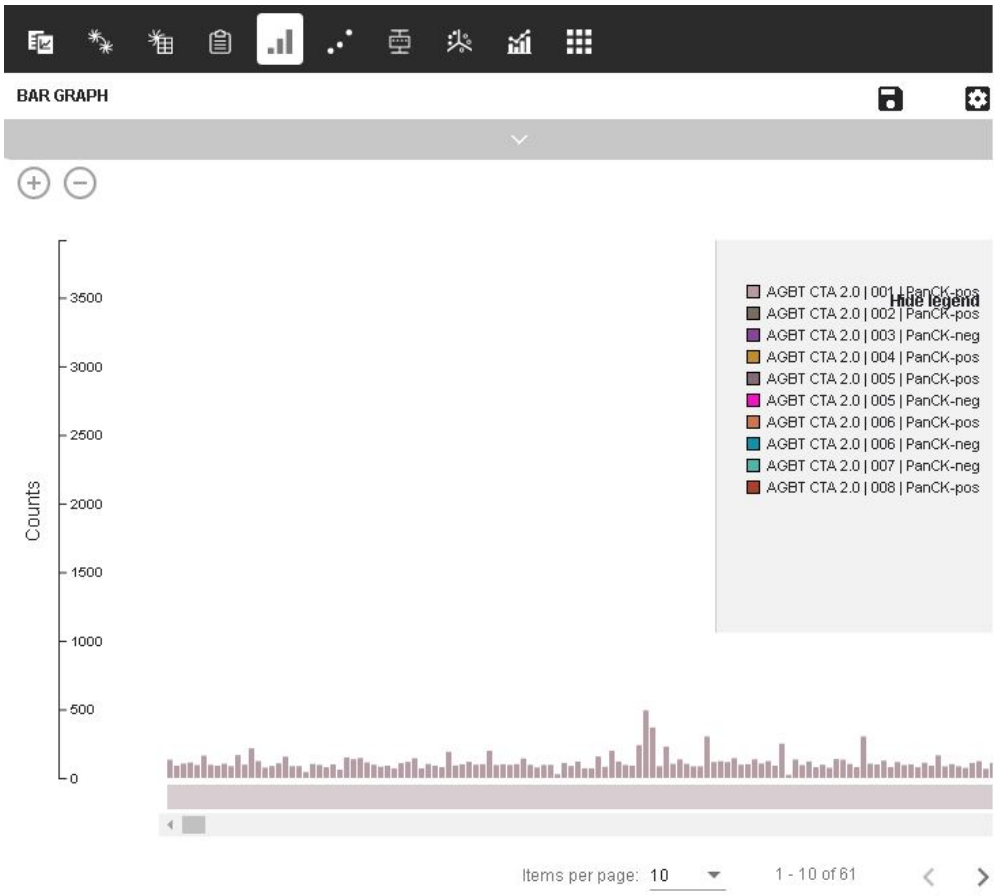


Figure 78: Bar graph

- Zoom in or out using the **+** and **-** buttons and scroll to view all results.
- Below the plots, select whether bars should be colored by **Counts**, **Probe**, **Segments**, or **Tags**. The resulting legend will be listed to the right of the plots.
- To name the visualization and/or the y-axis, select the **gear icon** (see Figure 79), enter the desired name, and select **Save**.
- To save the visualization, select the **Save** icon. Access the saved image later under the **Summary** tab. For more

Figure 79: Save and gear icons

information on accessing saved visualizations, exporting visualizations, or exporting data from visualizations, see **Saving & Exporting on page 99**.

Click the arrow to access the drop-down field **(see Figure 80)**.

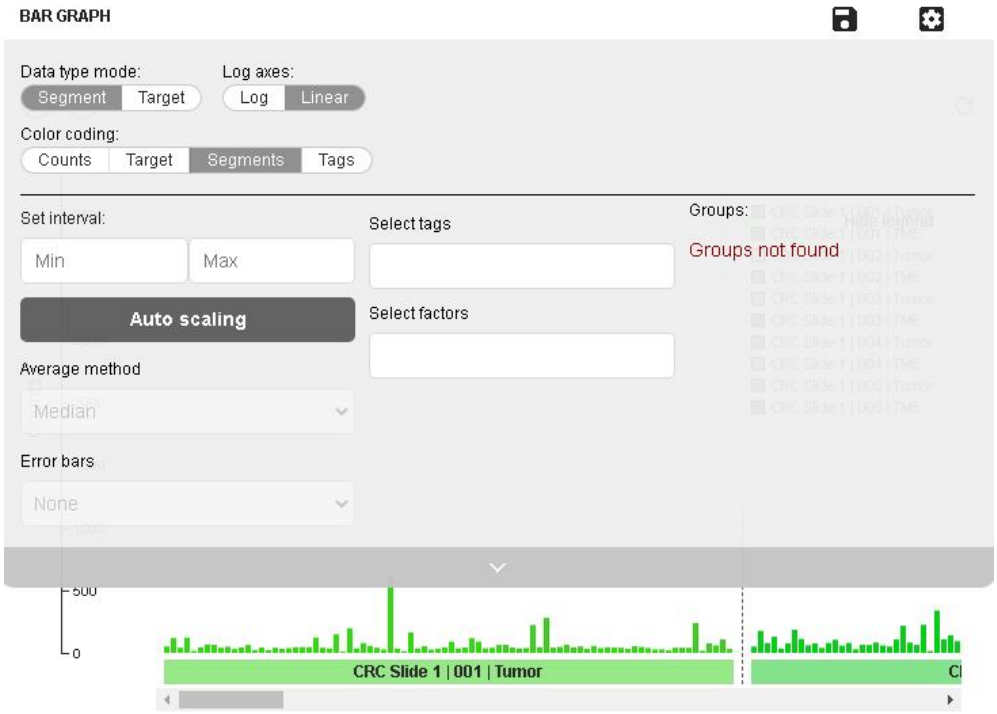


Figure 80: Bar graph drop-down field

Here, you may customize the graph by deciding whether to:

- View segments or probes on the x-axis using the **Segment/Probe** slider.
- View intensity data in linear or log space using the **Linear/Log** slider.
- View ratio data (if applicable) as **Ratios**, **Fold Changes**, or **Log_2** ratio.
- Apply auto-scaling by entering a **Min Count Value** and/or a **Max Count Value** (only available when viewing linear intensity data, not in log scale).
- Apply grouping by selecting **Tags**, **Factors**, **Average method** (median, geomean, average), and **Error bars** (SE, SD, none).

Correlation plot

The correlation plot (see Figure 81) is a visualization that plots one segment's results on the x-axis and a different segment's results on the y-axis. Alternatively, you may choose to plot one target's results vs. another target's results (see below).

