Sequencing Analysis Viewer v2.4

Software Guide

ILLUMINA PROPRIETARY

For Research Use Only. Not for use in diagnostic procedures.

Document # 15066069 v03

For Research Use Only. Not for use in diagnostic procedures.

2

Introduction

Sequencing Analysis Viewer (SAV) v2.4 is a software application that allows real-time viewing of quality metrics generated by the Real-Time Analysis (RTA) software on Illumina sequencing systems. SAV is compatible with all HiSeq™ systems, both NextSeq™ systems, MiSeq™, MiniSeq™, NovaSeq™, and HiScanSQ™.

> **NOTE**
> SAV v2.4 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 software guide provides installation instructions and describes the various metrics you can view with this tool. If SAV is installed on your sequencing instrument, viewing metrics does not interfere with the sequencing run.

Set Up Sequencing Analysis Viewer

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

> **NOTE**
> Contact your Illumina field representative for help installing the software on a sequencing system.

Requirements

Sequencing Analysis Viewer does not require an advanced personal computer because the instrument control computer running the Real-Time Analysis 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, Windows 10, or later

▶ Network access to the run data

▶ .Net framework 4.5.1

▶ Visual C++ Redistributable for Visual Studio 2015

Download the Installer

1   In your browser, open the Service Software support page.

2   Select Downloads.

3   Download the Service Software v2.4 Installer (*.zip) file.

Install Sequencing Analysis Viewer

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 Setup Wizard opens.

5   Select Next.

The installation starts, and the progress screen opens. When the installation is finished, a desktop shortcut is created and the setup complete screen appears.

6   Select **Finish**.

Test the Installation

When you have installed the SAV software, use the following instructions to make sure that the application was installed properly and that you have the proper permissions.

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

2   In the Run Folder field, copy the folder location or select **Browse** to navigate to a run folder. Select the run folder only, and not the parent folder or any folder or file inside the run folder.
    The Run Folder varies depending on your Illumina sequencing system and the operating system on your personal computer.

Figure 1   Example HiSeq Run Folder Structure
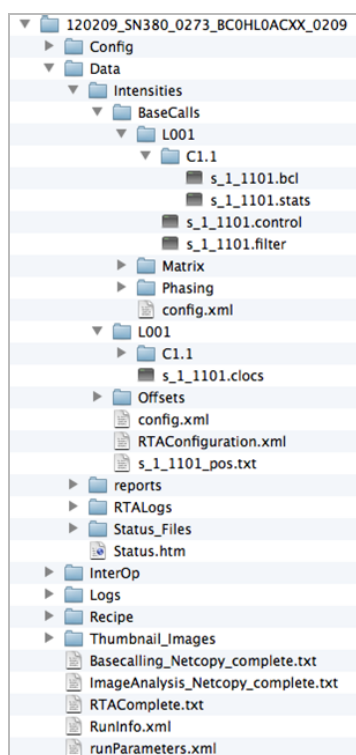
3   Select **Refresh**.
    The software starts loading data. When complet, it displays available quality metrics for the run.
    If metrics are not displayed or an error message appears, make sure that you selected a *valid* run folder and have permissios to view the data.

Sequencing Analysis Viewer Files

SAV uses the following files as input:

▶   RunInfo.xml

▶   RunParameters.xml

▶ The InterOp folder is in the directory: <run directory>\InterOp. The InterOp files provide the following information:
  ▶ Extraction metrics
  ▶ Quality metrics
  ▶ Error metrics
  ▶ Tile metrics
  ▶ Extended 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 details the metrics displayed in the Analysis and Imaging panels.

