Sequencing Analysis Software Version 5.4

Applied Biosystems Sequencing Analysis Software v5.4 analyzes, displays, edits, saves, and prints sample files that are generated from Applied Biosystems DNA analyzers and genetic analyzers. The Sequencing Analysis software:

- Uses a basecaller algorithm that performs basecalling for pure and mixed base calls
- Generates quality values to provide basecall accuracy information for pure and mixed base calls
- Generates analysis reports to help troubleshoot and provide easy assessment of data quality
- Can generate an audit trail of base changes
- Allows you to use electronic signatures.

Definitions in Sequencing Analysis Software v5.4

- Analysis Protocol – All the settings necessary for analysis and used to perform basecalling and post processing.
- Analysis Report – The report that shows the status of the data analysis. Use this to help troubleshoot and to easily assess data quality.
- Clear Range – The region of a DNA sequence that remains after excluding the low-quality or error-prone sequence at both the 5′ and 3′ ends; determined during post-processing.
- Length of Read (LOR) – The measure of the length of quality bases. The LOR is defined by the user in the Display Settings dialog box and is displayed in the analysis report.
- Quality Value (QV) – The per-base estimate of the basecalling accuracy.
- Sample Score – The average quality value of the bases in the clear range sequence for the sample.
- ABI basecaller – The previous-generation algorithm in Sequencing Analysis software that identifies only pure bases. This older algorithm is no longer under development and supports only older run modules.
- KB™ Basecaller – The latest-generation algorithm that provides accurate basecallling for pure and mixed bases, and a quality value (QV) for each base. It has lower error rates, longer length of read, and higher accuracy with PCR products than the ABI basecaller. Use of this algorithm is strongly recommended.

Input Sample Files

Sequencing Analysis Software is compatible with sample files that are generated from:

- Applied Biosystems 3730/3730*xl* DNA Analyzers and 3130/3130*xl* Genetic Analyzers
- ABI PRISM® 310 and 3100/3100-*Avant* Genetic Analyzers
- Applied Biosystems 3500/3500xL and 3500 Dx/3500xL Dx Genetic Analyzers.

Output Files for Sequencing Analysis

You can generate sample files as:

- Analyzed ABI files (.ab1)
- Text file of the sequence in ABI or FASTA format (.seq; all bases or only clear range bases)
- Phred (.phd.1) files
- Standard chromatogram format (.scf) files
- Analysis reports (.txt, HTML, PDF, or XML format)

Program Overview

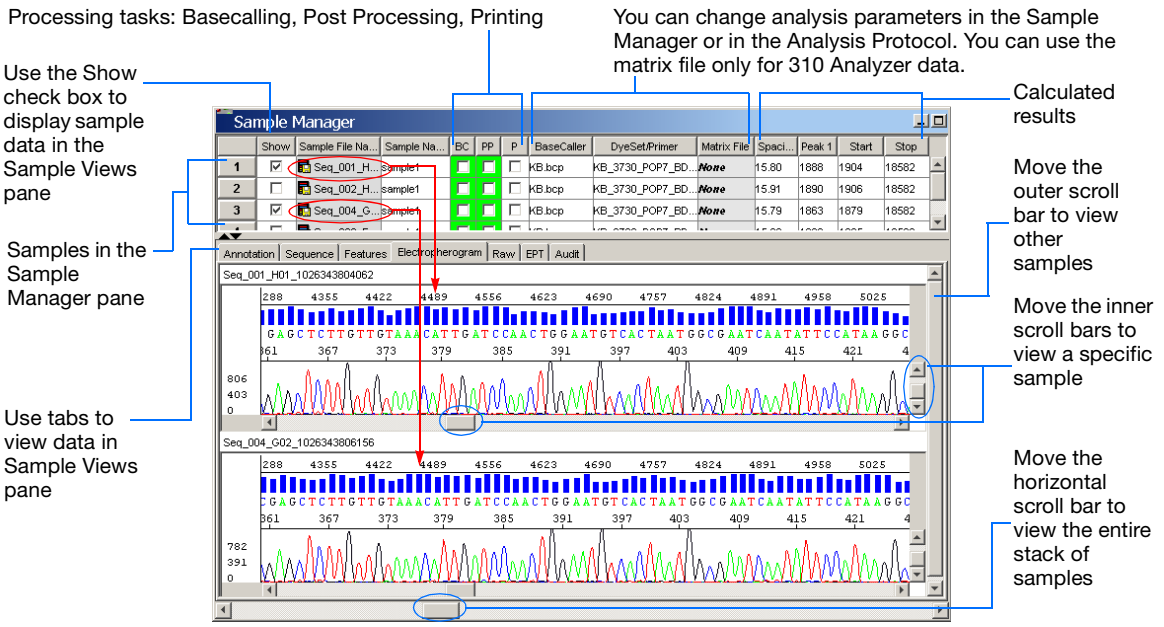
To analyze and review your samples, you:

- **Add sample(s) to the Sample Manager.**
- **Edit and apply an analysis protocol (optional).**
- **Analyze the samples.**
- Generate, review, and export an analysis report and review results.
- Display the sample data.
- Review low-quality samples and edit bases if necessary.

For more information, refer to the *Sequencing Analysis Software v5.4 User Bulletin* (PN 4378383) and the *Applied Biosystems DNA Sequencing Analysis Software v5.1 User Guide* (PN 4346366).

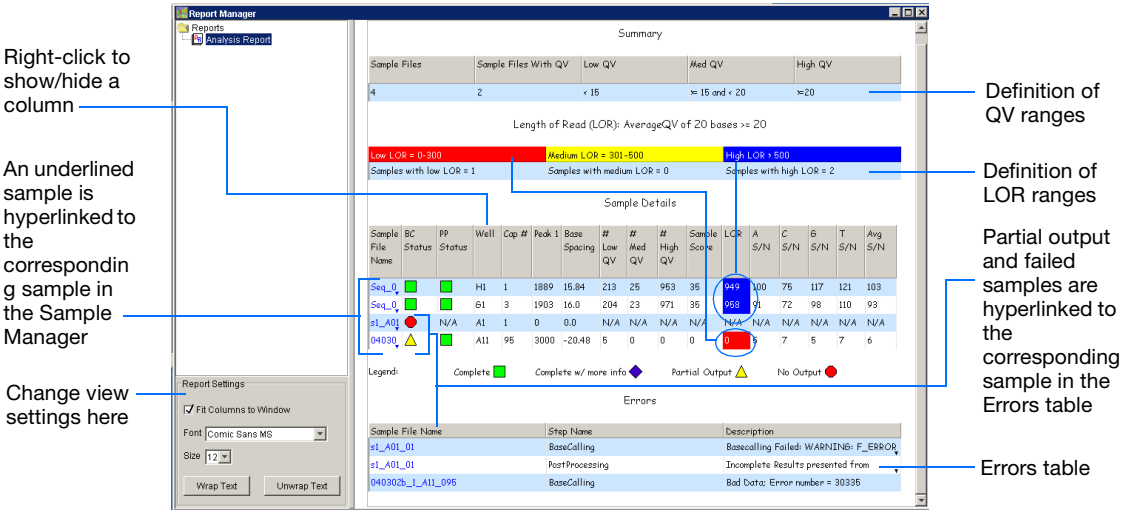
The Sample Manager

The Sample Manager displays sample files and all the analysis parameter values. All data are analyzed or reanalyzed using this window and displayed in single or multiple sample views.



The Analysis Report

An analysis report shows the status of data analysis. You can generate an analysis report for any samples added to the Sample Manager. If the data are analyzed, the report displays a summary of QVs and LORs, as well as individual sample information and errors. If the data are not analyzed, the report displays status information. You export a report as a tab-delimited file, then open it in Microsoft Excel® software to look for trends.



Data Analysis and Review

All analysis in Sequencing Analysis software occurs in the Sample Manager. You can perform an analysis and review of your sequencing sample files by following the steps below.

To launch the analysis software, double-click the Sequencing Analysis desktop icon.

Step 1: Add sample(s) to the Sample Manager.

1A. Click  (Add Sample) to open the Add Samples dialog box.