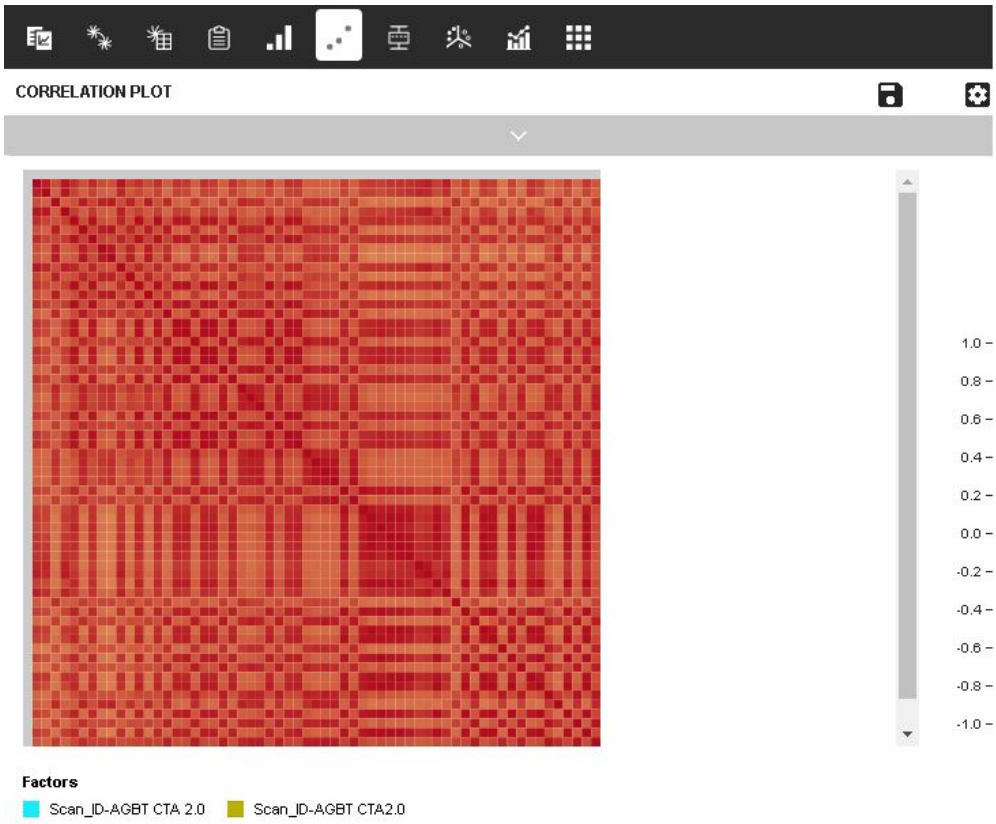


Figure 81: Correlation plot

- Note that the correlation plot is limited to 100x100 segments/targets. The correlation calculation will fail if a segment has the same count for all targets, for example if no targets have been detected. Remove these segments from the analysis.
- Hover over a cell on the correlation plot to see a pop-up displaying the represented segments and correlation coefficient.
- To name the visualization, x-axis, or y-axis, select the **gear icon**, enter the desired name(s), and select **Save**. Use the check box to **Display trendline** or show or hide the **x-** or **y-axis** labels here, as well.
- To save the visualization, select the **Save** icon. Access the saved image later under the **Summary** tab. For more information on accessing saved visualizations, exporting visualizations, or exporting data from visualizations, see Saving & Exporting on page 99.

Click the arrow to access the drop-down field (see Figure 82).

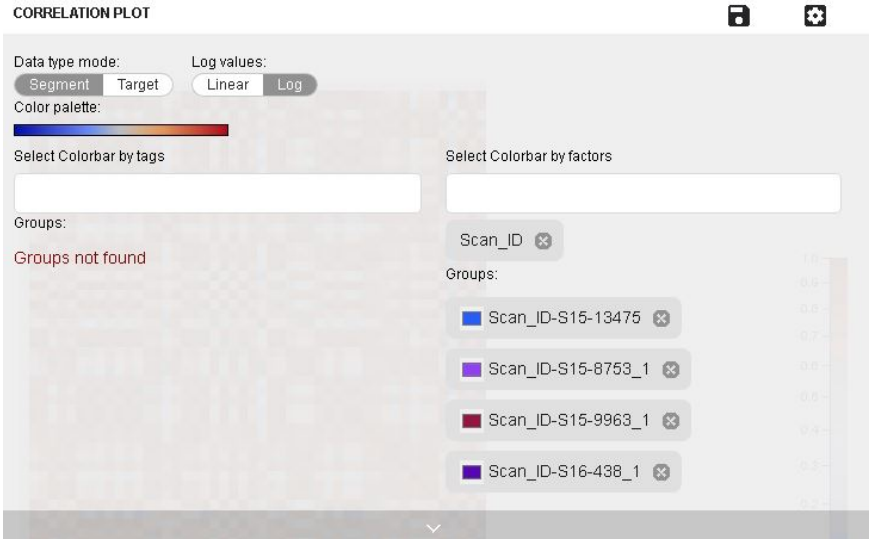


Figure 82: Correlation plot drop-down field

Here, you may customize the graph by deciding whether to:

- View segments or probes on the x-axis using the **Segment/Target** slider.
- View intensity data in linear or log space using the **Linear/Log** slider.
- Change the color scheme

Click and drag to select part or all of the correlation plot. A scatter plot option will display below. Search for x- and y- axis targets using the search fields. Right-clicking summons a menu (see Figure 83) with which you can:

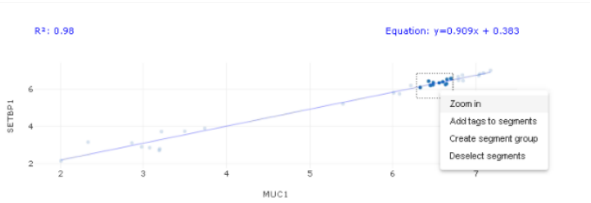


Figure 83: Right-click scatter plot menu

- **Zoom in**.
- **Add tags** to segments.
- **Create a target/segment group** from the selection. See Creating New Target Groups on page 75.
- **Deselect targets/segments** in the selection.

Box plot

Box plots (see Figure 84) 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. These visualizations display differences between subsets of an experiment without making any assumptions about the underlying statistical distribution; they are non-parametric.

