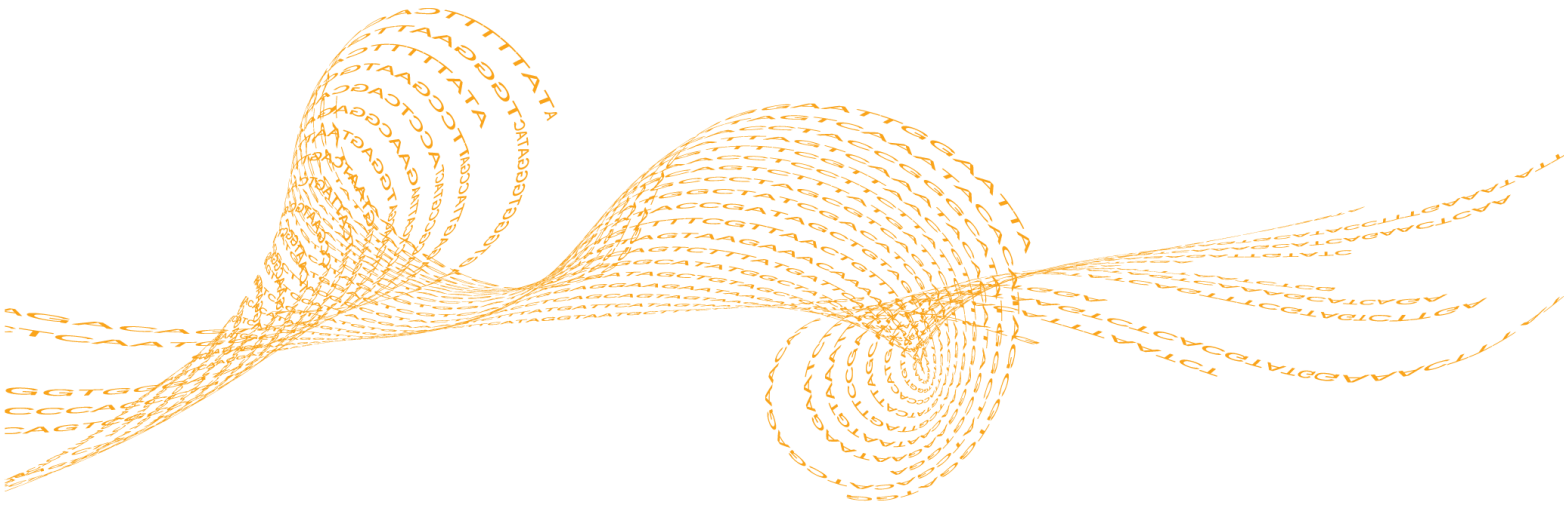
Sequencing Analysis Viewer Software

User Guide

FOR RESEARCH USE ONLY

Revision History	3
Introduction	4
Setting Up Sequencing Analysis Viewer Software	5
Data Availability	8
Loading Data	9
Analysis Tab	10
Imaging Tab	17
Summary Tab	19
Tile Status Tab	21
TruSeq Controls Tab	22
Indexing Tab	26
InterOp Files	28
Technical Assistance	



ILLUMINA PROPRIETARY
Part # 15020619 Rev F
October 2014

illumina®

Revision History

Part #	Revision	Date	Description of Change
15020619	F	October 2014	Added support for HiSeq X and InterOp files for RTA 1.1x and 2.x.
15020619	E	January 2014	Added support for NextSeq, and descriptions of InterOp files.
15020619	D	January 2013	Added figure of in-line controls in a successful TruSeq DNA PCR-Free run
15020619	C	October 2012	User Guide for Sequencing Analysis Viewer Software v1.8.20 and up, describes binned Q-scores.
15020619	B	October 2011	User Guide for Sequencing Analysis Viewer Software v1.8, which supports dual indexes and MiSeq metrics.
15020619	A	January 2011	Initial Release

Introduction

Sequencing Analysis Viewer Software (SAV) v1.8 is an application that allows you, in real time, to view important quality metrics generated by the Real-Time Analysis (RTA) software on the Illumina sequencing systems. SAV Software is compatible with all HiSeq systems, HiSeq X, NextSeq, MiSeq, GAIIX, and HiScanSQ.

The SAV Software can be installed on a personal computer to view quality metrics from a remote location. This user guide provides instructions to install the software on a personal computer, and describes the various metrics that can be viewed using this tool. If the SAV Software is installed on your sequencing instrument, viewing metrics does not interfere with the sequencing run.

Figure 1 Sequencing Analysis Viewer Software

Setting Up Sequencing Analysis Viewer Software

This section describes how to set up Sequencing Analysis Viewer Software on your personal computer. You must have admin privileges to install the software.

NOTE
If the SAV Software is not installed on your sequencing instrument, contact your FAS. The instructions in this section are not for installing the SAV Software on the sequencing instrument.

Requirements

Sequencing Analysis Viewer Software does not need an advanced personal computer, because the instrument control computer running the Real-Time Analysis (RTA) software does the heavy computational work. The following items are required to run the software:

- Desktop computer running 32-bit or 64-bit version Windows XP, Windows Vista, or Windows 7
- Network access to the run data
- .Net framework 4.0

Downloading Sequencing Analysis Viewer Software Installer

To download the SAV Software installer, perform the following:

1. Open your browser, and go to support.illumina.com/sequencing/sequencing_software/sequencing_analysis_viewer_sav.html.
2. Click **Downloads**.
3. Download the Sequencing Analysis Viewer (SAV) v1.8.x Installer (*.zip) file.

Installing Sequencing Analysis Viewer Software

To install the SAV Software, perform the following:

1. Navigate to the location where the software is saved.
2. Double-click the folder to unzip the installer (*.zip).
3. Double-click the installer folder.
4. Double-click the installer (*.msi).
   The Sequencing Analysis Viewer Software Setup Wizard opens.
5. Click **Next**.
   The installation starts, and the progress screen opens. When the installation is finished, a desktop shortcut is created, and the setup complete screen opens.
6. Click **Finish**.

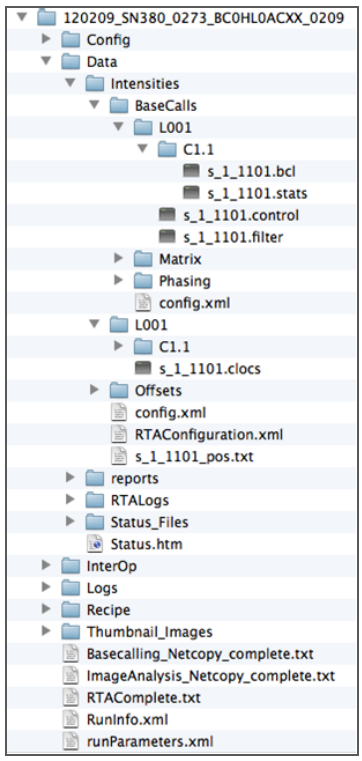
Testing Sequencing Analysis Viewer Software Installation

When you have installed the SAV Software, check whether the application was installed properly, and whether you have the proper permissions. Perform the following:

1 Double-click the Illumina Sequencing Analysis Viewer Software desktop shortcut, or go to C:\Illumina\Illumina Sequencing Analysis Viewer Software and double-click Sequencing Analysis Viewer Software.exe.
The SAV Software opens.

2 In the **Run Folder** field, copy the location of or click **Browse** to select a run folder. Make sure to highlight the run folder and not the parent folder or any folder/file inside the run folder. An example of a run folder from a HiSeq is shown here.

Figure 2 Connect to Run Folder



NOTE
The **Run Folder** screen varies depending on your Illumina sequencing system and the operating system on your personal computer.

