Sequencing Analysis Viewer v2.1
Software Guide

Introduction 3
Set Up Sequencing Analysis Viewer Software 3
Metric Definitions 5
Loading Data 6
Analysis Tab 7
Imaging Tab 9
Summary Tab 10
Indexing Tab 11
InterOp Files 11
Revision History 18
Technical Assistance
This document and its contents are proprietary to Illumina, Inc. and its affiliates ("Illumina"), and are intended solely for the contractual use of its customer in connection with the use of the product(s) described herein and for no other purpose. This document and its contents shall not be used or distributed for any other purpose and/or otherwise communicated, disclosed, or reproduced in any way whatsoever without the prior written consent of Illumina. Illumina does not convey any license under its patent, trademark, copyright, or common-law rights nor similar rights of any third parties by this document.

The instructions in this document must be strictly and explicitly followed by qualified and properly trained personnel in order to ensure the proper and safe use of the product(s) described herein. All of the contents of this document must be fully read and understood prior to using such product(s).

FAILURE TO COMPLETELY READ AND EXPLICITLY FOLLOW ALL OF THE INSTRUCTIONS CONTAINED HEREIN MAY RESULT IN DAMAGE TO THE PRODUCT(S), INJURY TO PERSONS, INCLUDING TO USERS OR OTHERS, AND DAMAGE TO OTHER PROPERTY.

ILLUMINA DOES NOT ASSUME ANY LIABILITY ARISING OUT OF THE IMPROPER USE OF THE PRODUCT(S) DESCRIBED HEREIN (INCLUDING PARTS THEREOF OR SOFTWARE).

© 2017 Illumina, Inc. All rights reserved.

Illumina, HiScan, HiSeq, HiSeq X, MiniSeq, MiSeq, NextSeq, NovaSeq, the pumpkin orange color, and the streaming bases design are trademarks of Illumina, Inc. and/or its affiliate(s) in the U.S. and/or other countries. All other names, logos, and other trademarks are the property of their respective owners.
Introduction

Sequencing Analysis Viewer Software (SAV) v2.1 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 is compatible with all HiSeq systems, HiSeq X, NextSeq, MiSeq, MiniSeq, NovaSeq, and HiScanSQ.

**NOTE**
This version does not support Genome Analyzer IIx; use SAV v1.8.37 instead.

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.

**NOTE**
Before installing SAV, check Requirements on page 3 to make sure that the target computer is up to date.

Set 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 64-bit version Windows XP, Windows Vista, Windows 7 or later
- Network access to the run data
- .Net framework 4.5.1
- Visual C++ Redistributable for Visual Studio 2015

Download the Sequencing Analysis Viewer Software Installer

To download the Sequencing Analysis Viewer Software installer, perform the following steps.

1. In your browser, open the Sequencing Analysis Viewer (SAV) support page.
2. Click Downloads.
3. Download the Sequencing Analysis Viewer (SAV) v2.1 Installer (*.zip) file.

Install Sequencing Analysis Viewer Software

To install the SAV Software, perform the following steps.

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.

Test the Sequencing Analysis Viewer Software Installation

When you have installed the SAV Software, make sure that the application was installed properly and that you have the proper permissions. Perform the following steps.

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 folder location 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.

   **Figure 1** Example of a HiSeq Run Folder

   ![Example of a HiSeq 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.
If the Sequencing Analysis Viewer Software does not display metrics, make sure that you pointed the software to a valid run folder and that 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
  - Image Metrics
  - Index Metrics
  - Empirical Phasing Metrics

For more information, see *InterOp Files* on page 11.