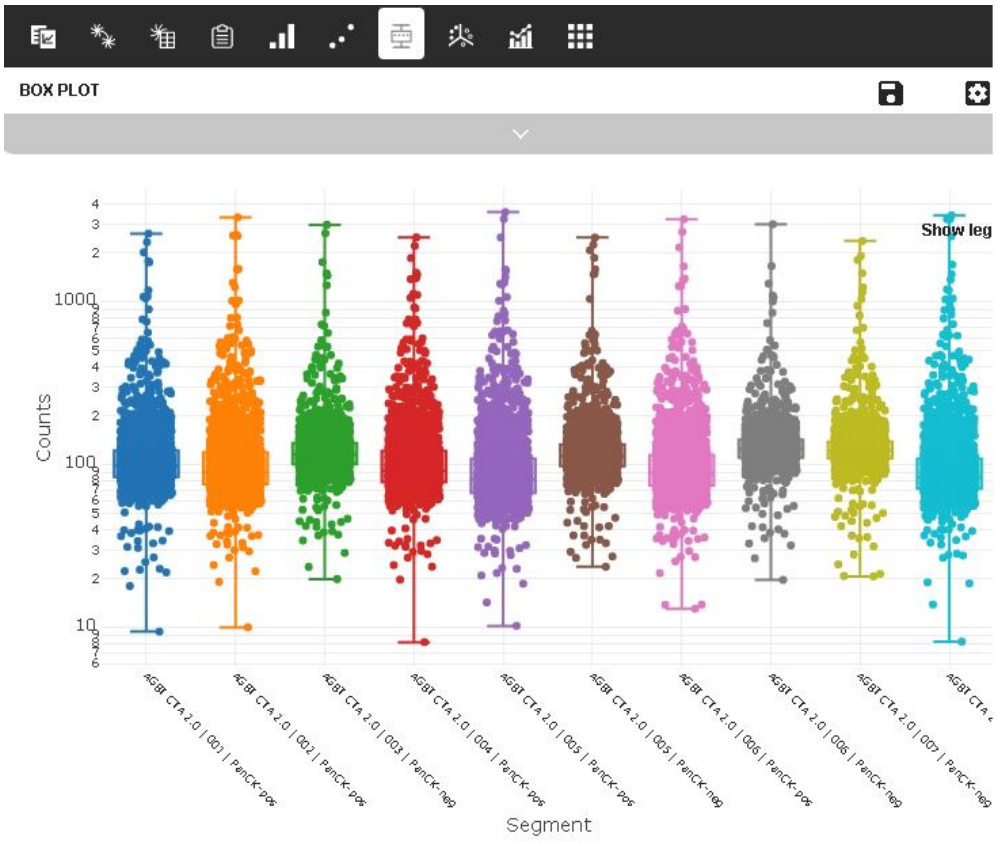


Figure 84: Box plot

- Hover over a box on the plot to see a pop-up displaying the segment, tags, and the values for the median, maximum, and first and third quartiles.
- Box plots with many segments or targets will load in "pages."
- Click and drag over an area of interest on the box plot to zoom in.
- The legend shows the color assigned to each plot and its corresponding label.
- Click a color box in the legend to either display or hide the plot.

- To name the visualization, x-axis, or y-axis, select the **gear icon** (see Figure 85) and enter the desired name (s). Check the respective boxes to show x- and/or y-axis labels. Select **Save**.

Figure 85: Save and gear icons

- To save the visualization, select the **Save** icon. Access the saved image later under the **Summary** tab. For more information on accessing saved visualizations, exporting visualizations, or exporting data from visualizations, see Saving & Exporting on page 99.

Click the arrow to access the drop-down field (see Figure 86).

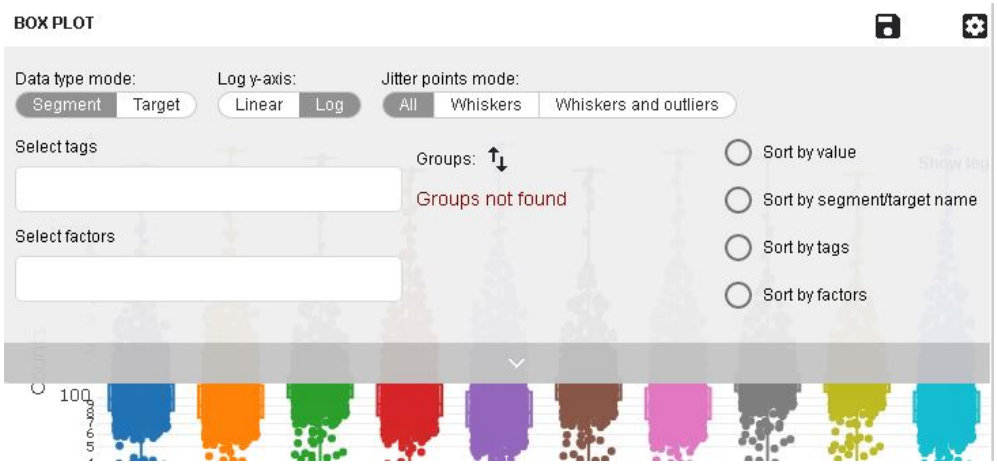


Figure 86: Box plot drop-down field

Here, you may customize the graph by deciding whether to:

- View segments or probes on the x-axis using the **Segment/Target** slider.
- View intensity data in linear or log space using the **Linear/Log** slider.
- **Jitter points** can be turned **On/Off** using the slider.
- **Group** by tags, if desired.
- Use the check boxes to **Sort by value, name, tag, or factor**.
- If sorting by value, you may choose to start from the **Min**, **Max**, or **Median** (the system will automatically change selections if **ascending** and **max** or **descending** and **min** are co-selected).

Forest Plot

The forest plot (see Figure 87) shows the distribution of ratio values for individual probes across all segments or groups of segments. Forest plots can only be generated from a ratio dataset. Fold changes are depicted as box and whisker plots along the horizontal axis against each probe name (listed vertically). A vertical axis is shown at ratio value equal to 1 (0 on a Log_2 scale).

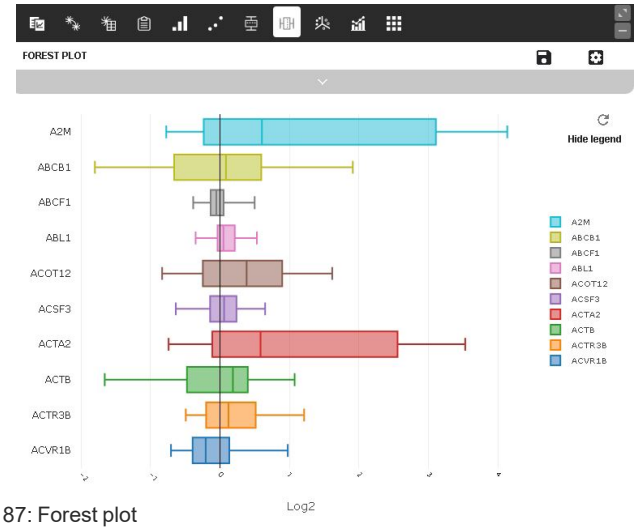


Figure 87: Forest plot

- Hover over a box to see a tooltip with the statistics for the distribution, depicted by the plot for each probe.
- The boxes span the first and third quartile of the distribution with a line indicating the median. Whiskers extend between the 95% confidence limits for the data.
- The legend shows the color assigned to each plot and its corresponding label.
- Click a color box in the legend to either display or hide the plot.
- Click and drag over an area of interest on the forest plot to zoom in.
- To name the visualization or x-axis, select the **gear icon** (see Figure 88), enter the desired name(s), and select **Save**.