3 Click **Refresh**.
The software starts loading data, and when completed it shows available quality metrics for that run.

Figure 3 Sequencing Analysis Viewer Software with Metrics Loaded

If the SAV Software does not display metrics, check whether you pointed the software to a valid run folder, and whether you have the proper permissions to view the data. If there is no valid run folder, the software shows an error message.

Sequencing Analysis Viewer Software Files

SAV Software uses the following files:

- runinfo.xml
- runparameters.xml
- The InterOp folder can be found in the directory: <run directory>\InterOp. The InterOp files provide the following information:
  - Extraction Metrics
  - Quality Metrics
  - Error Metrics
  - Tile Metrics
  - Corrected Intensity Metrics
  - Control Metrics

  For more information, see *InterOp Files* on page 28.
- Thumbnails (optional)

Data Availability

This table represents the earliest cycle at which data populates into the InterOp folders. Due to variations in RTA processing speeds, the following metrics may not be populated for a number of cycles after the cycle listed in the chart.

Metric	MiSeq v2	MiSeq v3	NextSeq 500	HiSeq	HiSeq X
FWHM	1	1	1	1	1
Template Generation	4	7	5	4	1
Intensity	4	7	5	5	1
Corrected Intensity	12	12	5	12	3
Signal to Noise	12	12	Not calculated	12	Not calculated
Cluster Density	20	7	5	5	1: Fixed at 2354k/mm^2
Clusters Passing Filter	25	25	25	25	25
Called Intensity	25	25	5	25	3
% Base	25	25	25	25	25
% Perfect Reads	25	25	Not calculated	25	Not calculated
%Q > 20, %Q > 30	25	25	25	25	25
% Phasing, % Prephasing	25	25	25	25	25
% Aligned	25	25	25	25	25
Error Rate	25	25	25	25	25
% PF	25	25	25	25	25

- Intensities are extracted after the template is generated.

  NOTE
  The image analysis catches up during template generation, so the template generation cycle takes significantly longer to finish. The analysis waits to end before proceeding with the next cycle. This behavior is normal.

- Corrected intensities are generated after phasing and prephasing are calculated. Base calls are generated after corrected intensities are generated.
- MiSeq, NextSeq, and HiSeq instruments running HCS 2.2.38 or higher generate phasing and prephasing estimates empirically for every cycle.

  TIP
  In any screen, click the **Refresh** button to gather the latest metrics.

- The phasing reported in the SAV Software is the tile median slope of the observed phasing corrections for cycles 1–25.

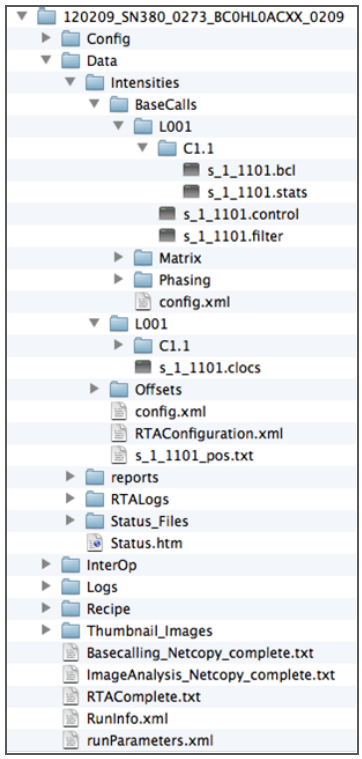
Loading Data

To load data, perform the following:

1 Double-click the Illumina Sequencing Analysis Viewer Software desktop shortcut, or go to C:\Illumina\Illumina Sequencing Analysis Viewer Software and double-click Sequencing Analysis Viewer Software.exe.
The Sequencing Analysis Viewer Software opens.

2 In the **Run Folder** field, copy the location of or click **Browse** to select a run folder. Make sure to highlight the run folder and not the parent folder or any folder/file inside the run folder. An example of a run folder from a HiSeq is shown here.

Figure 4 Connect to Run Folder



NOTE
The Run Folder screen varies depending on your Illumina sequencing system and the operating system on your personal computer.

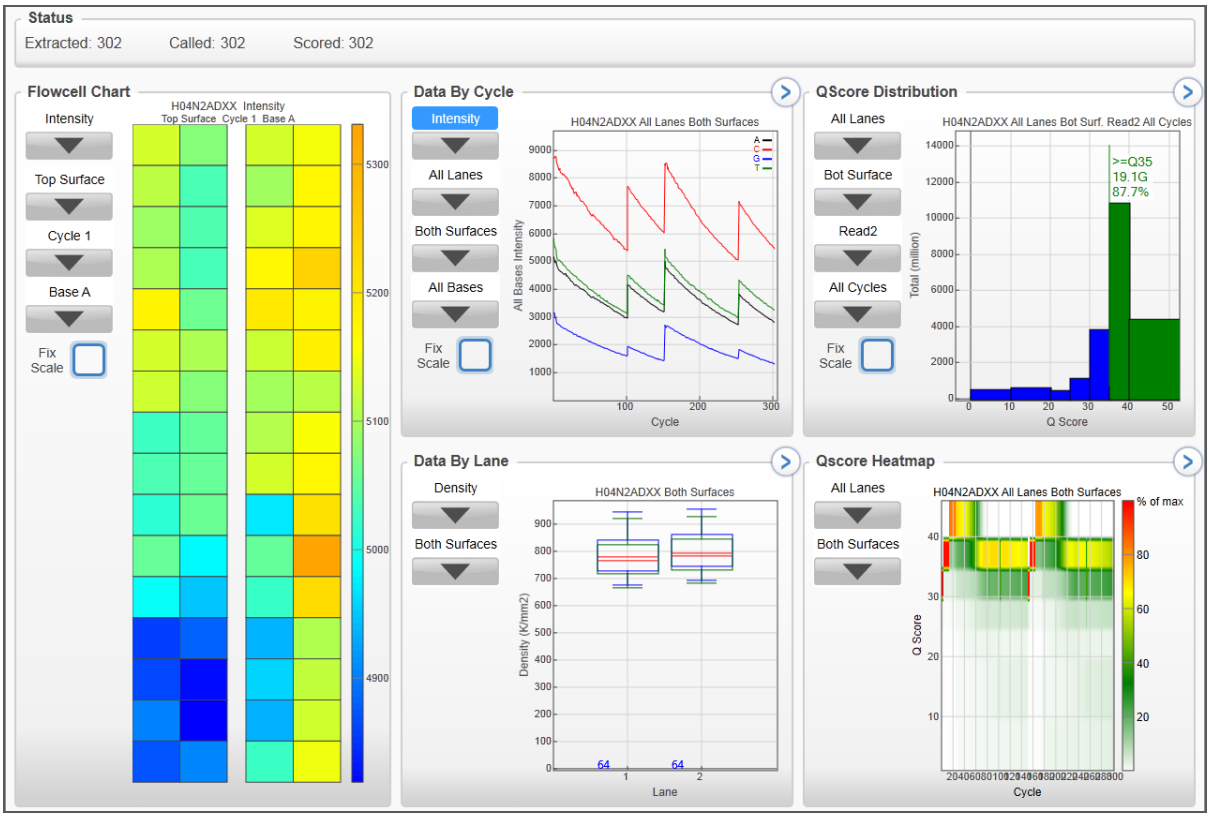
3 Click **Refresh**.
The SAV Software starts loading data showing quality metrics for that run.

Analysis Tab

The Analysis tab consists of six panes, which are described on the following pages:

- *Status Pane* on page 10
- *Flow Cell Chart* on page 11
- *Data by Cycle Plot* on page 12
- *Data by Lane Plot* on page 13
- *Q-Score Distribution Plot* on page 15
- *Q-Score Heat Map* on page 16

Figure 5 Analysis Tab



NOTE
If you are using the SAV Software to view metrics from the MiSeq sequencing instrument (which has only one lane), select 1 or All in the **Lane** drop-down list.