- Thumbnails (optional)

**Metric Definitions**

The following table provides descriptions for the metrics shown in the Analysis and Imaging Panels of SAV Software.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensity (also referred to as P90)</td>
<td>The 90% percentile extracted intensity for a given image (lane/tile/cycle/channel combination). On platforms using two-channel sequencing, only red and green are shown. On platforms using four-channel sequencing, 4 channels (A, C, G, and T) are shown.</td>
</tr>
<tr>
<td>FWHM</td>
<td>The average full width of clusters at half maximum (representing their approximate size in pixels). For a more detailed description, see the Wikipedia page for Full width at half maximum.</td>
</tr>
<tr>
<td>Corrected Intensity</td>
<td>HiSeqX, HiSeq 3000 and 4000, MiniSeq, NovaSeq — This metric has been deprecated. MiSeq, HiSeq 1000, 1500, 2000, and 2500 — 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.</td>
</tr>
<tr>
<td>Called Intensity</td>
<td>For a given base in a lane/tile/cycle, the average intensity for all clusters that were called as that base. Not reported for NovaSeq.</td>
</tr>
<tr>
<td>% No Calls</td>
<td>The percentage of clusters on a tile for which no base (N) has been called</td>
</tr>
<tr>
<td>% Base</td>
<td>The percentage of called (non-N) clusters for which the selected base has been called.</td>
</tr>
<tr>
<td>Signal to Noise</td>
<td>The signal to noise ratio is calculated as mean called intensity divided by standard deviation of noncalled intensities. Not calculated for NextSeq and MiniSeq two-channel sequencing, HiSeq 3000/4000, HiSeqX, or NovaSeq.</td>
</tr>
</tbody>
</table>
### Definition

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Rate</td>
<td>The calculated error rate, as determined by a spiked in PhiX control sample. If a PhiX control sample is not run in the lane or if the tile has not been selected for alignment, this number is not available.</td>
</tr>
<tr>
<td>% Q ≥ 20, %Q ≥ 30</td>
<td>The percentage of bases with a quality score of 20 or 30 or higher, respectively.</td>
</tr>
<tr>
<td>Median Q-Score</td>
<td>The median Q-Score for each tile over all bases for the current cycle. These charts are generated after the 25th cycle. This metric is best used to examine the Q-scores of your run as it progresses. The %Q30 plot can give an over simplified view due to its reliance on a single threshold.</td>
</tr>
<tr>
<td>Density</td>
<td>The density of clusters for each tile (in thousands per mm2).</td>
</tr>
<tr>
<td>Density PF</td>
<td>The density of clusters passing filter for each tile (in thousands per mm2).</td>
</tr>
<tr>
<td>Clusters</td>
<td>The number of clusters for each tile (in millions).</td>
</tr>
<tr>
<td>Clusters PF</td>
<td>The number of clusters passing filter for each tile (in millions).</td>
</tr>
<tr>
<td>% Pass Filter</td>
<td>The percentage of clusters passing filter.</td>
</tr>
<tr>
<td>% Phasing, % Prephasing</td>
<td>The average rate (percentage per cycle) at which molecules in a cluster fall behind (phasing) or jump ahead (prephasing) during the read. This metric is an estimate from 25 cycles of data. For an estimate calculated from the entire read, see the Phasing and Prephasing Slope and Offset columns.</td>
</tr>
<tr>
<td>% Phasing Weight, % Prephasing Weight (also referred to as Phasing and Prephasing)</td>
<td>The rate for a given cycle at which molecules in a cluster fall behind (phasing) or jump ahead (prephasing).</td>
</tr>
<tr>
<td>Phasing Offset, Prephasing Offset</td>
<td>The full-read estimate for the starting rate per cycle at which molecules in a cluster fall behind (phasing) or jump ahead (prephasing) during the read.</td>
</tr>
<tr>
<td>Phasing Slope, Prephasing Slope</td>
<td>The full-read estimate for the average rate per cycle at which molecules in a cluster fall behind (phasing) or jump ahead (prephasing) during the read.</td>
</tr>
<tr>
<td>% Aligned</td>
<td>The percentage of the passing filter clusters that aligned to the PhiX genome.</td>
</tr>
<tr>
<td>Minimum/Maximum Contrast</td>
<td>The 10th and 95.5th percentiles per channel of selected columns of the raw image, respectively</td>
</tr>
</tbody>
</table>

### Loading Data

To load data, perform the following steps.

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. Click the tab containing the appropriate query information.

3. In the Run Folder field, copy the folder location 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.

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

4. Click Refresh.

   The SAV Software starts loading data showing quality metrics for that run.

