applied biosystems

Sequencing Analysis Software 8

USER GUIDE

Publication Number MAN0019595

Revision A.0

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

ThermoFisher SCIENTIFIC

Life Technologies Holdings Pte Ltd | Block 33 | Marsiling Industrial Estate Road 3 | #07-06, Singapore 739256
For descriptions of symbols on product labels or product documents, go to thermofisher.com/symbols-definition.

Revision history: Pub. No. MAN0019595

Revision	Date	Description
A.0	6 September 2022	New document for Sequencing Analysis Software.

The information in this guide is subject to change without notice.

Contents

1 About Sequencing Analysis Software

About Sequencing Analysis Software

The Applied Biosystems™ Sequencing Analysis Software 8 performs the following tasks:

- Calls the bases
- Defines and displays mixed bases, if Smart Deep™ Basecaller or KB™ Basecaller is used
- Calculates and displays quality values, if Smart Deep™ Basecaller or KB™ Basecaller is used
- Calculates and displays the clear range
- Calculates sample score
- Creates output files in ABI (.seq), FASTA (.seq), Phred (.phd.1), and standard chromatogram (.scf) formats
- Generates an analysis report containing sample analysis statistics
- Prints data for each sample file
- Prints analysis reports
- Creates an audit trail to track all changes to bases and analysis settings, if enabled
- Compatible with Microsoft™Windows™ 10 Professional operating systems

Note: Do not install Sequencing Analysis Software 8 on computers with operating systems that are earlier than Microsoft™ Windows™ 10.

New features in Sequencing Analysis Software 8

- Enables basecalling using Smart Deep™ Basecaller v1.0, in addition to KB™ Basecaller and ABI™ Basecaller.
  For more information about basecallers and trace file compatibility, see “Supported data files” on page 10.
- Activates the Smart Deep™ Basecaller license with FlexNet.
- Enables enhanced trace visualization with Smart Deep™ Basecaller for plasmid data only.
  For more information about Smart Deep™ Basecaller enhanced trace visualization, see “Basecalling tab” on page 105.
- Smart Deep™ Basecaller analysis support in ABCall Utility 1.4.
  For more information, see the *ABCall Utility 1.4 User Guide* (Pub. No. MAN0028060).

Smart Deep™ Basecaller overview

Smart Deep™ Basecaller outperforms KB™ Basecaller in the following ways.

- Longer read length.
- More accurate pure and mixed base basecalls.
- Improved basecalling accuracy through heterozygous insertion deletion variants.
- More robust basecalling results through dye blob, mobility shift, and n-1 peak artifacts.
- Optional view for enhanced trace visual resolution at the 3’ end of plasmid sequencing traces.

Supported data files

Data files generated on the following instruments can be analyzed using the appropriate basecaller and mobility file. There is limited support for analysis of data files generated on earlier platforms and earlier versions of the data collection software.

Instrument	Supported basecaller
SeqStudio™ Genetic Analyzer	• Smart Deep™ Basecaller • KB™ Basecaller
SeqStudio™ Flex Series Genetic Analyzer	• Smart Deep™ Basecaller • KB™ Basecaller
3500/3500xL Genetic Analyzer	• Smart Deep™ Basecaller • KB™ Basecaller
3730/3730*xl* DNA Analyzer	• Smart Deep™ Basecaller • KB™ Basecaller • ABI™ Basecaller

(continued)

Instrument	Supported basecaller
3130/3130*xl* Genetic Analyzer	• KB™ Basecaller • ABI™ Basecaller
310 Genetic Analyzer	• KB™ Basecaller • ABI™ Basecaller

Note:

- Sequencing Analysis Software 8 must be installed on a computer with Microsoft™ Windows™ 10 or later operating systems. Although Sequencing Analysis Software 8 is compatible with .ab1 files that are generated from data collection software from computers with earlier Windows™ operating systems, you cannot install the software on computers with the older operating systems.
- You can use ABCall Utility 1.4 to run the Smart Deep™ Basecaller, KB™ Basecaller, and ABI™ Basecaller in batch mode. For more information, see the *ABCall Utility 1.4 User Guide* (Pub. No. MAN0028060).

About sample files

Sample files contain the target DNA sequence plus all the analysis information that is necessary to interpret the data and processing parameters. Sample files contain the following information about the DNA sequence:

- Raw data, as captured by the instrument before any post-collection processing
- Analysis results such as quality values and mixed bases
- Post processing results such as clear range
- The first sequence called by the Basecaller program
- Any edited basecalls that have been saved to the file
- Annotation information describing the instrument run and analysis conditions
- Processed (analyzed) electropherogram information that visually describes the intensity of each fluorescent signal
- Summary of electrophoresis conditions (voltage, temperature, current, power) during the run
- Audit trail of all actions that modify the end result sequence, if activated

All this information can be viewed in graphical and text formats.

The sample file name is the file that contains the sample information. The sample or database file name cannot be changed from within the **Sample Manager** window. It is created by using the data collection software.

The sample file name appears with the icon for the sample file when viewed from the hard disk. All sample files have the extension .ab1 and have a maximum character length of 255, including the .ab1 extension.

Note: The sample name is distinct from the sample file name. However, you may assign the same name to both, or you can change the sample name in the **Sample Manager** window.

Workflows

Add sample files to Sample Manager

1. **Start the software** (page 33)
2. **Create, edit, or select new Analysis Defaults** (page 35)
3. **Add sample files to the Sample Manager** (page 36)

(Optional) Edit Analysis Protocol settings

1. **Create or edit an analysis protocol** (page 39)
2. **Apply an analysis protocol to samples in the Sample Manager** (page 41)
3. **Edit a sample analysis protocol** (page 41)
4. **Apply the original pre-analysis settings to samples** (page 43)
5. **Delete an Analysis Protocol** (page 43)

Analyze the samples

Analyze the sample files (page 44)

Review processing parameters (BC, PP, and P check boxes) (page 45)

Review and edit base spacing, peak 1 location, and start and stop points (page 55)

Review the analysis parameters used in processing (page 60)

Edit Basecaller, DyeSet/Primer, and Matrix File selections (analysis parameters) (page 61)

Review and export reports

Review the Analysis Report (page 66)

Customize the data displayed in the Analysis Report (page 67)

Print the analysis report (page 68)

Export the analysis report (page 68)

2 Install the software

To perform the installation procedures in this chapter, you must have a Windows™ account with privilege to install programs on the computer.

Workflow

Install the software (most instruments)

Getting started

- "Computer requirements" on page 15
- "Before you begin" on page 16
- "(If needed) Upgrade from previous software versions" on page 16

Install the software (page 18)

Install Sequencing Analysis Software 8.

Install the software (3730/3730*xl* DNA Analyzer only)

Getting started

- "Computer requirements" on page 15
- "Before you begin" on page 16
- "(If needed) Upgrade from previous software versions" on page 16

Co-install the software with 3730/3730*xl* Data Collection Software versions (page 19)

- Launch the data collection software, then stop all services.
- Install the Sequencing Analysis Software 8, start the software, then activate the Smart Deep™ Basecaller license.
- Start all servies in the data collection software service console, then select the executable file.

Computer requirements

System Component	Requirements
CPU	Recommended: Core 2 Duo E8400/3.0 GHz (Core i7–4770S) Minimum: 2.0 GHz[1]
Operating system	Microsoft™ Windows™ 10[2]
Screen resolution	Minimum: 1280 × 1024 pixels Maximum: 2048 × 1152 pixels
Display scale	100%
Memory	Recommended: 4 GB (16 GB) Minimum: 2 GB[3]
Disk space	1 GB[4]

[1] If CPU is <3.0 GHz, the software displays a warning message during installation that software performance may be suboptimal.
[2] Do not install Sequencing Analysis Software 8 on computers with operating systems that are earlier than Microsoft™ Windows™ 10.
[3] If memory is <2 GB the software is not installed.
[4] If disk space is <2 GB the software prompts you to free up space. The software is not installed without sufficient space.

Before you begin

- Decide where to install the software:
  - On a computer that is connected to your genetic analyzer
  - On a computer that is dedicated to data analysis
- If you have an earlier version of the Sequencing Analysis Software installed, see "Uninstall the software" on page 16.
- Temporarily turn off any virus protection software.

  See the virus protection software user instructions for more information.
- *(3730/3730xl DNA Analyzer only)* Temporarily turn off Microsoft™ Defender.

  See the Microsoft™ Defender user instructions for more information.
- Exit all programs, except the data collection software for your genetic analyzer.

IMPORTANT! To properly install Sequencing Analysis Software 8 on computer that is connected to a genetic analyzer, the data collection software must be running. To install Sequencing Analysis Software 8 on computer that is connected to a 3730/3730*xl* DNA Analyzer, see "Co-install Sequencing Analysis Software 8 with 3730/3730xl Data Collection Software versions" on page 19.

(If needed) Upgrade from previous software versions

Before you install Sequencing Analysis Software 8, you must uninstall earlier versions of the software (see "Uninstall the software" on page 16).

Note: You can import analysis protocols from older versions of Sequencing Analysis Software (see "Export existing analysis protocols" on page 17 and "Import existing analysis protocols" on page 17). There is limited support for analysis of data files generated on older platforms and older versions of data collection software.

Sequencing Analysis Software 8 analyzes sample files generated on the following data collection software:

- SeqStudio™ Flex Series Instrument Software
- SeqStudio™ Data Collection Software
- 3130/3130xl Data Collection Software versions
- 3730/3730*xl* Data Collection Software versions
- 3500/3500xL Data Collection Software versions
- 310 Data Collection Software

Uninstall the software

IMPORTANT! If you are uninstalling the software from a computer connected to a SeqStudio™ Genetic Analyzer, 3500/3500xL Genetic Analyzer, 3730/3730*xl* DNA Analyzer, or 3130/3130*xl* Genetic Analyzer open the data collection software.

Close all other programs before running the uninstaller.

Note: If you want to import your analysis protocols from an earlier version of the DNA Sequencing Analysis Software, you must export and save existing analysis protocols (see "Export existing analysis protocols" on page 17). This action must be completed before you uninstall the software. Once Sequencing Analysis Software 8 is installed you can import your saved protocols (see "Import existing analysis protocols" on page 17).

1. Select **Start**, then navigate to your current version of the Sequencing Analysis Software.
2. Click **Uninstall**.
   The installer prompts the user to close any running application before proceeding with the uninstallation.
3. Click **OK**.
4. In the **Confirm Uninstall** dialog box, click **OK**.
   The files are uninstalled from the hard drive.
5. In the **Uninstall Complete** dialog box, click **Finish**.
   The system creates a backup folder, for example, **data_v7.0**. The backup folder retains the **DataStore** and **Logging** folders.

   IMPORTANT! The uninstall process may not delete files or folders that you have moved from their original installed location.

The system is now ready for installation of a new version of Sequencing Analysis Software.

Export existing analysis protocols

Analysis protocols should be present/created in the **Analysis Protocol Manager**.

1. In the homescreen, select **Analysis** ▸ **Analysis Protocol Manager**.
2. In the **Analysis Protocol Manager** window, select the protocol to be exported, then click **Export**.
   This will export the selected protocol as a .xml file and open the **Save** window for you to save the protocol for later use.

Import existing analysis protocols

After you have installed Sequencing Analysis Software 8, import your saved analysis protocols.

1. In the homescreen, select **Analysis** ▸ **Analysis Protocol Manager**.
2. In the **Analysis Protocol Manager** window, click **Import**, then select the exported analysis protocol in .xml file format.
3. Click **OK**
   The imported protocol will be listed in the **Analysis Protocol Manager**.

Install the software

To install the software on a computer connected to a 3730/3730*xl* DNA Analyzer, see “Co-install Sequencing Analysis Software 8 with 3730/3730xl Data Collection Software versions” on page 19.

The default installation destination folder is `C:\Applied Biosystems`.

1. Insert the installation USB.
2. Double-click **setup.exe** from the downloaded installer.
   The **InstallShield Wizard** launches automatically.
3. The installer prompts you to close any running applications before proceeding with the installation setup. Close the applications, then click **OK**.

   Note: Exit all programs, except 3730/3730*xl* or 3500/3500xL Data Collection Software, if applicable.
4. In the **Welcome to the InstallShield Wizard** window, click **Next**.
5. In the **License Agreement** window:
   a. Read the agreement.
   b. Select **I accept the terms of license agreement**.
   c. Click **Next**.
6. In the **Application Installation Location** window, click **Browse** to select a location, then click **Next**.
7. In the **Installation Parameters** window, in the **Current Settings** pane, verify the information, then click **Next**.
8. In the **InstallShield Wizard Complete** window, click **Finish**.

A Sequencing Analysis Software 8 shortcut is added to the desktop and to the **Start** menu.

Note: Turn the virus protection software on in case it was turned off during installation.

Co-install Sequencing Analysis Software 8 with 3730/3730*xl* Data Collection Software versions

Note: 3730/3730*xl* Data Collection Software 4.1 and 3730*xl* Data Collection Software 5 are only for use with the 3730/3730*xl* DNA Analyzer.

1. Launch the 3730/3730*xl* Data Collection Software.

   Note: The 3730/3730*xl* DNA Analyzer instrument computer needs to have 5 partitions.

   - If the 3730/3730xl data collection software is not installed, install the software on a Dell™ OptiPlex™ XE2 or XE3 computer running on Windows™ 10.

     IMPORTANT! Restart the 3730/3730*xl* DNA Analyzer, but do not open the service console.

     Note: See “Computer requirements” on page 15 for the recommended the screen resolution and display scale during installation.

   - If the 3730/3730*xl* Data Collection Software is installed, open the software, then stop all services.

     a. On your computer, navigate to one of the following:
        - *(3730/3730xl Data Collection Software 4.1)* `C:\Applied Biosystems\3730 Data Collection Software 4.1`
        - *(3730xl Data Collection Software 5)* `C:\Applied Biosystems\3730xl Data Collection Software 5`

        The **Service Console** window is displayed.

     b. In the **Service Console** window, click **Stop All**.

        The services will appear red. See Figure 1.

     c. Close the **Service Console** window.

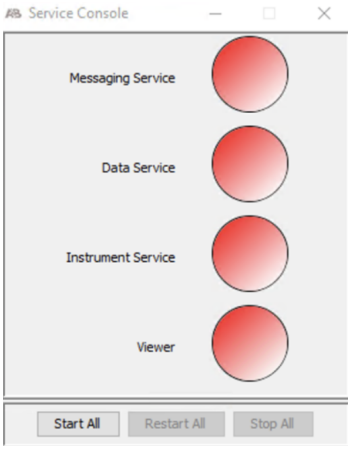


Figure 1 Stop all services

2. Install the Sequencing Analysis Software 8, start the software, then activate the Smart Deep™ Basecaller license.

   a. To install the Sequencing Analysis Software 8, see “Install the software” on page 18.

   b. To start the software and set up your account, see Chapter 3, “Get started with Sequencing Analysis Software 8”.

   c. To activate the Smart Deep™ Basecaller license, see Chapter 4, “Activate the Smart Deep™ Basecaller license”.

3. Open the 3730/3730*xl* Data Collection Software service console, then start all services.
   a. On your computer, navigate to one of the following:
      - *(3730/3730xl Data Collection Software 4.1)* `C:\Applied Biosystems\3730 Data Collection Software 4.1`
      - *(3730xl Data Collection Software 5)* `C:\Applied Biosystems\3730xl Data Collection Software 5`

      The **Service Console** window is displayed.

   b. In the **Service Console** window, click **Start All**.

      The services will appear green. See Figure 2.

4. On your computer, navigate to `C:\Applied Biosystems\SeqA\AppSeqA`, then double-click the .exe file, **SeqA_RegisterApp.exe**.

Note: **SeqA_RegisterApp** must be selected while all services are started.

The Sequencing Analysis Software 8 is co-installed with the 3730/3730*xl* Data Collection Software 4.1 or 3730*xl* Data Collection Software 5.

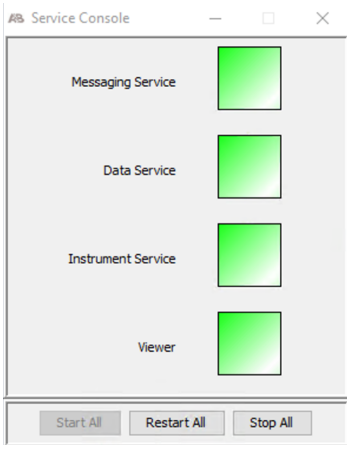


Figure 2 Start all services

Unregister Sequencing Analysis Software 8 co-installed with 3730/3730xl Data Collection Software versions

1. Open the 3730/3730*xl* Data Collection Software 4.1 or 3730*xl* Data Collection Software 5 service console, then start all services.
   a. On your computer, navigate to on of the following:
      - *(3730/3730xl Data Collection Software 4.1)* `C:\Applied Biosystems\3730 Data Collection Software 4.1`
      - *(3730xl Data Collection Software 5)* `C:\Applied Biosystems\3730xl Data Collection Software 5`

      The **Service Console** window is displayed.

b. In the **Service Console** window, click **Start All**.

   The services will appear green. See Figure 2.

2. On your computer, navigate to `C:\AppliedBiosystems\SeqA\AppSeqA`, then double-click the .exe file, **SeqA_UnRegisterApp.exe**.

Note: **SeqA_UnRegisterApp** must be selected while all services are started.

The Sequencing Analysis Software 8 is unregistered with the 3730/3730*xl* Data Collection Software 4.1 or 3730*xl* Data Collection Software 5.

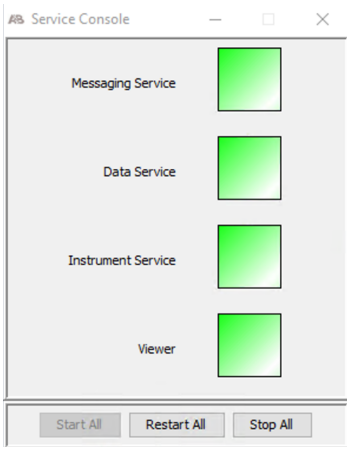


Figure 3 Start all services

3 Get started with Sequencing Analysis Software 8

Edit .ab1 file software association (if needed)

Note: If you do not have Primer Express™ software loaded on your computer, skip this section.

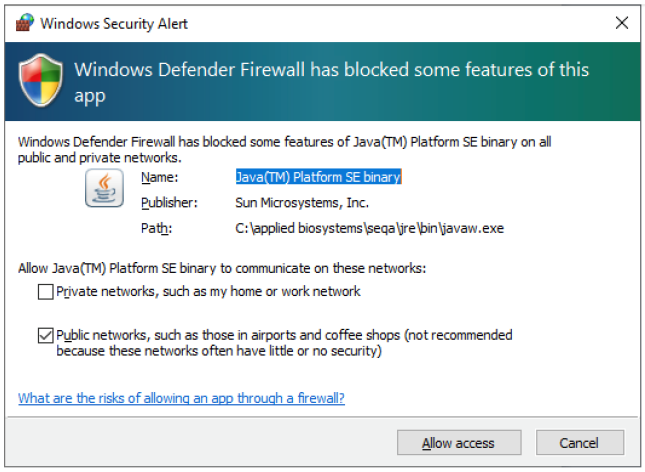
If you have both the Sequencing Analysis Software 8 and Primer Express™ software installed on your computer, then the AB1 file may be associated with either software, depending on the order of the installation. Use this procedure to ensure that the .ab1 file type associates with the Sequencing Analysis Software 8, enabling you to automatically open the software when you double-click a sample file.

1. On your computer desktop, select **Start** ▸ **Settings** ▸ **Apps**.
2. In the **Apps** menu on the left, click **Default apps**.
3. Click **Choose default apps by file type**.
4. Navigate to .ab1, then confirm that Sequencing Analysis Software 8 is the default app. If not, click the icon, then select the Sequencing Analysis Software 8 from the list.
5. Close the **Settings** window.

Start the software for the first time

Note: Before you start the software, confirm your screen resolution and display scale settings (see “Computer requirements” on page 15).

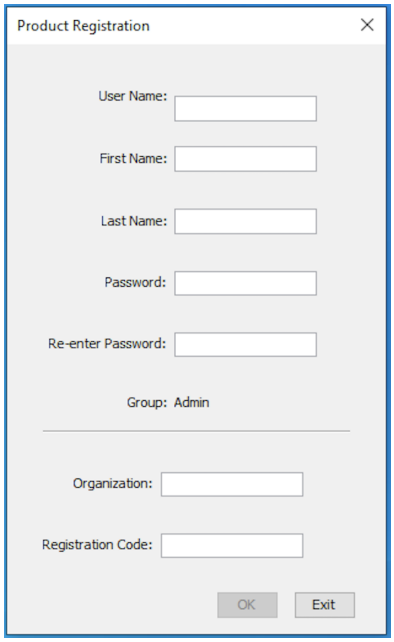
1. Start the software by double-clicking the desktop shortcut or select **Start ▸ SeqA 8**.
   a. Occasionally on new systems, Java may show a **Windows Security Alert**. Click **Allow access**, then the application will direct you to the **Product Registration** dialog box.



2. In the **Product Registration** dialog box, enter all the information in the text fields.
   The user name and password must be 6 to 15 characters long.

IMPORTANT! The user name must not contain spaces or invalid characters (\ / : * ? " < > |).

Note: The first user created is automatically assigned Administrator privileges.



3. Enter the registration code that you received with your software.
4. Click **OK**.
   While the program is loading, the splash screen opens, then the Log In dialog box opens.
5. In the **Log In** dialog box, enter your user name and password, then click **OK**.
6. In the **License Dialog** window, read the license agreement, then click **Accept**.
7. Read the license agreement and click **Accept**.

The software opens to the home screen. To create additional accounts for other users, see “Create a new user account (admin only)” on page 24.

File-naming convention

Some alphanumeric characters are not valid for user names or file names. The invalid characters include:

- Spaces
- \ / : * ? " < > |

An error message is displayed if you use any of the invalid characters. You must remove the invalid character to continue.

Create a new user account (admin only)

We recommend that you create a user profile for each individual who uses the software. There are three levels of users: administrator, scientist, and analyst. The **Users** tab allows exporting of user names and access privileges for these users.

IMPORTANT! The administrator is the only person who can set up and change the information in the **Users** tab. The selections in this tab are inactive for all other users.

1. In the software home screen, select **Tools** ▸ **Options**.
2. In the **Options** dialog box, in the **Users** tab, click **New**.
   The **User Management** dialog box is displayed.
3. Enter the user name, first and last name, and password in the appropriate fields.
4. In the **User Group** dropdown list, select the level of user.

Note: The user name should only contain alphanumeric characters. Do not use spaces or invalid characters (\ / : * ? " < > |) (see “File-naming convention” on page 24).

5. *(Optional)* Select the **Enable Electronic Signature** checkbox.

6. Click **OK**.
   The new user appears in the list in the **Users** tab.

Restart the software so that new users can log in. The users can change their passwords after logging into the software (see “Changing the user password” on page 99).

Set up authentication settings

The authentication settings control user lockout, software timeout, and password expiration.

Note: The Administrator is the only person who can set up and change the information in the **Authentication** tab. The selections in this tab are inactive for all other users.

If you have authentication settings that were exported from another instance of the software, click **Import** to import the settings, instead of completing a manual set up.

1. In the software home screen, select **Tools** ▸ **Options**.
   The **Options** window is displayed.

2. In the **Authentication** tab, edit the following settings.

Setting	Description
User lockout	• The **Lockout user after invalid login attempts** field indicates the number of times that a user can enter an incorrect password or user name before they are locked out. Enter the number or accept the default. • The **within minutes** field indicates the maximum number of login attempts that can occur within the defined time. Enter a number or accept the default. • The **Maintain lockout for minutes** field indicates the number of minutes that must elapse before the user can login again after being locked out of the software. Enter the number of minutes or accept the default.
Timeout	Select **Timeout Feature On** to enable the software to time out and force the user to sign in again. In the **Automatic timeout after minutes** field, enter the number of minutes or accept the default.
Password expiration	The **Change password every days** field indicates the number of days before the users must enter a new password. Enter a number of days or accept the default.

3. Click **OK** to save the authentication settings.

Set up audit reasons

The **Authentication** and **Audit** panes provide a way to track project changes such as base change, variants, or processes you want to track. You must select **Audit Trail On** for tracking to occur.

Note: The administrator is the only person who can set up and change the information in the **Audit** tab. The selections in this tab are inactive for all other users.

---

1. In the software home screen, select **Tools** ▸ **Options** ▸ **Audit**.
2. In the **Audit** tab, select the **Audit Trail On** check box.
   If the audit trail is on, then a dialog box will automatically open if one of the events in the **Reasons** list occurs.
3. To create a new audit reason, click **New**.
4. In the **Audit Reason Editor** dialog box, in the **Reason** field, enter an audit reason.
5. (*Optional*) Enter a description of the reason.
6. Click **OK**.
   The new reason appears in the **Reason** list.
7. In the **Options** dialog box, click **OK** to save the audit settings.

Whenever a change is made in any of the data views, the **Audit Reason Editor** dialog box opens, allowing you to select the reason for the change from a dropdown list.

Enable and set up electronic signature function

The electronic signatures function is turned off by default. When enabled, you can to electronically sign off when saving, printing, or exporting an analysis report. You can also sign at any other time.

Each electronic signature record is captured in the **Electronic Signature** tab of **Sample Manager** (see “About the Electronic Signature tab” on page 133).

Enable electronic signatures for the software

Electronic signatures must be enabled for the software, then for each user that needs to use the feature. The administrator is the only user who can enable electronic signatures.

1. Click **Tools** ▸ **Options**
2. In the **Electronic Signature** tab, select **Enable Electronic Signature**.
3. Select the actions that require an electronic signature.

   Note: Only users with the electronic signature function enabled will be prompted for a signature when completing the action.

   - **Save Sample(s)**
   - **Print**
   - **Export Report**

Enable electronic signatures for each user (see “Enable Electronic Signatures for a user” on page 27).

Enable Electronic Signatures for a user

Electronic signatures must be enabled for the software (see “Enable electronic signatures for the software” on page 27). The administrator is the only user who can enable electronic signatures.

1. Click **Tools** ▸ **Options**
2. In the **Users** tab, select a user, then click **Open**.
3. In the **User Management** dialog box, select **Enable Electronic Signature**
4. Click **OK** to close the **User Management** dialog box.
5. In the **Users** tab, click **OK** to save the changes.

Modify the electronic signature actions

Each time you electronically sign, you must select an action. The software comes with two default actions, **Review** and **Approve**. An administrator can modify the existing actions or add other actions.

1. Log in as an administrator.
2. Click **Tools** ▸ **Options**.
3. In the **Electronic Signature** tab:
   - To update an existing action, select an action from the **Actions List** dropdown list, then click **Open**.
   - To create a new action, click **New**.

4. In the **Electronic Signature Action** dialog box, edit the appropriate fields.

   Note: Only users with the electronic signature function enabled will be prompted for a signature when completing the action.

   - **Action**—Enter or edit the action name.
   - **Meaning**—Enter or edit the action meaning.
   - **Challenge**—Enter or edit the legal implications of the action.

5. (*Optional*) To inactivate an action, select **Inactive**.
   If selected, the action is grayed out in the **Electronic Signature** tab of the **Options** dialog box, and no longer appears in the list of actions for an electronic signature.

6. Click **OK** to save any changes.

View the configuration history for an electronic signature action

If a change is made to an electronic signature action, the software logs the change in the configuration history.

1. Click **Tools** ▸ **Options**

2. In the **Electronic Signature** tab, click **Show Configuration History**.

3. In the **Configuration History** window, review the list of active and inactive actions. If an action has been modified, the most recent version is at the bottom of the list.