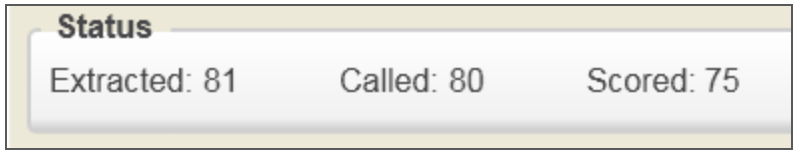
Status Pane

The Status pane reports the progress of the analysis. The following steps are reported:

- **Extracted**: the last complete cycle for which image analysis has been performed.
- **Called**: the last complete cycle that has been base-called.
- **Scored**: the last complete cycle that has been quality scored.

The statistics are updated as soon as there is new analysis information available, and you hit the **Refresh** button.

Figure 6 Status Pane

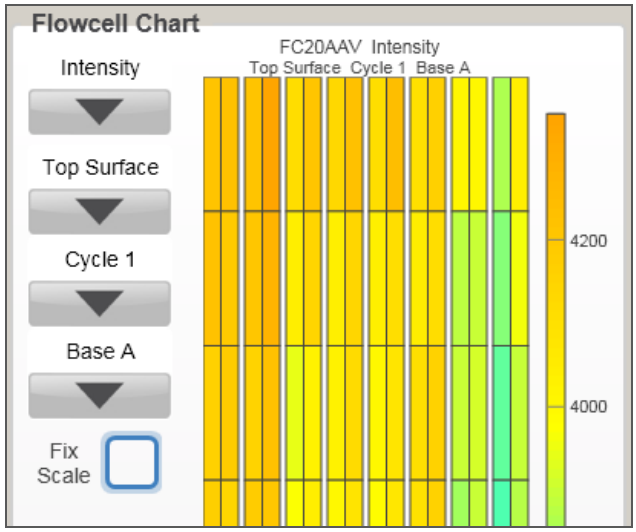


Flow Cell Chart

The Flow Cell Chart shows color-coded quality metrics per tile for the entire flow cell, and has the following features:

- You can select the displayed metric, surface (if your sequencer scans multiple surfaces), cycle, and base through the drop-down lists.
- The color bar to the right of the chart indicates the values that the colors represent. The color bar is also a slider: you can adjust the scale values by right-clicking; change the contrast using the mouse wheel.
  The chart is displayed with auto scaling by default, or can be fixed by checking the **Fix Scale** checkbox.
- Tiles that have not been measured or are not monitored are gray.
- The interactive tooltips provide the lane, tile, and value of the data point.
- Clicking a tile opens the Imaging tab with more detailed information for that tile (see *Imaging Tab* on page 17).
- By right-clicking an image, you can copy it to the clipboard.

Figure 7 Flow Cell Chart



You can monitor the following quality metrics with this chart:

- **Intensity**—This chart shows the intensity by color and cycle of the 90% percentile of the data for each tile. On platforms using two-channel sequencing, only red and green are shown. On platforms using four-channel sequencing, 4 channels are shown.
- **FWHM**—The average full width of clusters at half maximum (in pixels). For a more detailed description, see en.wikipedia.org/wiki/Full_width_at_half_maximum.
- **Corrected Intensity**:
  - MiSeq and HiSeq: The intensity corrected for cross talk between the color channels and phasing and prephasing.

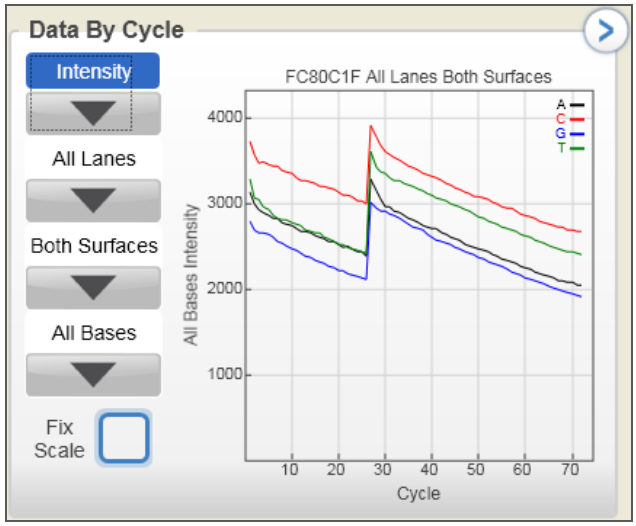
  - NextSeq: Intensities in NextSeq two-channel sequencing are calculated using a different method than in four-channel sequencing. The values presented here are not used in base calling, but can aid in assessing progress of a run. For NextSeq, called and corrected intensities graphs are identical.
- **Called Intensity**—The intensity for the called base.
- **% Base**—The percentage of clusters for which the selected base has been called.
- **Signal to Noise**—The signal to noise ratio is calculated as mean called intensity divided by standard deviation of non-called intensities. Not calculated for NextSeq two-channel sequencing or HiSeq X.
- **Error Rate**—The calculated error rate, as determined by a spiked in PhiX control sample. If a PhiX control sample is not run in the lane, this chart is not available.
- **% Perfect Reads**—The percentage of reads that align perfectly, as determined by a spiked in PhiX control sample. If a PhiX control sample is not run in the lane, this chart is not available. Not calculated for NextSeq two-channel sequencing.
- **%Q > 20**, **%Q > 30**—The percentage of bases with a quality score of 20 or 30 or higher, respectively. These charts are generated after the 25th cycle, and the values represent the current cycle. Click the **Accum** checkbox to view all cumulative data up to that point. Leave the box unchecked to display only the current cycle.
- **Median Q-Score**—The median Q-Score for each tile over all bases for the current cycle. These charts are generated after the 25th cycle. This plot is best used to examine the Q-scores of your run as it progresses. Bear in mind that the %Q30 plot can give an over simplified view due to its reliance on a single threshold.
- **Density**—The density of clusters for each tile (in thousands per mm^2).
- **Density PF**—The density of clusters passing filter for each tile (in thousands per mm^2).
- **Clusters**—The number of clusters for each tile (in millions).
- **Clusters PF**—The number of clusters passing filter for each tile (in millions).
- **% Phasing**, **% Prephasing**—The percentage of molecules in a cluster for which sequencing falls behind (phasing) or jumps ahead (prephasing) the current cycle within a read.
- **% Aligned**—The percentage of the sample that aligned to the PhiX genome.

Data by Cycle Plot

The Data by Cycle pane shows plots that allow you to follow the progression of quality metrics during a run. These plots have the following features:

- You can select the displayed metric, lane, surface, and base through the drop-down lists.
- The plots are displayed with auto scaling by default, or can be fixed by checking the **Fix Scale** checkbox.
- The chevron in the top right-hand corner toggles the plot between pane view and full screen view.
- You can pan the graph by clicking-and-dragging, zoom in by using the mouse wheel, and zoom in only on a particular axis by using the mouse wheel over that axis.
- By right-clicking an image, you can copy it to the clipboard.