1B. Select the files that you want to add to the **Samples To Add** section of the dialog box. To add:

- A single file—Select the file, then click **Add Selected Samples**.
- Multiple files—Shift-click to select continuous samples or Ctrl-click to select discontinuous samples, then click **Add Selected Samples**.
- All samples in a single folder—Select the folder, then click **Add Selected Samples**.

1C. Click **OK**.

The dialog box closes, and the selected files are added to the Sample Manager.

Step 2: Edit and apply an analysis protocol (optional).

To change the analysis protocol settings for samples:

2A. Select the sample rows of interest in the Sample Manager.

- Use Shift-click to select continuous samples.
- Use Ctrl-click to select discontinuous samples.

2B. Select **Analysis** ▸ **Analysis Protocol Manager**.

2C. In the Analysis Protocol column, select the protocol that you want to edit, then double-click the protocol name.

2D. Make changes in the General, Basecalling, Mixed Bases, and Clear Range tabs, as appropriate.

2E. Click **OK** to save the protocol and close the Sequence Analysis Protocol Editor.

2F. Click:

- **Apply to Selected Samples** to apply the protocol to the samples that you selected in step 2A. *or*
- **Apply to All Samples** to apply the protocol to all the samples in the Sample Manager.

2G. Click **Done** to close the Analysis Protocol Manager.

Step 3: Analyze the samples.

Click  (Start Analysis). The Sequencing Analysis software automatically performs basecalling, post processing, and printing according to the analysis protocol:

Basecalling

When you select **BC** (basecalling) in the Sample Manager, the basecaller performs calls the bases using the:

- KB™ basecaller
  - If you select the mixed base option in the analysis protocol, then mixed bases are called.
  - If you do not select the mixed base option, then pure bases are called (A, C, G, and T only).
  - If you select the mixed base option, then quality values (QVs) are calculated for pure and mixed bases.

  or
- ABI basecaller

  Assigns A, C, G, T, or N to every base (no mixed-base calling and QV options)

The Sequencing Analysis software then generates optional file formats (.seq, .phd.1, and/or .scf).

Post Processing

When you select PP (post processing) in the Sample Manager, the software calculates the clear range.

Printing

When you select P (printing) in the Sample Manager, the software prints the sample views after analysis and post processing.

Note: The views that are printed are specified in the Options dialog box. To change the defaults, select **Tools** ▸ **Options**, then select the **Printing** tab.

Step 4: Generate, review, and export an analysis report and review results.

4A. In the Sample Manager:

- Look for green, blue, yellow, or red boxes for the BC parameter. Green indicates that basecalling was successful, blue indicates success (but with some anomalies), yellow indicates poor quality data, and red indicates failure.

  Note: The blue and yellow results apply only to samples that are analyzed using the KB™ basecaller.

- Look for green or red boxes for the PP and/or P parameters. Green indicates that the process was successful; red indicates failure.
- Review the base spacing, peak 1 location, and start and stop points. A red value in the Base Spacing column indicates that the spacing could not be calculated, and that the default value was used for analysis.

4B. Click  (Analysis Report) to generate and display the report.

a. Review the data in the report.

b. To export the report, select **File ▸ Export Report**. You can export the file in tab-delimited text, HTML, PDF, or XML.

Step 5: Display the sample data.

5A. Use the Sample Manager (Figure 1) or Sample Navigator view (Figure 2) to display data for any of the following:

- A single sample—Double-click the sample file name or select the corresponding **Show** check box.
- Multiple samples—Shift-click, Shift-drag or Ctrl-click the sample row numbers to select the sample files, then click  .
- All samples—Click  (above the row numbers) or Shift-drag the sample row numbers to select all samples, then click  .