4. (*Optional*) Double-click an entry to view the description of the action in a dialog box.

5. (*Optional*) To display the **Challenge Text**, select **View** ▸ **Show Challenge Column**.

6. (*Optional*) To export the configuration history as a tab-delimited text file, select **File** ▸ **Export**.
   Select an export location, then click **Export**.

Set up automated KB/SDB basecalling

If the Sequencing Analysis Software 8 is installed on a computer connected to a 3730/3730*xl* DNA Analyzer, then you need to set up the data collection software for automatic KB/SDB basecalling. See “Co-install Sequencing Analysis Software 8 with 3730/3730xl Data Collection Software versions” on page 19.

Note: Automated KB/SDB basecalling is not supported for other instruments.

Note: *(SDB only)* Activate the Smart Deep™ Basecaller license prior to setting up automated basecalling (see Chapter 4, “Activate the Smart Deep™ Basecaller license”).

The following data collection softwares are supported:

- 3730*xl* Data Collection Software 5
  See *3730xl DNA Analyzer with 3730xl Data Collection Software 5 User Guide* (Pub. No. 100077621).
- 3730/3730*xl* Data Collection Software 4.1
  See *Applied Biosystems™ 3730/3730xl DNA Analyzers Getting Started Guide* (Pub. No. 4478016).

4 Activate the Smart Deep™ Basecaller license

The following sections describe how to activate the Smart Deep™ Basecaller license.

Smart Deep™ Basecaller is compatible with .ab1 files generated on the 3500/3500xL Genetic Analyzer, 3730/3730*xl* DNA Analyzer, SeqStudio™ Genetic Analyzer, and SeqStudio™ Flex Series Genetic Analyzer.

Smart Deep™ Basecaller is not compatible with 3130, 3100, or 310 instrument data.

Note:

- A 90-day Smart Deep™ Basecaller demonstration license is included with the Sequencing Analysis Software 8. To order the Smart Deep™ Basecaller license, contact your local sales office.
- The license is valid until the expiration date. If you need to uninstall, then reinstall the software on the same computer, use the same license until the expiration date. Each Smart Deep™ Basecaller license is tied to the computer ID.
- For questions about the license, send an email message to **seqa.support@thermofisher.com**

Add AB1 files and start Smart Deep™ Basecaller analysis

You must first install and register Sequencing Analysis Software 8 before you can add files and start sample analysis.

For instructions to back up analysis protocols, uninstall an earlier version of the software, and install Sequencing Analysis Software 8, see Chapter 2, “Install the software”.

1. Add one or more .ab1 files (see “Add sample files to the Sample Manager” on page 36).
2. In the **Sample Manager**, select **SDB.bcp** from the **BaseCaller** dropdown menu.

   Note: The appropriate SDB .mob files for trace files uploaded from all supported instruments are automatically updated in the **DyeSet/Primer** dropdown menu in **Sample Manager**.

   ---
3. In the **Sample Manager**, select the **BC** parameter checkbox.
4. Click (**Start Analysis**).

   The **Warning: Smart Deep Basecaller** window is displayed.
5. In the **Warning: Smart Deep Basecaller** window, click **Yes**.

   The license activation window appears, automatically displaying the Computer ID, which cannot be edited.

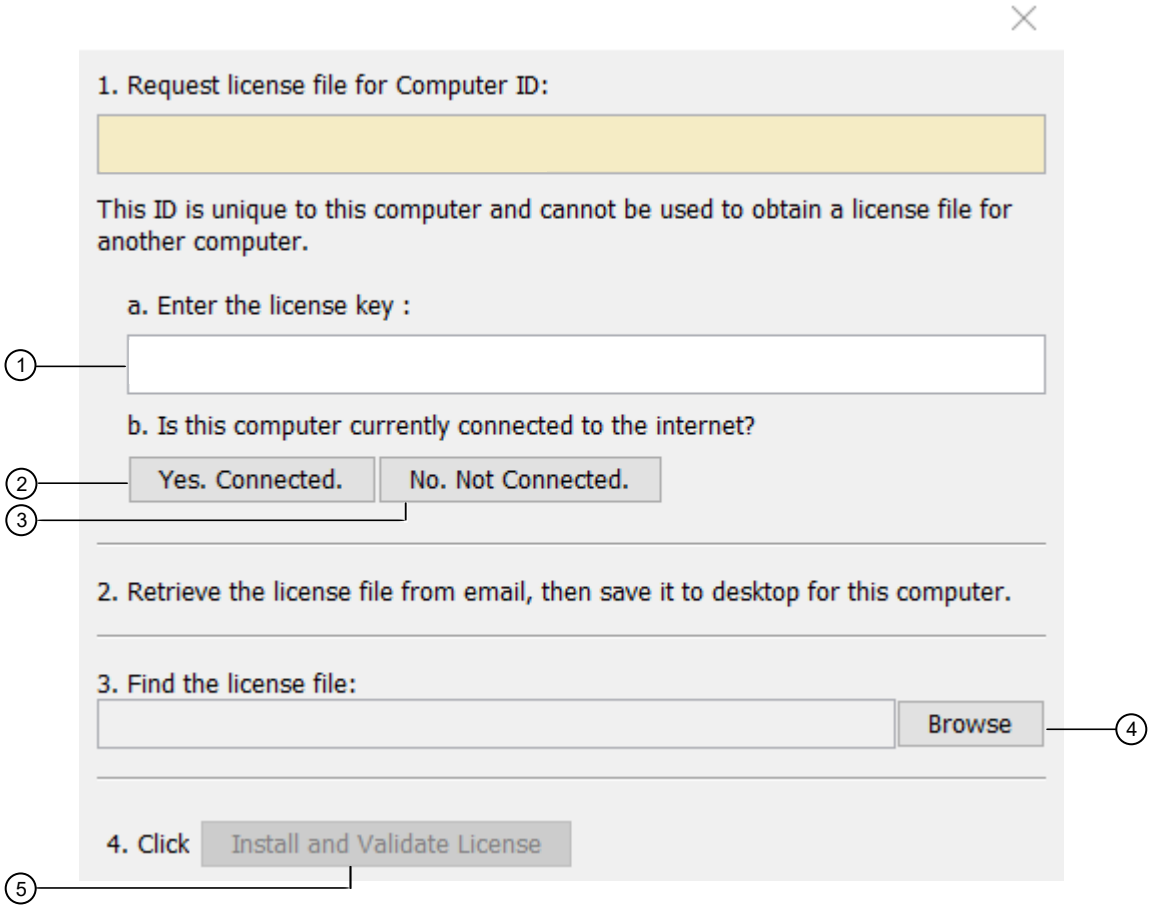


Figure 4 Activation window

① License key
② Connect to the internet (online activation)
③ Not connected to the internet (offline activation)
④ Browse for the **SampleRequest.bin** file (offline activation)
⑤ **Install and Validate License** (offline activation)

If	Go to
The computer is connected to the Internet	"Activate the Smart Deep™ Basecaller license online (preferred)" on page 31
The computer is not connected to the Internet	"Activate the Smart Deep™ Basecaller license offline" on page 32

Activate the Smart Deep™ Basecaller license online (preferred)

Follow these instructions if your computer is connected to the internet.

1. In the **Enter the license key** field (see Figure 4), enter the SDB license key that you received via email.

2. Click **Yes. Connected** (see Figure 4).

   The **Success: Smart Deep Basecaller** window appears, confirming successful application of the Smart Deep™ Basecaller license.

3. In the **Success: Smart Deep Basecaller** window, click **OK**.

The Smart Deep™ Basecaller license is activated. The Smart Deep™ Basecaller is ready to be used for analysis.

Note: The license is valid until its expiration date. If you need to uninstall, then reinstall the software on the same computer, use the same license until its expiration date.

Activate the Smart Deep™ Basecaller license offline

Follow these instructions if your computer is *not* connected to the internet.

1. In the **Enter the license key** field, enter the SDB license key that you received via email, then click **No. Not connected** (see Figure 4 on page 31).

   The **Save the request file** window appears.

2. Click **Save** to save the **SampleRequest.bin** file.

3. Email the **SampleRequest.bin** file to **seqa.support@thermofisher.com**.

4. Download the **CapabilityResponse.bin** file received via email.

5. In the license activation window:

   a. Click **Browse** to navigate to the **CapabilityResponse.bin** file (see Figure 4 on page 31).

   b. Click **Install and Validate License** (see Figure 4 on page 31).

   The **Success: Smart Deep Basecaller** window appears.

6. In the **Success: Smart Deep Basecaller** window, click **OK**.

The Smart Deep™ Basecaller license is activated. The Smart Deep™ Basecaller is ready to be used for analysis.

Note: The license is valid until its expiration date. If you need to uninstall, then reinstall the software on the same computer, use the same license until its expiration date.

---

5 Add sample files to the Sample Manager

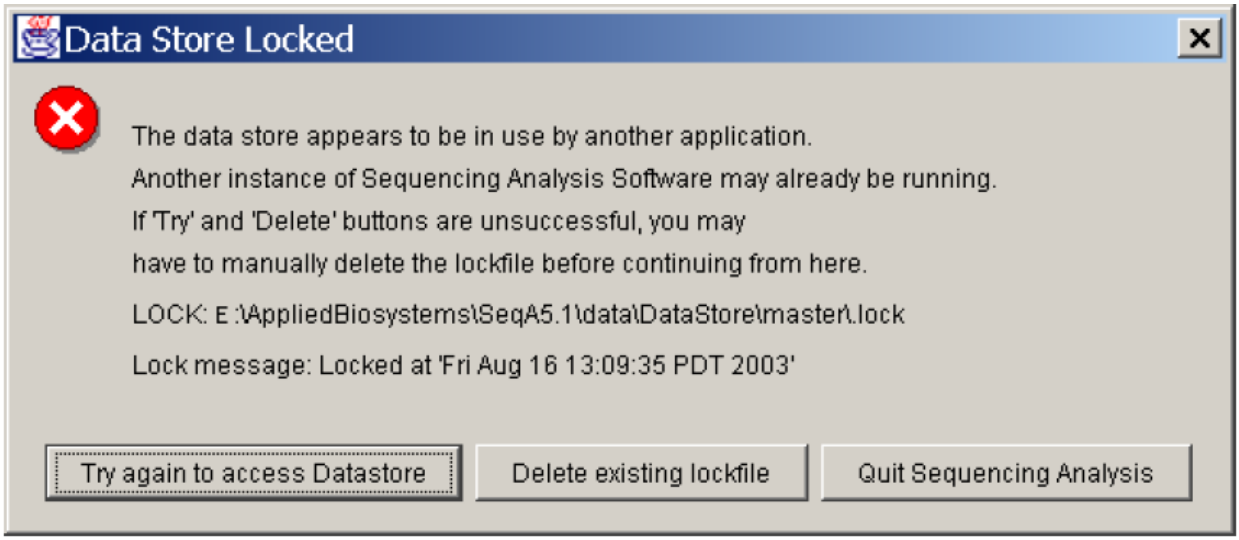
When a sample file is added to the **Sample Manager**, it passes through the Analysis Defaults (see “Analysis defaults” on page 101).

Start the software

1. Double-click the desktop shortcut or select **Start** ▸ **SeqA 8**.
2. Enter the user name and password, then click **OK**.

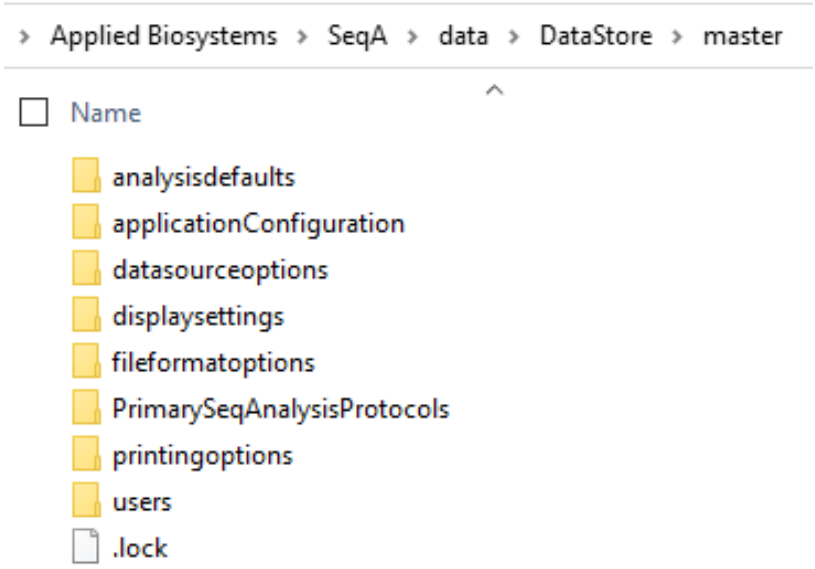
If the Sequencing Analysis Software 8 does not open

Sometimes when the software has been previously shut down improperly, the software cannot access the DataStore information. When this happens, the following dialog box opens instead of the Sequencing Analysis main window.



To remove the lockfile, try any of the following methods until one works:

- Click the **Delete existing lockfile** button.
- Click the **Try again to access Datastore** button.
- Locate and delete the lockfile using the path:
  `C:\AppliedBiosystems\SeqA\data\DataStore\master`



After the lockfile is removed, the Sequencing Analysis main window opens.

Create, edit, or select new Analysis Defaults

For more information about Analysis Defaults, see “Analysis defaults” on page 101.

1. In the menu bar, select **Analysis ▸ Analysis Defaults**.

   You can also edit the Analysis Defaults when adding sample files to the **Sample Manager** (see “Add sample files to the Sample Manager” on page 36).

2. In the **Analysis Defaults** window, in the **Add Samples Settings** pane, select an option from the **Analysis Protocol** dropdown list.

   Note: By default, the Analysis Protocol is set to **None**.

Option	Description
New...	Create a new Analysis Protocol. Click **New...**, then see “Create or edit an analysis protocol” on page 39.
Edit...	Edit an existing Analysis Protocol. Select **Edit...**, then see “Create or edit an analysis protocol” on page 39.
Select **None**	Do not apply an Analysis Protocol to incoming samples. **Note:** Sample files generated on a 310 Genetic Analyzer do not contain an Analysis Protocol. If an Analysis Protocol is not selected in the **Analysis Defaults**, then an alert message is displayed during import of these sample files.
Select an existing Analysis Protocol.	Select an existing Analysis Protocol from the dropdown list.

3. In the **Add Sample Settings** pane, select or deselect processing parameters.

   - **Base Calling (BC)**
   - **Post Processing (PP)**
   - **Print (P)**

   For more information about processing parameters, see “Processing parameters” on page 115.

4. In the **Sequence File Formats** pane, select one of the following options.

Option	Description
Use the settings in the sample's Analysis Protocol	If selected, the sequence file formats of the Analysis Protocol are applied to sample files.
Override the sample's Analysis Protocol and set to	If selected, the sequence file formats of the Analysis Protocol are overwritten. Allows for the creation of .seq file in the ABI or FASTA format, (.scf) files and/or Phred (.phd.1) files

5. Click **OK**.

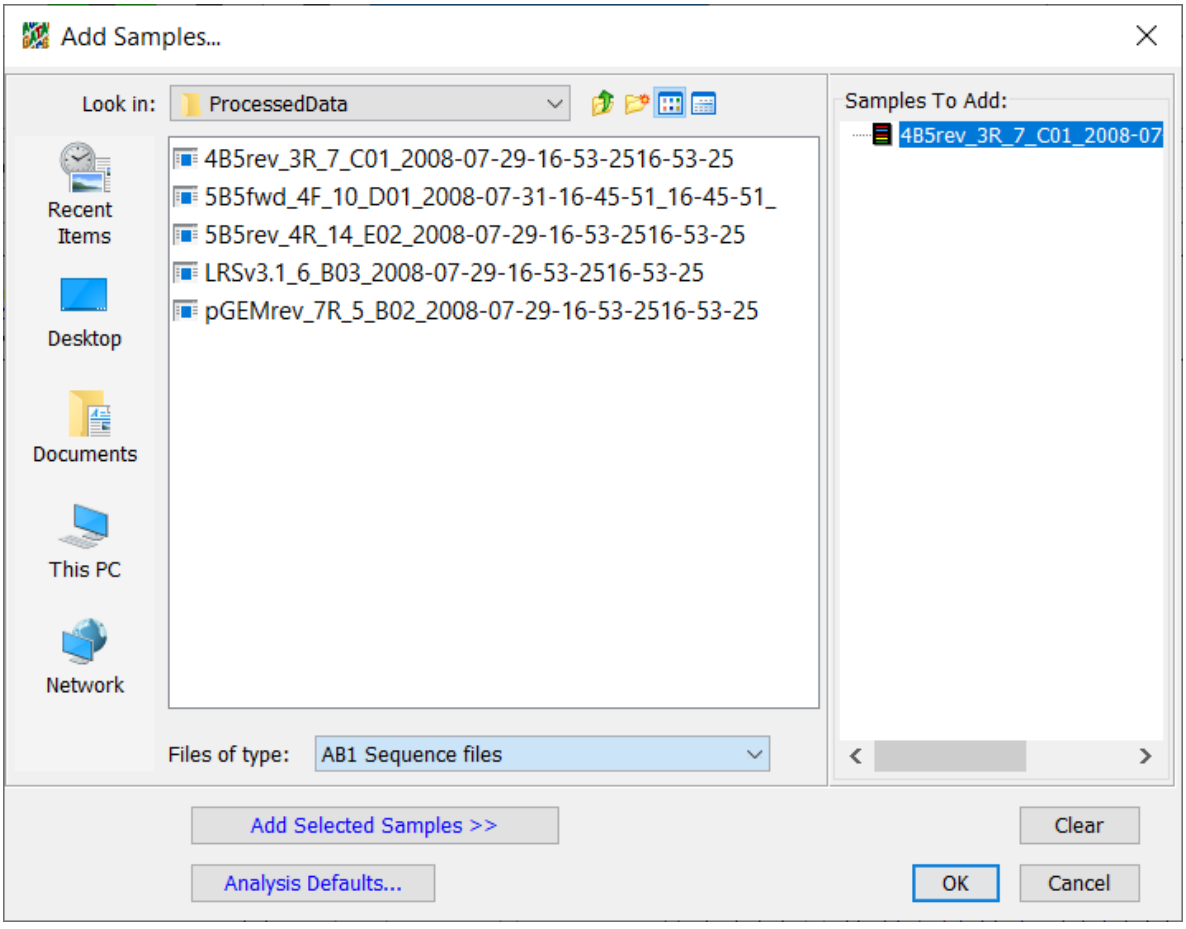
The Analysis Defaults settings are applied to samples at the following times:

Setting	Settings are applied to samples …
Analysis protocol (310 Genetic Analyzer sample files only)	After the samples are added to the Sample Manager.
Processing parameters	After the samples are added to the Sample Manager.
Sequence file formats parameters	After the samples are analyzed.

Add sample files to the Sample Manager

For shortcuts to open sample files, see “Additional methods to add sample files to Sample Manager” on page 37.

1. In the **Sample Manager** screen, select an option to add one or more samples:
   - In the toolbar, click .
   - In the menu bar, select **File** ▸ **Add Samples**.
2. In the **Add Samples** dialog box, in the **Files** pane, navigate to the sample files.



3. In the **Files** pane, select a folder, sample file or sample files, then click **Add Selected Samples**

Option	Action
A single file	Select the file, then click **Add Selected Samples**.
Multiple files	**Ctrl+Click** to select discontinuous samples, or **Shift+Click** to add continuous samples, then click **Add Selected Samples**. You can also add the files individually.
A folder containing sample files	Select the folder, then click **Add Selected Samples** **Note:** If you add a folder that contains sub-folders, any sample files located in the sub-folders are not added to the **Sample Manager**.

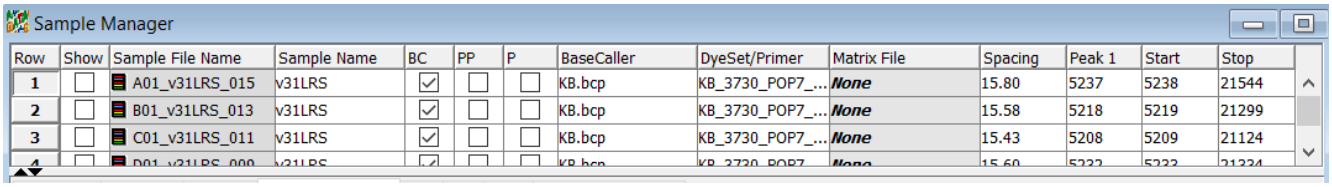
   The sample files populate the **Samples To Add** pane.

4. (*Optional*) To edit analysis defaults, click **Analysis Defaults** (see "Create, edit, or select new Analysis Defaults" on page 35).
   The analysis defaults can also be edited prior to adding the sample files.

5. When all of the sample files are added to the **Samples To Add** pane, click **OK**.

   Note: If a **Missing Analysis Defaults** alert displays while adding samples to **Sample Manager**, add analysis defaults to the samples (see "Create, edit, or select new Analysis Defaults" on page 35).

Files are listed in the **Sample Manager** in the order in which they were added to the **Samples to Add** pane.



Additional methods to add sample files to Sample Manager

Note: If the software is not already open, these methods also open the software.

- Double-click the sample file.
- Drag selected sample files to the software shortcut icon on the computer desktop.
- Select the sample files, right-click, then select **Open with SeqA8App**.

  Note: Limit the number of samples to 15.

Remove sample files from Sample Manager

Note: If a sample file is actively being analyzed by the software, it cannot be removed from **Sample Manager**. To skip analysis a sample without removing the sample file from **Sample Manager**, deselect the processing parameters columns (**BC**, **PP**, **P**) before reanalysis.

Select one of the following options to remove the sample files:

To remove …	Then …
All the samples	In the menu bar, select **File** ▸ **Remove All Samples**.
A single sample	1. In the **Sample Manager** pane, click the number in the **Row** column. 2. Hit **Delete**, click  in the tool bar, or select **File** ▸ **Remove All Samples**
Multiple samples that are not next to each other	1. In the **Sample Manager** pane, Ctrl+click the number in the **Row** column for each sample file. 2. Hit **Delete**, click  in the tool bar, or select **File** ▸ **Remove All Samples**
Multiple files that are next to each other	1. In the **Sample Manager** pane, click the number in the **Row** column for the first sample, then Shift+click the **Row**. or, Drag the mouse down the row number column. 2. Hit **Delete**, click  in the tool bar, or select **File** ▸ **Remove All Samples**

6 Edit analysis protocol settings

Edit the Analysis Protocol settings for samples in the **Sample Manager**.

Create or edit an analysis protocol

For more information about the parameters defined in the Analysis Protocol, see “Analysis protocols” on page 103.

To edit the analysis protocol for an individual sample, see “Edit a sample analysis protocol” on page 41.

1. In the homescreen, select **Analysis** ▸ **Analysis Protocol Manager**.
2. In the **Analysis Protocol Manager** window, select one of the following options:

Option	Action
Create a new analysis protocol.	Click **New**.
Create a new analysis protocol from an existing analysis protocol, then edit.	Select an existing analysis protocol from the list, then click **Save As**.
Edit an existing analysis protocol.	Select an existing analysis protocol from the list, then click **Edit**.

3. In the **General** tab, enter the analysis protocol name and description, then edit the sequence file format settings:
   - **Write .Seq File**

     Note: If **Write Clear Range in .Seq File** is selected, the software only writes the clear range into the .seq file.
   - **Write Standard Chromatogram Format (.scf)**
   - **Write Phred (.phd.1) File**

   For more information, see “General tab” on page 104.

4. In the **Basecalling** tab, configure the following settings:

   For more information, see "Basecalling tab" on page 105.

Parameter	
Basecalling	1. Select a basecalling algorithm from the dropdown list (see "Select the correct basecaller" on page 118). **Note:** There are many algorithm options for the old basecaller. KB™ Basecaller and Smart Deep™ Basecaller each have only one algorithm option. 2. Select a mobility file in the **DyeSet/Primer** dropdown list (see "Select a DyeSet/Primer (mobility) file" on page 119). If KB™ Basecaller is selected, select a .mob file beginning with KB and that matches your instrument, polymer, and dye set configuration. If Smart Deep™ Basecaller is selected, select a .mob file beginning with SDB and that matches your instrument, polymer, and dye set configuration. **Note:** If the mobility file name is displayed in bold italicized font, then the file is not present in the Mobility folder and the sample file cannot be processed with that file. To copy DyeSet/Primer files into the **Mobility** folder, see "Copy and paste mobility files for 310 Genetic Analyzer" on page 134. 3. (310 Genetic Analyzer only) Select a file in the **Matrix File** dropdown list (see "About matrix files (310 Genetic Analyzer only)" on page 120).
Processed Data	(Smart Deep™ Basecaller or KB™ Basecaller only) Select to process the data with a **True Profile** or **Flat Profile**.
Quality Threshold	(Smart Deep™ Basecaller or KB™ Basecaller only) Select an option for assigning N's to basecalls.
Ending Base	Select an option to select the ending base. Select **At PCR Stop** if the electropherogram is expected to contain a PCR stop.
SBD Trace Visualization	(Smart Deep™ Basecaller only) Select **Classic View** or **Enhanced View**. When **Enhanced View** is selected (Smart Deep™ Basecaller only), the trace resolution is improved relative to the **Classic View**. There is no difference in the bases called or the quality values between the **Enhanced View** and the **Classic View**. When **Classic View** is selected, the trace resolution is identical between KB™ Basecaller and Smart Deep™ Basecaller, but the basecalls and quality values may differ between KB™ Basecaller and Smart Deep™ Basecaller.

5. In the **Mixed Bases** tab, select **Use Mixed Base Identification**, then set a new secondary peak threshold or accept the default of ≥25%.

   Note: The mixed base identification option is disabled if you select **Enhanced View** in the **Basecalling** tab.
   For Smart Deep™ Basecaller, only the default value of 25% is available.

   For more information, see "Mixed Bases tab" on page 111

6. In the **Clear Range** tab, select one or more stop points for data analysis:
   - **Use clear range minimum and maximum**
   - (Smart Deep™ Basecaller and KB™ Basecaller) **Use quality values**
   - (ABI™ Basecaller) **Use identification of N calls**

   For more information, see “Clear range tab” on page 113.
7. Click **OK** to save the analysis protocol and close the **Sequence Analysis Protocol Editor** dialog box.
8. Click **Done** to close the **Analysis Protocol Manager**.

To apply the analysis protocol to samples, see “Apply an analysis protocol to samples in the Sample Manager” on page 41.

Apply an analysis protocol to samples in the Sample Manager

1. If applying the analysis protocol to select samples only, select the sample rows in **Sample Manager**.
   - Shift+click to select continuous samples.
   - Ctrl+click to select discontinuous samples.
2. In the menu bar, select **Analysis** ▸ **Analysis Protocol Manager**.
3. In the **Analysis Protocol Manager** window, select an analysis protocol.
4. (*Optional*) Edit the analysis protocol (see “Create or edit an analysis protocol” on page 39).
5. In the **Analysis Protocol Manager** window, click one of the following options:
   - **Apply to All Samples** to apply the protocol to all the sample files in the **Sample Manager**.
   - **Apply to Selected Samples** to apply the protocol to the sample files selected in the **Sample Manager**.
6. Click **Done** to close the **Analysis Protocol Manager**.

After applying an analysis protocol to samples, you must reanalyze the samples for the protocol settings to take affect (see “Analyze the sample files” on page 44).