Figure 8 Data by Cycle Plot



You can monitor the following quality metrics with this plot:

- **Intensity**—This plot shows the intensity by color of the 90% percentile of the data for each cycle.
- **FWHM**—The average full width of clusters at half maximum (in pixels). For a more detailed description, see en.wikipedia.org/wiki/Full_width_at_half_maximum.
- **Corrected Intensity**:
  - MiSeq and HiSeq: The intensity corrected for cross talk between the color channels and phasing and prephasing.
  - NextSeq: Intensities in NextSeq two-channel sequencing are calculated using a different method than in four-channel sequencing. The values presented here are not used in base calling, but can aide in assessing progress of a run. For NextSeq, called and corrected intensities graphs are identical.
- **Called Intensity**—The intensity for the called base.
- **% Base**—The percentage of clusters for which the selected base has been called.
- **Signal to Noise**—The signal to noise ratio is calculated as mean called intensity divided by standard deviation of non-called intensities. Not calculated for NextSeq two-channel sequencing or HiSeq X.
- **Error Rate**—The calculated error rate, as determined by a spiked in PhiX control sample. If a PhiX control sample is not run in the lane, this chart is not available.
- **% Perfect Reads**—The percentage of reads that align perfectly, as determined by a spiked in PhiX control sample. If a PhiX control sample is not run in the lane, this chart is not available. Not calculated for NextSeq two-channel sequencing.
- **%Q > 20**, **%Q > 30**—The percentage of bases with a quality score of 20 or 30 or higher, respectively. These charts are generated after the 25th cycle, and the values represent the current cycle. Click the **Accum** checkbox to view all cumulative data up to that point. Leave the box unchecked to display only the current cycle.

Data by Lane Plot

The Data by Lane pane shows plots that allow you to view quality metrics per lane. These plots have the following features:

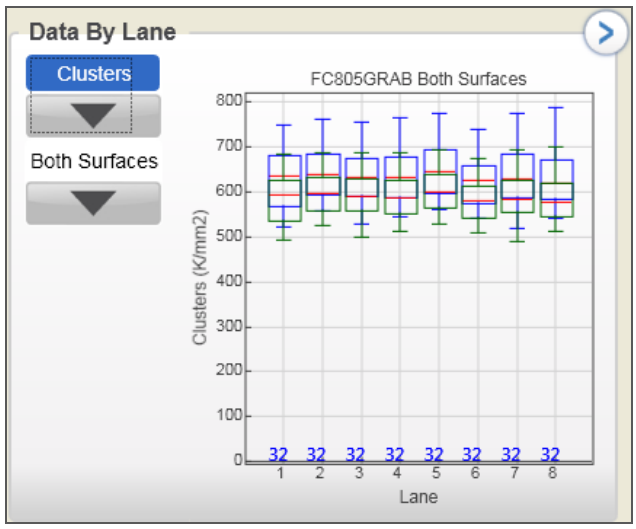
- You can select the displayed metric, surface, and read (when applicable) through the drop-down lists.

- The chevron in the top right-hand corner toggles the plot between pane view and full screen view.
- You can zoom in/out and pan the plots.
- By right-clicking an image, you can copy it to the clipboard.

The plots share a number of characteristics.

- The red line indicates the median tile value.
- The box outlines the interquartile range (the middle 50% of the data) for the tiles analyzed for the data point.
- The error bars delineate the minimum and maximum without outliers.
- The outliers are the values that are more than 1.5 times the interquartile range below the 25th percentile, or more than 1.5 times the interquartile range above the 75th percentile. Outliers are indicated as dots.
- The numbers above the X-axis indicate the number of analyzed tiles used for the data point.

Figure 9 Data by Lane Plot



You can monitor the following quality metrics with this plot (blue boxes are for raw clusters, green boxes for clusters passing filter):

- The density of clusters for each tile (in thousands per mm^2).
- The number of clusters for each tile (in millions).
- **% Phasing**, **% Prephasing**—The percentage of molecules in a cluster for which sequencing falls behind (phasing) or jumps ahead (prephasing) the current cycle within a read. The graphs are split out by read.

  For MiSeq, NextSeq, and HiSeq X, RTA generates phasing and prephasing estimates empirically for every cycle. The value displayed here is therefore not used in the actual phasing/prephasing calculations, but is an aggregate value determined from the first 25 cycles. For most applications, the reported value should be very close to the applied value. However, on MiSeq or NextSeq, for low diversity samples or samples with unbalanced base composition, the reported value might not accurately reflect the values being applied because the value changes from cycle to cycle. HiSeq X currently does not support low diversity samples or samples with unbalanced base composition.
- **% Aligned**—The percentage of the sample that aligned to the PhiX genome. The graphs are split out by read.

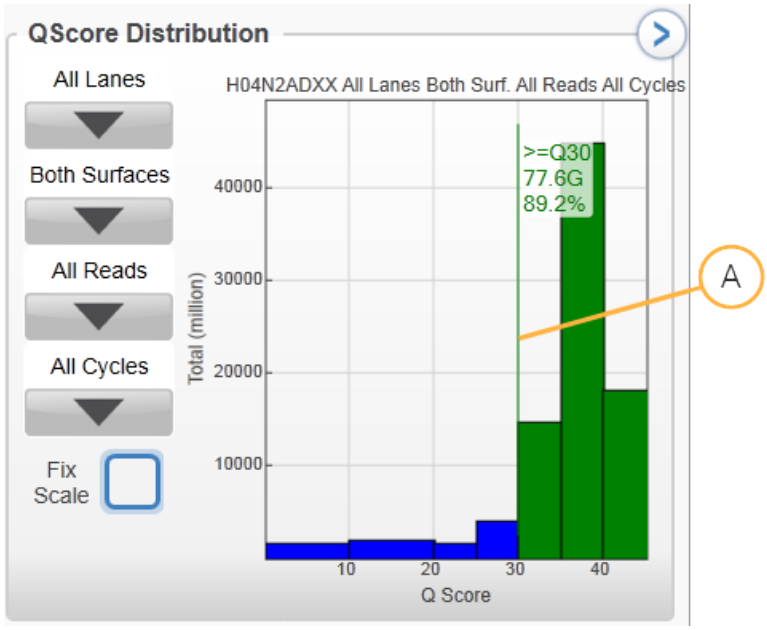
Q-Score Distribution Plot

The Q-score Distribution pane shows plots that allow you to view the number of reads by quality score. The quality score is cumulative for current cycle and previous cycles, and only reads that pass the quality filter are included. The quality scores are binned in groups: Q-score less than 20 in groups of 10, Q-scores 20–40 in groups of five, and one group of Q-scores above 40.

These plots have the following features:

- You can select the displayed lane, surface, read, and cycle through the drop-down lists.
- The **Read** drop-down list is used to select the first cycle for calculating the histogram.
- The **Cycle** drop-down list is used to determine the last cycle used for calculating the histogram. The last cycle for calculating the histogram is the minimum of the cycle selected in the Cycle drop-down box and the last cycle of read selected in **Read** drop-down list.
- The cutoff slider allows you to determine how many bases have a minimum Q-score or higher. Grab the slider with your mouse pointer, and drop it at the minimum Q-score. The SAV Software then calculates how many bases have that Q-score or higher.
- The chevron in the top right-hand corner toggles the plot between pane view and full screen view.
- You can pan the graph by clicking-and-dragging, zoom in by using the mouse wheel, and zoom in only on a particular axis by using the mouse wheel over that axis.
- By right-clicking an image, you can copy it to the clipboard.

Figure 10 Q-score Distribution Plot



A Cutoff Slider

The Q-score is based on the Phred scale. The following table lists Q-scores and the corresponding estimated base call error rate at that Q-score.