Metric	Definition
Intensity (or P90)	The 90% percentile extracted intensity for a given image (lane/tile/cycle/channel combination). On platforms using two-channel sequencing, only the two channels are shown. On platforms using four-channel sequencing, four channels (A, C, G, and T) are shown.
FWHM	The average full width of clusters at half maximum (representing their approximate size in pixels).
Corrected Intensity	• HiSeq X, HiSeq 3000 and 4000, MiniSeq, NovaSeq 6000—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 500 and 550—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 help assess run progress. The called and corrected intensities graphs are identical.
Called Intensity	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 6000.
% No Calls	The percentage of clusters on a tile for which no base (N) has been called.
% Base	The percentage of called (non-N) 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 noncalled intensities. Not calculated for NextSeq and MiniSeq two-channel sequencing, HiSeq 3000, HiSeq 4000, HiSeq X, and NovaSeq 6000.
Error Rate	The calculated error rate, as determined by a PhiX spike-in. If a PhiX control is not run in the lane or if the tile has not been selected for alignment, this number is not available.
% Q ≥ 20, %Q ≥ 30	The percentage of bases with a quality score of 20 or 30 or higher, respectively.
Median Q-score	The median Q-score for each tile over all bases for the current cycle. These charts are generated after cycle 25. This metric is best used to examine the Q-scores of a run as it progresses. Because it relies on one threshold, the %Q30 plot can give an oversimplified view.
Density	The density of clusters for each tile (in thousands per mm^2).

Metric	Definition
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).
% Pass Filter	The percentage of clusters passing filter.
Legacy Phasing Rate, Legacy Prephasing Rate	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. Previous versions of SAV called this metric % Phasing and % Prephasing. For an estimate calculated from the entire read, see the Phasing and Prephasing Slope and Offset columns.
Phasing Weight, Prephasing Weight	The rate for a given cycle at which molecules in a cluster fall behind (phasing) or jump ahead (prephasing).
Phasing Offset, Prephasing Offset	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.
Phasing Slope, Prephasing Slope	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.
% Aligned	The percentage of the passing filter clusters that aligned to the PhiX genome.
Minimum/Maximum Contrast	The 10th and 99.5^{th} percentiles per channel of selected columns of the raw image, respectively
Occupied Count (k)	The total number of patterned wells (in thousands) that are occupied by DNA sequence
% Occupied	The percentage of patterned wells that are occupied by DNA sequence.

Load Data

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

2 In SAV, select the tab containing the appropriate query information.

3 [Optional] Load one of the 10 most recently loaded runs using the Recent Runs drop-down list.

4 In the Run Folder field, copy the folder location or select **Browse** to navigate to a run folder. Select the run folder only, and not the parent folder or any folder or file inside the run folder.

5 Select **Refresh**.
The software starts loading data. When complete, it displays available quality metrics for the run.

Analysis Tab

The Analysis tab includes six panes, which are described in this section.

NOTE

If you are using the SAV to view metrics from the MiSeq system (which has only one lane), select **1** or **All** from the Lane drop-down list to display lane information.

Flow Cell Chart

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

For Research Use Only. Not for use in diagnostic procedures.

▶ You can select the displayed metric, surface (if your system scans multiple surfaces), cycle, and base using the drop-down lists.

▶ The color bar to the right of the chart indicates the values that each color represents.

    ▶ The color bar is also a slider: right-click to adjust the scale values or use the mouse wheel to change the contrast.

    ▶ 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 appear gray.

▶ Interactive tooltips provide the lane, tile, and value of each data point.

▶ Clicking a tile opens the Imaging tab with more detailed information for that tile. See *Imaging Tab* on page 8.

▶ By right-clicking an image, you can copy it to the clipboard.

▶ The chevron displayed in the top-right corner switches the plot between pane view and full-screen view.

Data by Cycle Plot

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

▶ You can select the displayed metric, lane, surface, and base using the drop-down lists.

▶ The plots are displayed with auto-scaling by default, or they can be fixed by selecting 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 displayed in the top-right corner switches the plot between pane view and full-screen view.

▶ Pan the graph by dragging. Zoom in using the mouse wheel, or zoom in on a specific axis by hovering over the axis and spinning the mouse wheel.

▶ By right-clicking an image, you can copy it to the clipboard.