Edit a sample analysis protocol

Edit the analysis protocol for an individual sample in the **Sample Manager**.

1. In the **Sample Manager**, select a sample row.
2. Click  or select **Analysis** ▸ **Analysis Protocol**.

3. In the **General** tab, enter the analysis protocol name and description, then edit the sequence file format settings:

   - **Write .Seq File**

     Note: If **Write Clear Range in .Seq File** is selected, the software only writes the clear range into the .seq file.

   - **Write Standard Chromatogram Format (.scf)**
   - **Write Phred (.phd.1) File**

   For more information, see “General tab” on page 104.

4. In the **Basecalling** tab, configure the following settings:

   For more information, see “Basecalling tab” on page 105.

Parameter	
Basecalling	1. Select a basecalling algorithm from the dropdown list (see “Select the correct basecaller” on page 118). 2. Select a mobility file in the **DyeSet/Primer** dropdown list (see “Select a DyeSet/Primer (mobility) file” on page 119). **Note:** If the mobility file name is displayed in bold italicized font, then the file is not present in the Mobility folder and the sample file cannot be processed with that file. To copy DyeSet/Primer files into the **Mobility** folder, see “Copy and paste mobility files for 310 Genetic Analyzer” on page 134. 3. (310 Genetic Analyzer only) Select a file in the **Matrix File** dropdown list (see “About matrix files (310 Genetic Analyzer only)” on page 120).
Processed Data	(Smart Deep™ Basecaller or KB™ Basecaller only) Select to process the data with a **True Profile** or **Flat Profile**.
Quality Threshold	(Smart Deep™ Basecaller or KB™ Basecaller only) Select an option for assigning N's to basecalls.
Ending Base	Select an option to select the ending base. Select **At PCR Stop** if the electropherogram is expected to contain a PCR stop.
SBD Trace Visualization	(Smart Deep™ Basecaller only) Select **Classic View** or **Enhanced View**. When **Enhanced View** is selected (Smart Deep™ Basecaller only), the trace resolution is improved relative to the **Classic View**. There is no difference in the bases called or the quality values between the **Enhanced View** and the **Classic View**. When **Classic View** is selected, the trace resolution is identical between KB™ Basecaller and Smart Deep™ Basecaller, but the basecalls and quality values may differ between KB™ Basecaller and Smart Deep™ Basecaller.

5. In the **Mixed Bases** tab, select **Use Mixed Base Identification**, then set a new secondary peak threshold or accept the default of ≥25%.

Note: The mixed base identification option is disabled if you select **Enhanced View** in the **Basecalling** tab.

For Smart Deep™ Basecaller, only the default value of 25% is available.

For more information, see “Mixed Bases tab” on page 111

6. In the **Clear Range** tab, select one or more stop points for data analysis:
   - **Use clear range minimum and maximum**
   - (Smart Deep™ Basecaller and KB™ Basecaller) **Use quality values**
   - (ABI™ Basecaller) **Use identification of N calls**

   For more information, see “Clear range tab” on page 113.

7. Click **OK** to save the analysis protocol.

Note: The change affects the protocol for the selected sample only.

After changing the sample analysis protocol, you must reanalyze the samples for the protocol settings to take affect (see “Analyze the sample files” on page 44).

Apply the original pre-analysis settings to samples

1. Select the sample rows in **Sample Manager**.
   - Shift+click to select continuous samples.
   - Ctrl+click to select discontinuous samples.
2. In the menu bar, select **Analysis** ▸ **Apply Last Saved Settings**.
3. Reanalyze the samples.
4. If you want to save the changes, save the sample files (see “Save sample files” on page 61).

After applying pre-analysis settings to samples, you must reanalyze the samples for the settings to take affect (see “Analyze the sample files” on page 44).

Delete an Analysis Protocol

1. In the menu bar, select **Analysis** ▸ **Analysis Protocol Manager** .
2. Select the Analysis Protocol, then click **Delete**.
3. In the **Deletion Confirmation** dialog box, click **Yes**.
4. Click **Done** to close the **Analysis Protocol Manager**.

7 Analyze the samples and review overall results

Review the overall results before you begin to work with the analyzed data.

To edit or update your analysis protocols, see “Analysis protocols” on page 103.

Analyze the sample files

1. In the **Sample Manager**, in the **Show** column, select the checkbox of the sample(s) you want to analyze.
2. Select the checkbox of the processing parameter(s) (**BC**, **PP**, **P**) you want to use.
   For more information on the processing parameters, see “Processing parameters” on page 115.
3. Click (Start Analysis), or select **Analysis** ▸ **Start Analysis**.

Review processing parameters (BC, PP, and P check boxes)

For more information on processing parameters, see “Processing parameters” on page 115.

1. In the **Sample Manager** view, review the **BC** checkbox for the basecalling status of each sample.
   For more information, see “Basecalling status indicators” on page 115.

Color	Indicates
Green	Basecalling analysis was successful.
Blue	(Smart Deep™ Basecaller and KB™ Basecaller only) Basecalling analysis was successful, with some anomalies. Review the analysis report, then examine the data in more detail.
Yellow	(Smart Deep™ Basecaller and KB™ Basecaller only) Basecalling analysis was performed, but the data is low-quality. The software saves the results to the sample file, but the called sequence is replaced with a placeholder sequence of five N’s instead of basecalled data. Review the analysis report, then examine the examine the data in more detail.
Red	Basecalling analysis failed.

Note: Review the base spacing, peak 1 location, and the start and stop points. A red value in the **Spacing** column means the spacing could not be calculated and the default value was used for analysis.

2. Review the **PP** checkbox for the post processing status of each sample.
   PP enables the calculation of the clear range.

Color	Indicates
Green	Post processing was successful.
Red	Post processing failed.
No color	• Post processing has not been started since the sample was added to the **Sample Manager**. • Post processing was completed previously and is still in the **Sample Manager**.

3. Review the **P** checkbox for the printing status of each sample.

Color	Indicates
Green	Printing was successful.
Red	Printing failed. Check printer connections and see the printer user guide if necessary.
No color	Printing has not been started since the sample was added to the **Sample Manager**.

Edit processing parameters

Select an option to edit the processing parameters:

- In the **Sample Manager**, select or deselect the **BC**, **PP** or **P** check boxes as appropriate.

Parameter	
BC	Select to enable basecalling.
PP	Select the checkbox to enable post processing. Post processing only occurs after basecalling and enables the calculation of the clear range. **Note:** Unanalyzed data cannot be post processed until the data is basecalled.
P	Select to enable printing. If BC and/or PP parameters are selected, printing is done after basecalling and post processing are complete.

- In the menu bar, select **Analysis** ▸ **Analysis Defaults**.
  In the **Add Samples Settings** pane, select or deselect the **BC**, **PP**, or **P** checkboxes as appropriate.

Edit sequence file formats

Select an option to edit the sequence file formats:

- Edit in the **Analysis Defaults** dialog box (see “Create, edit, or select new Analysis Defaults” on page 35).
- Edit the file format in the sample analysis protocol (see “Edit a sample analysis protocol” on page 41).
- Edit in the **File Format** tab in the **Options** dialog box (see “Edit file format using Options” on page 47).

Edit file format using Options

Option	Description
Use the settings in the sample's analysis protocol	When selected, the sequence file formats of the analysis protocol or analysis defaults are used.
Override the sample's analysis protocol	When selected, the new sequence file formats selections *override the analysis protocol and analysis default settings*. When selected, software creates the following during analysis: • A .seq file for printing the sequence as text file or for using the file in other software. – ABI format is used with Applied Biosystems™ software. – FASTA format is used with other software • Standard chromatogram format (.scf) files • Phred (.phd.1) files

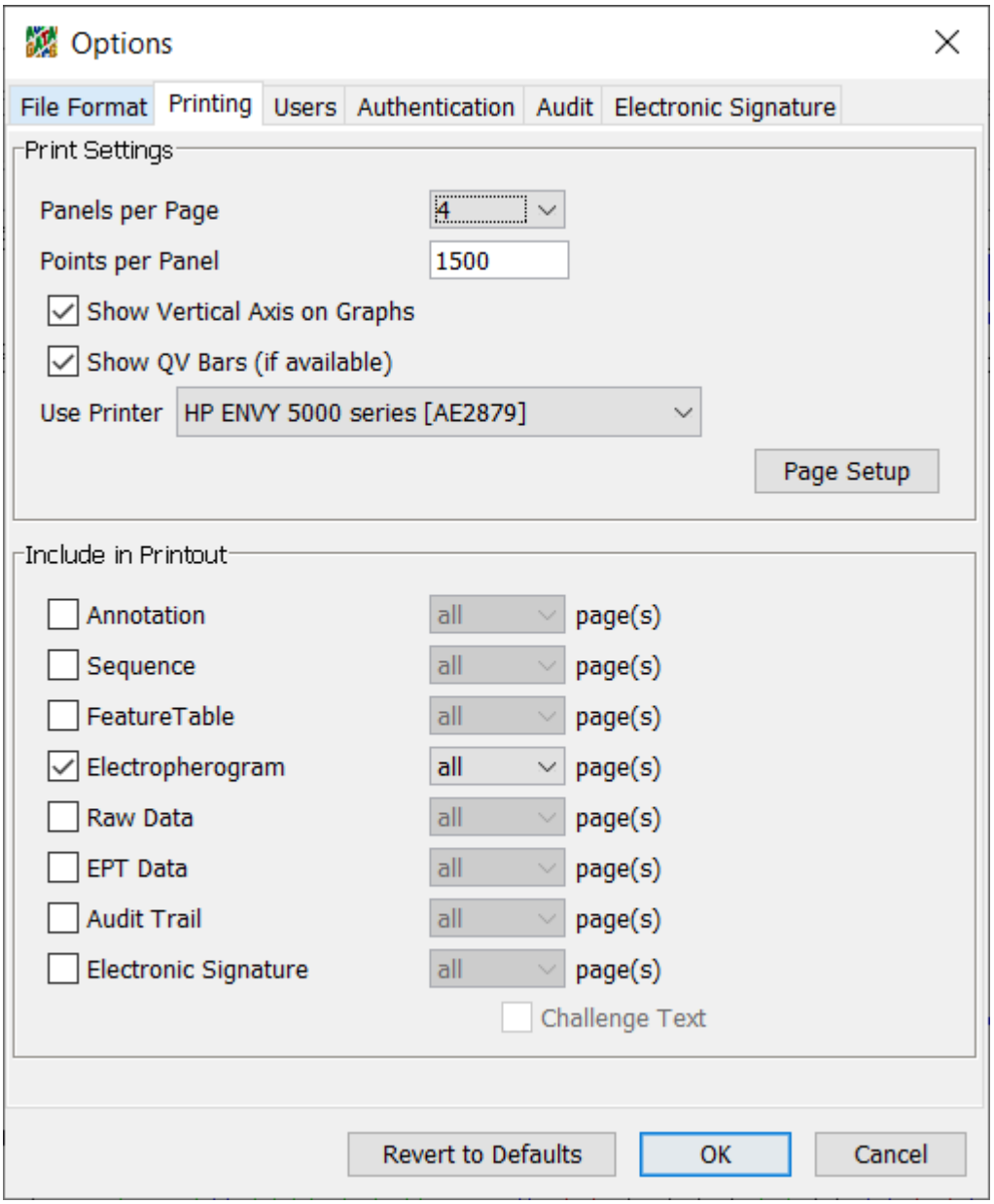
1. Select **Tools** ▸ **Options**, then select the **File** tab.
2. Select to override the sample's analysis protocol, then select the desired file formats.
3. Click **OK**.
4. Add samples to the **Sample Manager**, reanalyze the sample to create the new file formats, then save the samples.

Automate printing of the sample file

To automate the print process:

1. Add the samples you want to print to the **Sample Manager**.
2. Select **Tools** ▸ **Options**.
   The **Options** dialog box is displayed.

3. Select the **Printing** tab.



In the **Printing** tab:

a. Select a value in the **Panels Per Page** dropdown list. The range is 1 to 15, and the default is 4.

Note: As a general rule, if the number of panels per page is set to a value greater than approximately 8, printing of quality values may be suppressed. This maximum number of panels per page that can be printed with quality values will vary, depending upon paper size, margins and page orientation.

---

b. Select a value in the **Points Per Panel** value box. The range is 100 to 12000 and the default is 1500 (about 120 bases).

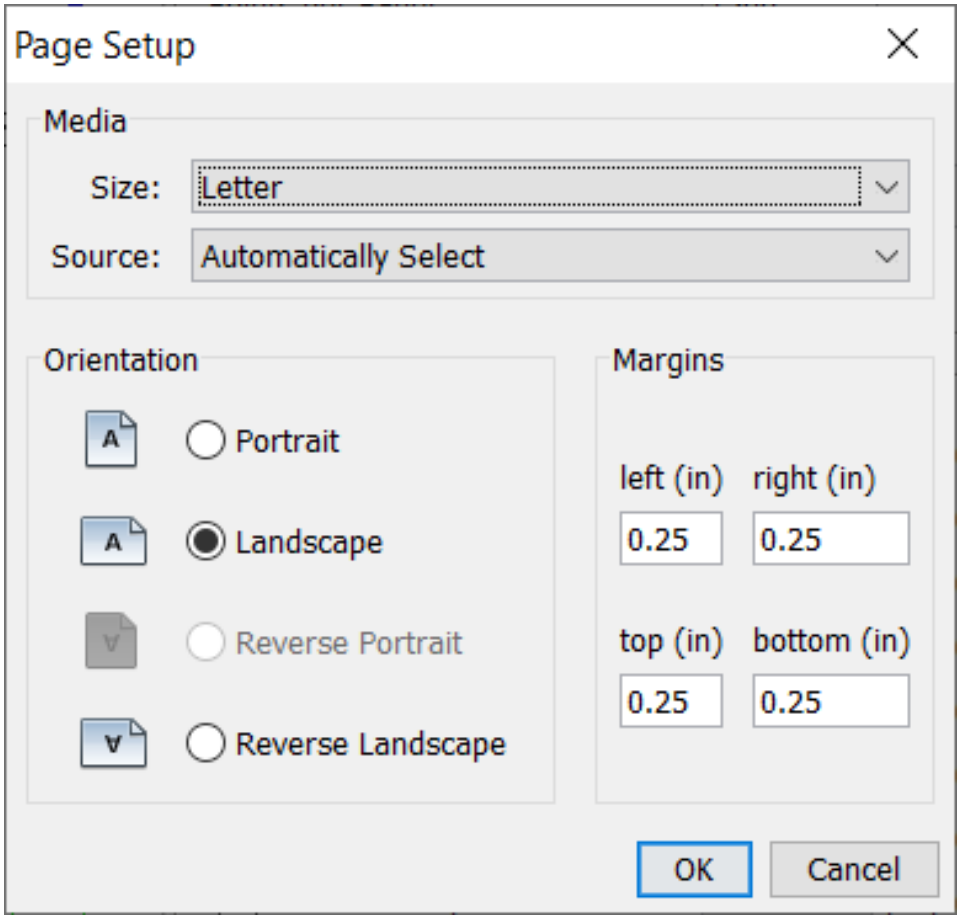
c. Select or deselect **Show Vertical Axis on Graphs** to show or hide the vertical axis when printing electropherogram, raw, and EPT data.

d. Select or deselect **Show QV Bars (if available)** to show or hide QV bars when printing sequence and electropherogram data.

e. Select a printer from the **Use Printer** dropdown list.

f. In the **Include in Printout** section, select checkboxes of the view(s) you want to print.
Use the dropdown list to select and the number of pages for each view.

4. Click **Page Setup**.
The **Page Setup** dialog box is displayed.



a. In the **Media** section, select paper and source from the drop-down lists.

b. In **Orientation** section, select the paper orientation for your print outs.

c. In **Margins** section, change the paper margins, if necessary for the printer you are using.

d. Click **OK**.

5. Click **OK**.

   Note: After the printing settings are configured, it is not necessary to repeat the process, unless you want to change the automated print settings.

6. For all samples you want to print, select the **P** (Printing) parameter checkbox in the **Sample Manager**.

   Note: The **P** check box prints only the options that you have selected to print in the **Printing** tab of **Options** dialog box.

   Note: Deselect the **BC** and **PP** check boxes if you do not want reanalysis of the data to occur before printing.

7. Click  .
   Printing will begin.

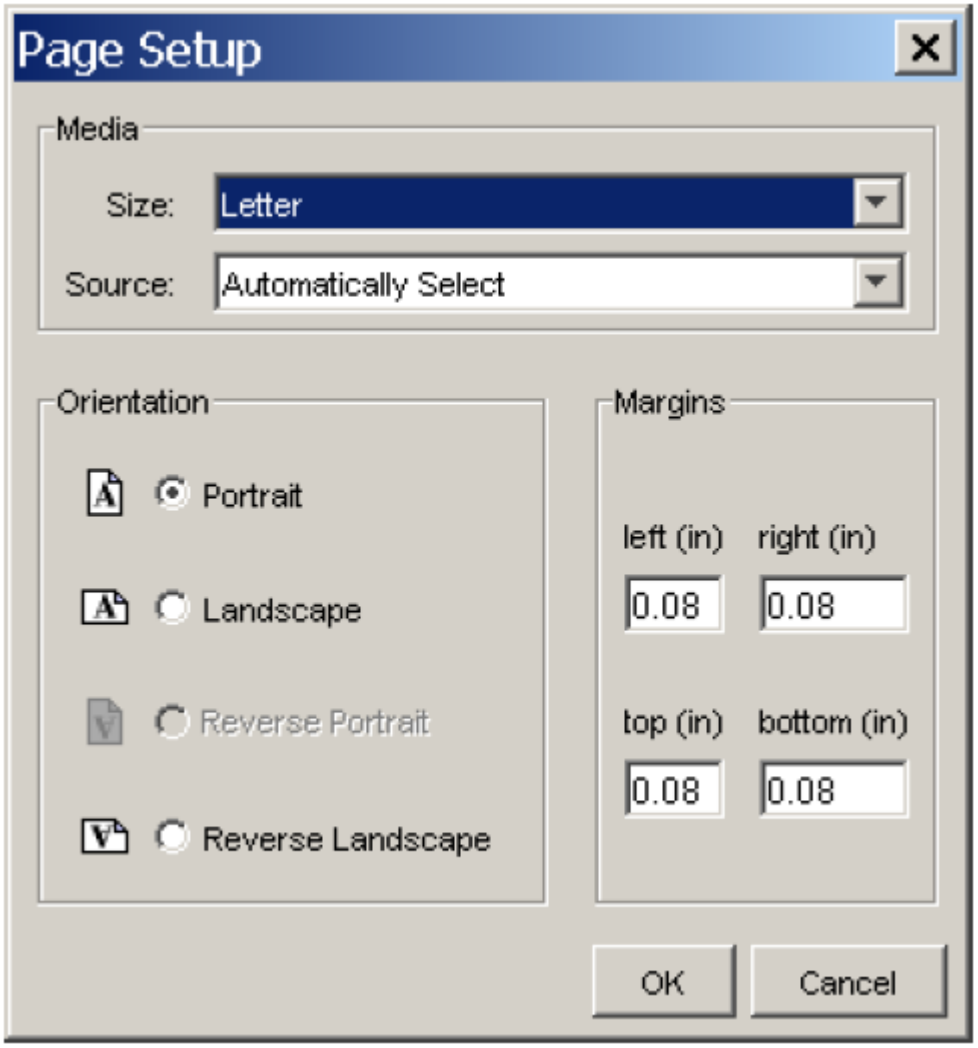
After the printing process is completed, a color code status is displayed in the **P** column. Green indicates success and red indicates failure.

By default, the selected checkbox in the **P** column is unchecked.

Manually print the sample file

To print the data manually:

1. Add the sample(s) you want to print to the **Sample Manager** window.
2. In the **Row** column, select one or more samples to print.
   - Click a row number to select one sample.
   - Shift+drag to select continuous sample rows.
   - Ctrl+click to select discontinuous sample rows.
3. Select **File** ▸ **Page Setup**.
   The **Page Setup** dialog box is displayed.



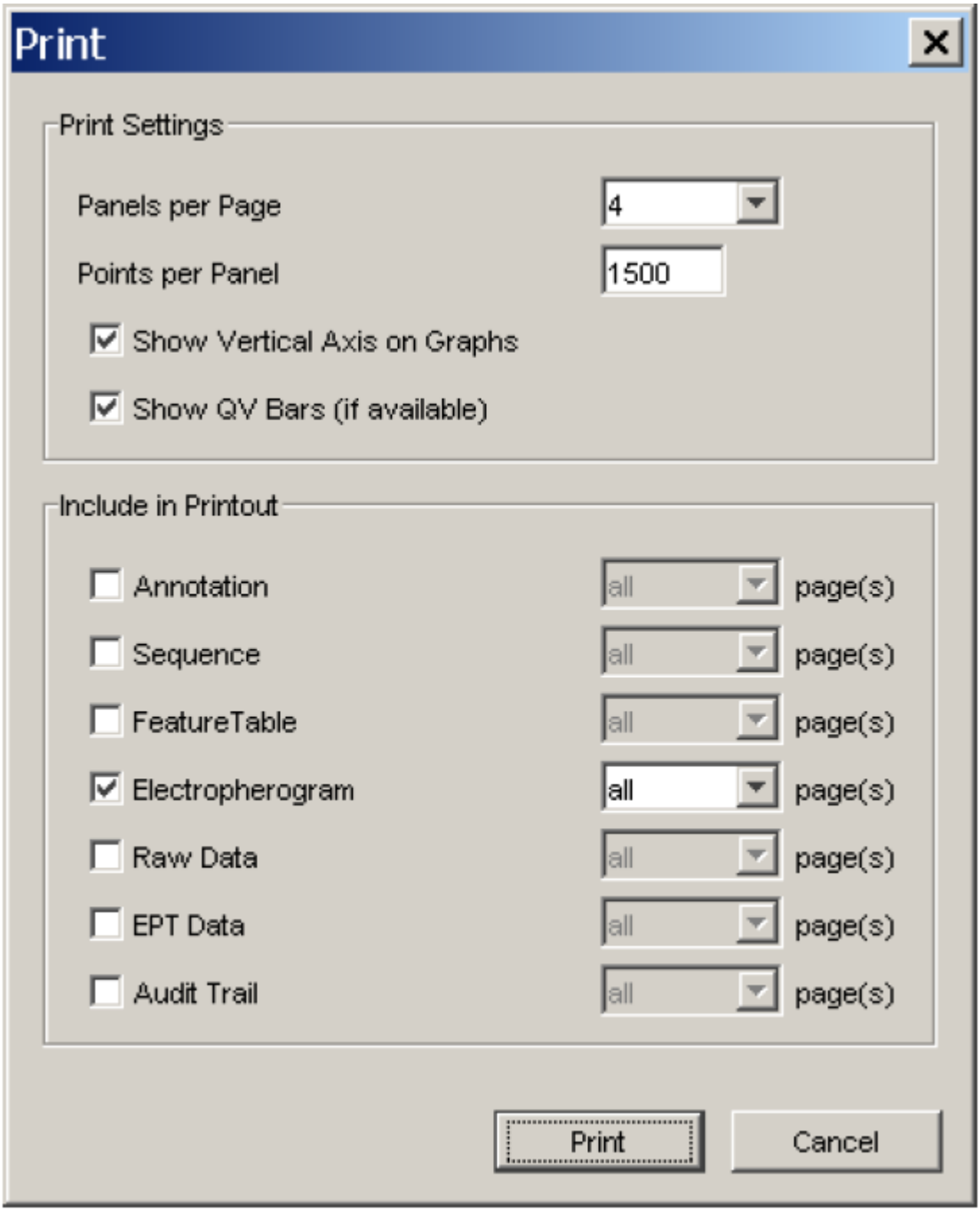
a. In the **Media** section, select paper size and source from the dropdown lists.

b. In **Orientation** section, select the paper orientation for your printouts.

c. In **Margins** section, change the paper margins, if necessary for the printer you are using.

d. Click **OK**.

4. Select **File** ▸ **Print**, or click .
   The **Print** dialog box is displayed.



5. In the **Print** dialog box:
   a. Select a value in the **Panels Per Page** dropdown list. The range is 1 to 15, and the default is 4.

   Note: As a general rule, if the number of panels per page is set to a value greater than approximately 8, printing of quality values may be suppressed. This maximum number of panels per page that can be printed with quality values will vary, depending upon paper size, margins and page orientation.

b. Select a value in the **Points Per Panel** value box. The range is 100 to 12000 and the default is 1500 (about 120 bases).

c. Select or deselect **Show Vertical Axis on Graphs** to show or hide the vertical axis when printing electropherogram, raw, and EPT data.

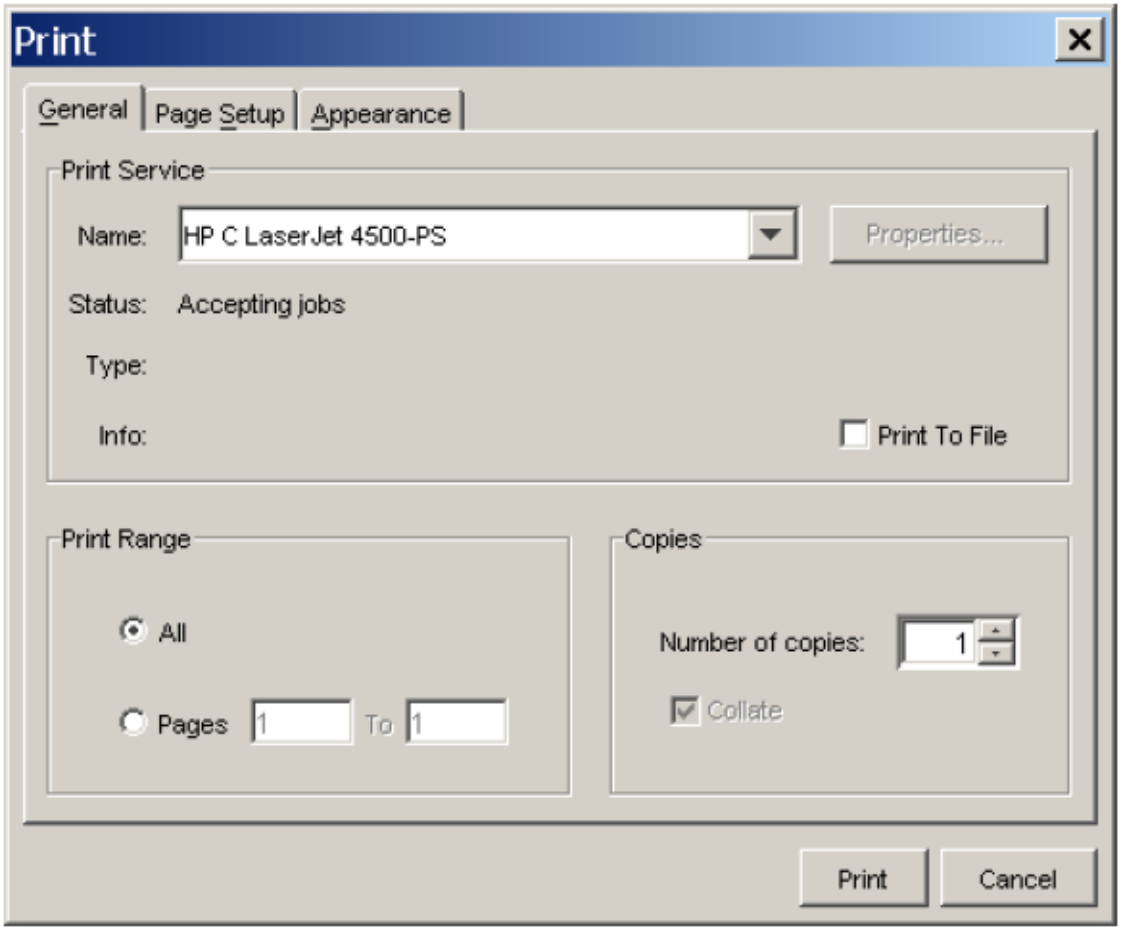
d. Select or deselect **Show QV Bars (if available)** to show or hide QV bars when printing sequence and electropherogram data.

e. In the **Include in Printout** section, select checkboxes of the view(s) you want to print.