Figure 88: Save and gear icons

- To save the visualization, select the **Save** icon. Access the saved image later under the **Summary** tab. For more information on accessing saved visualizations, exporting visualizations, or exporting data from visualizations, see Saving & Exporting on page 99.

Click the arrow to access the drop-down field (see Figure 89).

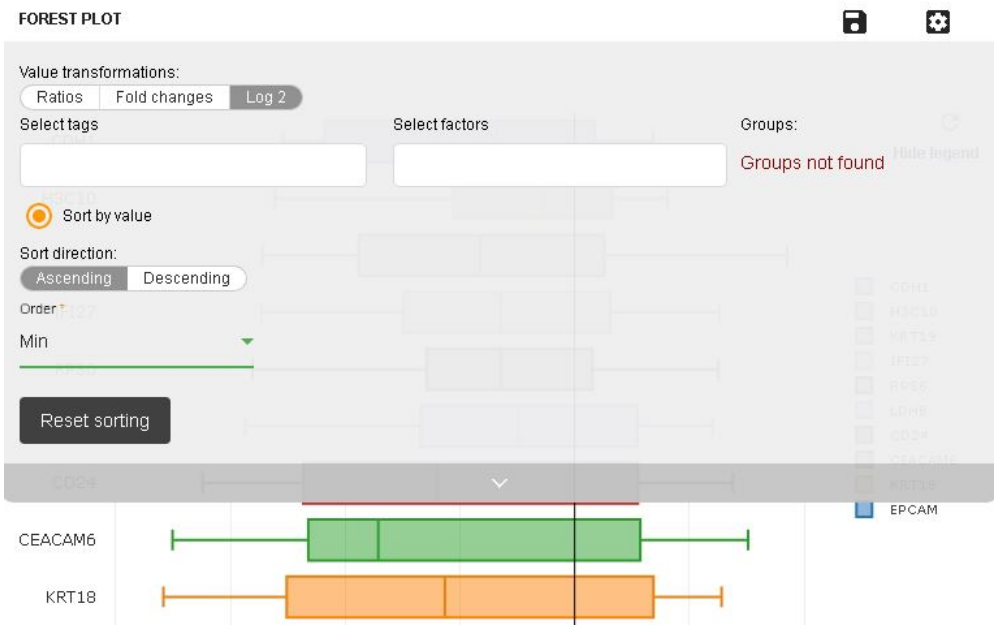


Figure 89: Forest plot drop-down field

Here, you may customize the graph by deciding whether to:

- View intensity data as **Ratios**, **Fold changes**, or in **Log_2**.
- Stratify and color data by **grouping by tags** and tag combinations.
- Color data points by tags:
  - Select one or more segment tag for coloring (combination groups will be created, as well).
  - Boxes not included in tag groups will be shown as gray.
  - To change the color of the grouping, click on the group's color box in the drop-down field; a color selection window will appear.
- Click the **Sort by value** button to sort data in **ascending** or **descending** order. You may choose to start from the **Min**, **Max**, or **Median** (the system will automatically change selections if **ascending** and **max** or **descending** and **min** are co-selected).

Principal Component Analysis (PCA) Plot

The PCA plot (see Figure 90) depicts the first three principal components for the selected dataset along the x-axis, y-axis, and z-axis of a three-dimensional plot.

PCA is a dimensionality reduction technique for visualizing variation within a dataset. Users should view PCA when they want to get a sense of how similar a segment is to the others. This technique can be used to identify clusters of similar segments. Each principal component (PC1-3, plotted as x,y,z axes by default) explains a certain percentage of the variation in the data (PC1 explains the most variation, PC2 second, etc). It's also possible using PCA data to determine which targets contribute to each PC, that is, which are responsible for variation between segments. PCA input data is scaled before analysis.

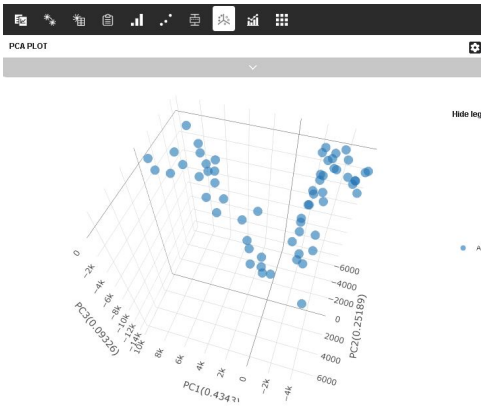


Figure 90: PCA plot

- Click on the plot and drag to rotate along the x, y, or z axis to view the plot in different axes' perspectives.
- Click on a data point to automatically highlight the segment in the segments pane and the scan image viewer.
- Hover over a datapoint on the plot to see a pop- up (see Figure 91) displaying the segment name it represents, associated tags, and each of its coordinates and to see its three dimensions defined.

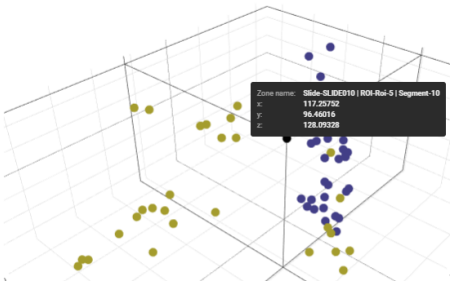


Figure 91: PCA plot hover-over pop-up

- To name the visualization, select the **gear icon**, enter the desired name, and select **Save**. Use the check boxes to **Export Columns** and **Display Legend** here, as well.

- To save the visualization, select the **Save** icon. Access the saved image later under the **Summary** tab. For more information on accessing saved visualizations, exporting visualizations, or exporting data from visualizations, see **Saving & Exporting on page 99**.

Click the arrow to access the drop-down field **(see Figure 92)**

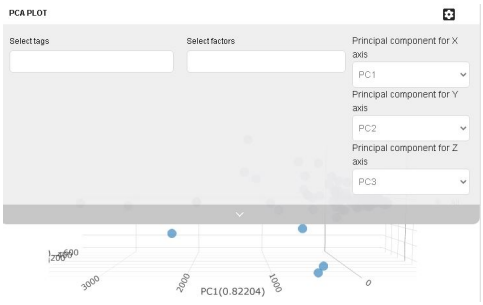


Figure 92: PCA drop-down field

Here, you may customize the graph to:

- Color segments by tags/factors
- Change the Principal Component axis names

Volcano Plot

A volcano plot **(see Figure 93)** is a scatter plot showing the measure of significance (-log10 of p-values) on the y-axis vs the difference in geometric means of probe expression between two groups of segments compared in a statistical test. This visualization is available for datasets which contain t-tests (p-values).