   **NOTE**
   
   You can also load 1 of the 10 most recently loaded runs using the Recent Runs drop-down list.
Analysis Tab

The Analysis tab consists of 6 panes, which are described on the following pages:

- Flow Cell Chart on page 1
- Data By Cycle on page 1
- Data By Lane on page 1
- Q-Score Distribution on page 1
- Q-Score Heat Map on page 1

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

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. 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. For some metrics (% ≥ Q20 and % ≥ Q30), you can monitor the metrics for a single cycle by default. Alternatively, you can monitor the metrics for the entire run (up to that cycle) by selecting the Accum 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 9).
- By right-clicking an image, you can copy it to the clipboard.
- The chevron in the top right-hand corner toggles the plot between pane view and full screen view.

You can monitor the following metrics in this chart: Intensity, FWHM, Corrected Intensity, Called Intensity, % Base, % No Calls, Signal to Noise, Error Rate, % ≥ Q20, % ≥ Q30, Median Q-Score, Density, Density PF, Clusters, Clusters PF, Phasing, Prephasing, % Phasing, % Prephasing, and % Aligned. For more information, see Metric Definitions on page 5.

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. For some metrics (% ≥ Q20 and % ≥ Q30), you can monitor the metrics for a single cycle by default. Alternatively, you can monitor the metrics for the entire run (up to that cycle) by selecting the Accum 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.

You can monitor the following metrics in this chart: Intensity, FWHM, Corrected Intensity, Called Intensity, % Base, % No Calls, Phasing, Prephasing, Median Q-Score, Signal to Noise, Error Rate, % ≥ Q20, and % ≥ Q30. For more information, see Metric Definitions on page 5.

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.
- For some metrics (Density and Clusters), there are 2 box plots present per lane.
  - Blue boxes — Represent the raw cluster metric
  - Green boxes — Represent the equivalent PF cluster metric

The plots share various 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.

You can monitor the following metrics in this chart: Density/Density PF, Clusters/Clusters PF, % Phasing, % Prephasing, and % Aligned. For more information, see Metric Definitions on page 5.

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 chastity filter are included. The quality scores are binned in groups.

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 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.

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 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.

### Imaging Tab

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

You can select the displayed cycle, lane, surface, swath, read, tile number, and section through the drop-down lists, and 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.

You can monitor the following metrics in this chart: Density, Density PF, Clusters, Clusters PF, % Pass Filter, % Aligned, % Phasing, % Prephasing, Error Rate, % ≥ Q20, % ≥ Q30, P90, % No Calls, % Base, FWHM, Corrected Intensity, Called Intensity, Signal to Noise, Minimum Contrast, Maximum Contrast, % Phasing Weight, % Prephasing Weight, Phasing Slope, Phasing Offset, Prephasing Slope, and Prephasing Offset. For more information, see Metric Definitions on page 5.

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

- Select all rows.
- Copy selected rows to clipboard.
- Export Imaging table to .txt file.
- 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 Cycle Status above the top table shows a summary of the last cycle extracted, basecalled, Q-scored, and error rated for each tile.