   Use the dropdown list to select and the number of pages for each view.

f. Click **Print**.

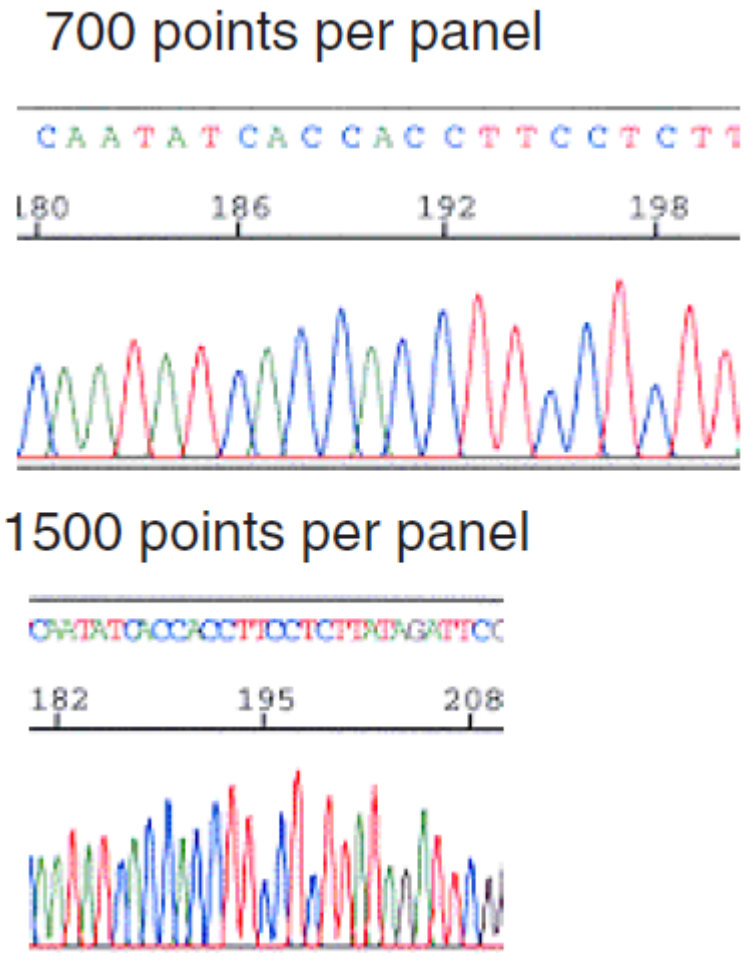
   A second **Print** dialog box opens.



6. Verify and/or change the settings as needed, then click **Print**. The dialog boxes will close, then printing will begin.

Edit printing options

Select the printing options, as described in the table below:

Option	Function
Print Settings pane	
Panels Per Page dropdown list	Select the number of panels to print on each page of graphical (Electropherogram, Raw Data, EPT) views. The default is four panels of 1500 points. The range is 1 to 15 panels.
Points Per Panel field	Enter the number of data points to display in each panel. The default is 1500 data points (~120 bases) per panel. If you decrease the number of data points per panel, the peaks are broader, with fewer bases per panel. The range is 100 to 12000 points. 
Show Vertical Axis on Graphs checkbox	Select/deselect to show /hide the vertical axis in graphs.
Show QV Bars (if available) checkbox	Select/deselect to show/hide the QV bars in the electropherogram and sequence views.
Use Printer dropdown list	Select a printer.

(continued)

Option	Function
Include in Printout	
Content checkboxes	Select the content to be printed and the number of pages of that view. The default is **Electropherogram** and all pages. The page range is 1-5.

Review and edit base spacing, peak 1 location, and start and stop points

When you analyze the sample files, the software calculates the base spacing, peak 1 location, and start and stop points (see “Base spacing, peak 1 location, and start and stop points” on page 121).

1. In the **Sample Manager** view, review the results in the **Spacing Values** column, then recalculate if needed.
   a. Review the results.

Text color	Indicates
Red	The spacing value could not be calculated and the default value in the basecaller file was used.
Black	The spacing value was successfully calculated.

   b. (*Optional*) To recalculate the spacing value for a sample, set the value to 0, then click ▶ to reanalyze.

2. Review the results in the **Peak 1 Location** column, then recalculate or change the value if needed.

Note: To start the analysis further along than the actual location of the first base peak (peak 1 location), change the value for the start point, not for the peak 1 location. Changing the peak 1 location value affects the way the DyeSet/Primer file is applied to correct for mobility shifts.

Option	
Recalculate peak 1 location	Reset the value to 0, then click ▶ to reanalyze.
Change peak 1 location	If the peak 1 location value is wrong due to low signal or other aberrations, enter a new value, then click ▶ to reanalyze. **Note:** If the peak 1 location value is wrong, the data can show bad spacing or strange mobility shifts. For more information about determining the peak 1 location data point, see “Determine data point values for spacing, peak 1, start and stop parameters” on page 56.

Note: If you recalculate or change the peak 1 location, the start point, and stop point are also recalculated.

3. Review the results in the **Start** column, then recalculate or change the value if needed.

Option	
Recalculate start point	Reset the **Peak 1 Location** column value to 0, then click ▶ to reanalyze.
Change start point	Enter a new value, then click ▶ to reanalyze. For more information about determining the start point, see “Determine data point values for spacing, peak 1, start and stop parameters” on page 56.

4. Review the results in the **Stop** column, then recalculate or change the value if needed.

IMPORTANT! For optimal analysis of PCR products that are shorter than the run, it is essential to reanalyze the sample with a stop point that encompasses only the true data peaks.

Option	
Recalculate stop point	Reset the **Peak 1 Location** column value to 0, then click ▶ to reanalyze.
Change stop point	Enter a new value, then click ▶ to reanalyze. For more information about determining the stop point , see “Determine data point values for spacing, peak 1, start and stop parameters” on page 56.

Determine data point values for spacing, peak 1, start and stop parameters

Specific data point values are used to set the spacing, peak 1, start and stop parameters. In the **Electropherogram**, **Raw**, or **EPT** view, use the crosshair feature to determine data point values.

1. In the **Sample Manager** or **Sample Navigator** view, select sample data to display (see “Display the sample data” on page 70).

2. Select the **Electropherogram**, **Raw**, or **EPT** tab.

3. In the plot, click near the point of interest to display the crosshair locator lines.

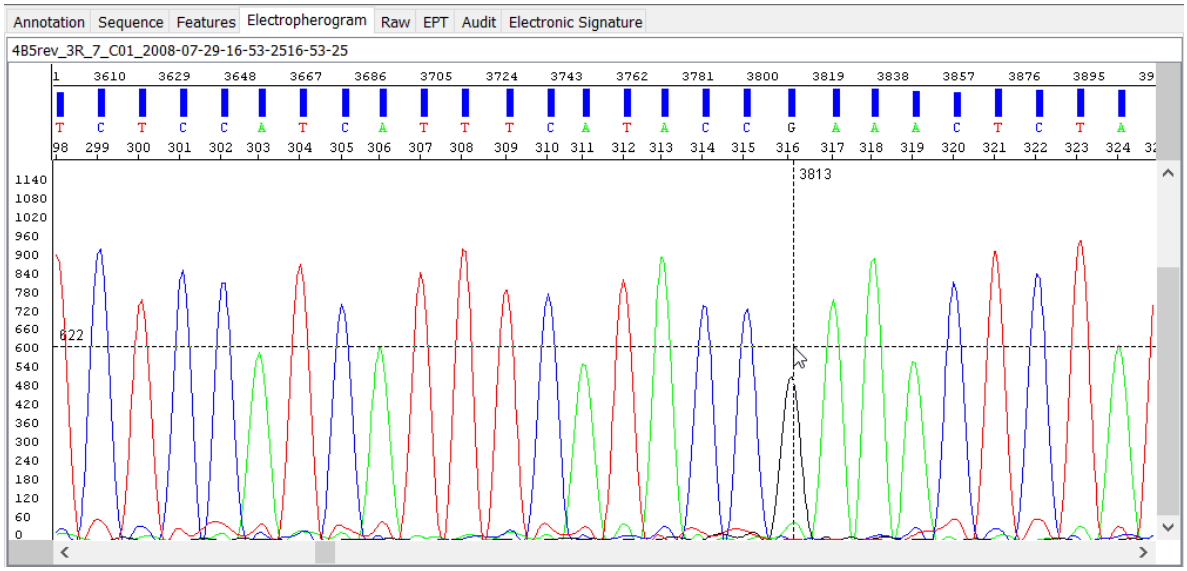
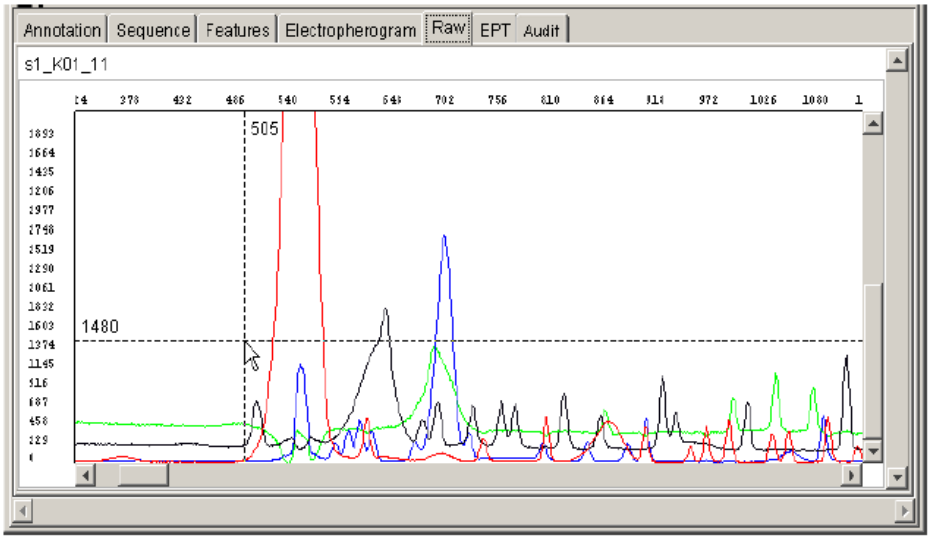
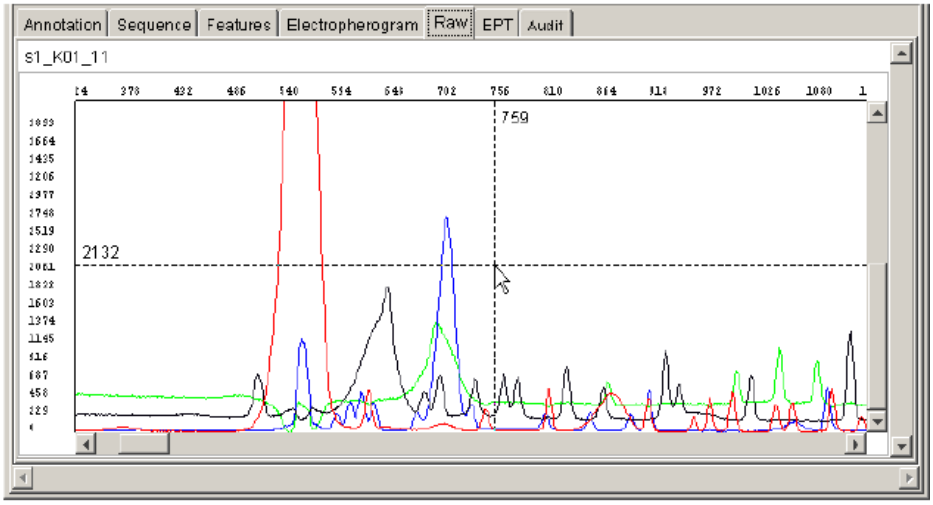


Figure 5 Electropherogram plot displaying crosshair locator lines

4. Drag the cursor across the plot until the locator lines intersect the point of interest. Use the following guidelines to identify correct data point for the analysis.

Value to determine	Procedure
Spacing	To manually determine the spacing value, use the cross-hair cursor to determine the scan numbers at the tops of two adjacent peaks, then subtract the smaller number from the larger number. **Note:** Do not use the first 100 bases or the last 200 bases.
Peak 1 Location	Place the cross-hair cursor at the beginning of the first base peak, then note the scan number. For example, in this raw data the cursor is positioned at the Peak 1 location. The scan number is 505 and represents where the data starts. 

(continued)

Value to determine	Procedure
Start Point	Place the cross-hair cursor at the beginning of the first basecalling peak to determine the start point scan number. **Note:** The start point value must be greater than or equal to the peak 1 location. For example, in this raw data plot the cursor is positioned at the Start Point. The scan number is 759 and represents where basecalling analysis starts. 
Stop Point	Place the cross-hair cursor at the end of the last peak to determine the end point scan number. **Note:** For optimal analysis of PCR products that are shorter than the run, it is essential to reanalyze the sample with a Stop Point that encompasses only the true data peaks.

Review the analysis parameters used in processing

In the **Sample Manager** view, review the selections in the following columns:

- **BaseCaller**
- **DyeSet/Primer**
- **Matrix File** (310 Genetic Analyzer only)

Note: If the name of the selected file appears as bold, italic text, then the software could not find the file in the expected location.

For analysis to continue, you must edit the analysis parameters. Select files that are present in the corresponding software folder:

File type	File location
BaseCaller file	`C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Params`
DyeSet/Primer file	`C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Mobility`
Matrix file	`C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Matrix`

To edit the analysis parameters, see “Edit Basecaller, DyeSet/Primer, and Matrix File selections (analysis parameters)” on page 61.

Edit Basecaller, DyeSet/Primer, and Matrix File selections (analysis parameters)

This procedure describes how to edit the analysis parameters in the **Sample Manager** window.

To edit analysis parameters in the **Analysis Protocol Manager** window, see “Create or edit an analysis protocol” on page 39.

To edit the analysis parameter selections for an individual sample, see “Edit a sample analysis protocol” on page 41.

Note: If you change the analysis parameter file selection, then reanalyze the samples, the basecaller recalculates the peak 1 location, start point, stop point, and spacing.

---

1. In the **Sample Manager**, select one or more samples.
2. In the **Basecaller** column, select a basecalling algorithm in the dropdown list (see “Select the correct basecaller” on page 118).
3. In the **DyeSet/Primer** column, select a mobility file in the dropdown list (see “Select a DyeSet/Primer (mobility) file” on page 119).

   Note: If the mobility file name is displayed in bold italicized font, then the file is not present in the **Mobility** folder and the sample file cannot be processed with that file. To copy DyeSet/Primer files into the **Mobility** folder, see “Copy and paste mobility files for 310 Genetic Analyzer” on page 134.

   ---

4. (310 Genetic Analyzer only) In the **Matrix File** column, select a file in the dropdown list.
5. Click  to reanalyze the samples.
   The basecaller recalculates the peak 1 location, start point, stop point, and spacing.
6. Save the sample files (see “Save sample files” on page 61).

Save sample files

The data in a sample file is not automatically saved after editing, basecalling and/or post processing.

Note: If a .seq file was created when the sample file was analyzed, then both the sample file and .seq file are updated when you save the sample file.

- To save selected sample files, select one or more sample files in the **Sample Manager**, then click  , or select **File ▸ Save Sample(s)**.
- To save all sample files, click  or select **File ▸ Save All Sample(s)**.

If you try to save changes to sample files that are read-only, the samples files will save as new files. You will be prompted to select a location to save the new sample files.

Use the electronic signature function

The electronic signature function must be enabled by an administrator (see "Enable and set up electronic signature function" on page 26).

Electronically sign a sample

1. In the **Sample Manager**, click a row header to select a sample then select **Tools ▸ Electronic Signature ▸ Sign**.

   If an action is set up to require an electronic signature, the **Electronic Signature Verification** dialog box will automatically open when the action is performed.

2. In the **Electronic Signature Verification** dialog box, choose an option in the **Action** dropdown list.

3. (*Optional*) In the **Comment** field, enter any comments.

4. Enter your user ID and password.

   Note: If you do not have electronic signature privileges, another user with signature privileges may sign with his or her user ID and password.

5. Click **OK**.
   A dialog box indicates that the electronic signature was successfully verified, and the electronic signature is logged in the **Electronic Signature** tab in the **Sample Manager**.

To view the electronic signatures for a sample, see "View electronic signatures for a sample" on page 62.

View electronic signatures for a sample

1. In the **Sample Manager**, in the **Show** column, select the checkbox for the sample you want to display.

2. In the **Sample Manager**, in the **Electronic Signature** tab, review the electronic signature records.

Verify an electronically signed sample

If a sample has been electronically signed, then you can verify whether or not the sample has had changes since the signature.

1. In the **Sample Manager**, click the row number of the sample.

2. Select **Tools ▸ Electronic Signature ▸ Verify**.
   The **Electronic Signature** dialog box is displayed with a message indicating if the sample is **VALID** or **INVALID**.
   - **VALID**—The sample has not changed since the electronic signature.
   - **INVALID**—The sample has changed since the electronic signature.

8 Review the Analysis Report

About the analysis report

An analysis report shows the success and/or failure of data analysis.

If the data has been analyzed, the report displays a summary of the individual sample information and errors, as well as QVs and length of reads (LORs). LORs are the usable ranges of high-quality or high-accuracy bases, as determined by quality value.

If the data is unanalyzed, the report displays status information. The report is used to help troubleshoot and provide easy assessment of data quality.

The report is used to help troubleshoot and provide easy assessment of data quality. It can be exported as a tab-delimited file and opened in Microsoft™ Excel™ software for trend analysis.

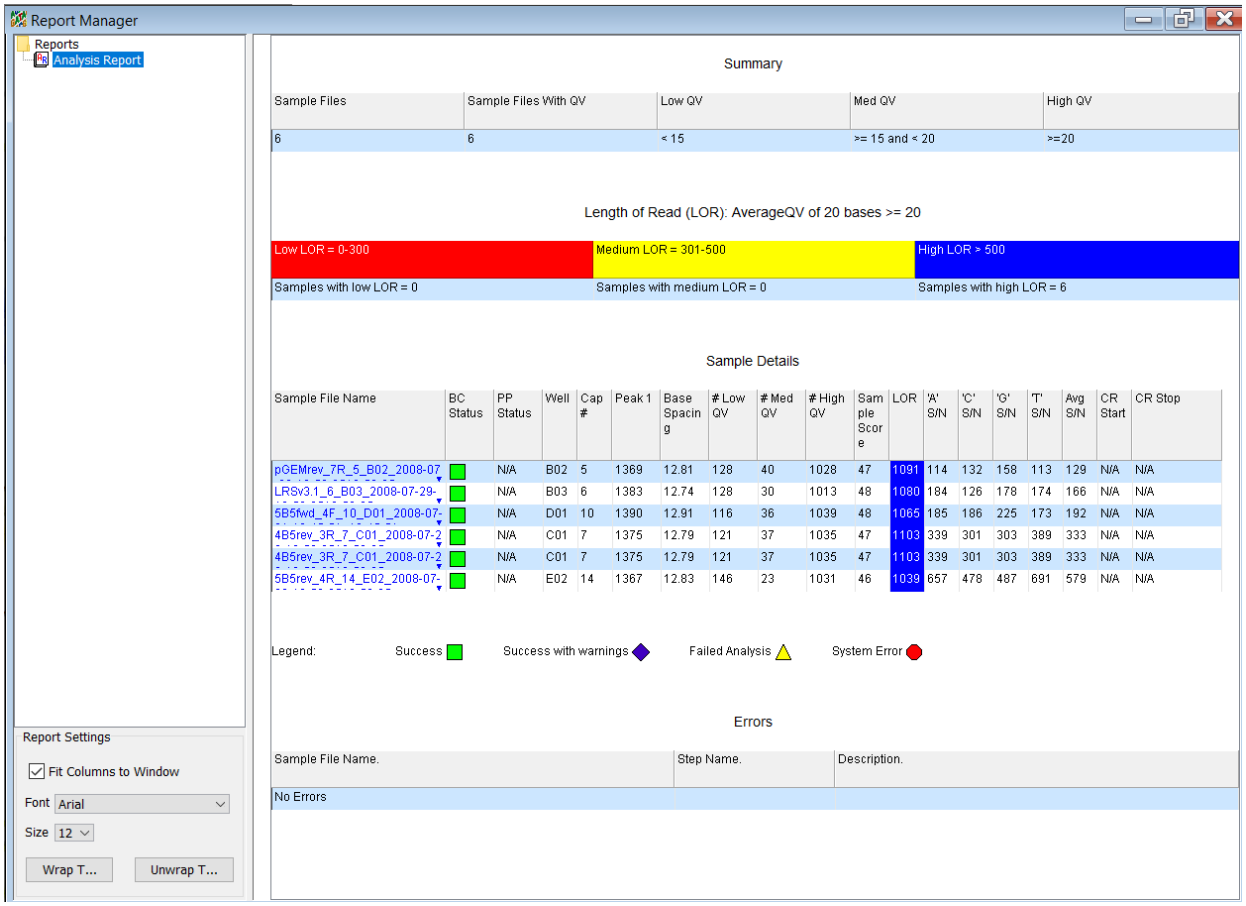


Figure 6 Analysis report example

Table 1 Parts of the analysis report

Table	Description
Summary	
Sample Files	The total number of samples in the report. The number depends on the number of samples in the Sample Manager.
Sample Files With QV	(Smart Deep™ Basecaller and KB™ Basecaller only) The total number of analyzed samples that contain Quality Values (QVs).
Low QV, **Medium QV**, or **High QV**	(Smart Deep™ Basecaller and KB™ Basecaller only) The low, medium, and high ranges for the QVs. **Note:** To edit the QV ranges, see “Edit the display settings” on page 87.
Length of Read (LOR) (Smart Deep™ Basecaller and KB™ Basecaller only)	
Low LOR, **Medium LOR**, or **High LOR**	Displays the LOR ranges, and the number of samples within the ranges. The LOR is the usable range of high-quality or high-accuracy bases, as determined by QVs. **Note:** To edit the LOR ranges, see “Edit the display settings” on page 87.

Table 1 Parts of the analysis report *(continued)*

Table	Description
Sample Details (partial output and failed samples are hyperlinked to the specific sample in the Errors table)	
Sample File Name	The name of the sample. The files are hyperlinked to the sample in the **Sample Manager**.
BC Status	Basecalling status • Green—Passed • Blue—Passed, with some anomalies • Yellow—Poor quality data, partial output (for data analyzed with Smart Deep™ Basecaller and KB™ Basecaller only) • Red—Failed • N/A—No analysis or not available Samples with red, yellow, or blue icons are hyperlinked to the specific sample in the Error table.
PP Status	Pass, fail, and not analyzed indicator for post processing • Green icon = passed • Red icon = failed • N/A = no post processing or not available in this session A sample with a red icon is hyperlinked to the specific sample in the Error table.
Well	The plate well number that contained the sample.
Cap #	The number of the capillary that the sample was run on.
Peak 1	The scan number that represents the first data point in the file that is from the sample, not including primer peaks.
Base Spacing	The value represents the calculated base spacing for the sample. Base spacing is defined by the number of scan points from the crest of one peak to the crest of the next peak.
# Low QV, **# Medium QV**, or **# High QV**	The number of bases where the QVs are in the low, medium, or high range.
Sample Score	The average quality value of the bases in the clear range sequence for that sample.
LOR	The usable range of high-quality or high-accuracy bases, as determined by quality values.
'A' S/N, **'G' S/N**, **'C' S/N**, or **'T' S/N**	The value represents the average signal/average noise of all the 'A', 'G', 'C,' or 'T' base in a sample.
Avg S/N	The value represents the average signal/noise value of all the bases in a sample.
CR Start	The start of the clear range.
CR Stop	The end of the clear range.

Table 1 Parts of the analysis report *(continued)*

Table	Description
Errors	
Sample file errors	Displays the errors that occurred during analysis and post processing; each sample has a hyperlink to the specific sample in the **Sample Manager**.

Error due to mismatched basecaller and DyeSet/Primer files

- If you select a KB™ DyeSet/Primer file and an ABI™ Basecaller for analysis, the basecalling is successful (green BC box) but the analyzed data is not usable.
- If you select a DT DyeSet/Primer file and an KB™ Basecaller for analysis, the basecalling fails. The following error message displays in the Error table.

File Name	Step Name	Description
_01	BaseCalling	Basecalling Failed: WARNING: F_ERROR -10023: ABIF access failure - tag not found: CNAM1 WARNING: F_ERROR 50305: Invalid mobility file F_ERROR -10023: ABIF access failure - tag not found: CNAM1 F_ERROR 50305: Invalid mobility file Analysis error: Error number = -15231
_01	PostProcessing	Incomplete Results presented from previous stage

Review the Analysis Report

We recommend that after analysis, you review the report before examining each sample file.

For more information about the content of the analysis report, see “About the analysis report” on page 63.

IMPORTANT! Any action that renders the report to be invalid, such as adding more samples to the **Sample Manager**, forces the analysis report to close.

---

1. To generate the analysis report, select one of the following options:
   - In the toolbar, click  (analysis report).
   - In the menu bar, select **Analysis** ▸ **Analysis Report**.

   The analysis report opens in the **Report Manager** window.

   For more information about the contents of the analysis report, see “About the analysis report” on page 63.