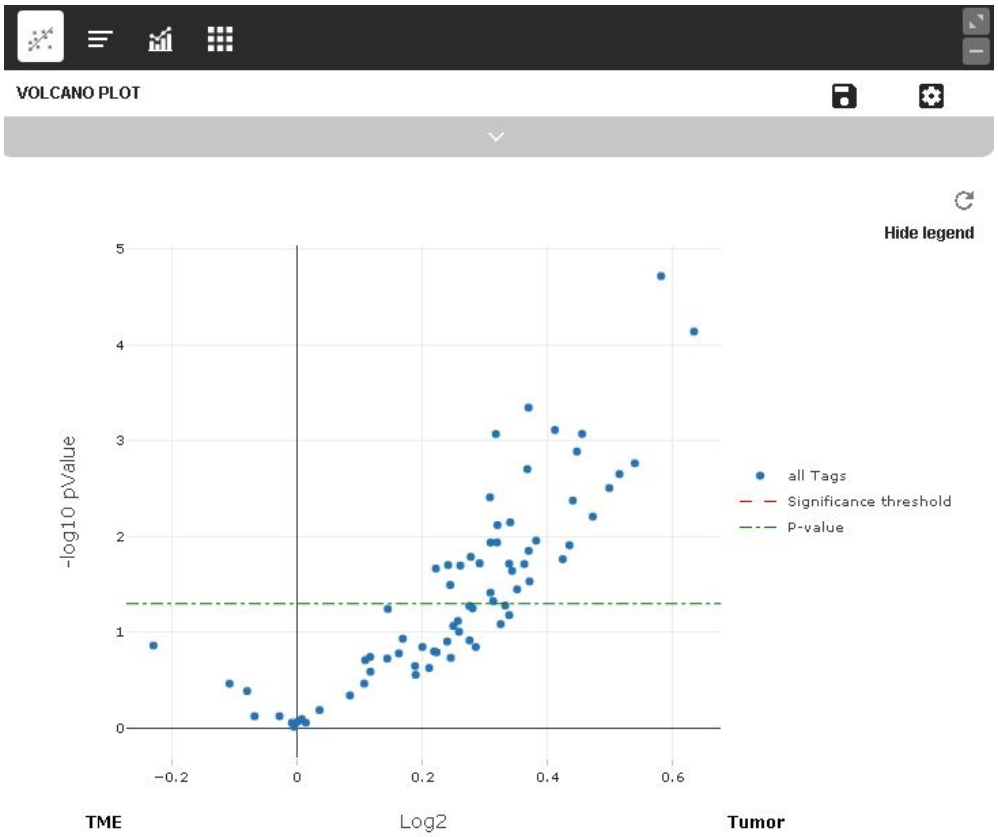


Figure 93: Volcano plot

- If there is more than one statistical test associated with a dataset, select the test of interest at the top of the window.
- Hover over a data point on the plot to see a pop-up displaying the target, p-value, and ratio value associated with that point.
- To name the visualization, x-axis, or y-axis, select the **gear icon** **(see Figure 94)**, enter the desired name(s), and select **Save**.
- To save the visualization, select the **Save** icon. Access the saved image later under the **Summary** tab. For more

Figure 94: Save and gear icons

information on accessing saved visualizations, exporting visualizations, or exporting data from visualizations, see **Saving & Exporting on page 99**.

Click the arrow to access the drop-down field **(see Figure 95)**.

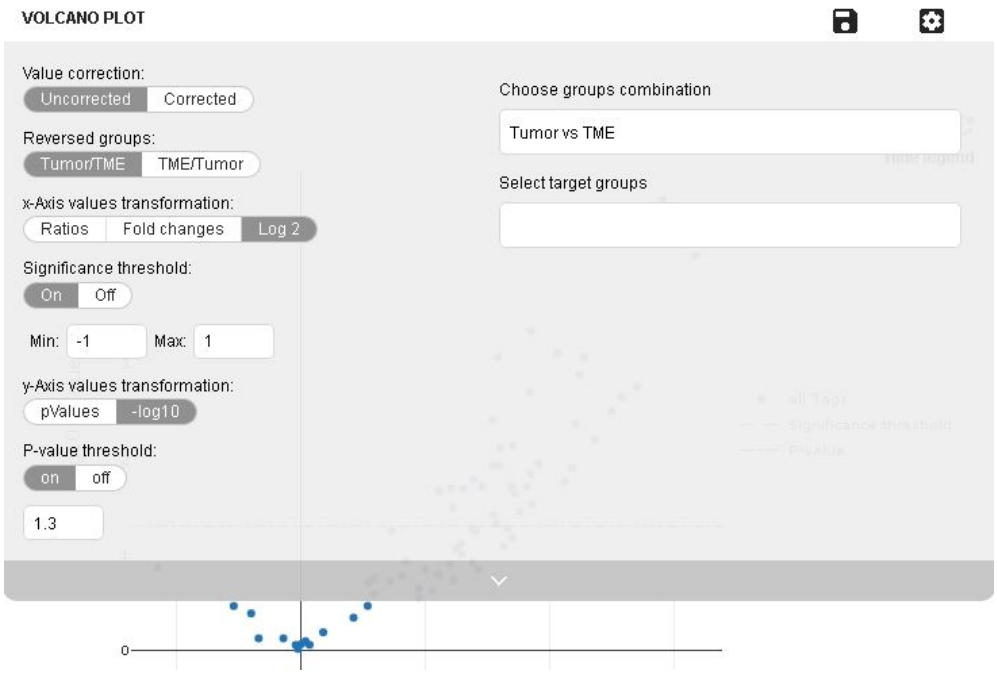


Figure 95: Volcano plot drop-down field

Here, you may customize the graph by deciding whether to:

- Display corrected or uncorrected data using the **Uncorrected/Corrected** slider.
- Select the **ratio** (e.g., Tumor/Immune or Immune/Tumor) to use in the plot by toggling the **Reversed groups**.
- View ratio data as **Ratio**, **Fold changes**, or **Log2 ratio** on the x- and y-axes
- Establish **Significance** or **P-value** threshold settings.

Click on any datapoint in the volcano plot to see a detailed box and scatter plot of the data behind it **(see Figure 96)**.

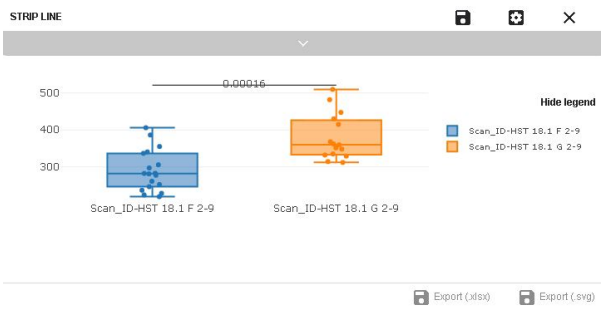


Figure 96: Box and scatter plot

Trend Plot

Trend plots (see Figure 97) show line graphs for all selected probes in the dataset. Segments are ordered along the x-axis, while probe counts are along the y-axis.