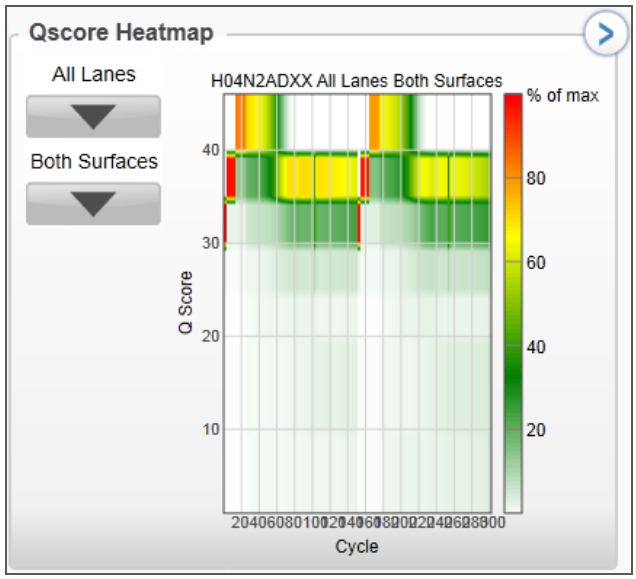
Q-score	Chance of Wrong Base Call
Q10	10%
Q20	1%
Q30	0.1%
Q40	0.01%

Q-Score Heat Map

The Q-score heat map shows plots that allow you to view the Q-score by cycle. These plots have the following features:

- You can select the displayed lane and surface through the drop-down lists.
- The color bars to the right of each chart indicate the values that the colors represent. The charts are displayed with auto scaling; the scale is always 0 to 100% of maximum value. Right-clicking allows you to change the color scheme.
- The chevron in the top right-hand corner toggles the plot between pane view and full screen view.
- By right-clicking an image, you can copy it to the clipboard.
- The quality score values are binned in groups: Q-score less than 20 in groups of 10, Q-scores 20–40 in groups of five, and one group of Q-scores above 40.

Figure 11 Q-score heat map

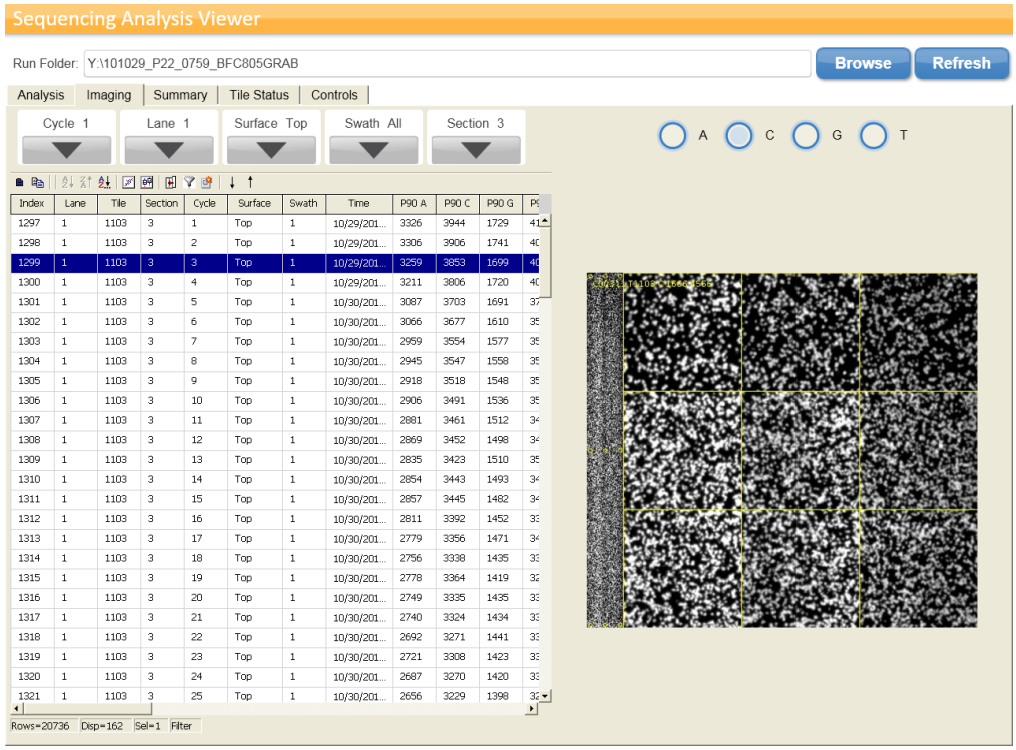


Imaging Tab

The Imaging tab lists detailed data and metrics for the run.

You can select the displayed cycle, lane, surface, swath, and section through the drop-down lists, while the checkboxes can be used to select the displayed bases. In addition, the imaging tab contains available thumbnails for the selected tile, highlighted in blue in the table. You can also use a mouse wheel or the arrow keys on a keyboard to scroll through images. Right-clicking on the top of a column provides more info.

Figure 12 Imaging Tab



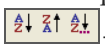
You can track the following metrics:

- **P90**—Shows the intensity by color of the 90% percentile of the data for each cycle.
- **Error Rate**—The calculated error rate, as determined by the PhiX alignment.
- **% Base**—The percentage of clusters for which the selected base has been called. Includes the percentage non-called (%NC).
- **FWHM**—The average full width of clusters at half maximum (in pixels). For a more detailed description, see en.wikipedia.org/wiki/Full_width_at_half_maximum.
- **Density**—The density of clusters detected by image analysis (in thousands per mm^2).
- **Density PF**—The density of clusters passing filter (in thousands per mm^2).
- **% Aligned**—The percentage of the sample that aligned to the PhiX genome, Read 1 and Read 2.
- **% Phasing/Prephasing**—The value used by RTA for the percentage of molecules in a cluster for which sequencing falls behind (phasing) or jumps ahead (prephasing) the current cycle within a read, Read 1 and Read 2.
- **Corrected Base**— MiSeq and HiSeq: The intensity corrected for cross talk between the color channels and phasing and prephasing. NextSeq: Intensities in NextSeq 2-channel sequencing are calculated using a different method than in 4-channel sequencing. The values presented here are not used in base calling, but can aide in

assessing progress of a run. For NextSeq, called and corrected base graphs are identical.

- **Called Base**—The intensity of the clusters that are called that base, after correction for cross talk and phasing/prephasing.
- **Signal to noise**—The signal to noise ratio is calculated as mean called intensity divided by standard deviation of non-called intensities. Not calculated for NextSeq two-channel sequencing.
- %**Q>=30**—The percentage of bases with a quality score of 30 or higher. This chart is generated after the 25th cycle, and the values represent the current cycle.
- **%PF**—The percentage of clusters passing filter.

The buttons above the data table allow you to do the following:

-  —Select all rows.
-  —Copy selected rows to clipboard.
-  —Sort columns in ascending, descending, or custom fashion.
-  —Create custom scatter or box plot.
-  —Choose columns to display.
-  —Filter columns, or clear filter.
-  —Move thumbnail selection up or down.

Summary Tab

The Summary tab leads to tables with basic data quality metrics summarized per lane and per read. All the statistics are given as means and standard deviations over the tiles used in the lane.

The following metrics are displayed in the top table, split out by read and total:

Yield Total	The number of bases sequenced, which is updated as the run progresses.
Projected Total Yield	The projected number of bases expected to be sequenced at the end of the run.
Yield Perfect	The number of bases in reads that align perfectly, as determined by alignment to PhiX of reads derived from a spiked in PhiX control sample. If a PhiX control sample is not run in the lane, this chart is not available.
Yield <=3 errors	The number of bases in reads that align with three errors or less, as determined by a spiked in PhiX control sample. If a PhiX control sample is not run in the lane, this chart is not available, and shows a zero value. This value is not calculated for NextSeq two-channel sequencing, and the value shown is always zero.
Aligned	The percentage of the sample that aligned to the PhiX genome, which is determined for each level or read independently.
% Perfect [Num Usable Cycles]	The percentage of bases in reads that align perfectly, as determined by a spiked in PhiX control sample, at the cycle indicated in the brackets. If no PhiX control sample is run in the lane, this chart shows 0% and the number of cycles used. This value is not calculated for NextSeq two-channel sequencing, and the value shown is always zero.
% <=3 errors [Num Usable Cycles]	The percentage of bases in reads that align with three errors or less, as determined by a spiked in PhiX control sample, at the indicated cycle. If no PhiX control sample is run in the lane, this chart shows 0% and the number of cycles used. This value is not calculated for NextSeq two-channel sequencing, and the value shown is always zero.
Error Rate	The calculated error rate of the reads that aligned to PhiX.
Intensity Cycle 1	The average of the A channel intensity measured at the first cycle averaged over filtered clusters.
% Intensity Cycle 20	The corresponding intensity statistic at cycle 20 as a percentage of that value at the first cycle. 100%x(Intensity at cycle 20)/(Intensity at cycle 1).
%Q>=30	The percentage of bases with a quality score of 30 or higher, respectively. This chart is generated after the 25th cycle, and the values represent the current cycle.