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

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>The sequencing read level.</td>
</tr>
<tr>
<td>Yield Total</td>
<td>The number of bases sequenced, which is updated as the run progresses.</td>
</tr>
<tr>
<td>Projected Total Yield</td>
<td>The projected number of bases expected to be sequenced at the end of the run, which is updated as the run progresses.</td>
</tr>
<tr>
<td>Aligned</td>
<td>The percentage of the passing filter clusters that aligned to the PhiX genome, which is determined for each level or read independently.</td>
</tr>
<tr>
<td>Error Rate</td>
<td>The calculated error rate of the reads that aligned to PhiX.</td>
</tr>
<tr>
<td>Intensity Cycle 1</td>
<td>The average of the A channel intensity measured at the first cycle averaged over filtered clusters. For the MiniSeq, NextSeq 500, and NovaSeq System, the red channel is used.</td>
</tr>
<tr>
<td>%Q \geq 30</td>
<td>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.</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiles</td>
<td>The number of tiles per lane.</td>
</tr>
<tr>
<td>Density</td>
<td>The density of clusters (in thousands per mm²) detected by image analysis, +/- 1 standard deviation.</td>
</tr>
<tr>
<td>Clusters PF</td>
<td>The percentage of clusters passing filtering, +/- 1 standard deviation.</td>
</tr>
<tr>
<td>Phas./Prephas.</td>
<td>The value used by RTA for the rate (in terms of percentage per cycle) at which molecules in a cluster fall behind (phasing) or jump ahead (prephasing) during a read. This metric is an estimate from the first 25 cycles of each read.</td>
</tr>
<tr>
<td>Phasing Equation, Prephasing Equation</td>
<td>The best-fit slope and offset of the phasing/prephasing corrections, calculated from the entire read.</td>
</tr>
<tr>
<td>Reads</td>
<td>The number of clusters (in millions).</td>
</tr>
<tr>
<td>Reads PF</td>
<td>The number of clusters (in millions) passing filtering.</td>
</tr>
<tr>
<td>%Q \geq 30</td>
<td>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.</td>
</tr>
<tr>
<td>Yield</td>
<td>The number of bases sequenced which passed filter.</td>
</tr>
<tr>
<td>Cycles Err Rated</td>
<td>The number of cycles that have been error-rated using PhiX, starting at cycle 1.</td>
</tr>
<tr>
<td>Aligned</td>
<td>The percentage that aligned to the PhiX genome.</td>
</tr>
<tr>
<td>Error Rate</td>
<td>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.</td>
</tr>
<tr>
<td>Intensity Cycle 1</td>
<td>The average of the A channel intensity measured at the first cycle averaged over filtered clusters.</td>
</tr>
</tbody>
</table>

The bottom of the page contains the Copy to Clipboard button, so you can copy the data to your computer, and a Zip My Run button, which zips InterOp files and logs to help troubleshoot debugging.
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, MiniSeq, HiSeq 3000/4000, and HiSeq X (uses RTA 2.x) and the NovaSeq (uses RTA 3.x), 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 the following information.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Reads</td>
<td>The total number of reads for this lane.</td>
</tr>
<tr>
<td>PF Reads</td>
<td>The total number of passing filter reads for this lane.</td>
</tr>
<tr>
<td>% Reads Identified (PF)</td>
<td>The total fraction of passing filter reads assigned to an index.</td>
</tr>
<tr>
<td>CV</td>
<td>The coefficient of variation for the number of counts across all indexes.</td>
</tr>
<tr>
<td>Min</td>
<td>The lowest representation for any index.</td>
</tr>
<tr>
<td>Max</td>
<td>The highest representation for any index.</td>
</tr>
</tbody>
</table>

Further information is provided regarding the frequency of individual indexes in both table and graph form. The table includes the following columns.

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

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

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. For NovaSeq, the files are broken up by cycle number. Sequencing Analysis Viewer Software shows the information of several of these files, which are described here.

Further information and open source software regarding the InterOp format can be found at https://github.com/Illumina/interop.

Extraction Metrics (ExtractionMetricsOut.bin)
This file contains extraction metrics such as FWHM scores and raw intensities.

NovaSeq
NovaSeq produces the ExtractionMetricsOut.bin file in version 3 format:

Part # 15066069 v03
For Research Use Only. Not for use in diagnostic procedures.
- byte 0: file version number (3)
- byte 1: L, the length of each record
- byte 2: ChannelCount, the channel count
- bytes \((N \times L + 3) - (N + 1) \times L + 2\): record:
  - 2 bytes: lane number (uint 16)
  - 4 bytes: tile number (uint32)
  - 2 bytes: cycle number (uint 16)
  - 4 x ChannelCount bytes: fwhm scores (float) for channels in the order listed in the RunInfo (or \([A, C, G, T]\) for legacy platforms).
  - 2 x ChannelCount bytes: intensities (uint 16) for channel in the order listed in the RunInfo (or \([A, C, G, T]\) for legacy platforms).