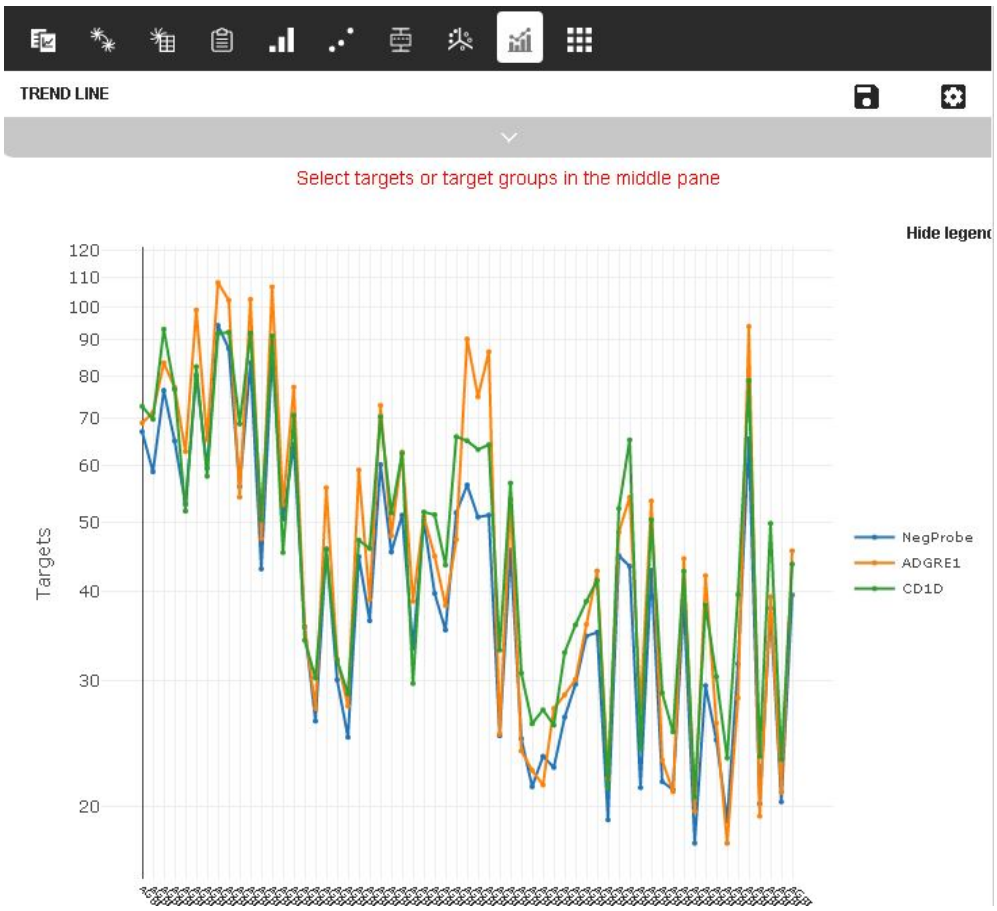


Figure 97: Trend plot

For NGS readout: By default, no probes are selected. Select the probes of interest to see the trend plot materialize.

For nCounter readout: Probes of the selected dataset are displayed by default.

- Hover over lines to see a pop-up with probe names and p-values.
- To name the visualization, select the **gear icon** (see Figure 98), enter the desired name, and select **Save**. Use the check boxes to **Export Columns** and **Display Legend** here, as well.

Figure 98: Save and gear icons

- To save the visualization, select the **Save** icon. Access the saved image later under the **Summary** tab. For more information on accessing saved visualizations, exporting visualizations, or exporting data from visualizations, see **Saving & Exporting on page 99**.

Click the arrow to access the drop-down field **(see Figure 99)**.

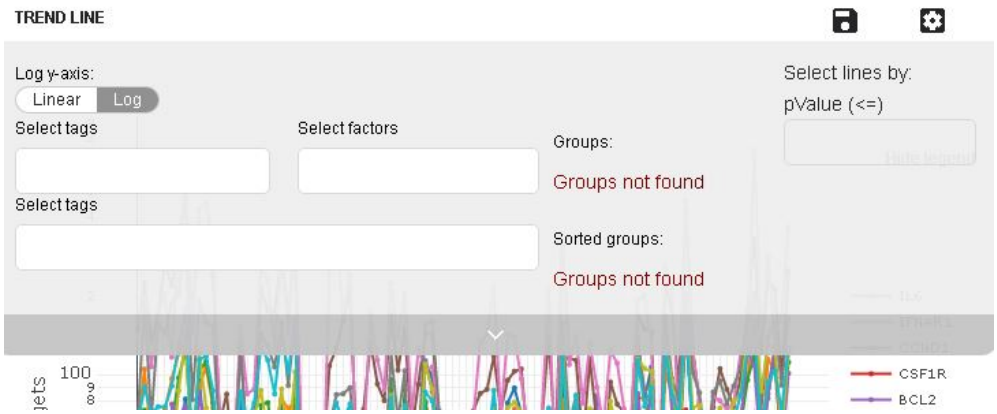


Figure 99: Trend line drop-down field

Here, you may customize the graph by deciding whether to:

- Use the slider to switch between **linear** and **log** values for the y-axis.
- Choose **Segment grouping by tags** or **factors** or to **sort segments by tags**.
- You can select lines by **establishing the maximum p-value**. Any lines that are representing results with that p-value or better will be selected.

Heatmap

The heatmap **(see Figure 100)** is an image that depicts counts by color. Segments are aligned along the x-axis and targets on the y-axis.

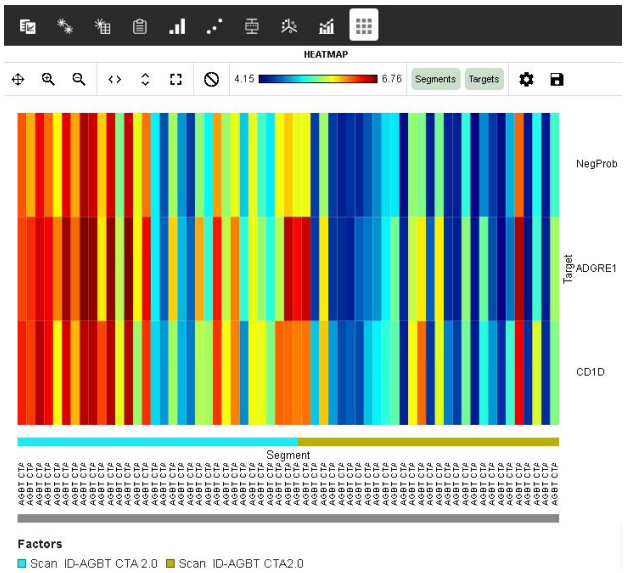


Figure 100: Heatmap

- Hover over an area of the heatmap to see a pop-up displaying the segment, target, count, and any tags.
- Use the toolbar at the top to zoom, fit to page (vertically, horizontally, or both), reset view, change the color scheme, or show/hide segments and targets.
- Use the scroll bar at right to scroll the heatmap through long lists of targets.
- Click the gear icon to name the visualization, choose linear or log data, and establish coloring by tags or factors.
- To save the visualization, select the **Save** icon. Access the saved image later under the **Summary** tab. For more information on accessing saved visualizations, exporting visualizations, or exporting data from visualizations, see **Saving & Exporting on page 99**.