The following metrics are available in the Read tables, split out by lane:

Tiles	The number of tiles per lane.
Density	The density of clusters (in thousands per mm^2) detected by image analysis, +/- one standard deviation.
Clusters PF	The percentage of clusters passing filtering, +/- one standard deviation.
Phas./Prephas.	The value used by RTA for the percentage of molecules in a cluster for which sequencing falls behind (phasing) or jumps ahead (prephasing) the current cycle within a read. For MiSeq and NextSeq, RTA generates phasing and prephasing estimates empirically for every cycle. The value displayed here is therefore not used in the actual phasing/prephasing calculations, but is an aggregate value determined from the first 25 cycles. For most applications, the value reported is very close to the value that is applied. For low diversity samples or samples with unbalanced base composition, the reported value can diverge from the values being applied because the value changes from cycle to cycle.
Reads	The number of clusters (in millions).
Reads PF	The number of clusters (in millions) passing filtering.
%Q>=30	The percentage of bases with a quality score of 30 or higher, respectively. This chart is generated after the 25^{th} cycle, and the values represent the current cycle.
Yield	The number of bases sequenced which passed filter.
Cycles Err Rated	The number of cycles that have been error rated using PhiX, starting at cycle 1.
Aligned	The percentage that aligned to the PhiX genome.
Error Rate	The calculated error rate, as determined by the PhiX alignment. Subsequent columns display the error rate for cycles 1–35, 1–75, and 1–100.
Intensity Cycle 1	The average of the A channel intensity measured at the first cycle averaged over filtered clusters.
%Intensity Cycle 20	The corresponding intensity statistic at cycle 20 as a percentage of that value at the first cycle. 100%x(Intensity at cycle 20)/(Intensity at cycle 1).

The bottom of the page contains two buttons, Copy to Clipboard and Generate IVC Plots, so you can copy the data to your computer, and generate your IVC plots at any time during the run.

Tile Status Tab

The Tile Status tab shows the progress of the analysis. The colors indicate the status of the image analysis for a tile for a particular cycle, as indicated in the legend to the right.

The tiles are organized by lane, and organized the following way:

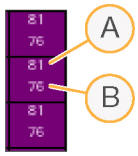
- The HiSeq Rapid Run has two lanes. Each lane has four columns. The first two columns are for the top surface of that lane, swath 1 and swath 2 respectively. A swath is a column of tiles in one lane, and there are two swaths per lane surface. The last two columns are for swath 1 and swath 2 of the bottom surface.
- The HiSeq High Output has eight lanes. Each lane has six columns representing the three swaths per surface.
- The HiSeq X has eight lanes. Each lane has four columns representing two swaths per surface.
- The MiSeq has one lane. The two columns represent a single swath per surface.
- The NextSeq has four lanes. Each lane is divided into three sections, which are imaged independently in three swaths per lane.

Clicking the **Refresh** button provides the most recent data.

Figure 13 Tile Status

In every tile, two numbers indicate the progress of base calling and quality scoring. For every tile, the top number indicates the last cycle that has been base called, and the bottom number indicates the last cycle for which quality has been scored.

Figure 14 Base Calling and Quality Scoring Progress



A Last Base-Called Cycle for Tile
B Last Quality-Scored Cycle for Tile

TruSeq Controls Tab

Sequencing Analysis Viewer Software reports the results of the analysis of the in-line controls on the TruSeq Controls tab. You can select the lane using the drop-down lists.

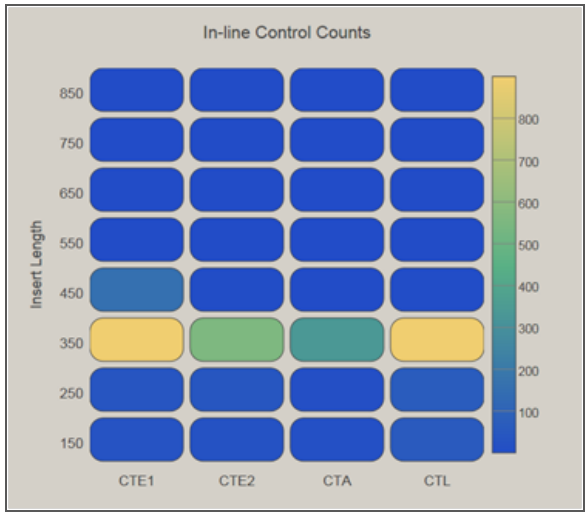
NOTE
This tab is not available for NextSeq or HiSeq X.

The in-line controls are added during the sample prep to check whether the sample prep steps are performed properly. RTA fishes them out of the samples and provides the data that the SAV Software can report. The following controls are included:

Step	Reaction	Enzyme	Control
End-Repair	Chew-back of 3′ overhangs/ fill-in of 5′ overhangs	DNA Polymerase	CTE1
End-Repair	Phosphorylation of 5′ OH groups	Polynucleotide Kinase	CTE2
A Tailing	Add 3′ A overhang	DNA Polymerase	CTA
Adapter ligation	Join adapters to inserts	DNA Ligase	CTL

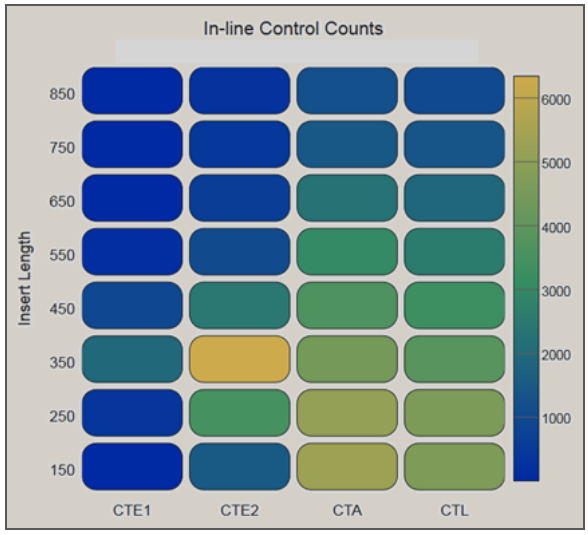
A successful run should have a positive signal for all four controls at one or two adjacent sizes. The reported size should correspond to the excised fragment. The absolute number of reads for each control does not matter much, but there should be a clear peak near the expected size. An example of a successful run for a TruSeq DNA 350 bp gel excised fragment is shown here.

Figure 15 In-Line Controls in a Successful Run



Libraries prepared with the TruSeq DNA PCR-Free Sample Prep kit show a narrow size distribution for CTE1 and CTE2. However, CTA and CTL size distributions are broad as the size selection step occurs upstream of the addition of these controls. An example of a successful run with a 350 bp size-selected insert is shown here.