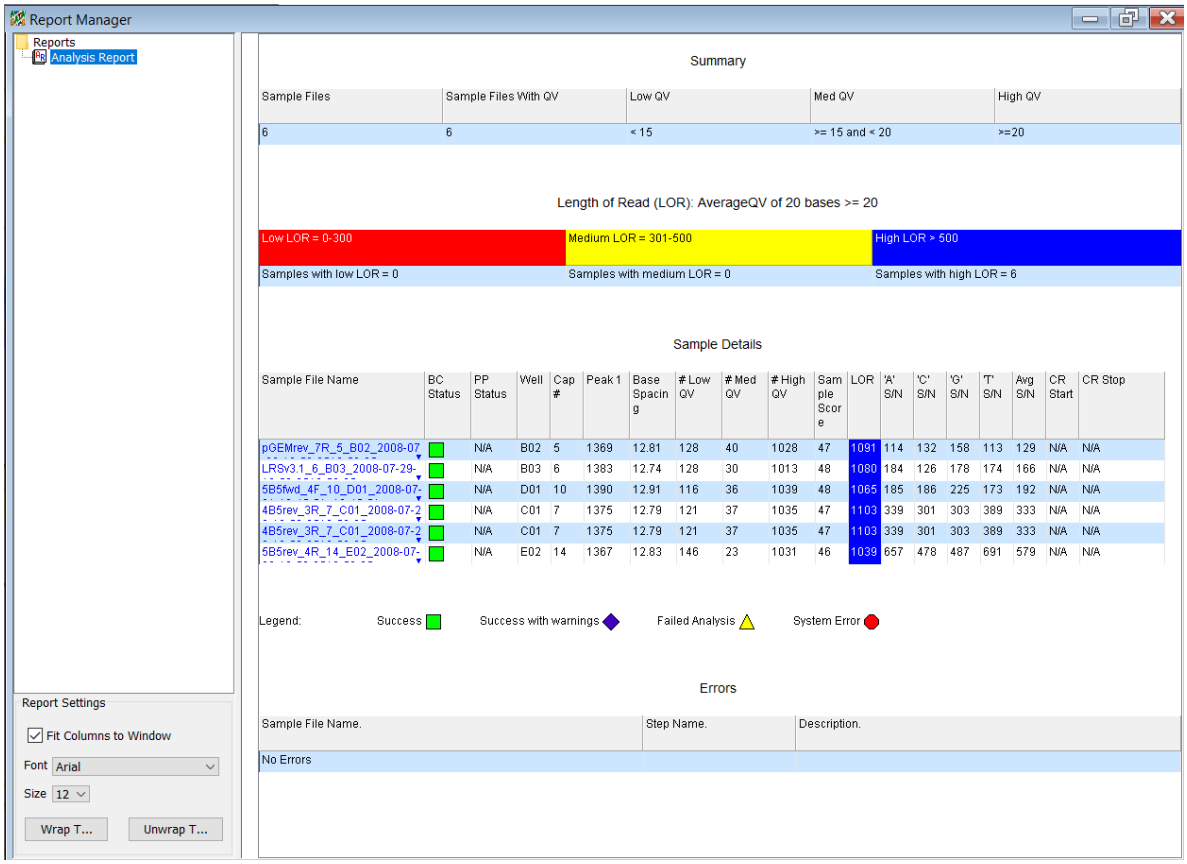


Figure 7 Analysis report example

2. (*Optional*) To edit the analysis report display settings, select the appropriate options in the **Report Settings** pane.

3. To view a sample in the **Sample Manager**, click the sample hyperlink in the analysis report.

Note: Hyperlinks move only from the analysis report to the **Sample Manager**, not vice versa. Click  to return to the analysis report.

Customize the data displayed in the Analysis Report

1. Right-click in any column heading of a table.

   A list of the column headings in the table is displayed.

2. To hide a column, deselect the column heading.

3. Repeat steps 1 and 2 to deselect additional headings.

4. To redisplay a column, right-click in any column heading, then select the column heading.

5. To sort the data A to Z or Z to A in a Sample Details or Errors table column, double-click in the column heading. Double-click again to sort in the opposite direction.

6. To change the order of the columns in any table, Ctrl+drag the column heading to a new location and release.

Print the analysis report

1. Open the analysis report. See step 1.

2. Select **File** ▸ **Print**, or click (print).
   The **Print** dialog box is displayed.

3. Select the appropriate print options, then click **Print**.

Export the analysis report

The analysis report can be exported in the following formats:

- Tab-delimited text
- HTML
- PDF
- XML

The tab-delimited file format can be opened in Microsoft™ Excel™ software (or any application that reads this type of format) for trend analysis.

1. Open the analysis report.

2. Select **File** ▸ **Export Report**.

3. In the **Export Analysis Report** dialog box, navigate to a folder location, enter a file name, then click **Export**.

9 Review sample data

Display the sample data

The software does not automatically show the sample file data. Select samples to display sample data in the **Sample Manager** or **Sample Navigator** view.

1. Click (toggle) to select the **Sample Manager** (default) or **Sample Navigator** view.
2. Select one or more samples to display the data:

Sample selection	Directions
A single sample	Double-click the sample file name or select the corresponding **Show** check box. **Note:** Double-clicking a sample file name deselects the **Show** check box for all other samples.
Multiple continuous samples	Shift+click or click-drag the sample row numbers to select the sample files, then click **Show**.
Multiple discontinuous samples	Ctrl+click the sample row numbers to select the sample files, then click **Show**.
All samples	Select the empty box above row number 1 or Shift+drag the sample row numbers to select all samples, then click **Show**.

The data display pane opens.

For more information about the data displays, see Chapter 13, “Sample data displays”.

Review the EPT plot

Review the EPT (ElectroPherogram Telemetry) plot to identify instrument performance issues that can affect sequencing quality.

Review the raw data

The raw data plot is used to determine the start and stop points for analysis, review peak and spacing features, and identify problems with the data that can affect the quality of the results.

For more information about the raw data plot, see “About the Raw tab” on page 131.

1. In the **Sample Manager** or **Sample Navigator** view, select sample data to display (see “Display the sample data” on page 70).

2. Select the **Raw** tab, then review the following data:
   - Scan numbers used by the software to start and stop basecalling.
   - Relative true peak intensities and peak resolution before the smoothing applied by the software.
   - Estimate base spacing by measuring the scan points that define two adjacent peaks.
   - Problems or noise in the baseline (for example, electronic spikes in the data or unusual baseline levels) that could result in poor basecalling or could indicate instrument problems.

Action	Procedure
Zoom in or out to see the data at different magnifications	See “Zoom in or out” on page 89.
Change the colors of the trace lines that represent the fluorescent dyes, or hide one or more trace lines	See “Edit the display settings” on page 87.
Display cross hairs and the coordinates for the current cursor location	Click in the data area of the window.
Print the window contents	See “Manually print the sample file” on page 50.

Reviewing the analyzed data

Table 2 Steps to review the analyzed data

Step	Description	
Review the spacing values.	Review the spacing values in the Sample Manager window. If a value is displayed in red text, then the basecaller failed while calculating the value, and used a default spacing value.	
Review the files used in processing.	Review the files specified for use during processing. If the name of a file appears as bold, black, italic text, then the software could not find the file in the expected location.	
	For analysis to proceed, you must specify a Basecaller that is present in the same folder as the Sequencing Analysis software, and a DyeSet/Primer file that is present in the Mobility folder.	
	To specify a …	**Use the path …**

Table 2 Steps to review the analyzed data *(continued)*

<table>
<tr><th>Step</th><th colspan="2">Description</th></tr>
<tr><td rowspan="3">Review the files used in processing.</td><td>Basecaller</td><td>C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Params</td></tr>
<tr><td>DyeSet/Primer file</td><td>C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Mobility</td></tr>
<tr><td>Matrix™ file</td><td>C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Matrix</td></tr>
<tr><td rowspan="8">Search for low, medium, and high QVs in the electropherogram (for data analyzed with the KB™ basecaller).</td><td colspan="2">You can search for QVs as follows:</td></tr>
<tr><td>To move …</td><td>Press …</td></tr>
<tr><td>right to the next low QV</td><td>F6</td></tr>
<tr><td>left to the next low QV</td><td>Shift+F6</td></tr>
<tr><td>right to the next medium QV</td><td>F7</td></tr>
<tr><td>left to the next medium QV</td><td>Shift+F7</td></tr>
<tr><td>right to the next high QV</td><td>F8</td></tr>
<tr><td>left to the next high QV</td><td>Shift+F8</td></tr>
<tr><td>Scroll through the electropherogram.</td><td colspan="2">Scroll through the length of the data in Electropherogram view.
Look for peaks at discrete locations, with no gaps or overlaps and very little noise.
Scroll towards the end of the window and look for well-resolved peaks.  ① Well-resolved peaks
② Poorly-resolved peaks</td></tr>
</table>

Table 2 Steps to review the analyzed data *(continued)*

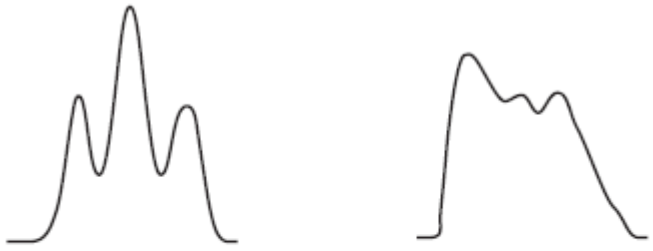
Step	Description	
Check basecalls in the electropherogram.	Look at the basecalls in the Electropherogram view. If two peaks are close together, or the peak is low, or the background noise level is high, then compare each peak to the bases called for that peak. If necessary, edit incorrect basecalls manually.	
Search for Ns in the electropherogram.	You can search for Ns as follows:	
	To …	**Press …**
	move forward	Tab key
	move backward	Shift+Tab
	If you can visually determine the correct basecall at an N location, manually change the N to the correct character.	

Review data quality in the Electropherogram

Review Quality Values (QVs) and peak resolution in the Electropherogram. For more information about QVs, see “About Quality Values” on page 95.

Note: Only data analyzed with Smart Deep™ Basecaller or KB™ Basecaller have Quality Values (QVs).

1. In the **Sample Manager** or **Sample Navigator** view, select sample data to display (see “Display the sample data” on page 70).
2. Select the **Electropherogram** tab.
3. Scroll through the length of the data in the Electropherogram, looking for well-resolved peaks at discrete locations, with no gaps or overlaps and very little noise.



① Well-resolved peaks
② Poorly-resolved peaks

4. (*Smart Deep™ Basecaller or KB™ Basecaller only*) If not already displayed, display the sample Quality Values (see “Display sample Quality Values” on page 75).
   The QVs are displayed as bars above each base in the sample.

5. (*Smart Deep™ Basecaller or KB™ Basecaller only*) Review the bases for low, medium, and high QVs:

Action	Description
Review bases with low QVs	Move right to the next low QV—Press **F6**
	Move left to the next low QV—Press **Shift+F6**
Review bases with medium QVs	Move right to the next medium QV—Press **F7**
	Move left to the next medium QV—Press **Shift+F7**
Review bases with high QVs	Move right to the next high QV—Press **F8**
	Move left to the next high QV—Press **Shift+F8**

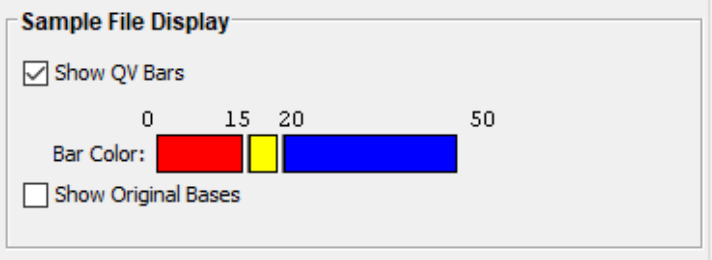
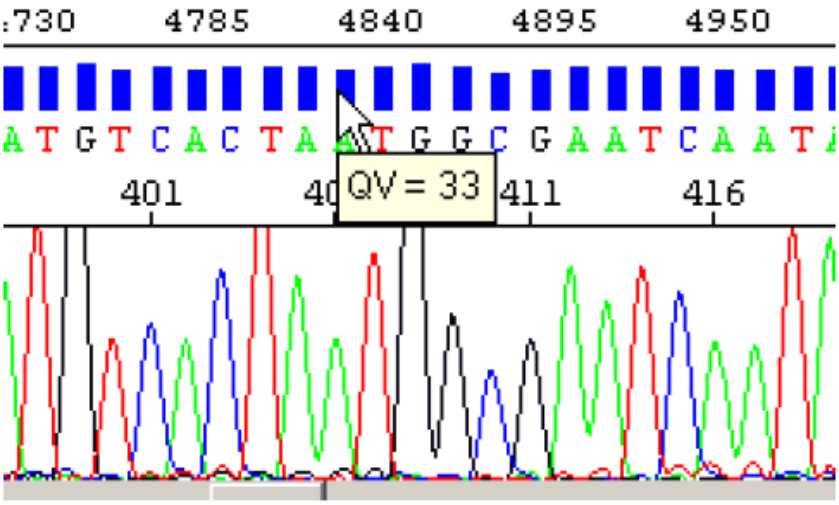
To edit base calls for bases with Ns or low and medium QVs, see “Edit base calls” on page 77.

Display sample Quality Values

Sample Quality Values (QVs) can be displayed in the **Electropherogram** or **Sequence** tabs.

1. In the **Sample Manager** or **Sample Navigator** view, select the sample data to display (see “Display the sample data” on page 70).

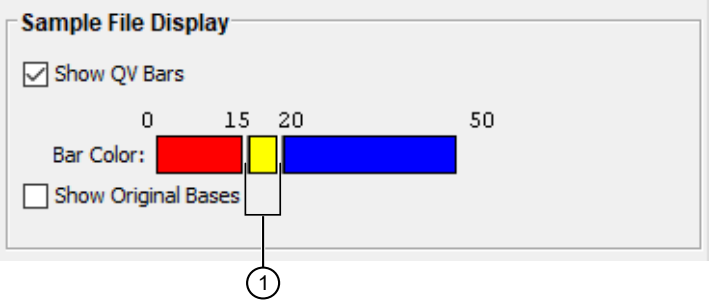
2. Select an option to display the QV bars in the **Electropherogram** and **Sequence** tabs.

Option	Description
Edit **Display Settings**	1. Click or select **Analysis** ▸ **Display Settings**. 2. In the **Sample File Display** section, select **Show QV Bars**, then click **OK**. 
Select **Show Quality Values**	1. Click or select **View** ▸ **Show Quality Values**. 2. To display the numerical value for a particular QV bar, hover over the QV bar for 2 seconds. 

Adjust Quality Value ranges

Quality Value (QV) ranges and the associated color can be adjusted.

1. Click or select **Analysis ▸ Display Settings**.
2. In the **Sample File Display** pane, drag the color bars to adjust the ranges.



① Adjustable color bar

The following table lists recommended QV ranges for pure base data. If using QV values for data quality review, we recommend that you use ranges as determined by your laboratory.

QV range	Default color	Description
Low—0 to 14	Red	Data quality is not acceptable.
Medium—15 to 19	Yellow	Data needs manual review.
High—20 or higher	Blue	Data quality is acceptable.

3. (*Optional*) To change the color of the QV bar, double-click the bar, then select a new color.
4. Click **OK** to apply the new settings.

Edit base calls

Bases can be edited in the **Electropherogram** or **Sequence** tab. If a base is edited in one tab, the base is immediately updated to match in the other tab.

1. In the **Sample Manager** or **Sample Navigator** view, select sample data to display (see "Display the sample data" on page 70).
2. In the **Electropherogram** or **Sequence** tab, edit bases, as needed.

Note: Edited bases are displayed in lowercase letters.

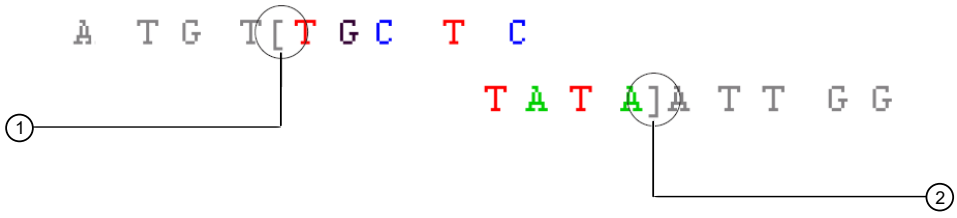
Option	Action
Add a base to the sequence.	1. Place the insertion point at the position in the sequence where you want to add one or more bases. 2. Enter the character(s) you want to insert. You can add any base-identification character that is recognized by the software, including IUPAC/IUB code (see "IUPAC/IUB codes" on page 163). **Note:** If you add a base, the new base does not have a QV value (see "About edited bases and QVs" on page 95).
Delete a base from the sequence.	Single-click the base you want to delete, then press the **Delete** key. **Note:** If you delete a base, the base QV is also deleted (see "About edited bases and QVs" on page 95).
Change a base in the sequence.	Single-click the base you want to change, then enter the new character for that position. **Note:** If you change a base, the numerical of the QV value for the base will stay the same, but the QV bar color will change to gray (see "About edited bases and QVs" on page 95).

3. Save the sample file (see "Save sample files" on page 61).

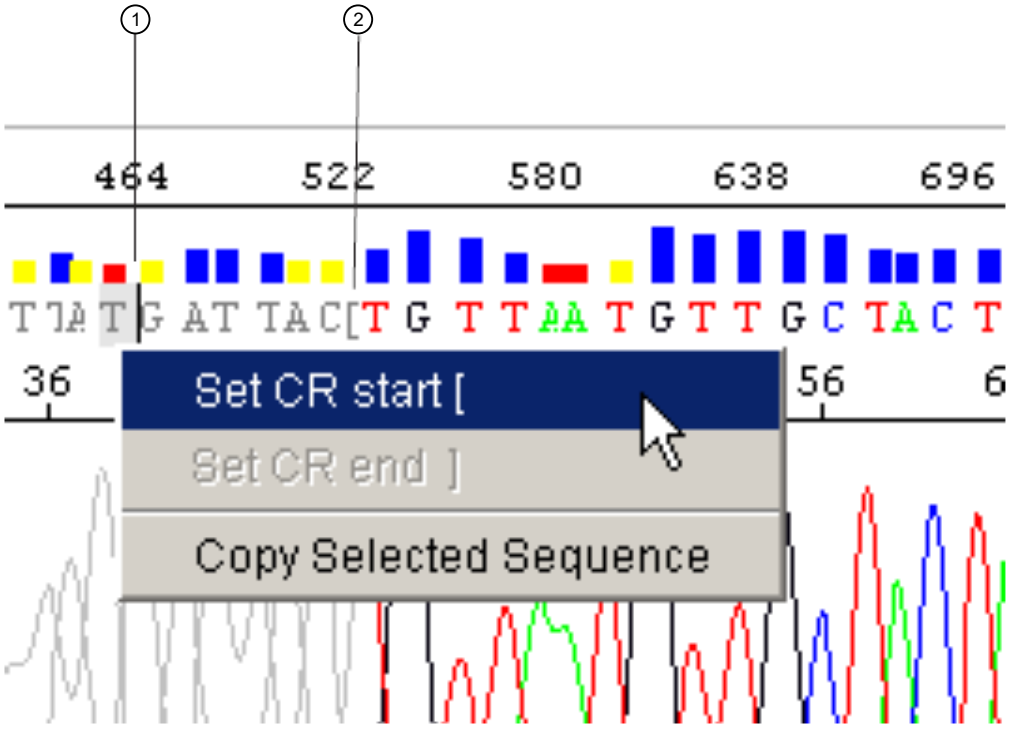
Note: If a .seq file was created when the sample file was analyzed, then both sample and the .seq files are updated when you save the sample file.

Edit the clear range

- Select an option to edit the clear range:

Option	Description
Edit using the clear range widget	**1.** In the **Electropherogram** tab, locate and select the 5′ (CR start) or 3′ (CR end) widget. **Note:** The widget turns from gray to black when selected.  ① 5' (CR start) widget ② 3' (CR end) widget **2.** Drag the widget along the bases to the right or left, the release the cursor. The new clear range is displayed. **3.** Repeat to define an new clear range for the opposite end of the data.

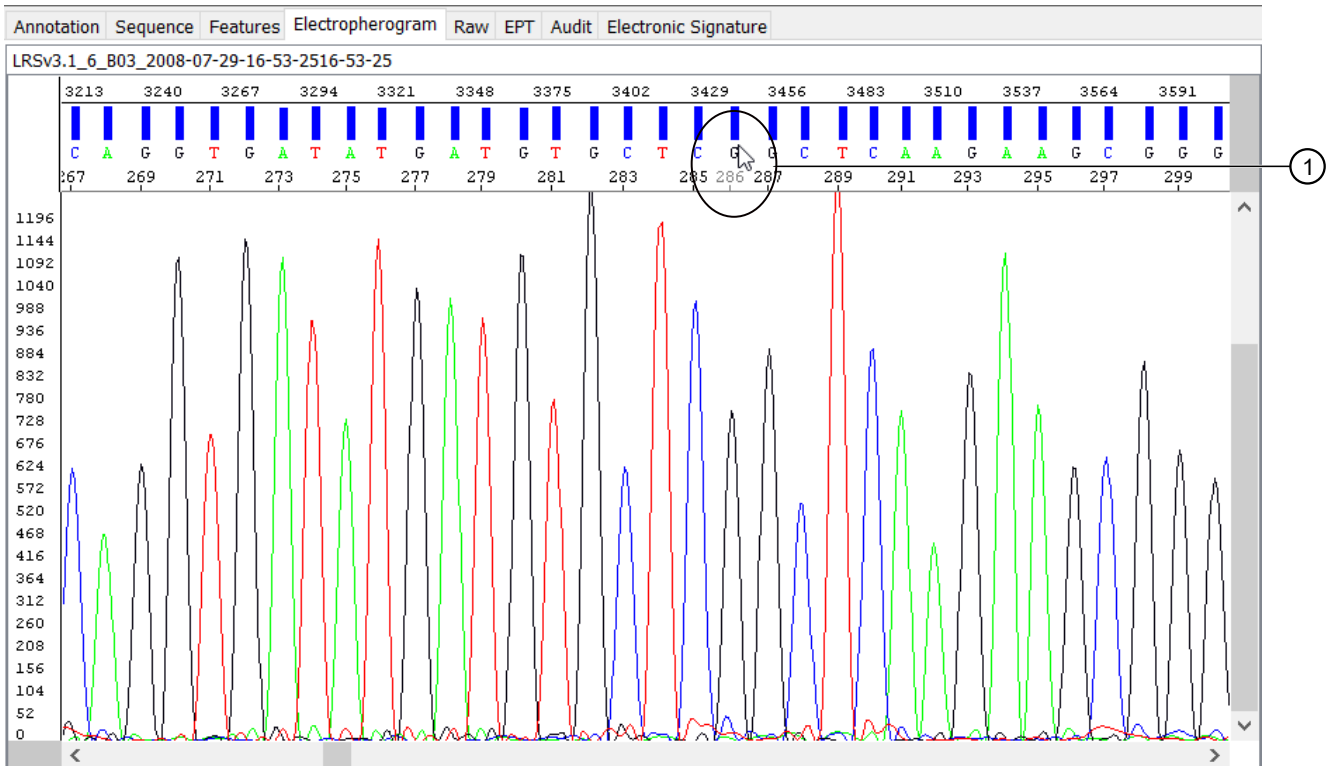
(continued)

Option	Description
Edit using the mouse	**1.** Place the cursor between two bases that represents the new location for the 5′ (CR start) or 3′(CR end) widget then right-click. The following dialog box opens:  ① New widget position ② Original widget position **2.** Select **Set CR start [** or **Set CR end]**, as appropriate. **3.** Repeat the process to define a new CR widget position for the opposite end.
Edit using the **Tools** menu	**1.** In the **Sample Manager**, in the **Show** column, select the checkbox of the sample(s) you want to display. **2.** In the **Electropherogram** or **Sequence** tab, determine the new beginning and ending base pair numbers. **3.** In the menu, select **Tools** ▸ **Set Clear Range**. **4.** In the **Set Clear Range** dialog box, enter the new beginning and ending base pair numbers, then click **OK**.

Display the base number

1. Select the **Electropherogram** tab.
2. Place the cursor over the base of interest.

   The base number is displayed. If you drag the cursor across the bases, a base number is displayed for each base.



① Base number displayed

Search for a pattern in a sequence

You can search for patterns using base-character (G, A, T, C) and IUPAC / IUB Codes in the **Sequence** or **Electropherogram** tabs.

1. In the **Sample Manager** or **Sample Navigator** view, select the sample data to display (see “Display the sample data” on page 70).
2. In the **Sequence** or **Electropherogram** tab, click the position in the sequence where you want to start the search.

   Note: The search begins at the cursor position and automatically wraps around. It is not case sensitive.

3. In the menu bar, select **Edit** ▸ **Find**, or press Ctrl+F.
   The **Find** dialog box opens.

4. In the **Find** dialog box, enter the pattern in the **Search for** field, then select the search type in the **Using** pane.

Search Type	Description
Literal String	Searches for the base-character (G, A, T, C) patterns as entered in the **Find** field.
IUPAC / IUB Codes	Searches for all possible matches for base-character (G, A, T, C) patterns that include IUPAC/IUB characters. For example, if you search for the pattern TAR, then the search locates either TAG or TAA. To view IUPAC / IUB codes, see "IUPAC/IUB codes" on page 163 or select **Help ▸ IUPAC Codes**.

5. Click **Find** to start the search.

   The software highlights the first instance of the specified pattern and marks its position in the summary graphic at the top of the **Sequence** or **Electropherogram** tab.

6. Select **Edit ▸ Find Again**, or press **Ctrl+G**, to find other occurrences of the same pattern.

Show the reverse complement data

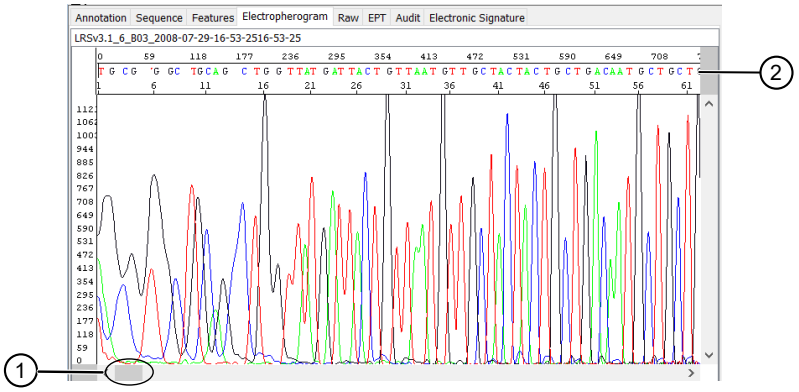
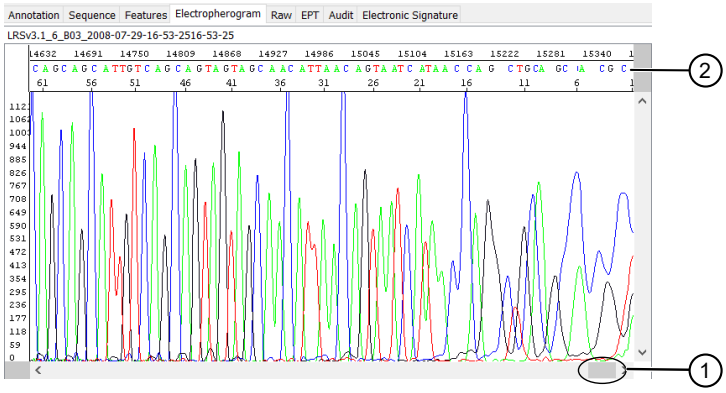
You can display the reverse complement of the data in the **Sequence**, **Electropherogram**, or **Raw** data tabs. This alters the display, changing each base to its complement and rewriting it in a 5′ to 3′ direction. When the reverse complement data is displayed, a check mark appears next to the command in the menu.

1. In the **Sample Manager** or **Sample Navigator** view, select the sample data to display (see "Display the sample data" on page 70).

2. Select the **Sequence**, **Electropherogram**, or **Raw** tab to view the data.

View	Changed items
Electropherogram	Basecalls and direction
Sequence	
Raw Data	Direction

3. In the menu bar, select **Tools ▸ Reverse Complement**.
   The data is reversed.

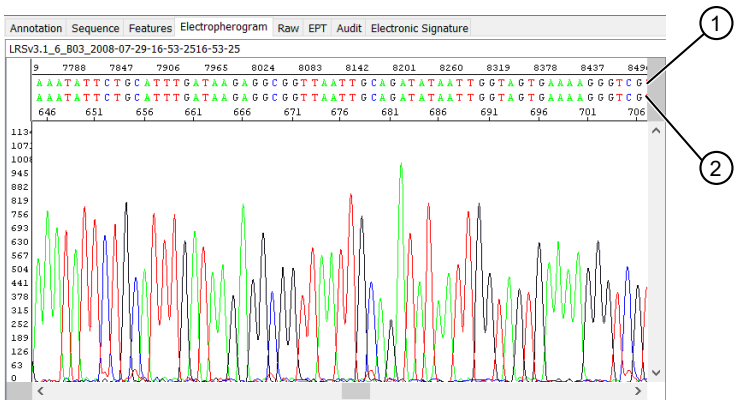
Direction	Example
5′ to 3′	 ① Beginning of data ② Bases called
3′ to 5′ (Reverse complement)	 ① Beginning of data ② Reverse complement of bases

To switch back to the original data, select **Tools ▸ Reverse Complement**.