Figure 101: Save and gear icons

Click and drag to select part or all of the heatmap. Right-clicking within this selected area summons a menu (see Figure 102) with which you can:

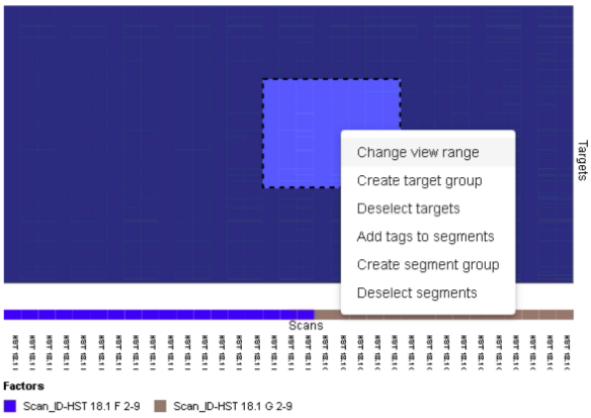


Figure 102: Right-click visualizations menu

- Create a probe group comprised of the selected probes.
- Deselect the selected probes from the current analysis.
- Add tags to the selected segments.
- Create a segment group comprised of the selected segments.
- Deselect the selected segments from the current analysis.

Saving & Exporting

Save and Export functions are primarily featured in the **Visualizations pane**.

Save the present visualization using the **Save** icon (see Figure 103) and access it later under the **Summary** tab.

Establish a name for the present visualization using the **gear** icon.

Figure 103: Save and gear icons

Export the present visualization in *.svg* format using the **Export(.svg)** button (see Figure 104).

Export the data upon which this visualization is based in .xlsx format using the **Export(.xlsx)** button (see Figure 104).

Figure 104: Export data from visualizations (.xlsx) or visualization images (.svg)

Heatmap, Cluster, and Bar graphs also have the option to be exported as scalable .png files.

Export the whole dataset in .xlsx format from the Datasets pane (see Manage Datasets on page 71).

Saving a visualization

1. Select the **Save** icon to save the present visualization.
2. Give this visualization a name, description, and tags (if desired) (see Figure 105).
3. Select **Save**.

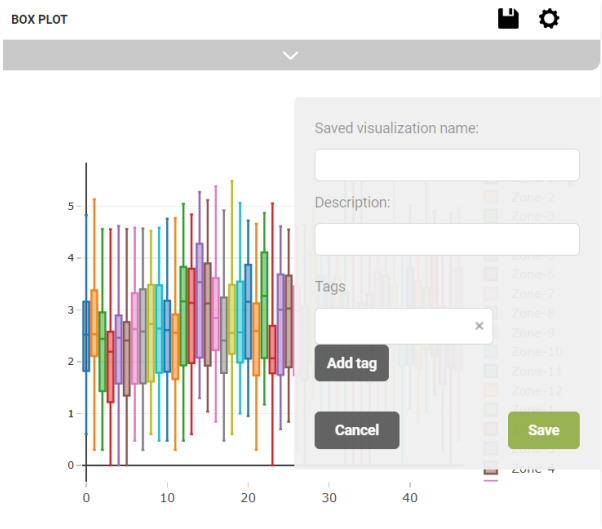


Figure 105: Saving a visualization

Accessing saved visualizations

1. Select the **Summary** tab in the **Visualizations pane**.
2. Select **Saved Visualizations** (see Figure 106).
3. Select the visualization of interest from the list.
4. You may select another visualization (if available) to compare to the present image by selecting it at the bottom of the window (see Figure 107).
5. You may review the segments and probes that make up this image by selecting the arrow at the top of the window.

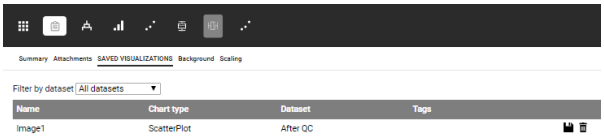


Figure 106: Saved visualizations tab under Summary on Visualizations pane

Figure 107: Comparing saved visualizations

Exporting visualizations

1. Select the **Export(.svg)** or **Export(.png)** (not available for all visualizations) button at the bottom of the **Visualizations pane** to export the present visualization.
2. The visualization will be exported from the DSP system in respective format. Exports in .png are scalable.

Exporting data on which a visualization is based

1. Select the **Export(.xlsx)** button at the bottom of the **Visualizations pane** to export the data upon which the present visualization is based.
2. The spreadsheet will be exported from the DSP system in .xlsx format.

Algorithm Details

Bar graph

Bar graphs are displayed as linear or $\log_2$ transformed values. When grouping: Median. Error bars are either the Standard deviation of values in a group, or standard error.

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. These visualizations display differences between subsets of an experiment without making any assumptions about the underlying statistical distribution; they are non-parametric. The box extends between 25th and 75th quartile. The whiskers extend to minimum and maximum.

Cluster

The cluster tool displays an unsupervised hierarchical clustering where the selected data represents log-transformed and z-scored values. Clustering is performed based on correlation to determine the position in the clustering dendrogram. The cluster heatmap will plot each segment-probe cell according to determined position using color to represent z-scores. When exporting data from the visualization pane, you will be exporting only the values you see, which are z-scores.

Forest plots

The forest plot shows the distribution of ratio values for individual probes across all segments or groups of segments. Forest plots can only be generated from a ratio data dataset. Fold changes are depicted as box and whisker plots along the horizontal axis against each probe name (listed vertically). A vertical axis is shown at ratio value equal to 1 (0 on a $\log_2$ scale). Also displayed are: median, q1, q3, and 95% confidence limits.

Heatmap

Heatmaps are displayed as linear counts or $\log_2$ transformed values.

Minimum nuclei count

Segments will be flagged when the minimum nuclei count is less that the provided value.

Minimum surface area

Segments will be flagged when the minimum surface area is less than the value provided (in μm^2).

Normalization

Normalization factor for each segment shall be calculated as the ratio of the average of all geometric mean or arithmetic mean of selected normalization probes across all segments to the geometric mean or arithmetic mean of selected normalization probes for that segment.

Pathway Analysis (NGS readout only)

Gene set enrichment analysis (GSEA) is implemented with the fGSEA package from Bioconductor using pathway groups from the Reactome database. GeoMx DSP software v2.5.1 uses Reactome database v78.

Korotkevich G, Sukhov V, Sergushichev A (2019). “Fast gene set enrichment analysis.” bioRxiv. doi: 10.1101/060012, http://biorxiv.org/content/early/2016/06/20/060012.

QC

See next page.

Scaling

When scaling to surface area is selected, probe count values for a particular segment shall be multiplied by a surface area scaling factor. Scaling factor for each segment shall be calculated as the ratio of the calculated geometric mean, average, median, maximum or minimum surface area across all actively selected segments in the dataset to the surface area of that segment.