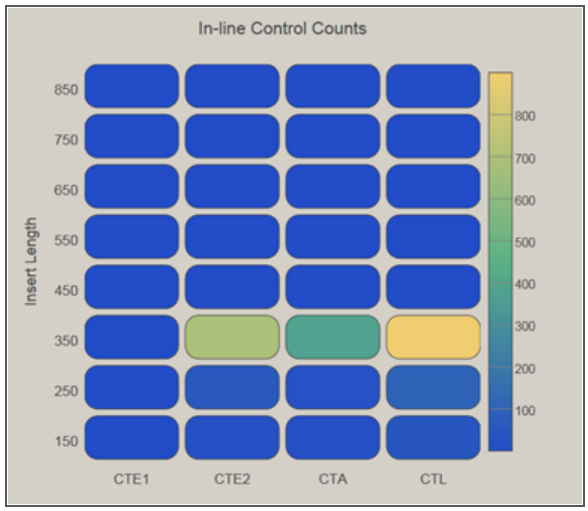
Figure 16 In-Line Controls in a Successful TruSeq DNA PCR-Free Run



If a run was unsuccessful due to an inefficient step in the sample prep, the controls would look like one of the following examples:

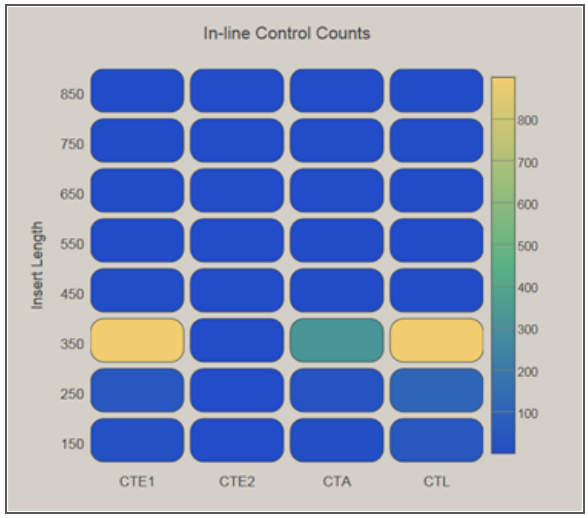
Inefficient End-Repair due to a Problem with DNA Polymerase

Figure 17 In-Line Control CTE1 Problem



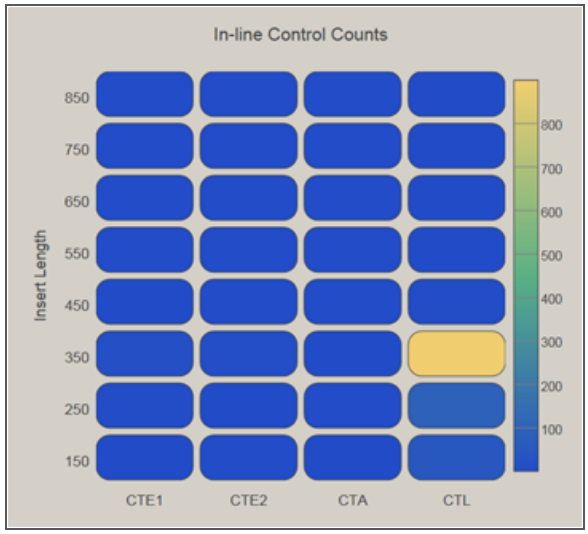
Inefficient End-Repair due to a Problem with Polynucleotide Kinase

Figure 18 In-Line Control CTE2 Problem



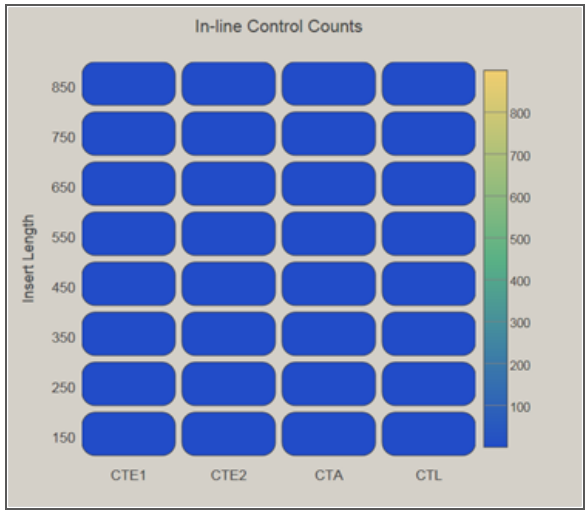
Inefficient A Tailing due to a Problem with DNA Polymerase

Figure 19 In-Line Control CTA Problem



Inefficient Adapter Ligation due to a Problem with DNA Ligase

Figure 20 In-Line Control CTL Problem



Indexing Tab

The Indexing tab lists count information for indexes used in the run. The Indexing tab is only available if a sample sheet was supplied to the instrument control software at the start of the run and the run is an index run.

NOTE
On the NextSeq and HiSeq X using RTA 2.0, this tab is populated if SAV Software is pointed to the InterOp folder after running bcl2fastq2. Otherwise, this tab is not populated. If the run was sent to BaseSpace, this info is available in BaseSpace.

You can select the displayed lane through the drop-down list.

The first table provides an overall summary of the indexing performance for that lane, including:

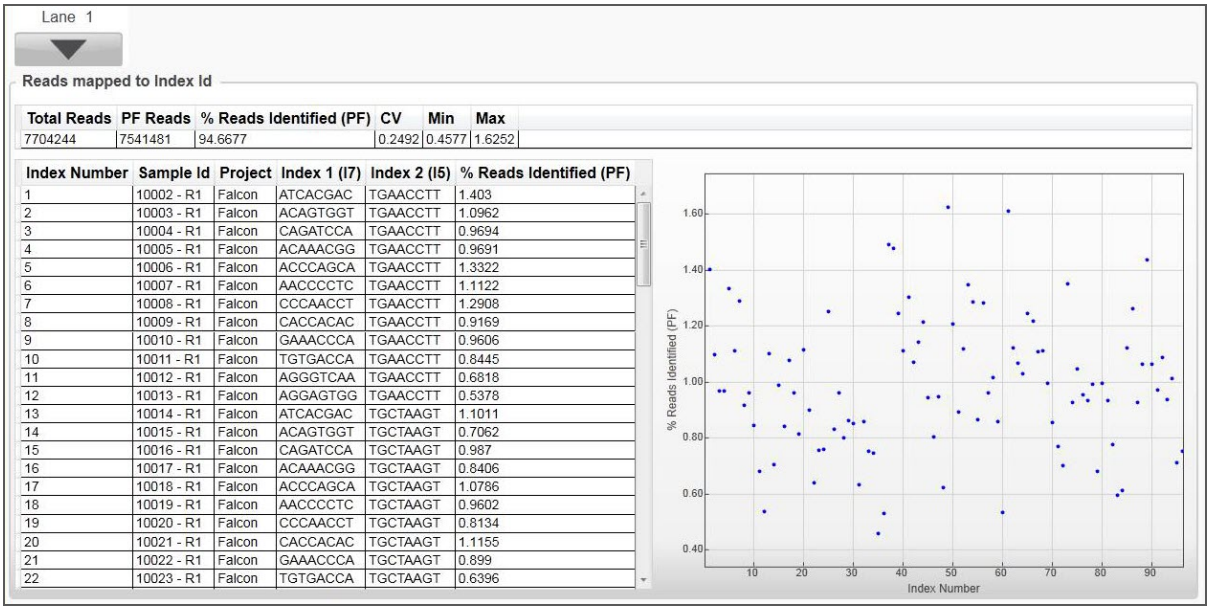
Total Reads	The total number of reads for this lane.
PF Reads	The total number of passing filter reads for this lane.
% Reads Identified (PF)	The total fraction of passing filter reads assigned to an index.
CV	The coefficient of variation for the number of counts across all indexes.
Min	The lowest representation for any index.
Max	The highest representation for any index.