Data by Lane Plot

The Data by Lane pane shows plots for viewing quality metrics per lane. These plots have the following features:

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

▶ The chevron displayed in the top-right corner switches the plot between pane view and full-screen view.

▶ Zoom in and out, and pan the plots.

▶ By right-clicking an image, you can copy it to the clipboard.

▶ For some metrics (Density and Clusters), each lane has two box plots.

    ▶ **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.

For Research Use Only. Not for use in diagnostic procedures.

▶ 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 values that exceed 1.5 times the interquartile range below the 25th percentile, or exceed 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.

Q-Score Distribution Plot

The Q-score Distribution pane shows plots that for viewing 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. On most systems, quality scores are binned in groups.

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%

These plots have the following features:

▶ You can select the displayed lane, surface, read, and cycle using the drop-down lists.

▶ The Read drop-down list selects the first cycle for calculating the histogram.

▶ The Cycle drop-down list determines the last cycle used for calculating the histogram.

▶ The cutoff slider determines how many bases have a minimum Q-score or higher. Move the slider to the minimum Q-score. The software calculates how many bases have that Q-score or higher.

▶ The chevron displayed in the top-right corner switches the plot between pane view and full-screen view.

▶ You can pan the graph by dragging. Zoom in using the mouse wheel, or zoom in on a specific axis by hovering over the axis and spinning the mouse wheel.

▶ You can copy an image to the clipboard by right-clicking it.

Q-Score Heat Map

The Q-score heat map shows plots for viewing the Q-score by cycle. These plots have the following features:

▶ You can select the displayed lane and surface using the drop-down lists.

▶ The color bars to the right of each chart indicate the values that each color represents. The charts are displayed with auto scaling. The scale is always 0% to 100% of maximum value. Right-click to change the color scheme.

▶ The chevron displayed in the top-right corner switches 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 a run.

- ▶ You can select the displayed cycle, lane, surface, swath, read, tile number, and section using the drop-down lists.
- ▶ Use the checkboxes to select the displayed bases.
- ▶ The imaging tab includes the available thumbnails for the selected tile, highlighted in blue.
- ▶ You can also use a mouse wheel or the keyboard arrow keys to scroll through images.
- ▶ Right-clicking the top of a column provides more information.

The buttons above the data table provide the following features:

Button	Description
	Select all rows.
	Copy selected rows to clipboard.
	Export Imaging table to a *.txt file.
	Sort columns in ascending, descending, or custom order.
	Create a custom scatter or box plot.
	Choose which columns to display.
	Filter columns, or clear a filter.
	Move the thumbnail selection up or down.

Summary Tab

The Summary tab provides tables with basic data quality metrics summarized per lane and per read.

All statistics are provided as means and standard deviations over the tiles used in the lane. Cycle-based metrics, such as yield, error rate, and % Q ≥ 30, are calculated over fully corrected, "usable" cycles. The last cycle of every read is removed from all computations.

The Cycle Status above the top table displays a summary of the last cycle extracted, base called, Q-scored, and error rated for each tile.

The top table displays the following metrics, divided by read and total.

Metric	Description
Level	The sequencing read level.
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, which is updated as the run progresses.
Aligned	The percentage of the passing filter clusters that aligned to the PhiX genome, which is determined for each level or read independently.
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. For the MiniSeq, NextSeq 500, and NovaSeq System, the red channel is used.
%Q ≥ 30	The percentage of bases with a quality score of 30 or higher, respectively. This chart is generated after cycle 25, and the values represent the current cycle.

The following metrics are available in the read tables, divided by lane.