Show or hide the original basecalling data

You can display the original sequence data in the **Electropherogram** tab. Displaying the original data can be helpful if you are editing bases in this view.

1. In the **Sample Manager** or **Sample Navigator** view, select the sample data to display (see "Display the sample data" on page 70).
2. Select the **Electropherogram** tab.
3. Click in the tool bar or select **View ▸ Show Original Sequence** in the menu.
   The original data is displayed.



① Original, uneditable data
② Editable copy of the data

To hide the original data, click in the tool bar or select **View ▸ Hide Original Sequence** in the menu.

Viewing, editing, and printing

Action	Procedure or Key Sequence
Zoom in or out to see the data at different magnifications	"Zoom in or out" on page 89.
Scale Y-axis data	"Scale the Y-axis for the raw data and electropherograms" on page 86
Edit the bases	"Edit base calls" on page 77.
Search for a base character, a range of bases, or a specified base pattern using the Find command (Ctrl+F) and Find Again (Ctrl+G)	"Search for a pattern in a sequence" on page 80.
Edit the sequence using any of the standard Windows™-based computer commands	"Edit base calls" on page 77.
Display reverse complement	"Show the reverse complement data" on page 81.

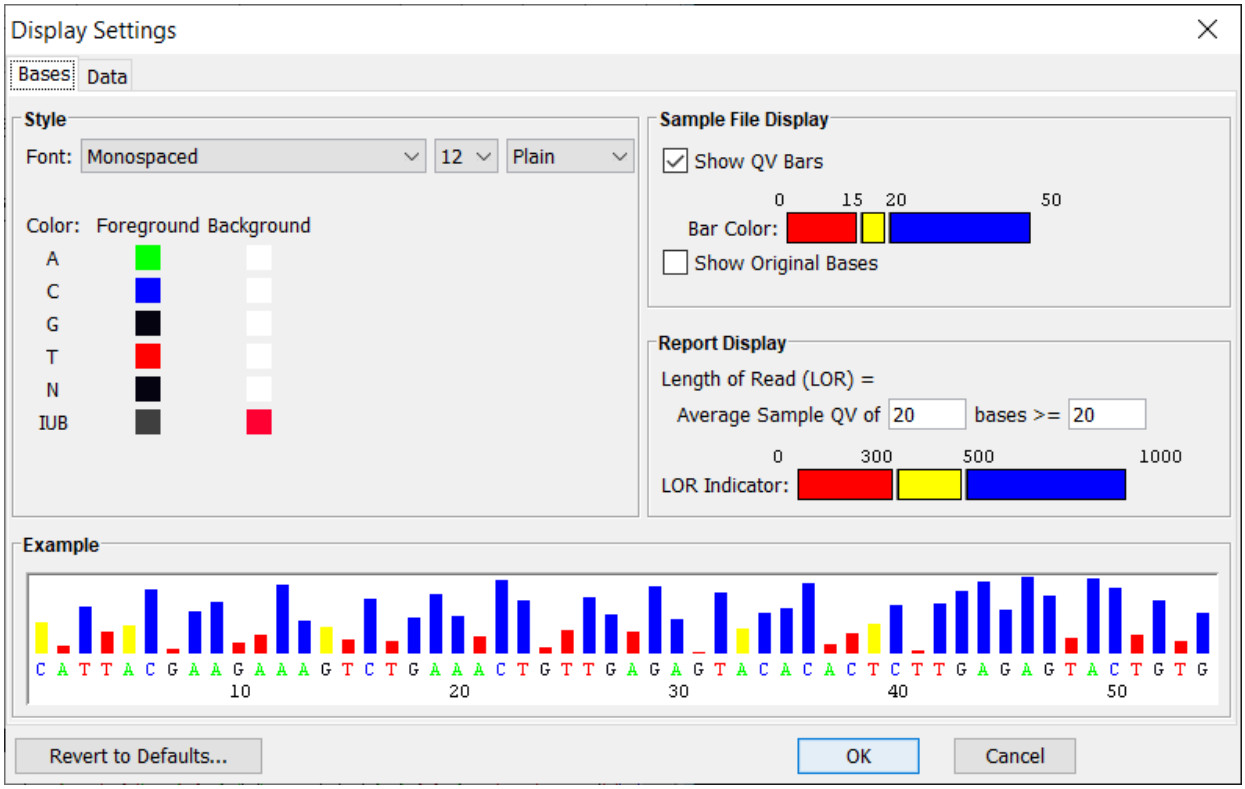
(continued)

Action	Procedure or Key Sequence
Display quality values and numbers	"Display sample Quality Values" on page 75.
Display the original, unedited basecalls while you edit the bases	"Show or hide the original basecalling data" on page 83.
Display cross hairs and the coordinates for the current cursor location	Click in the data area of the window.
Print the window contents	"Manually print the sample file" on page 50.
Move to next base	Right arrow key
Move to previous base	Left arrow key
Find next occurrence of N to the right	Tab key
Find next occurrence of N to the left	Shift+Tab keys
Move pointer 10 bases to the right	F5 key
Move pointer 10 bases to the left	Shift+F5 keys
Move pointer right to the next low QV	F6 key
Move pointer left to the next low QV	Shift+F6 keys
Move pointer right to the next medium QV	F7 key
Move pointer left to the next medium QV	Shift+F7 keys
Move pointer right to the next high QV	F8 key
Move pointer left to the next high QV	Shift+F8 keys

Customize plot and data views

About display settings

For the **Electropherogram**, **Raw**, and **EPT** views of the sample window, you can use the **Display Settings** to do the following:



Tabs	In this tab you can…
Bases	• Determine which color is used to represent each kind of data. • Change the font type, size and style, and color for the displayed bases. • Select to display/hide QV bars, and original bases. • Change the color and ranges for the LOR. • Select colors to represent QV and Length of Read (LOR) ranges. • Turn on the display of quality values (QVs) and original bases, selectively.
Data	• Change the type of scaling used for the display. • Change the colors of the trace lines to make them easier to see on screen. • Turn off one or more trace line(s) selectively. • Select data attributes to display in the electropherogram, raw, and EPT data. • Selectively turn off one or more trace lines.

IMPORTANT! Any change you make in this dialog box affects all displays of the selected view and remains in effect until you change the setting again in this dialog box.

To open the **Display Settings** dialog box, select **Analysis** ▸ **Display Settings**, or click .

Scale the Y-axis for the raw data and electropherograms

You can scale the Y-axis for the raw data and/or electropherograms in the **Sample Manager** so that all plots use the same scale. This makes comparisons between samples easier.

For more information about the **Raw** view, see “About the Raw tab” on page 131. For more information about the Electropherogram view, see “About the Electropherogram tab” on page 129.

1. To change the Y-axis scaling, select **View** ▸ **Y Axis Scale**, then select one of the following options.

   Note: The selected setting appears with a check mark next to it.

Option	Description
Scale Individually	This uses the range in the plot to set the Y-axis range. This is the default.
Scale All to Max and Min Y	For all currently displayed files, this finds the overall minimum and maximum Y-axis values, then sets the Y-axis range for every plot to the overall minimum and maximum.
Scale All to Custom Settings	Scales every plot using the values set in the **Display Settings** dialog box.
Scale To	Opens the **Set Y-Scale** dialog box.

2. If you select **Scale To**, the **Set Y Scale for all Electropherograms and Raw Data** dialog box is displayed.

   Perform one of the following actions:

   - Edit the **From** and **to** values for the **Electopherogram View Y Data Range** and/or **Raw Data View Y Data Range** plots.

     This defines the range for the Y-axis scale. The values must be at least 100 units apart.
   - Click **Get Autoscale Values**.

     This will update the dialog box with the minimum and maximum Y-axis values for the electropherogram and the raw data of the files currently displayed in the **Sample Manager** (those files that have **Show** selected).
   - Click **Get Custom Values** to update the minimum and maximum Y-axis values to those set in the **Display Settings** (see “Edit the display settings” on page 87).

3. Click **Apply** to rescale the plots using the new values and leave the dialog box open.

   Click **OK** to rescale the plots using the new values and close the dialog box.

Note: If you show or hide files in the **Sample Manager**, the Y-axis is not rescaled. Select **View ▸ Y Axis Scale** again to rescale the plots. To set the Y-axis scaling in the **Display Settings**, see "Edit the display settings" on page 87.

Edit the display settings

You can edit the display settings for the electropherogram, raw data, and EPT plots. Any changes made to the display settings will apply to all of the plots. For more information, see "About display settings" on page 85.

1. Select **Analysis ▸ Display Settings** or click  .
   The **Display Settings** dialog box opens.

2. In the **Bases** tab, edit the display settings.

Note: The **Example** pane shows you what the data display will look like with the selected settings.

Option	Description
Style	Select the font type, size, and style from the dropdown lists. For each base, click the color squares to edit foreground and background colors.
Sample File Display	Select to show or hide QV bars and original bases. To edit the QV values ranges, see "Adjust Quality Value ranges" on page 76.
Report Display	Length of Read (LOR) is the usable range of high quality or high accuracy bases, as determined by quality values. You can view the LOR in the analysis report, instead of opening the sample file (see "Analysis report screen description" on page 144). The short, medium, and long ranges and the color associated with a LOR can be modified using the LOR Indicator bar. Use the two sliders to define the short, medium, and long ranges. The sliders move in increments of 5 bases. To edit the color indicator for the short, medium, and long ranges, click the color bar, select a new color, then click **OK**.

Note: To revert to the default settings, in the **Display Settings** dialog box, click **Revert to Defaults**.

3. In the **Data** tab, edit the following settings:

Attribute	Function
Pane Height, Single Sample (cm or Full Screen)	Used to enter an editable alphanumeric value to set the height of the pane that displays the electropherogram, raw, and EPT data. The value is either Full Screen or the height of the row in cm. The default value is Full Screen.
Pane Height, Multiple Samples (cm or Full Screen)	Used to enter an editable alphanumeric value to set the height of the pane that displays the electropherogram, raw, and EPT data. The value is either full screen or the height of the row in cm. The smaller the value the more panes of data are displayed in one screen.
Vertical Scale (%)	Used to enter an editable, numeric integer to set the scaling of the data relative to the pane height. The range is 1 to 100. For EPT data, the scaling is set to the highest peak in each category.
Horizontal Scale (%)	Used to enter an editable, numeric integer to set the scaling of the data relative to the pane height. The range is 1 to 100.
Custom Y-Scale Values Min and **Max**	Used to enter values for the Y scale. Minimum and Maximum values must be at least 100 units apart.
Show A, C, G, and T Data	When the checkboxes are selected, the electropherogram and raw data traces for the A, C, G, and T are displayed. All checkboxes are selected by default.
Show Volts/100, **Show µAmps**, **Show mWatts * 10**, and **Show Degrees C**	When the checkboxes are selected, the data traces for the volts, µAmps, mWatts, and temperature are displayed. All checkboxes are selected by default and the colors represent: • Volts/100 = Blue • µAmps = Green • mWatts = Black • Degrees C = Red

Note: To revert to the default settings, click **Revert to Defaults**.

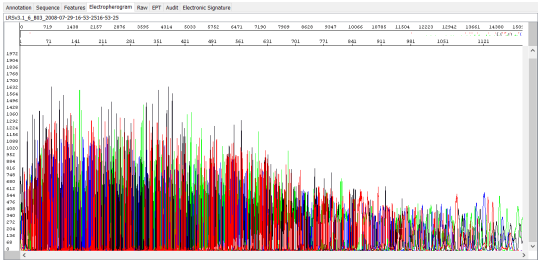
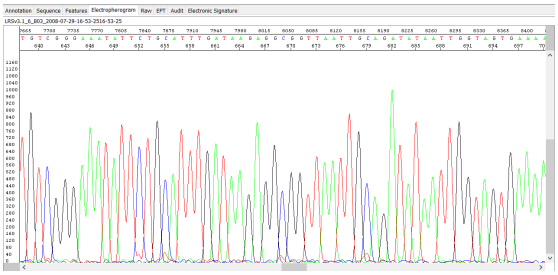
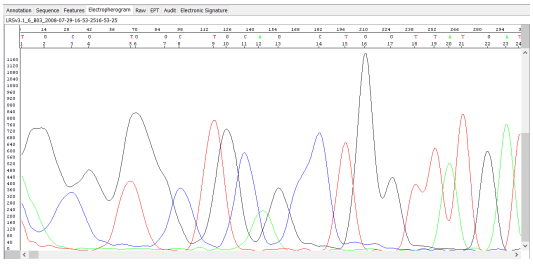
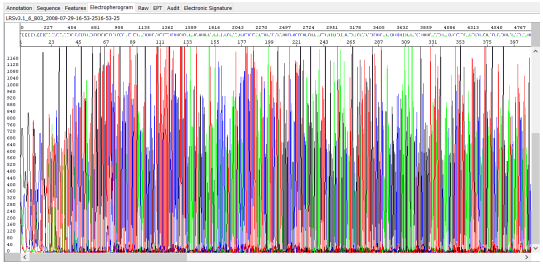
4. Click **OK**.

5. To rescale the plots using the new custom Y-scale values, select **View ▸ Y Axis Scale ▸ Scale All to Custom Settings**.

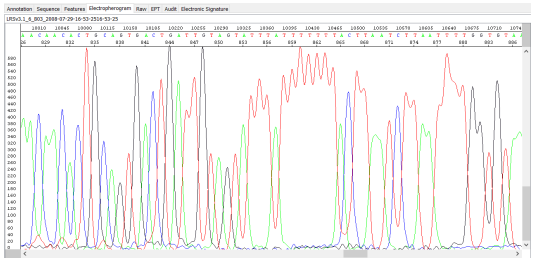
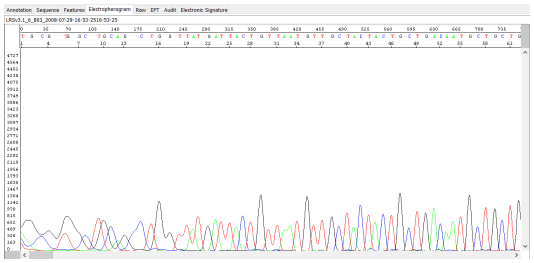
The display changes are applied to sample files in the **Sample Manager** and new samples added to the **Sample Manager**. The changes remain in effect until you change the settings again in this dialog box.

Zoom in or out

Use the following zoom functions to change the data display.

View Command	Description	Procedure
Full View 	Displays all the data in a standard size window.	• Click  • Select **View** ▸ **Full View** • Press Ctrl+[
Default Size 	Restores the display to the initial/default zoom factor. The initial zoom factor is determined by the display settings.	• Click  • Select **View** ▸ **Default Size** • Press Ctrl+]
Zoom In Horizontal 	Enlarges the view horizontally so that more detail is visible.	• Click  • Select **View** ▸ **Zoom In Horizontal** or • Press Ctrl+=
Zoom Out Horizontal 	Reduces the view horizontally so that a larger area is visible.	• Click  • Select **View** ▸ **Zoom Out Horizontal** • Press Ctrl+Minus

(continued)

View Command	Description	Procedure
Zoom In Vertical 	Enlarges the view vertically so that more detail is visible.	• Click  • **View ▸ Zoom In Vertical** • Press Ctrl+Shift+=
Zoom Out Vertical 	Reduces the view horizontally so that a larger area is visible.	• Click  • Select **View ▸ Zoom Out Vertical** • Press Ctrl+Shift+Minus

View printed electropherograms

A printed electropherogram shows a four-color view of analyzed data, with peaks representing the bases.

To set the number of panels displayed on each page:

1. Select **Tools ▸ Options**.
2. In the **Printing** tab, select the numbers of panels from the **Panels per page** dropdown list.

For more information on panels per page, see “Manually print the sample file” on page 50 or “Edit printing options” on page 54.

Why print an electropherogram

Printed electropherograms have the following advantages over electropherograms viewed on the screen:

- Include information from the Plate Record that is visible on screen only in Annotation view.
- Display several panels of data on one page; on the screen, you can see only one section of the data at a time.

Trace and base colors

In the electropherogram, the colors represent individual bases in the sequence, as indicated below. The letter above each peak is colored, according to the color of the corresponding base.

Base	Color
C	Blue
A	Green
G	Black
T	Red

Printout header and footer

The header and footer on the printed electropherogram contains information about the run and can be useful for troubleshooting. The following figures and tables explain the header and footer contents.

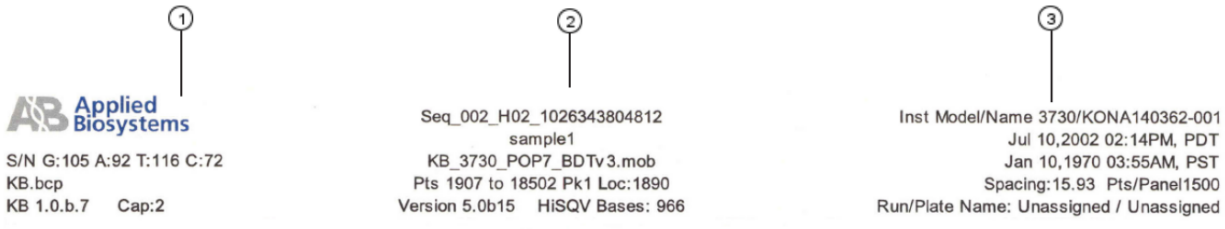


Figure 8 Example of printout header information

① First Column
② Second Column
③ Third Column

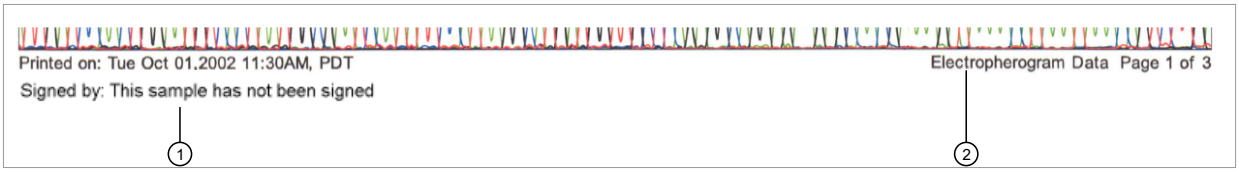


Figure 9 Example of printout footer information

① Print date and electronic signature
② View type, and page x of x

Table 3 Printout Header Information

Field	Description
First Column	
Signal/Noise	The signal-to-noise ratio is the average of the signal strength of each base divided by the average of the noise of that base.
Basecaller name	Name of the basecaller used to analyze the data.
Basecaller version, and capillary number	• Version of the basecaller used to analyze the data. • Number of the capillary use to generate the data.
Second Column	
Sample file name	Name used for the sample file. The name in the File Name column in the Sample Manager window, as saved on the hard drive.
Sample name	Name entered in the Sample Name column of the Plate Record.
DyeSet/Primer file	DyeSet/Primer file used during analysis to adjust for mobility shifts.
Points, Peak 1 Loc:	• Range of the collected data points that were used to analyze the data. • Peak 1 is the data point where the analyzed data starts.
SeqA version, HighSQV bases	• Version of the Sequencing Analysis used to analyze the data. • Number of bases with QV values in the high range.
Third Column	
Instrument model	Instrument model used to collect the data.
Date and time of analysis	Date and time the analysis took place.
Date and time of collection	Date and time the data collection took place.
Spacing, and points/panel	• Base spacing as calculated by the basecaller. • Number of data points used to display the data per each panel.
Run/Plate name	• Name of the run. • Name of the plate.

Table 4 Printout Footer Information

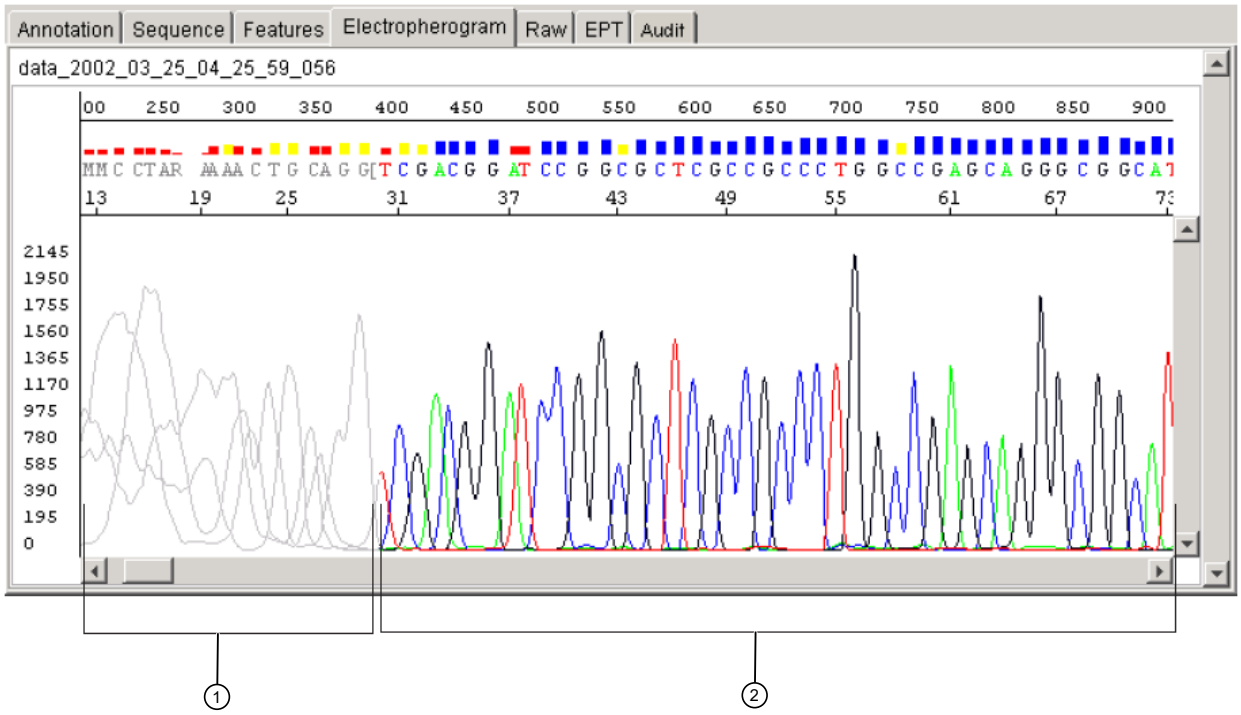
Field	Description
Print date and electronic signature	Date and time of printing.
View type, and page x of x	• Type of view: annotation, sequence, electropherogram, raw data, or EPT. • Page number for this page and the total number of pages.

10 Review low-quality samples

About the clear range

The clear range is the region of sequence that remains after excluding the low-quality or error-prone sequence at both the 5′ and 3′ ends. In the Electropherogram and Sequence views, the excluded data is displayed in gray. The bases outside the clear range cannot be edited.

IMPORTANT! The clear range display is not available if the clear range option was not selected in the analysis protocol, and the sample was not post-processed by selecting **PP** in **Sample Manager**.



① Data outside the clear range
② Data within the clear range

To edit the clear range, see “Edit the clear range” on page 78.

About edited bases and QVs

Changing, deleting, or inserting a base will affect the Quality Value (QV) displayed for the base. For instructions on editing base calls, see “Edit base calls” on page 77.

Action	Result
Change a base	The new base is in lowercase and the QV has the same value but be displayed as a gray bar.
Change a base back to the original call	The base displays in uppercase and the QV bar color is restored.
Insert a base	The inserted base displays in lowercase and it has no QV.
Delete a base	The QV for the deleted base disappears.
Reinsert a deleted base	The reinserted base displays in lowercase and it has no QV.

About Quality Values

IMPORTANT! Only samples analyzed with KB™ Basecaller or Smart Deep™ Basecaller have Quality Values (QVs).

The basecaller assigns QVs for each base, including pure and mixed bases. The QV is a per-base estimate of the basecaller accuracy.

Mixed base calls yield lower QVs than pure base calls. For information on selecting the mixed bases option for data analysis, see “Mixed Bases tab” on page 111.

The QVs are displayed as bars above each base in the sample. The height and color of the bar indicates its value. The taller the bar, the higher the QV. QV ranges and bar colors are editable (see “Adjust Quality Value ranges” on page 76).

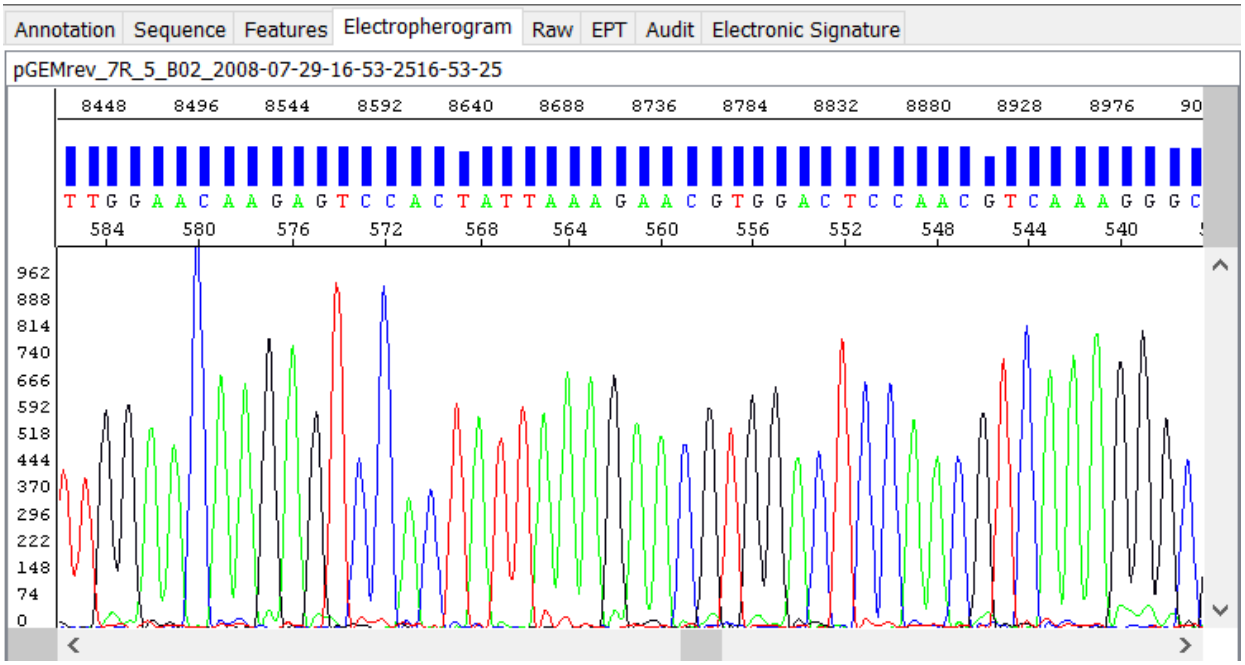


Figure 10 Example of analyzed data with QV values

Interpretation of Quality Values

Per-base QVs are calibrated on a scale corresponding to:

QV = –10log$_{10}$(*Pe*)

where *Pe* is the probability of error.

Smart Deep™ Basecaller and KB™ Basecaller generate QVs from 1 to 99, with 1 being low confidence and 99 being high confidence.