Further information is provided regarding the frequency of individual indexes in both table and graph form. The table contains several columns, including

Index Number	A unique number assigned to each index by AS for display purposes.
Sample ID	The sample ID assigned to an index in the sample sheet.
Project	The project assigned to an index in the sample sheet.
Index 1 (I7)	The sequence for the first Index Read.
Index 2 (I5)	The sequence for the second Index Read.
% Reads Identified (PF)	The number of reads (only includes Passing Filter reads) mapped to this index.

This information is also displayed in graphical form. In the graphical display, indexes are ordered according to the unique Index Number assigned by AS.

Figure 21 Indexing Tab



InterOp Files

The InterOp files can be found in the directory: <run directory>\InterOp. These file formats are the same for all Illumina sequencing systems except where noted. Sequencing Analysis Viewer Software shows the information of several of these files, which are described here.

Extraction Metrics (ExtractionMetricsOut.bin)

Contains extraction metrics such as FWHM scores and raw intensities.

- byte 0: file version number (2)
- byte 1: length of each record
- bytes (N * 38 + 2) - (N *38 + 39): record:
  - 2 bytes: lane number (uint16)
  - 2 bytes: tile number (uint16)
  - 2 bytes: cycle number (uint16)
  - 4 x 4 bytes: fwhm scores (float) for channel [A, C, G, T] respectively
  - 2 x 4 bytes: intensities (uint16) for channel [A, C, G, T] respectively
  - 8 bytes: date/time of cif creation

  Where N is the record index

Quality Metrics (QualityMetricsOut.bin)

Contains quality score distribution.

- byte 0: file version number (4)
- byte 1: length of each record
- bytes (N * 206 + 2) - (N *206 + 207): record:
  - 2 bytes: lane number (uint16)
  - 2 bytes: tile number (uint16)
  - 2 bytes: cycle number (uint16)
  - 4 x 50 bytes: number of clusters assigned score (uint32) Q1 through Q50

  Where N is the record index

NOTE
MiSeq and HiSeq use RTA version 1.1x
NextSeq and HiSeq X use RTA version 2.x

Error Metrics (ErrorMetricsOstarut.bin)

Contains cycle error rate as well as counts for perfect reads and read with 1–4 errors.

- byte 0: file version number (3)
- byte 1: length of each record
- bytes (N * 30 + 2) - (N *30 + 11): record:
  - 2 bytes: lane number (uint16)
  - 2 bytes: tile number (uint16)
  - 2 bytes: cycle number (uint16)
  - 4 bytes: error rate (float)
  - 4 bytes: number of perfect reads (uint32)
  - 4 bytes: number of reads with one error (uint32)
  - 4 bytes: number of reads with two errors (uint32)
  - 4 bytes: number of reads with three errors (uint32)
  - 4 bytes: number of reads with four errors (uint32)

Where N is the record index

Tile Metrics (TileMetricsOut.bin)

Contains aggregate or read metrics by tile.

- byte 0: file version number (2)
- byte 1: length of each record
- bytes (N * 10 + 2) - (N *10 + 11): record:
  - 2 bytes: lane number (uint16)
  - 2 bytes: tile number (uint16)
  - 2 bytes: metric code (uint16)
  - 4 bytes: metric value (float)

  Where N is the record index and possible metric codes are:
  - code 100: cluster density (k/mm^2)
  - code 101: cluster density passing filters (k/mm^2)
  - code 102: number of clusters
  - code 103: number of clusters passing filters
  - code (200 + (N – 1) * 2): phasing for read N
  - code (201 + (N – 1) * 2): prephasing for read N
  - code (300 + N – 1): percent aligned for read N

Corrected Intensity Metrics (CorrectedIntMetricsOut.bin)

Contains base call metrics.

- byte 0: file version number (2)
- byte 1: length of each record
- bytes (N * 48 + 2) - (N *48 + 49): record:
  - 2 bytes: lane number (uint16)
  - 2 bytes: tile number (uint16)
  - 2 bytes: cycle number (uint16)
  - 2 bytes: average intensity (uint16)
  - 2 bytes: average corrected int for channel A (uint16)
  - 2 bytes: average corrected int for channel C (uint16)
  - 2 bytes: average corrected int for channel G (uint16)
  - 2 bytes: average corrected int for channel T (uint16)
  - 2 bytes: average corrected int for called clusters for base A (uint16)
  - 2 bytes: average corrected int for called clusters for base C (uint16)
  - 2 bytes: average corrected int for called clusters for base G (uint16)
  - 2 bytes: average corrected int for called clusters for base T (uint16)
  - 20 bytes: number of base calls (float) for No Call and channel [A, C, G, T] respectively
  - 4 bytes: signal to noise ratio (float)

Control Metrics (ControlMetricsOut.bin)

Contains pull out information for Illumina in-line sample controls. Not generated in RTA 2.x.

- byte 0: file version number (1)
- bytes (variable length): record:
  - 2 bytes: lane number (uint16)
  - 2 bytes: tile number (uint16)
  - 2 bytes: read number (uint16)

- 2 bytes: number bytes X for control name(uint16)
- X bytes: control name string (string in UTF8Encoding)
- 2 bytes: number bytes Y for index name(uint16)
- Y bytes: index name string (string in UTF8Encoding)
- 4 bytes: # clusters identified as control (uint32)

Image Metrics (ImageMetricsOut.bin)

Contains min max contrast values for image. Not generated in RTA 2.x.

- byte 0: file version number (1)
- byte 1: length of each record
- bytes (N * 12 + 2) - (N *12 + 13): record:
  - 2 bytes: lane number (uint16)
  - 2 bytes: tile number (uint16)
  - 2 bytes: cycle number (uint16)
  - 2 bytes: channel id (uint16) where 0=A, 1=C, 2=G, 3=T
  - 2 bytes: min contrast value for image (uint16)
  - 2 bytes: max contrast value for image (uint16)

Index Metrics t (IndexMetricsOut.bin):

Report the indexes count. Not generated in RTA 2.x.

- Byte 0: file version (1)
- Bytes (variable length): record:
  - 2 bytes: lane number(unint16)
  - 2 bytes: tile number(unint16)
  - 2 bytes: read number(unint16)
  - 2 bytes: number of bytes Y for index name(unint16)
  - Y bytes: index name string (string in UTF8Encoding)
  - 4 bytes: # clusters identified as index (uint32)
  - 2 bytes: number of bytes V for sample name(unint16)
  - V bytes: sample name string (string in UTF8Encoding)
  - 2 bytes: number of bytes W for sample project(unint16)
  - W bytes: sample project string (string in UTF8Encoding)

Notes

Technical Assistance

For technical assistance, contact Illumina Technical Support.

Table 1 Illumina General Contact Information

Address	5200 Illumina Way San Diego, CA 92122 USA
Website	www.illumina.com
Email	techsupport@illumina.com

Table 2 Illumina Customer Support Telephone Numbers

Region	Contact Number	Region	Contact Number
North America	1.800.809.4566	Italy	800.874909
Austria	0800.296575	Netherlands	0800.0223859
Belgium	0800.81102	Norway	800.16836
Denmark	80882346	Spain	900.812168
Finland	0800.918363	Sweden	020790181
France	0800.911850	Switzerland	0800.563118
Germany	0800.180.8994	United Kingdom	0800.917.0041
Ireland	1.800.812949	Other countries	+44.1799.534000

Safety Data Sheets

Safety data sheets (SDSs) are available on the Illumina website at support.illumina.com/sds.html.

Illumina
San Diego, California 92122 U.S.A.
+1.800.809.ILMN (4566)
+1.858.202.4566 (outside North America)
techsupport@illumina.com
www.illumina.com