Metric	Description
Tiles	The number of tiles per lane.
Density	The density of clusters (in thousands per mm^2) detected by image analysis, +/- 1 standard deviation.
Clusters PF	The percentage of clusters passing filtering, +/- 1 standard deviation.
Legacy Phasing/Prephasing Rate	The value used by Real-Time Analysis for the rate (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.
Phasing Slope/Offset, Prephasing Slope/Offset	The best-fit slope and offset of the phasing/prephasing corrections, calculated from the entire read.
Reads	The number of clusters (in millions).
Reads PF	The number of clusters (in millions) passing filter.
%Q ≥ 30	The percentage of bases with a quality score of 30 or higher. This chart is generated after cycle 25, and the values represent the current cycle.
Yield	The number of bases sequenced thatpassed 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. The base unit is the error rate for a single tile. SAV averages the per-cycle error rate across all the relevant cycles for a tile before calculating these statistics.
Intensity Cycle 1	The average of the A channel intensity measured at the first cycle averaged over filtered clusters. For the MiniSeq, NextSeq 500, and NovaSeq 6000, the red channel is used.

The bottom of the page displays the following buttons:

Button	Description
Copy to Clipboard	Helps you copy the data to your computer.
Zip My Run	Zips InterOp files and logs to help troubleshoot debugging.

Indexing Tab

The Indexing tab lists count information for indexes used in the run. This tab is only available when a sample sheet is uploaded to the control software during run setup. You also must conduct an index run.

> **NOTE**
>
> On all systems except MiSeq, this tab is populated when SAV connects to the InterOp folder after running bcl2fastq2 software. Otherwise, this tab is not populated. If the run is connected to BaseSpace Sequence Hub, the indexing information is available there.

Use the Lane drop-down list to select a lane and view the indexing information. The option for all lanes aggregates indexing information from any samples with the same index sequence and sample ID.

The top table summarizes the indexing performance for your selection:

Metric	Description
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.

The bottom table provides information on the frequency of individual indexes:

Column	Description
Index Number	A unique number assigned to each index by SAV 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 by the unique index number SAV assigned.

InterOp Files

The InterOp files are in the directory: <run directory>\InterOp. These file formats are the same for all Illumina sequencing systems except where noted. For NovaSeq 6000, cycle numbers divide different files. Sequencing Analysis Viewer shows the information of several of these files, which are described in this section.

For more information and open source software regarding the InterOp format, see github.com/Illumina/interop.

Extraction Metrics (ExtractionMetricsOut.bin)

The extraction metrics file contains extraction metrics such as FWHM scores and raw intensities.

NovaSeq 6000

NovaSeq 6000 produces the ExtractionMetricsOut.bin file in version 3 format:

▶ byte 0: file version number (3)

▶ byte 1: L, the length of each record

▶ byte 2: ChannelCount, the channel count

▶ bytes (N * L+ 3) - ((N + 1) * L + 2): record:
  ▶ 2 bytes: lane number (uint16)
  ▶ 4 bytes: tile number (uint32)
  ▶ 2 bytes: cycle number (uint16)
  ▶ 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 (uint16) for channel in the order listed in the RunInfo (or [A, C, G, T] for legacy platforms).

Where N is the record index.

HiSeq X, HiSeq, NextSeq, MiSeq, MiniSeq, HiScanSQ

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

▶ 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 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)

The quality metrics file contains the quality score distribution.

NovaSeq 6000

NovaSeq 6000 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 (uint16)
   ▶ 4 bytes: tile number (uint32)
   ▶ 2 bytes: cycle number (uint16)
    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)

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

MiniSeq, HiSeq X, and HiSeq 3000/4000 instruments running RTA v.2.7.1 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)

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

▶ 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)

The error metrics file contains cycle error rate. Version 3 format contains counts for perfect reads and read with 1–4 errors.

NovaSeq 6000

NovaSeq 6000 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, HiScanSQ

HiSeq, HiSeq X, NextSeq, MiSeq, MiniSeq, HiScanSQ 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)

The tile metrics file contains aggregate or read metrics by tile.

NovaSeq 6000

NovaSeq 6000 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)
    - ▶ 4 bytes: PF 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, HiScanSQ

HiSeq, HiSeq X, NextSeq, MiSeq, MiniSeq, HiScanSQ 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

Extended Tile Metrics (ExtendedTileMetricsOut.bin)

The extended tile metrics file contains aggregate or read metrics by tile.