Where \(N\) is the record index

**HiSeq, HiSeq X, NextSeq, MiSeq, MiniSeq, HiScan SQ**

HiSeq, HiSeq X, NextSeq, MiSeq, MiniSeq, HiScan SQ produce the ExtractionMetricsOut.bin file in version 2 format:

- byte 0: file version number (2)
- byte 1: length of each record
- bytes \((N \times 38 + 2) - (N \times 38 + 39)\): record:
  - 2 bytes: lane number (uint 16)
  - 2 bytes: tile number (uint16)
  - 2 bytes: cycle number (uint 16)
  - 2 bytes: cycle number (uint 16)
  - 4 x 4 bytes: FWHM scores (float) for image channels (padded with zeroes for NextSeq systems)
  - 2 x 4 bytes: intensities (uint16) for image channels (padded with zeroes for NextSeq systems)
  - 8 bytes: date/time of cif creation

Where \(N\) is the record index

**Quality Metrics (QMetricsOut.bin)**

This file contains the quality score distribution.

**NovaSeq**

NovaSeq produces the QMetricsOut.bin file in version 7 format:

- byte 0: file version number (7)
- byte 1: length of each record
- byte 2: quality score binning (byte flag representing if binning was on)
  - if (byte 2 == 1) // quality score binning on
  - byte 3: number of quality score bins, numBins
  - for all bins B from 1 to numBins:
    - byte 4 + 3*(B-1): lower boundary of bin B
    - byte 5 + 3*(B-1): upper boundary of bin B
    - byte 6 + 3*(B-1): remapped scores of bin B
- The remaining bytes are for the records, with each record in this format:
  - 2 bytes: lane number (uint 16)
  - 4 bytes: tile number (uint32)
  - 2 bytes: cycle number (uint 16) if (byte 2 == 1)
4 x numBins bytes: number of clusters assigned to Q-score bins 1 – numBins (uint32) else
4 x 50 bytes: number of clusters assigned score Q1 through Q50 (uint32)

Where N is the record index

**MiniSeq, HiSeq X and HiSeq 3000/4000 running RTA v2.7.1 or later**

MiniSeq, HiSeq X and HiSeq 3000/4000 instruments running RTA v2.7.1 or newer produce the QMetricsOut.bin file in version 6 format:

- byte 0: file version number (6)
- byte 1: length of each record, L
- byte 2: quality score binning (byte flag representing if binning was on)
  if(byte 2 == 1) // quality score binning on
  - byte 3: number of quality score bins, B
  - bytes 4 – (4+B-1): lower boundary of quality score bins
  - bytes (4+B) – (4+2*B-1): upper boundary of quality score bins
  - bytes (4+2*B) – (4+3*B-1): remapped scores of quality score bins
- The rest of the file is composed of records of length L:
  - 2 bytes: lane number (uint16)
  - 2 bytes: tile number (uint16)
  - 2 bytes: cycle number (uint16) if(byte 2 == 1)
    - 4 x B bytes: number of clusters assigned to Q-score bins 1 – B (uint32) else
    - 4 x 50 bytes: number of clusters assigned score Q1 through Q50 (uint32)

Where N is the record index

**NextSeq, HiSeq X, and HiSeq running RTA v1.18.64 and newer, or RTA v2.1.x–v2.6.x**

NextSeq, HiSeq X, and HiSeq instruments running RTA v1.18.64 and newer, or RTA v2.1.x–v2.6.x produce the QMetricsOut.bin file in version 5 format:

- byte 0: file version number (5)
- byte 1: length of each record
- byte 2: quality score binning (byte flag representing if binning was on)
  if(byte 2 == 1) // quality score binning on
  - byte 3: number of quality score bins, B
  - bytes 4 – (4+B-1): lower boundary of quality score bins
  - bytes (4+B) – (4+2*B-1): upper boundary of quality score bins
  - bytes (4+2*B) – (4+3*B-1): remapped scores of quality score bins
- 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

MiSeq and other HiSeq platforms

MiSeq and other HiSeq platforms produce the QMetricsOut.bin file in version 4 format:

- byte 0: file version number (4)
- byte 1: length of each record

For Research Use Only. Not for use in diagnostic procedures.
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, HiSeq 3000/4000, and HiSeq X use RTA version 2.x

Error Metrics (ErrorMetricsOut.bin)
This file contains cycle error rate. Version 3 format contains counts for perfect reads and read with 1–4 errors.