When scaling to nuclei count is selected, probe count values for a particular segment shall be multiplied by a nuclei scaling factor. The scaling factor is calculated as the ratio of the calculated geometric mean, average, median, maximum or minimum nuclei count across all actively selected segments in the dataset to the nuclei count of that segment.

Scatter plot

R^2 value is calculated as the Pearson correlation coefficient; **RSQ** formula in Excel. Equation is also per slope and the intercept formula in Excel.

Segment Summary

Scan dimensions are included in the scan summary. **newCenterX** and **newCenterY** are the ROI coordinates relative to the exported image, with 0,0 at the top left of the image.To convert the ROI coordinates to match your exported scan image using the height, width, and offset:

xRatio = exportedImage.width / scanDimensions.width

newCenterX = (centerX - offsetX) * xRatio

yRatio = exportedImage.height / scanDimensions.height

newCenterY = (centerY - offsetY) * yRatio

Strip plots

Strip plots depict one probe per visualization. Dots on the strip plot represent each value. The line shows the median for the group. Min, max, and +/- 1.5 inter-quartile range are also displayed.

QC

For NGS readout:

Biological Probe QC

Grubbs test looks for outlier probes either low or high within a target and ROI. Probes that are outliers in a proportion of ROIs above the user-set threshold are removed from all ROIs and considered "global" outliers. Probes that are outliers in a proportion of ROIs below the user-set threshold are removed only from those ROIs in which they are an outlier and are considered "local" outliers.

The limit of quantitation (LOQ) for each AOI is calculated as {user-defined value} geometric standard deviations above the geometric mean of the negative probes. This is calculated after the exclusion of outlier probes.

For nCounter readout:

FOV registration QC

Each individual lane scanned on an nCounter system is divided into a few hundred imaging sections, called Fields of View (FOV). The system images these FOVs separately, and sums the barcode counts (representing the GeoMx DSP probe counts) of all FOVs. Finally, the system reports the number of FOVs successfully imaged as FOV Counted. Significant discrepancy between the number of FOV for which imaging was attempted (FOV Count) and for which imaging was successful (FOV Counted) may indicate an issue with imaging performance. Recommended percentage of registered FOVs (i.e., FOV Counted over FOV Count) is 75%. Lanes will be flagged if this percentage is lower.

Binding density QC

The imaging unit only counts the barcodes that are unambiguously distinguishable. It simply will not count barcodes that overlap within an image. Too many overlapping codes in the image, however, will create a condition called image saturation in which significant data loss could occur. To determine the level of image saturation, the nCounter instrument calculates the number of optical features per square micron for each lane as it processes the images. This is called the Binding Density. The Binding Density is useful for determining whether data collection has been compromised due to image saturation. Lanes will be flagged when Binding Density values is not within the

	range of 0.1 and 1.8 (for nCounter SPRINT Profiler) or 2.25 (nCounter MAX/FLEX instrument). **Positive control normalization QC** The GeoMx Hyb Codes contain synthetic ERCC spike-ins. Counts for these spike-ins may be used to correct for platform-associated sources of (technical) variation (e.g., automated purification, hybridization conditions, etc.). Normalization is carried out using a scaling approach. The geometric mean count values of the spike-ins for each segment are multiplied with a scaling or normalization factor to bring them to the reference value which is set to the arithmetic mean of these geometric mean values. Lanes will be flagged when this Positive ERCC normalization factor is outside of the range 0.3 and 3.

Troubleshooting

Suggested actions to resolve certain issues are listed below **(see Table 1)**. Contact GeoMxSupport@nanostring.com for additional support.

Table 1: Troubleshooting data analysis

Issue	Possible cause	Suggested action
Display does not show latest work	Software is lagging slightly behind.	Click F5 or Assist (wrench) > Refresh Page. However, if waiting for a plot to load, avoid clicking F5 as it begins the loading sequence over again.
"Study is locked by another user"	The software thinks the study is open in more than one place.	• Make sure only 1 person is trying to access the study or has it open in a remote browser at any given time. • Clear the browser cache to remove any old, open versions of the study in the cache. At the GeoMx instrument, clear the cache by logging out with Ctrl Alt Del and logging in again. • Restart the Data Analysis auxiliary server, if applicable. Access to your auxiliary server may require your local IT group.
"Unable to get analysis ID from DA" or "Error Creating Study"	The GeoMx software may not be communicating with an installed auxiliary server.	Reboot the GeoMx DSP and then the auxiliary server. Access to your auxiliary server may require your local IT group. Then, try again to create the study.
Can't upload Annotations file or other problems with Manage Annotations	Common cause is that there is a formatting issue or typo in the text of the annotations file being uploaded.	• Column headers should be one word and may not begin with "ROI" or the column will be mistaken for ROI_IDs by the software. • Check that you do not have leading or trailing spaces in the sample name or slide name. • Ensure that you have not changed the Segment names, ROI names, or Scan names in the Annotation file, or the software will not recognize the information.

Issue	Possible cause	Suggested action
nCounter study: no FOV counts and very low binding density	Cartridge may have been tilted in nCounter instrument.	Rescan the cartridge after confirming it is seated in the scanning spot.
NGS study: low % aligned in all segments	Incorrect or missing pkc file associated with the run; software cannot align barcodes to pkc library.	Confirm that all the correct pkc files were associated with the run.
	Short-read masking during BCL to FASTQ conversion changes all nucleotides to "N"	Open FASTQ file to check that sequences are output. Rerun BCL to FASTQ conversion without N-masking parameters (refer to Illumina platform-specific information about BCL to FASTQ conversion).
NGS study: low % aligned in a subset of segments	Contamination with another NGS library	Check BioAnalyzer or TapeStation trace for correctly sized library. Ensure best practices for NGS library prep.
	PCR 'drop-outs' or failed reactions	Repeat library prep taking care to follow PCR setup best practices.
NGS study: low sequencing saturation	May be due to a high complexity or amount of input to PCR.	• Sequence libraries further until desired sequencing saturation level is achieved. Reads from multiple sequencing runs or Illumina platforms may be combined to increase sequencing depth. • Evaluate experimental conditions that may have led to low sequencing saturation.

Symbols & Definitions

	Manufacturer
	Authorized Representative in the European Community (safety)
	Safety CE Mark
	UK Conformity Assessed Mark
	Catalogue or Reference Number
	Batch code / Lot number
	Serial number
	Temperature range storage conditions
	Lower limit of temperature storage conditions
	Upper limit of temperature storage conditions
	For Use by / Expiry Date

NanoString Technologies, Inc.
530 Fairview Ave North
Seattle, Washington 98109 USA
www.nanostring.com

CONTACT US
info@nanostring.com
Tel: +1 888 358 6266
Fax: +1 206 378 6288

SALES CONTACTS
United States: us.sales@nanostring.com
EMEA: europe.sales@nanostring.com
Asia Pacific & Japan: apac.sales@nanostring.com
Other regions: info@nanostring.com