- Typical high quality pure base QVs range from 20 to 50
- Typical high quality mixed base QVs range from 10 to 30
- KB™ Basecaller mixed base QVs range from 1 to 20
- Smart Deep™ Basecaller mixed base QVs range from 1 to 30
- Size and color of QVs bars are identical for QVs 50 to 99

See the following table for the probability of a basecalling error for QVs ranging from 1 to 99.

Table 5 Quality Values and probability of error

QV	*Pe*	QV	*Pe*	QV	*Pe*
1	79%	21	0.79%	41	0.0079%
2	63%	22	0.63%	42	0.0063%
3	50%	23	0.50%	43	0.0050%
4	39%	24	0.39%	44	0.0039%

Table 5 Quality Values and probability of error *(continued)*

QV	*Pe*	QV	*Pe*	QV	*Pe*
5	31%	25	0.31%	45	0.0031%
6	25%	26	0.25%	46	0.0025%
7	20%	27	0.20%	47	0.0020%
8	15%	28	0.15%	48	0.0015%
9	12%	29	0.12%	49	0.0012%
10[1]	10%	30[1]	0.10%	50[1]	0.0010%
11	7.9%	31	0.079%	60	0.0001%
12	6.3%	32	0.063%	70	0.00001%
13[1]	5.0%	33	0.050%	80	0.000001%
14[1]	4.0%	34	0.040%	90	0.0000001%
15[1]	3.2%	35	0.320%	99	0.00000012%
16	2.5%	36	0.025%	—	—
17[1]	2.0%	37	0.020%	—	—
18	1.6%	38	0.016%	—	—
19	1.3%	39	0.013%	—	—
20[1]	1.0%	40[1]	0.010%	—	—

[1] Commonly used cut-off values for quality values

About sample scores

A sample score is the average quality value of the bases in the clear range sequence for that sample.

11 Manage user accounts

Create a new user account (admin only)

We recommend that you create a user profile for each individual who uses the software. There are three levels of users: administrator, scientist, and analyst. The **Users** tab allows exporting of user names and access privileges for these users.

IMPORTANT! The administrator is the only person who can set up and change the information in the **Users** tab. The selections in this tab are inactive for all other users.

1. In the software home screen, select **Tools** ▸ **Options**.
2. In the **Options** dialog box, in the **Users** tab, click **New**.
   The **User Management** dialog box is displayed.
3. Enter the user name, first and last name, and password in the appropriate fields.
4. In the **User Group** dropdown list, select the level of user.

   Note: The user name should only contain alphanumeric characters. Do not use spaces or invalid characters (\ / : * ? " < > |) (see "File-naming convention" on page 24).
5. *(Optional)* Select the **Enable Electronic Signature** checkbox.
6. Click **OK**.
   The new user appears in the list in the **Users** tab.

Restart the software so that new users can log in. The users can change their passwords after logging into the software (see "Changing the user password" on page 99).

Edit user account information (admin only)

IMPORTANT! The administrator is the only person who can set up and change the information in user account information in the **Users** tab. The selections in this tab are inactive for all other users.

1. In the software home screen, select **Tools** ▸ **Options**.
2. In the **Options** dialog box, in the **Users** tab, double-click the user name to edit.
3. In the **User Management** dialog box, change or correct the user information, then click **OK**.
4. Click **OK** to close the **Options** dialog box.

Export user account settings

Note: It is possible to import or export user settings from one computer to another. .

1. In the software home screen, select **Tools** ▸ **Options**.
2. In the **Options** dialog box, in the **Authentication** tab, click **Export** to export the application configuration settings and/or settings for a single user or multiple users in a zipped .ctf format.
3. Enter the path for exporting files in the **Export Application Configuration** dialog box, then click **Export**.
4. Click **OK** to close the **Options** dialog box.

Changing the user password

Users in all groups (administrator, scientist and analyst) can change their account password.

1. Log in to the software with your account.
2. In the software home screen, select **Tools** ▸ **Change Password**.
3. Enter your current password.
4. Enter your new password, then reenter it again.
5. Click **OK**.

Authentication tab

IMPORTANT! The administrator is the only person who can set up and change the information in the Authentication tab. The selections in this tab are inactive for all other users.

Table 6 Parts of the Authentication tab

Item	Function
Lockout user after __ invalid login attempts	Used to lock out users from the software if they input the incorrect user or password after the specified attempts
within __ minutes	Used to define the time.
Maintain lockout for __ minutes	Used to determine the amount of time the user is locked out of the software.
Timeout Feature On checkbox	Used to turn off or on the timeout feature. The default is off.
Automatic timeout after __ minutes	Used to determine the number of minutes of program is inactivate before the user is required to enter a password to reactivate the program.
Change password every __ days	Used to force all user groups to change their password at the specified interval.

Audit tab

IMPORTANT! The administrator is the only person who can set up and change the information in the Authentication & Audit tab. The selections in this tab are inactive for all other users.

Table 7 Parts of the Authentication and Audit Tab

Item	Function
Audit Trail On checkbox	Used to turn off or on the audit trail feature. The default is off. If the Audit trail is on, information is created whenever a user takes an action that modifies the end result sequence. This information can be viewed and printed, but not edited.
Audit Reason	Used to define reasons for base changes, insertions, deletions, and other actions taken with the data.
New button	Used to create additional reasons.
Open button	Used to modify or verify a reason.

12 Sample analysis

All analysis is done in the **Sample Manager**.

Analysis defaults

When a sample file is added to the **Sample Manager**, it passes through the **Analysis Defaults**. The **Analysis Defaults** define the following settings:

Setting	Description	When the setting is applied
Add Samples Settings		
Analysis Protocol	See "Analysis protocols" on page 103. The default Analysis Protocol is set to **None**.	When samples are added to the **Sample Manager**, the default Analysis Protocol is only applied to samples that do not have an **Analysis Protocol**, such as files created on the 310 Genetic Analyzer. The DyeSet/Primer and matrix files defined in the data collection sample sheet are used, and all other settings from the **Analysis Protocol** (including the basecaller) are applied. **IMPORTANT!** The DyeSet/Primer file must match the chemistry and basecaller type that you are using in both data collection and the **Analysis Protocol**. **Note:** If a sample does not contain an **Analysis Protocol** and a default is not selected, then an alert message is displayed during import to **Sample Manager**.

(continued)

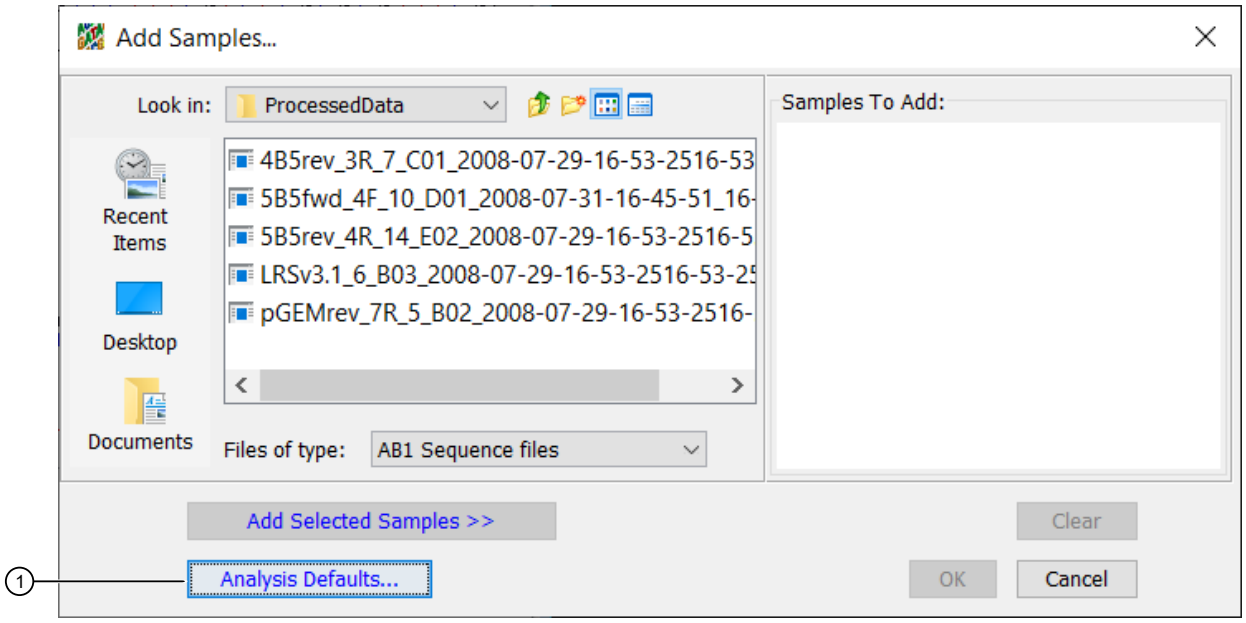
Setting	Description	When the setting is applied
Processing parameter (PP) settings	• **Basecalling (BC)** • **Post processing (PP)** • **Printing (P)**	PP settings are applied when the samples are added to the **Sample Manager** and analyzed. The clear range is calculated during post processing.
Sequence File Formats		
File format	Select one of the following options: • Use the settings in the sample's Analysis Protocol • Override the sample's Analysis Protocol and set to: .seq, .scf, and/or .phd.1	File format settings are applied when the samples are added to the **Sample Manager** and analyzed.

For information about viewing or editing the Analysis Defaults, see “Create, edit, or select new Analysis Defaults” on page 35.

Ways to access analysis defaults

There are two ways to access the Analysis Defaults dialog box.

- Select **Analysis** ▸ **Analysis Defaults**
- Click **File** ▸ **Add Sample(s)** ▸ **Analysis Defaults**



① **Analysis Defaults**

Analysis protocols

To navigate to the analysis protocol manager window, click **Analysis** ▸ **Analysis Protocol Manager**, then select **New** or **Edit**.

The **Sequence Analysis Protocol Editor** dialog box is displayed.

An analysis protocol is saved in a sample file and defines the basecalling and post-processing settings for that sample, including:

Tab	Description
General	Name and describe the Analysis Protocol, and define the sequence file formats to be used (see "General tab" on page 104).
Basecalling	Select the basecaller, DyeSet/Primer file, matrix, and analysis stop point. For Smart Deep™ Basecaller, select the **SDB Trace Visualization** option (see "Basecalling tab" on page 105).
Mixed Bases	Specify whether or not to use mixed base identification, and if so, define the percent value of the second highest to the highest peak, for consideration as a mixed base by the KB Basecalling algorithm. Reaching the threshold is a necessary but not sufficient condition for arriving at a mixed base determination (see "Mixed Bases tab" on page 111).
Clear Range	Specify the clear range based on base positions, quality values, or number of ambiguities (Ns) (see "Clear range tab" on page 113).

There are two types of Analysis Protocols:

Type	Description
Master analysis protocol	The software includes default master analysis protocols in the **Analysis Protocol Manager**. You can also create new master analysis protocols in the **Analysis Protocol Manager** ("Create or edit an analysis protocol" on page 39). The settings defined in a master analysis protocol are copied, then assigned to a sample using one of the following options: • Manually (see "Apply an analysis protocol to samples in the Sample Manager" on page 41) • By the analysis default when adding samples to the **Sample Manager** (see "Create, edit, or select new Analysis Defaults" on page 35)
Sample analysis protocol	This analysis protocol is saved in the sample file, and is specific to that sample. Editing a sample analysis protocol only effects the selected sample (see "Edit a sample analysis protocol" on page 41). A sample analysis protocol cannot be directly applied to other samples.

General tab

To navigate to the **General** tab, click **Analysis** ▸ **Analysis Protocol Manager**, then select **New** or **Edit**.

The **Sequence Analysis Protocol Editor** dialog box is displayed.

The **General** tab includes the analysis protocol name and description. The tab also includes settings for **Sequence File Formats**:

Option	Function
Write .Seq File	When selected, the software creates a .seq file for printing the sequence as text file or for using the file in other software. If **Write Clear Range in .Seq File** is selected, the software only writes the clear range into the .seq file. • ABI format is used with Applied Biosystems™ software. • FASTA format is used with other software
Write Standard Chromatogram Format (.scf)	When selected, the software creates a .scf file that can be used with other software. When created, the .scf extension is not appended to the file name.
Write Phred (.phd.1) File	When selected and either KB™ Basecaller or Smart Deep™ Basecaller is used, the software creates a .phd.1 file that can be used with other software.

Basecalling tab

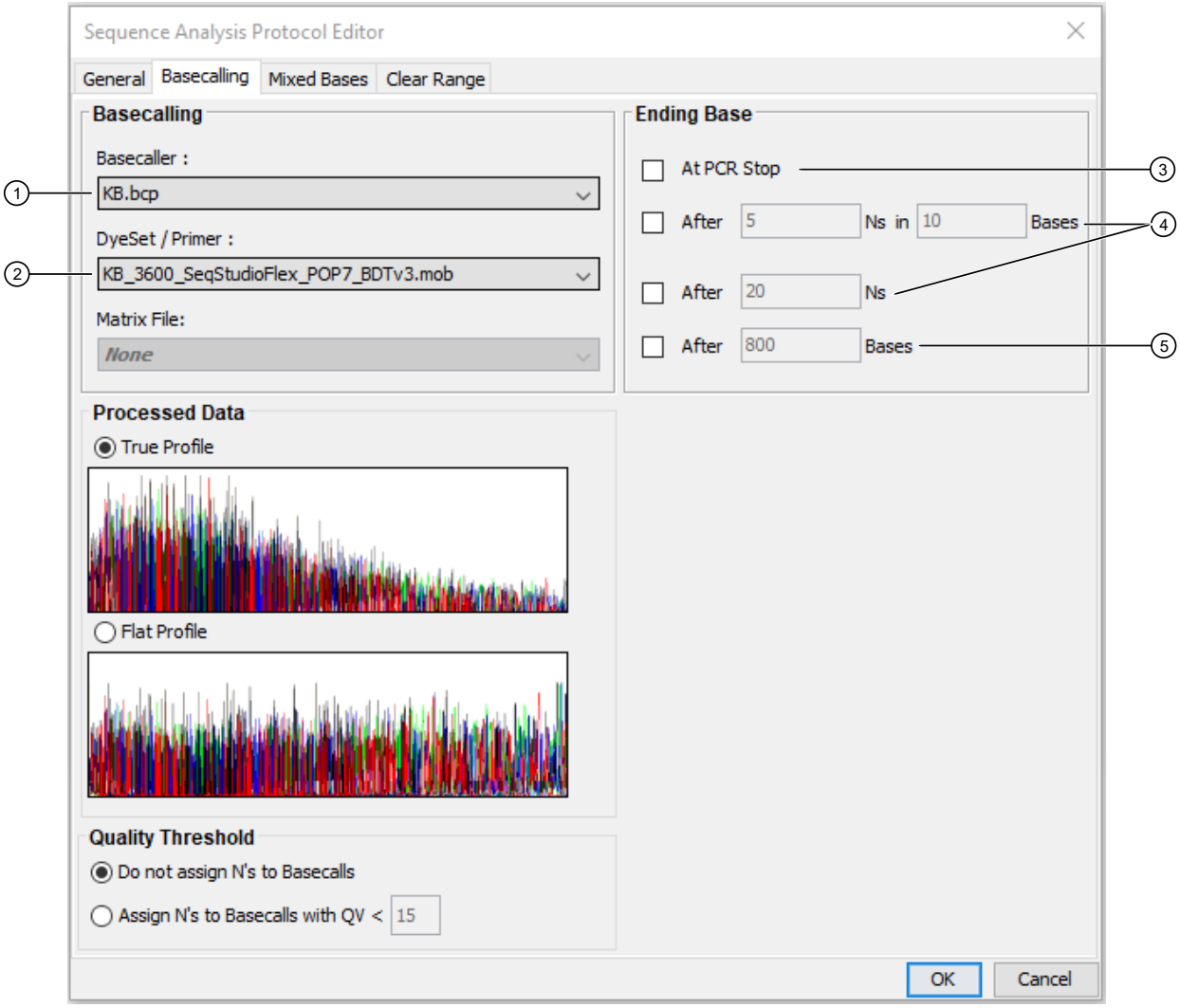


Figure 11 Basecalling tab (using KB and ABI)

① Basecaller selection, use with KB and ABI (also use with SDB, see Figure 12)
② .mob file, use with KB and ABI (also use with SDB, see Figure 12)
③ Use with KB and ABI (also use with SDB, see Figure 12)
④ Use with KB and ABI (also use with SDB, see Figure 12)
⑤ Use with KB and ABI (also use with SDB, see Figure 12)

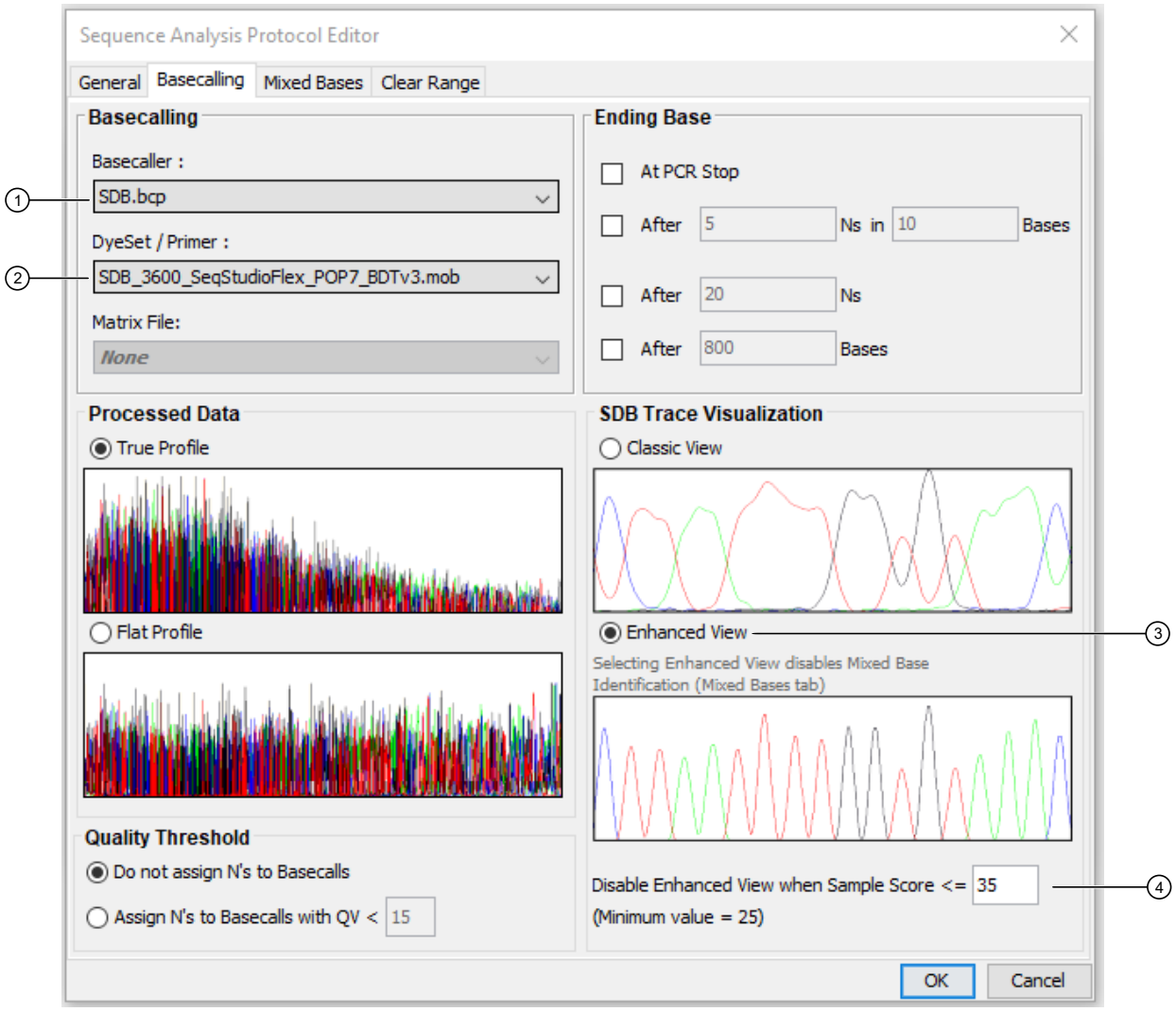


Figure 12 Basecalling tab (using SDB)

① SDB basecaller
② SDB .mob file
③ Enhanced trace visualization
④ Disable the enhanced view

Table 8 Parts of the Basecalling Tab

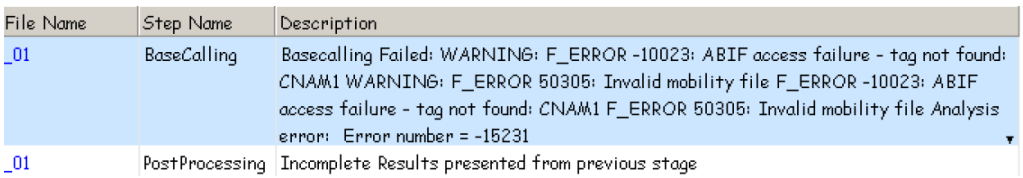
Item	Function
Basecalling	
Basecaller drop-down list	Used to select a basecaller. These files contain algorithms used to call the bases in a sample file. • KB™ Basecaller—Algorithm calculates mixed or pure bases, and quality values. • Smart Deep™ Basecaller—Algorithm calculates mixed or pure bases, and quality values. • ABI™ Basecaller—Algorithm used in previous versions of ABI PRISM™ sequencing analysis software. Refer to Appendix C, “Basecallers and DyeSet/Primer files” for a list of basecallers sorted by instrument.
DyeSet/Primer dropdown list	Used to select a DyeSet/Primer file. These files contain algorithms used to correct for mobility shifts and color code changes due to the type of chemistry used. Refer to Appendix C, “Basecallers and DyeSet/Primer files” for a list of DyeSet/Primer files sorted by instrument and basecaller. **IMPORTANT!** The DyeSet/Primer file type must match the basecaller type and the chemistry used. • If you select a KB™ DyeSet/Primer file and an ABI™ Basecaller for analysis, the basecalling is successful (green BC box) but the analyzed data is not usable. • If you select a DT DyeSet/Primer file and an KB™ Basecaller for analysis, the basecalling fails. The following error message displays in the Error table.  File Name | Step Name | Description _01 | BaseCalling | Basecalling Failed: WARNING: F_ERROR -10023: ABIF access failure - tag not found: CNAM1 WARNING: F_ERROR 50305: Invalid mobility file F_ERROR -10023: ABIF access failure - tag not found: CNAM1 F_ERROR 50305: Invalid mobility file Analysis error: Error number = -15231 _01 | PostProcessing | Incomplete Results presented from previous stage
Matrix File dropdown list	Used to select a matrix file for 310 data. The column contains ***None***, if the data has been generated on a 3100, 3100-Avant™, 3700 or 3730/3730*xl* instrument. The matrix column is: • Used for 310 data because the matrix is applied to the data during basecalling. • Not used for 3100/3100-Avant™, 3700, or 3730/3730*xl* data because the matrix is applied to the data during data collection.

Table 8 Parts of the Basecalling Tab *(continued)*

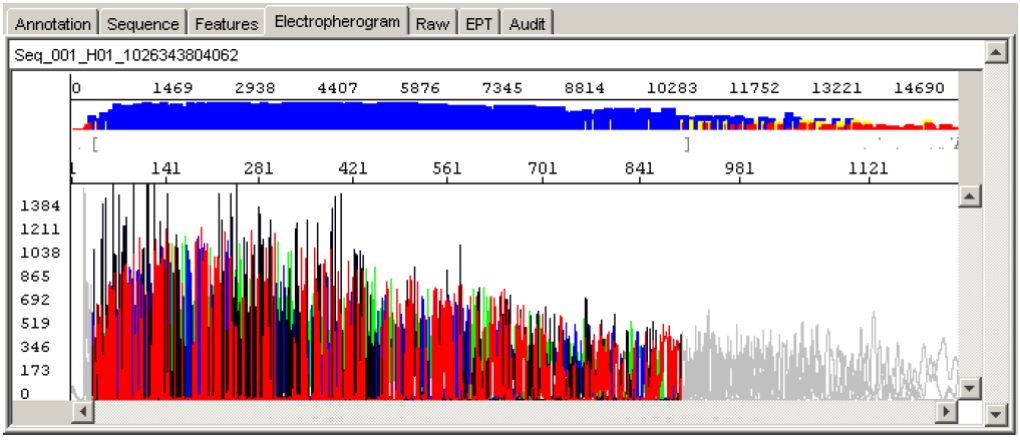
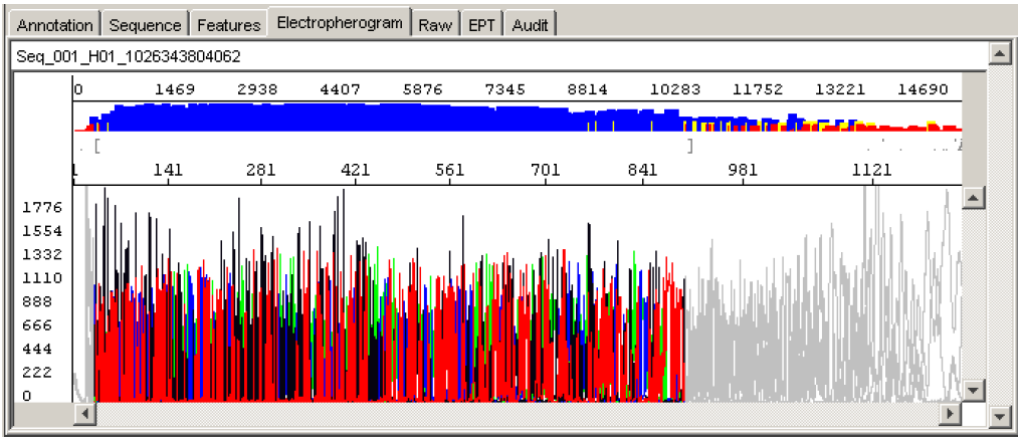
Item	Function
Processed Data	
True Profile	Used to display data as processed traces scaled uniformly so that the average height of peaks in the region of strongest signal is about equal to a fixed value. The profile of the processed traces will be very similar to that of the raw traces. 
Flat Profile	Used to display the data as processed traces scaled semi-locally so that the average height of peaks in any region is about equal to a fixed value. The profile of the processed traces will be flat on an intermediate scale (> about 40 bases). **Note:** This option is applied to data that is analyzed with the KB™ Basecaller and Smart Deep™ Basecaller only. If you use the ABI™ Basecaller the profile option reverts to **True Profile**. 

Table 8 Parts of the Basecalling Tab *(continued)*

Item	Function
Ending Base	
At PCR Stop checkbox	Sets the analysis endpoint at the end of the PCR fragment. The software determines the endpoint by locating the large peak that is characteristic of the end of a short PCR fragment when sequenced with dye primer chemistry. If the endpoint peak is not sufficiently large, the software may fail to recognize the PCR stop point. **Note:** If there is noise after the PCR data, the noise is considered as signal, and the stop point is incorrectly calculated to be after the noise.
After __ Ns in __bases checkbox	Sets the analysis endpoint after a certain number of Ns occur within a certain number of bases (for example, after 5 Ns are detected within a range of 10 bases).
After __ Ns checkbox	Sets the analysis endpoint after a certain number of Ns occur (for example, after 20 Ns are detected).
After __ Bases checkbox	Sets the analysis endpoint after a certain number of bases (for example, after 800 bases are detected).
Quality Threshold	
Do not assign N's to Basecalls	When using the KB™ Basecaller or Smart Deep™ Basecaller, use this setting assign no Ns to basecalls.
Assign N's to Basecalls with QV <	When using the KB™ Basecaller or Smart Deep™ Basecaller, use this setting assign Ns to base with QVs less than the set point. The QV will still be displayed.
SDB Trace Visualization[1] (see Figure 12)	
Classic View	When you select **Classic View**, the trace resolution is identical between KB™ Basecaller and Smart Deep™ Basecaller, but the basecalls and quality values may differ between KB™ Basecaller and Smart Deep™ Basecaller. When you select **Classic View**, you can select the **Use Mixed Base Identification** checkbox in the **Mixed Bases** tab (see “Mixed Bases tab” on page 111). The **Mixed Bases** checkbox is disabled if you select **Enhanced View**.