NovaSeq
NovaSeq produces the ErrorMetricsOut.bin file in version 4 format:
   - byte 0: file version number (4)
   - byte 1: length of each record (12)
   - bytes (N * 12 + 2) - (N * 12 + 13): record:
     - 2 bytes: lane number (uint16)
     - 4 bytes: tile number (uint32)
     - 2 bytes: cycle number (uint16)
     - 4 bytes: error rate (float)
Where N is the record index

HiSeq, HiSeq X, NextSeq, MiSeq, MiniSeq, HiScan SQ
HiSeq, HiSeq X, NextSeq, MiSeq, MiniSeq, HiScan SQ produce the ErrorMetricsOut.bin file in version 3 format:
   - 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)
This file contains aggregate or read metrics by tile.
**NovaSeq**

NovaSeq produces the TileMetricsOut.bin file in version 3 format:

- byte 0: file version number (3)
- byte 1: length of each record (15)
- bytes 2-5: area of a tile in mm^2 (float)
- bytes (N * 15 + 6) - (N * 15 + 20): record:
  - 2 bytes: lane number (uint16)
  - 4 bytes: tile number (uint32)
  - 1 byte: metricCode, the metric code (char)
  - if(metricCode == 't')
    - 4 bytes: cluster count (float)
  - else if(metricCode == 'r')
    - 4 bytes: read number (uint32)
    - 4 bytes: % aligned (float)
  - else if(metricCode == '0')
    - 8 bytes: 0

where N is the record index

---

**HiSeq, HiSeq X, NextSeq, MiSeq, MiniSeq, HiScan SQ**

HiSeq, HiSeq X, NextSeq, MiSeq, MiniSeq, HiScan SQ produce the TileMetricsOut.bin file in version 2 format:

- 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)**

This file contains base call metrics.

---

**NovaSeq**

NovaSeq produces the CorrectedIntMetricsOut.bin file in version 4 format:

- byte 0: file version number (4)
byte 1: length of each record (28)
byte (N * 28 + 2) - (N * 28 + 29): record:
  2 bytes: lane number (uint16)
  4 bytes: tile number (uint32)
  2 bytes: cycle number (uint16)
  4 bytes: number of No Call base calls (uint32)
  4 bytes: number of A base calls (uint32)
  4 bytes: number of C base calls (uint32)
  4 bytes: number of G base calls (uint32)
  4 bytes: number of T base calls (uint32)

MiniSeq, HiSeq X and HiSeq 3000/4000 running RTA v2.7.1 or later

MiniSeq, HiSeq X and HiSeq 3000/4000 instruments running RTA v2.7.1 or newer produce the CorrectedIntMetricsOut.bin file in version 3 format:

byte 0: file version number (3)
byte 1: length of each record
byte (N * 34 + 2) - (N * 34 + 35): record:
  2 bytes: lane number (uint16)
  2 bytes: tile number (uint16)
  2 bytes: cycle number (uint16)
  2 bytes: average corrected int for called clusters for channel A (uint16)
  2 bytes: average corrected int for called clusters for channel C (uint16)
  2 bytes: average corrected int for called clusters for channel G (uint16)
  2 bytes: average corrected int for called clusters for channel T (uint16)
  4 bytes: number of N (no call) calls (uint32)
  4 bytes: number of A base calls (uint32)
  4 bytes: number of C base calls (uint32)
  4 bytes: number of G base calls (uint32)
  4 bytes: number of T base calls (uint32)

Other HiSeq, HiSeq X, MiSeq, and NextSeq platforms

All other platforms produce the CorrectedIntMetricsOut.bin file in version 2 format:

byte 0: file version number (2)
byte 1: length of each record
byte (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)

**Image Metrics (ImageMetricsOut.bin)**

This file contains minimum and maximum contrast values for image. NovaSeq does not produce the ImageMetricsOut.bin file.