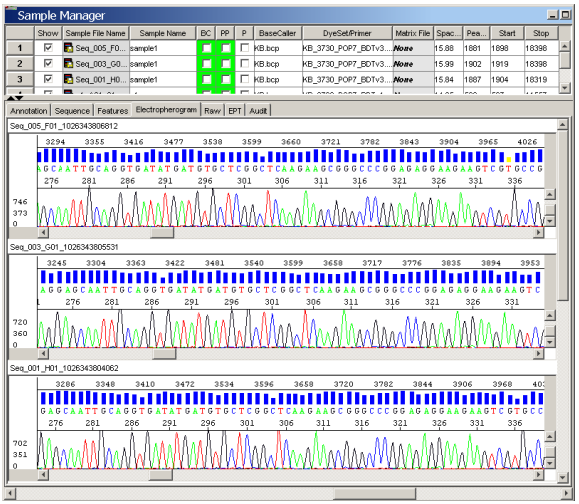


Figure 1 Sample data in the Sample Manager view

5B. To show the data in the Sample Navigator view, select **View ▸ Sample Navigator** or click  .

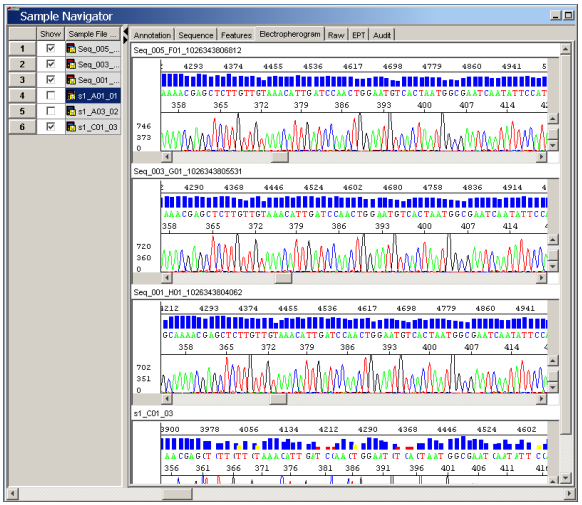


Figure 2 Sample data in the Sample Navigator view

Step 6: Review low-quality samples and edit bases if necessary.

6A. Select a sample file with a low sample score.

6B. Review your results in the sample file:

a. Review the raw, analyzed, and EPT data.

b. Review low-quality basecalls, and check for errors.

6C. Edit the bases, as needed.

- Navigate through the data, using the keyboard shortcuts in Table 2.
- Make changes to bases as appropriate. If you:
  - Insert a base – No QV is added
  - Delete a base – QV is deleted
  - Change a base – QV has the same value but is displayed as a gray bar

6D. Save the sample file. The .seq file that is created when the data are analyzed is updated when you save the sample file.

Quality Values

The QV is a per-base estimate of the basecaller accuracy. The QVs are calibrated on a scale corresponding to:

$$QV = -10 \log_{10}(Pe)$$

where *Pe* is the probability of error.

The KB™ basecaller generates QVs from 1 to 99 (Table 1).

Table 1 % Probability that a basecall is incorrect

Quality Value	% Probability the basecall is incorrect
10	10
20	1
30	0.1
40	0.01
50	0.001

- Typical high-quality pure bases have QVs of 20 to 50
- Typical high-quality mixed bases have QVs of 10 to 50
- Size and color of QVs bars for QVs from 50 to 99 are identical

You can display QVs as bars above each base in the Sequence and Electropherogram views. The QVs are also displayed in the analysis report.

To change the colors and values of the QV bars:

1. Select (Display Settings) ▸ **Bases** tab.
2. In the Sample File Display section (Figure 3), change the range for the QVs. Use the two sliders to specify the low, medium, and high ranges.

QV Bar	Default Color and Range	Set the range to identify data that…
Low	Red: 0 to 14	Are not acceptable
Medium	Yellow: 15 to 19	Need manual review
High	Blue: 20	Are acceptable