NovaSeq 6000

NovaSeq 6000 produces the ExtendedTileMetricsOut.bin file in both the version 2 and 3 formats.

Format for version 3:

- ▶ byte 0: file version number (3)
- ▶ byte 1: length of each record (18)
- ▶ bytes (N * 18 + 2) - (N * 18 + 19): record:
  - ▶ 2 bytes: lane number (uint16)

▶ 4 bytes: tile number (uint32)

▶ 4 bytes: occupied cluster count (float)

▶ 4 bytes: upper left fiducial location in X (float)

▶ 4 bytes: upper left fiducial location in Y (float)

Where N is the record index.

Format for version 2:

▶ byte 0: file version number (2)

▶ byte 1: length of each record (10)

▶ bytes (N * 10 + 2) - (N * 10 + 11): record:

   ▶ 2 bytes: lane number (uint16)

   ▶ 4 bytes: tile number (uint32)

   ▶ 4 bytes: occupied cluster count (float)

Where N is the record index.

Corrected Intensity Metrics (CorrectedIntMetricsOut.bin)

The corrected intensity metrics file contains base call metrics.

NovaSeq 6000

NovaSeq 6000 produces the CorrectedIntMetricsOut.bin file in version 4 format:

▶ byte 0: file version number (4)

▶ byte 1: length of each record (28)

▶ bytes (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)

   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 v.2.7.1 or newer produce the CorrectedIntMetricsOut.bin file in version 3 format:

▶ byte 0: file version number (3)

▶ byte 1: length of each record

▶ bytes (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)

Where N is the record index.

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

▶ 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)

Where N is the record index.

Image Metrics (ImageMetricsOut.bin)

The image metrics file contains minimum and maximum contrast values for images.

NovaSeq 6000

NovaSeq 6000 produces the ImageMetricsOut.bin file in version 3 format:

▶ byte 0: file version number (3)

▶ 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)

  ▶ 4 bytes: tile number (uint32)

  ▶ 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)

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 v.2.7.1 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)

  Where N is the record index.

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)

  Where N is the record index.

Index Metrics (IndexMetricsOut.bin):

The index metrics 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(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)

Revision History

Part #	Date	Description of Change
Document # 15066069 v03	November 2017	• Revised document for release of Sequencing Analysis Viewer v2.4, which supports the NovaSeq 6000 Sequencing System and NovaSeq Control Software v1.2.0.42. • Removed the Status Pane section. • Minor text edits.
Document # 15066069 v02	February 2016	• Revised document for release of Sequencing Analysis Viewer 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 v1.10, which supports HiSeq 3000 and HiSeq 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

Region	Toll Free	Regional
North America	+1.800.809.4566	
Australia	+1.800.775.688	
Austria	+43 800006249	+43 19286540
Belgium	+32 80077160	+32 34002973
China	400.066.5835	
Denmark	+45 80820183	+45 89871156
Finland	+358 800918363	+358 974790110
France	+33 805102193	+33 170770446
Germany	+49 8001014940	+49 8938035677
Hong Kong	800960230	
Ireland	+353 1800936608	+353 016950506
Italy	+39 800985513	+39 236003759
Japan	0800.111.5011	
Netherlands	+31 8000222493	+31 207132960
New Zealand	0800.451.650	
Norway	+47 800 16836	+47 21939693
Singapore	+1.800.579.2745	
Spain	+34 911899417	+34 800300143
Sweden	+46 850619671	+46 200883979
Switzerland	+41 565800000	+41 800200442
Taiwan	00806651752	
United Kingdom	+44 8000126019	+44 2073057197
Other countries	+44.1799.534000	

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**.

Illumina
5200 Illumina Way
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

For Research Use Only. Not for use in diagnostic procedures.

illumına®