**MiniSeq, HiSeq X and HiSeq 3000/4000 running RTA v2.7.1 or later**

MiniSeq, HiSeq X and HiSeq 3000/4000 instruments running RTA v2.7.1 or newer produce the ImageMetricsOut.bin file in version 2 format:
- byte 0: file version number (2)
- byte 1: length of each record (recordLength, depends on number of channels)
- byte 2: number of channels
- bytes (N * recordLength + 3) - ((N+1) * recordLength + 2): record:
  - 2 bytes: lane number (uint16)
  - 2 bytes: tile number (uint16)
  - 2 bytes: cycle number (uint16)
  For each channel in the channel set
  - 2 bytes: min contrast value for image (uint16)
  For each channel in the channel set
  - 2 bytes: max contrast value for image (uint16)

**MiSeq and HiSeq running RTA 1.x**

MiSeq and HiSeq instruments running RTA 1.x produce the ImageMetricsOut.bin file in version 1 format:
- 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 (IndexMetricsOut.bin):**

This file reports the indexes count and is not generated in RTA 2.x or RTA 3.x.
- Byte 0: file version (1)
- Bytes (variable length): record:
  - 2 bytes: lane number(uint16)
  - 2 bytes: tile number(uint16)
  - 2 bytes: read number(uint16)
2 bytes: number of bytes Y for index name (uint16)
Y bytes: index name string (string in UTF8Encoding)
4 bytes: # clusters identified as index (uint32)
2 bytes: number of bytes V for sample name (uint16)
V bytes: sample name string (string in UTF8Encoding)
2 bytes: number of bytes W for sample project (uint16)
W bytes: sample project string (string in UTF8Encoding)

Revision History

<table>
<thead>
<tr>
<th>Part #</th>
<th>Date</th>
<th>Description of Change</th>
</tr>
</thead>
</table>
| Document # 15066069 v03 | February 2017 | • Revised document for release of Sequencing Analysis Viewer Software v2.1, which supports the NovaSeq System.   
• Removed the Status Pane section.                                      |
| Document # 15066069 v02 | February 2016 | • Revised document for release of Sequencing Analysis Viewer Software v1.11, which supports the MiniSeq System.   
• Moved Revision History section to back of the guide.                          |
| Document # 15066069 v01 | October 2015 | • Revised document for release of Sequencing Analysis Viewer Software v1.10, which supports HiSeq 3000/4000.  
• Removed Data Availability section.  
• Added Metrics Definition section.                                           |
| Part # 15066069 Rev. A | December 2014 | Initial release for use with Sequencing Analysis Viewer Software v1.8.46, which supports HCS v2.2.58 and HiSeq X Control Software v3.1.  
Derived from Sequencing Analysis Viewer User Guide 15020619 Rev. F.               |
Technical Assistance

For technical assistance, contact Illumina Technical Support.

Website: www.illumina.com
Email: techsupport@illumina.com

Illumina Customer Support Telephone Numbers

<table>
<thead>
<tr>
<th>Country</th>
<th>North America</th>
<th>Germany</th>
<th>Singapore</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.800.809.4566</td>
<td>0800.180.8994</td>
<td>1.800.579.2745</td>
</tr>
<tr>
<td>Australia</td>
<td>1.800.775.688</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>0800.296575</td>
<td>1.800.812949</td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>0800.81102</td>
<td>800.874909</td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>400.635.9898</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>80882346</td>
<td>0800.0223859</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>0800.918363</td>
<td>0800.451.650</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>0800.911850</td>
<td>800.180.480</td>
<td></td>
</tr>
<tr>
<td>Hong Kong</td>
<td>800960230</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>1.800.812949</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>800.874909</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>0800.111.5011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>0800.0223859</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td>0800.451.650</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>800.16836</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>900.812168</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>020790181</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>0800.563118</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taiwan</td>
<td>00806651752</td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0800.917.0041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other countries</td>
<td>+44.1799.534000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

Product documentation—Available for download in PDF from the Illumina website. Go to support.illumina.com, select a product, then select Documentation & Literature.