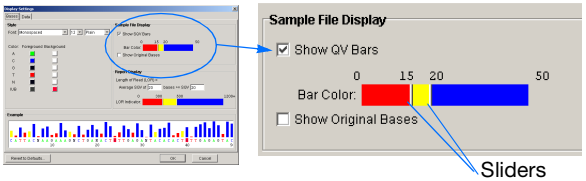


Figure 3 Display settings: Bases tab

3. Click **OK** to save the new settings and close the dialog box.

Shortcuts for Working with Data

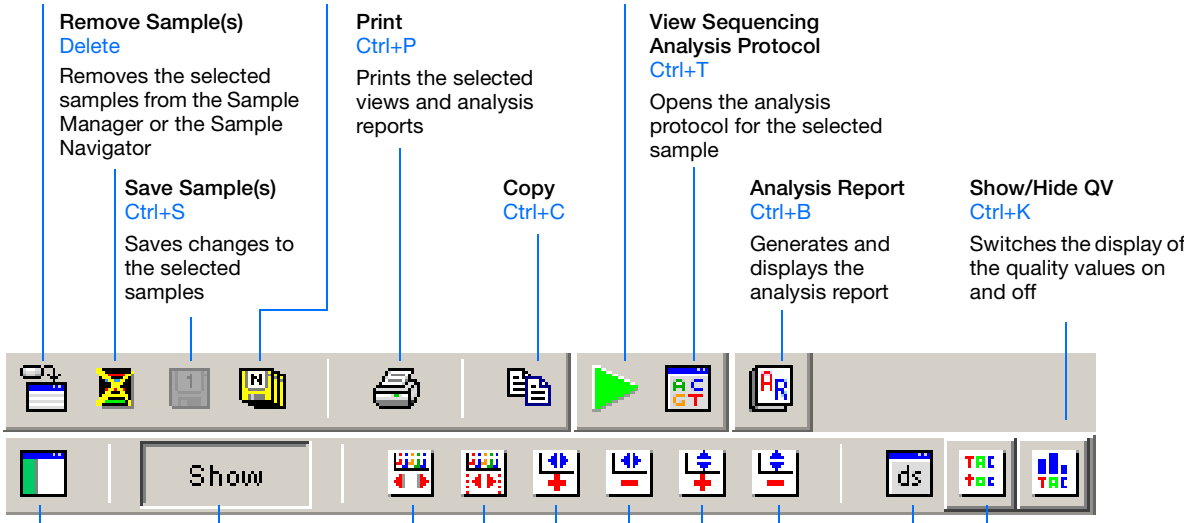
Use the following shortcuts when you work with data in the Sequence and Electropherogram views.

Table 2 Keystrokes for navigating through data

To move to the …	Press …
Next base	Right arrow (→)
Previous base	Left arrow (←)
Next or previous N	Tab or Shift+Tab
Next or previous 10 bases	F5 or Shift+F5
Next or previous low QV	F6 or Shift+F6
Next or previous medium QV	F7 or Shift+F7
Next or previous high QV	F8 or Shift+F8

Toolbar

The toolbar displays buttons for the most commonly used functions. See the graphic below for the name, keyboard shortcut, and description of each toolbar button.



Add Sample(s)
Ctrl+I
Opens the Add Sample(s) dialog box

Save All Sample(s)
Ctrl+Shift+S
Saves changes to all samples

Start Analysis
Ctrl+R
Starts the selected analysis, post processing, and printing tasks

Remove Sample(s)
Delete
Removes the selected samples from the Sample Manager or the Sample Navigator

Print
Ctrl+P
Prints the selected views and analysis reports

View Sequencing Analysis Protocol
Ctrl+T
Opens the analysis protocol for the selected sample

Save Sample(s)
Ctrl+S
Saves changes to the selected samples

Copy
Ctrl+C

Analysis Report
Ctrl+B
Generates and displays the analysis report

Show/Hide QV
Ctrl+K
Switches the display of the quality values on and off

Toggle
Ctrl+N
Switches between the Sample Navigator and the Sample Manager views

Show/Hide
Ctrl+U
Displays the selected sequence file data

Full View
Ctrl+[
Displays all data in a standard-size window

Actual Size
Ctrl+]
Restores the display to the initial default zoom factor

Zoom In Horizontal
Ctrl+=
Enlarges the view horizontally

Zoom Out Horizontal
Ctrl+Minus
Reduces the view horizontally

Zoom In Vertical
Ctrl+Shift+=
Enlarges the view vertically

Zoom Out Vertical
Ctrl+Shift+Minus
Reduces the view vertically

Display Settings
Ctrl+Y
Opens the Display Settings dialog box

Show/Hide Original Sequence
Ctrl+J
Switches the display of the original sequence on and off

4/2009