Table 8 Parts of the Basecalling Tab *(continued)*

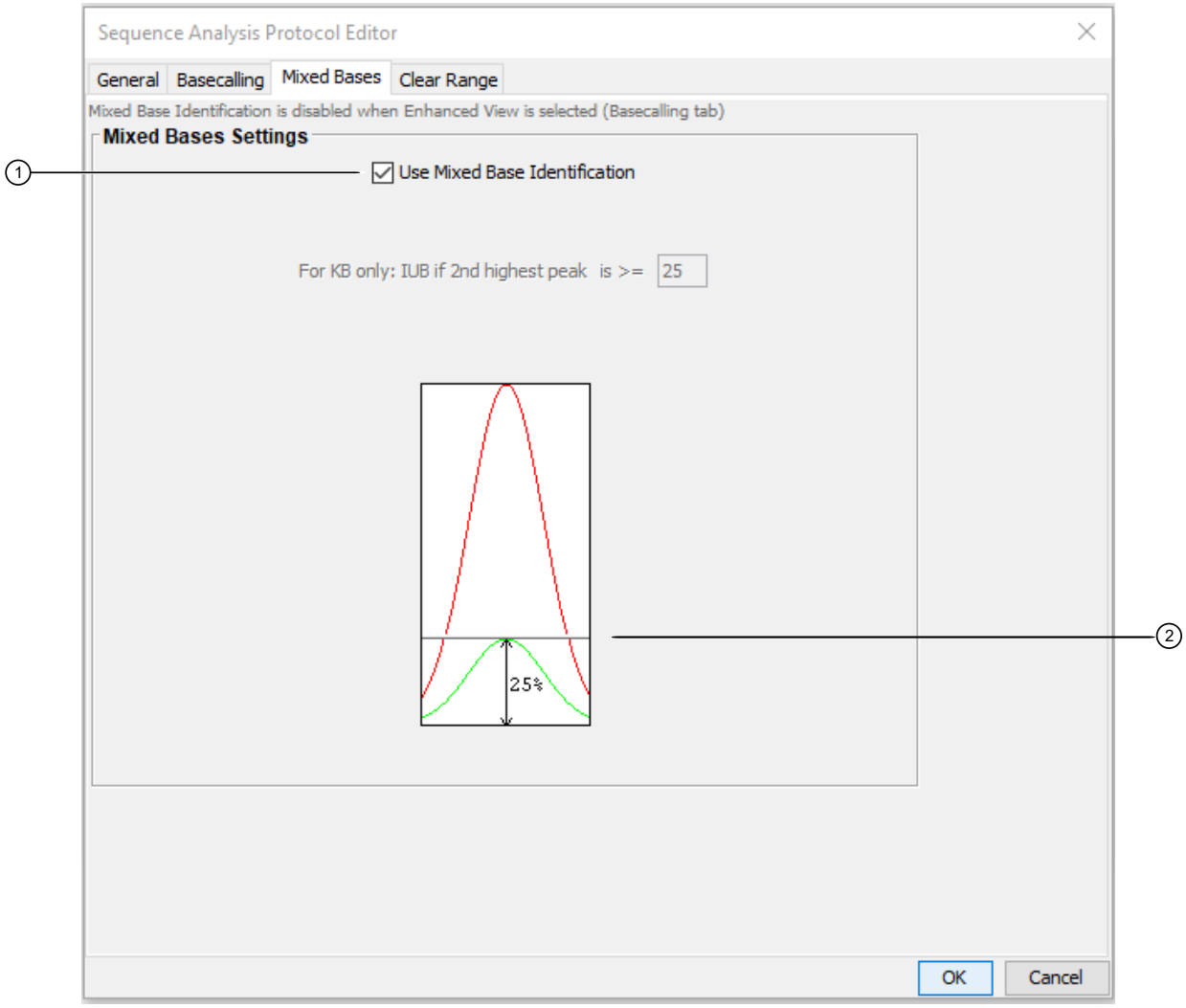
Item	Function
Enhanced View	When **Enhanced View** is selected (Smart Deep™ Basecaller only), the trace resolution is improved relative to the **Classic View**, and the baseline noise is reduced. There is no difference in the bases called or the quality values between the **Enhanced View** and the **Classic View**. By default, if the sample score is below 35, the enhanced trace visualization is disabled. However, you can lower this limit to a minimum sample score of 25. **Note:** There is no difference in the bases called between **Enhanced View** and **Classic View**. When **Enhanced View** is selected, you cannot select the **Use Mixed Base Identification** checkbox in the **Mixed Bases** tab (see “Mixed Bases tab” on page 111).
Disable Enhanced View when Sample Score <=	This feature is availble when **Enhanced View** is selected. This feature disables **Enhanced View** if the sample score is less than the input value (defalut sample score is 35). **Note:** The minimum sample score you can input is 25.

[1] Only available if you select the SDB basecaller in the basecalling tab.

Mixed Bases tab

Mixed bases are one-base positions that contain two bases. These bases are assigned the appropriate two-letter IUB code.

Note: This option should be used with the KB™ Basecaller and Smart Deep™ Basecaller only. The **Use Mixed Base Identification** checkbox is disabled if **Enhanced View** is selected in the **Basecalling** tab for **SDB Trace Visualization.**



1 Checkbox for mixed base identification
2 Slider line

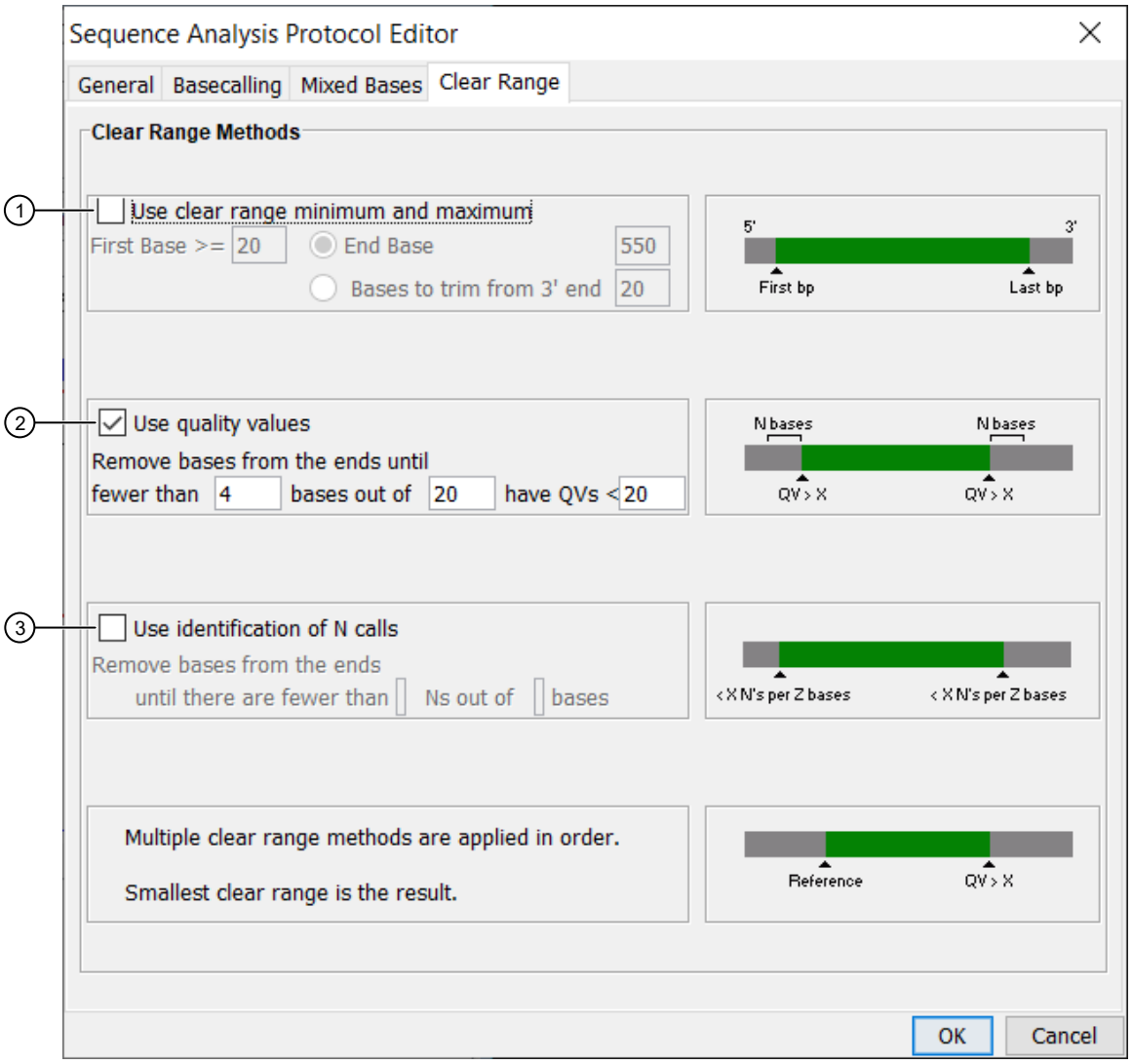
Table 9 Parts of the Mixed Bases Settings section

Item	Description
Use Mixed Base Identification checkbox	• Select the **Use Mixed Base Identification** checkbox for mixed base data. The KB™ Basecaller and Smart Deep™ Basecaller assign A, C, G, T, or an IUB code (R, Y, K, M, S, or W) to every base. • Deselect the Mixed Bases option for pure base data. The KB™ Basecaller and Smart Deep™ Basecaller assign A, C, G, or T to every base. **Note:** The QVs indicate the quality of the basecalls.
Call IUB if 2nd highest peak is >= __	The secondary peak threshold, which is set as a percentage of the primary peak, is used for consideration as a mixed base by the basecalling algorithm. Reaching this threshold is a necessary but not sufficient condition for arriving at a mixed base determination. For KB™ Basecaller, the percentage limit can be set by typing in a value or moving the slider line on the graphic up or down. The default is 25%. For Smart Deep™ Basecaller, the limit is set at 25% and cannot be changed.

Clear range tab

The clear range is the region of sequence that remains after excluding the low-quality or error-prone sequence at both the 5′ and 3′ ends.

Note: You must select post processing (PP) in the **Sample Manager** before you click **(Start Analysis)** for clear range methods to take effect.



① Use with KB, SDB, and ABI
② Use with KB and SDB
③ Use with KB, SDB, and ABI

Table 10 Parts of the Clear Range Methods section

Item	Function
Use clear range minimum and maximum checkbox	When selected, the clear range defined is by the starting base number and last base defined or x number of bases from the 3' end.
Use quality values checkbox	When selected, the range is defined by the QVs when quality values are available. This sets a window with a particular number of allowed low quality bases.
Use identification of N calls checkbox	When selected, the range is between the first and last base defined by a certain number of Ns. This sets a window with a particular number of allowed ambiguous base calls (Ns).

Note: More than one method can be used at a time. The clear range methods are applied in order. The smallest clear range is the result.

Sharing master analysis protocols between data collection software and Sequencing Analysis Software

Master analysis protocols can also be created in the 3730/3730*xl* DNA Analyzer, 3130/3130*xl* Genetic Analyzer, and 3500/3500xL Genetic Analyzer data collection softwares.

The conditions for sharing a master analysis protocol (MAP) between the instrument data collection software and the Sequencing Analysis Software are described in the following table:

Conditions	Result	Recommended action
Sequencing Analysis software installed while data collection was open (proper installation)[1]		
• MAP created in Sequencing Analysis Software • Data collection software is open	MAP is registered in both applications, and is available for use in the data collection software.	—
• MAP created in data collection software • Sequencing Analysis Software is open	MAP is registered in both applications, and is available for use in the Sequencing Analysis Software.	—

(continued)

Conditions	Result	Recommended action
Sequencing Analysis software installed while data collection was closed (improper installation)		
• MAP created in Sequencing Analysis Software or in data collection software • Other software is open or closed	Sequencing Analysis Software was never registered in the Data Service. There is no communication between the software.	1. Uninstall the Sequencing Analysis Software. 2. Open the data collection software. 3. Reinstall, then register the Sequencing Analysis Software.

[1] For proper installation of the Sequencing Analysis Software on a computer that is connected to a 3730/3730*x*/ DNA Analyzer, 3130/3130*x*/ Genetic Analyzer, or 3500/3500xL Genetic Analyzer, the instrument data collection software must be running. For software installation information, see Chapter 2, "Install the software".

Processing parameters

Parameter	Description
BC (Basecalling)	The **BC** parameter is a check box option that performs basecalling on samples when you click (Start Analysis). When you add a sample to the Sample Manager window then the software sets this check box to match the **BC** checkbox in the analysis defaults.
PP (Post processing)	The **PP** parameter is a check box option that performs the post processing on basecalled samples when you click you click (Start Analysis). When the **PP** parameter (post processing) is selected, there is a calculation of the clear range. The clear range is the region of sequence that remains after excluding the low-quality or error-prone sequence at both the 5′ and 3′ ends. If the KB™ Basecaller or Smart Deep™ Basecaller were used for analysis the clear range is calculated from the QVs. If an ABI™ Basecaller was used, the range is calculated from the Ns in the data and/or trim by the number of bases at the start and end of the data.
P (Printing)	The **P** parameter is a checkbox option in the **Sample Manager** window that prints the information, selected in the printing options, for samples after all processing is complete.

Basecalling status indicators

Depending on the basecaller you use, either one of two or one of four of status indicators are displayed in the analysis report to indicate sequence quality.

- ABI™ Basecaller—Green and/or red
- KB™ Basecaller—Green, blue, yellow, or red
- Smart Deep™ Basecaller—Green, blue, yellow, or red

To view the **Analysis Report**, click **Analysis** ▸ **Analysis Report**, or click .

Note: If you use the ABI™ Basecaller, you should carefully review all sequences against the processed electropherograms.

Table 11 Basecalling status indicators

Symbol	Indicates
(Green box)	Successful analysis and data output.
(Blue diamond)	The sample data are sub-optimal or contain anomalous characteristics, but otherwise can still be called. Consult the analysis report for detailed diagnostic messages, and examine the processed electropherogram and basecalls to determine the severity of the problem. Samples in this category still contain long regions of high-quality calls, as indicated by the per base quality values.
(Yellow triangle)[1]	The sequencing experiment fails or is severely compromised in a way that prevents a successful analysis. The software saves the results to the sample file, but the called sequence is replaced with a placeholder sequence of five N's instead of basecalled data. A common reason for this status is a failed sample, where few or no DNA peaks are evident in the raw data. **Note:** if such a failed sample is basecalled using the ABI™ Basecaller, the green status is likely to be returned, but the data are unreliable.
(Red octagon)	The red status typically indicates a software configuration error or invalid input to the basecalling algorithm. No analysis results will be written to the sample file and the file remains in the unanalyzed state. Use the analysis report to diagnose the problem or contact technical support.

[1] The partial output file is available for poor data analyzed with the KB™ basecaller. A partial output file contains five Ns instead of basecalled data.

Checkbox color status

The color of this checkbox indicates the post processing status. The color status is cleared at the start of each new sample processing.

Color	Description
Green	Succeeded.
Red	Failed.
No color	• Has not been started since the sample was added to the **Sample Manager** window. • Was completed previously and is still in the **Sample Manager** window.

P (printing) parameter check box color status

The color of this check box indicates the printing status. The color status is cleared at the start of each new sample processing

Color	Description
Green	Printing succeeded.
Red	Printing failed. Check your printer connections, referring to your printer manual if necessary.
No color	Has not been started since the sample was added to the **Sample Manager** window.

Analysis parameters

About the basecaller file

The basecaller parameter is used to identify bases during the most recent sample analysis. There are three basecallers available for analysis:

- Smart Deep™ Basecaller—Algorithm that calculates mixed or pure bases, and quality values.
- KB™ Basecaller—Algorithm that calculates mixed or pure bases, and quality values.
- ABI™ Basecaller—Algorithm used in previous versions of ABI PRISM™ Sequencing Analysis software.

Note: ABI™ Basecaller is an older algorithm. We recommend that you use the KB™ Basecaller and Smart Deep™ Basecaller algorithms for analysis.

About basecalling

When the **BC** (basecalling) parameter is selected, the selected basecaller performs the following tasks:

- KB™ Basecaller and Smart Deep™ Basecaller
  - If the mixed base option is selected, mixed bases are called.
    Mixed bases are one-base positions that contain two bases. The basecaller assigns A, C, G, T, or an IUB code (R ,Y, K, M, S, or W) to every base.
  - Calls pure bases, if the mixed base option is not selected.
    The basecaller assigns A, C, G, or T to every base.
  - Calculates quality values (QVs) for pure and mixed bases.
  - Call Ns when the quality threshold is not met (if selected).
  - Processes data with true or flat profile.
  - For SDB only, the option to improve trace resolution and baseline with the Enhanced View is available.
- ABI™ Basecaller: Assigns A, C, G, T, or N to every base.

Select the correct basecaller

Basecallers are selected based on a number of factors:

- Instrument model you are using
- Length of your capillary or plates, and the run speed
- Polymer or gel type

Use the table below to determine the correct basecaller file for your data. The basecaller files are listed by instrument model.

Instrument	See
SeqStudio™ Flex Series Genetic Analyzer	"SeqStudio™ Flex Series Genetic Analyzer files" on page 152
SeqStudio™ Genetic Analyzer	"SeqStudio™ Genetic Analyzer files" on page 153
3500/3500xL Genetic Analyzer	"3500/3500xL Genetic AnalyzerRUO files" on page 154
3730/3730*xl* DNA Analyzer	"3730/3730xl DNA Analyzer files" on page 155
3130/3130*xl* Genetic Analyzer	"3130/3130xl Genetic Analyzer files" on page 157
310 Genetic Analyzer	"310 Genetic Analyzer files" on page 157

About the DyeSet/Primer parameter

A DyeSet/Primer file corrects for mobility shifts and color-code changes, depending on which chemistry was used. The default DyeSet/Primer is the file specified in the data collection software.

DyeSet/Primer files are sometimes known as mobility or .mob files. All DyeSet/Primer files have the extension .mob.

IMPORTANT! The DyeSet/Primer file must match the basecaller, genetic analyzer, polymer, and chemistry that you are using.

Note: DyeSet/Primer files are filtered based on the selected basecaller and instrument model.

The DyeSet/Primer File names use a combination of characters to indicate the basecaller, instrument, chemistry, and polymer type. The abbreviations are as follows:

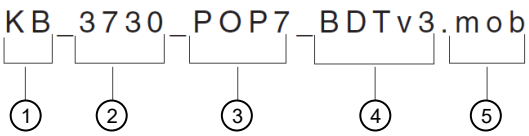


Figure 13 Examples of DyeSet/Primer file naming convention

① Basecaller
② Instrument
③ Type of polymer or gel
④ Chemistry
⑤ File extension

Abbreviation	For Runs Using …
Basecaller	
SDB	Smart Deep™ Basecaller
KB	KB™ Basecaller
DP	Dye primer chemistry and ABI™ Basecaller
DT	Dye terminator chemistry and ABI™ Basecaller
Type of polymer or gel	
POP1	POP–1™ polymer
POP4	POP-4™ polymer
POP6	POP-6™ polymer
POP7	POP-7™ polymer
Chemistry	
BDTv3	v3.1 Terminator chemistry
BDv3	
BDv1	v1.1 Terminator chemistry
BD	
BDv1	
BDTv3direct	BigDye™ Direct chemistry
dRhod	dRhodamine Terminator chemistry
-21M13	Dye primer chemistry – the -21M13 primer is labeled
M13Rev	Dye primer chemistry – the M13Rev primer is labeled

Select a DyeSet/Primer (mobility) file

The basecaller algorithm needs the DyeSet/Primer information to be able to apply the proper mobility shift corrections. If you specified the wrong DyeSet/Primer (mobility) file in the Data Collection software or used a different chemistry than the selected file, then you can select a new mobility file for each affected sample from the drop-down list. After you select a new mobility file, the sample files must be reanalyzed.

DyeSet/Primer files are selected based on a number of factors:

- Basecaller selected
- Instrument model you are using
- Polymer or gel type
- The chemistry you are using (BDTv1.1, BDTv3.1, or BigDye™ Direct)

See the following table for information about selecting the mobility file:

Instrument[1]	See
SeqStudio™ Flex Series Genetic Analyzer	"SeqStudio™ Flex Series Genetic Analyzer files" on page 152
SeqStudio™ Genetic Analyzer	"SeqStudio™ Genetic Analyzer files" on page 153
3500/3500xL Genetic Analyzer	"3500/3500xL Genetic AnalyzerRUO files" on page 154
3730/3730*xl* DNA Analyzer	"3730/3730xl DNA Analyzer files" on page 155
3130/3130*xl* Genetic Analyzer	"3130/3130xl Genetic Analyzer files" on page 157
310 Genetic Analyzer	"310 Genetic Analyzer files" on page 157

[1] Smart Deep™ Basecaller is only available for SeqStudio™ Flex Series Genetic Analyzer, SeqStudio™ Genetic Analyzer, 3500/3500xL Genetic Analyzer, and 3730/3730*xl* DNA Analyzer.

If the file is not present

The **DyeSet/Primer** drop-down list in the **Sample Manager** displays all the DyeSet/Primer files in the Mobility folder.

If the DyeSet/Primer file is not present in the Mobility folder, then analysis using that file is not possible. If the file name is displayed in bold, italicized font in the DyeSet/Primer file field, the file is not present in the Mobility folder and the sample file cannot be processed with that file. To copy DyeSet/Primer files into the Mobility folder of the analysis software, see "Copy and paste mobility files for 310 Genetic Analyzer" on page 134.

The path to the Mobility folder is:

```
C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Mobility
```

About matrix files (310 Genetic Analyzer only)

Although the dyes in a dye set fluoresce at different wavelengths, there is some overlap in the emission spectra of the dyes used. This spectral overlap must be eliminated for proper data analysis.

A matrix file contains a mathematical description of the overlap of the dyes in a given dye set.

Matrix files are only used with the 310 Genetic Analyzer.

If the matrix file is not present in the **Matrix** folder, then analysis using that file is not possible. If the file name is displayed in bold italicized font in the **Matrix** field of the **Sample Manager**, then the file is not in the **Matrix** folder. Refer to and/or to copy the files in the **Matrix** folder of the analysis folder.

The path to the matrix folder is:

```
C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Matrix.
```

Base spacing, peak 1 location, and start and stop points

When you analyze the sample files, the software calculates the base spacing, peak 1 location, and start and stop points. The calculated values can be reviewed and edited in the **Sample Manager** table (see “Review and edit base spacing, peak 1 location, and start and stop points” on page 55).

Parameter	Description
Spacing	Spacing is defined by the number of scan points from the crest of one peak to the crest of the next peak. During basecalling, a spacing calibration curve is applied to the data to determine a base spacing value. If the spacing value could not be determined, then the default value in the basecaller file was used.
Peak 1 location	The first data point that is from the sample not including primer peaks in dye primer chemistries. The value is initially calculated by the software and is the reference point for the spacing and mobility corrections performed by the basecalling algorithm. If the peak 1 location value is wrong due to low signal or other aberrations, the data can show bad spacing or strange mobility shifts.
Start point	The raw data point where the basecalling starts in the sample file. The **Start** point is normally the same as the beginning of the first base peak (peak 1 location). If the raw data immediately after the Peak 1 location is clearly unusable or if you want to analyze only a portion of the raw data, you can change the start point to later in the data. The start point can never be less than the peak 1 location value.
Stop point	The **Stop** point parameter specifies the last raw data point to be included in the basecalling. If the default **Stop** point is used, this endpoint is the last data point in the file. You can basecalling before the last data point if there is clearly unusable raw data at the end of the sample or if you want to analyze only a portion of the raw data in the sample.

Basecalling errors

Sometimes poor results are due to sample file basecalling errors. Common examples of errors that affect basecalling are:

- Incorrect basecaller and/or dyeset/primer used for basecalling
- Wrong peak 1 location and start point calculated by the software
- Incorrect stop point selected
- Bad base spacing
- Poor quality data

IMPORTANT! Results from Smart Deep™ Basecaller are likely to be different from results from KB™ Basecaller. The differences are likely minor, and are observed at the 5' and 3' ends and through regions with dye blobs, N-1 peaks, mobility shifts, and heterozygous insertion deletion variants.

13 Sample data displays

View or edit the sequence data in the **Data Displays** pane in **Sample Manager**.

Tab	Description
Annotation	Summary of the sample information written by the data collection and analysis software (see "About the Annotation tab" on page 123).
Sequence	The nucleotide (base) sequence text called for the data. Gray colored sequence text represents the trimmed bases. This view is available only after basecalling is done (see "About the Sequence tab" on page 126).
Features	The features that were found in the sequence by the post processing (clear range) (see "About the Features tab" on page 128).
Electropherogram	A four-color picture of analyzed data, with peaks representing the bases. The original bases, edited bases or complementary bases can be displayed. This is the default view that is displayed when an sample files are shown and is available only after basecalling is done (see "About the Electropherogram tab" on page 129).
Raw (data)	The raw data collected by the instrument (see "About the Raw tab" on page 131).
EPT	A plot of run voltage, current, power and temperature values (see "About the EPT tab" on page 133).

(continued)

Tab	Description
Audit	Information about modifications to the data (base change, deletion, insertion, change in analysis settings, sample name change). This window contains data only if the Audit Trail feature is activated in the Authentication and Audit tab of the Options dialog box (see "About the Audit tab" on page 133).
Electronic Signature	You can to electronically sign off when saving, printing, or exporting an analysis report. You can also sign at any other time. Each electronic signature record is captured in the **Electronic Signature** tab of **Sample Manager**. See "About the Electronic Signature tab" on page 133.

About the Annotation tab

The **Annotation** tab displays the following data collection and analysis information:

Note: The information in the window can be viewed but not edited.

Field	Description
Trace Identification	
Sample name	Name entered in the **Sample Name** column of the plate record of data collection.
Sample file name	Name of the module file used to run the sample.
Capillary Number	Capillary number used to electrophorese the sample.
Well/Tube Position	Well position of the sample.
Instrument and Data Collection Software	
Instrument Name	Name of the instrument.
Instrument Model	Instrument model used to collect the data.
Data Collection SW Version	Version of software used to collect data.
Insrument Firmware Version	
Data Analysis	
Number of Scans Collected	Range of the collected data points that were used to analyze the data.
Basecaller Name	Name of the basecaller used to analyze the data.
Basecaller Version	Version of the basecaller used to analyze the data.
Dyeset/Primer	DyeSet/Primer file used during analysis to adjust for mobility shifts.

(continued)

Field	Description
Number of Bases Detected	Total number of bases in the sample.
Clear range start	The 5' and the 3' end of the clear range.
Clear range end	
Sample Score (also known as Trace Score)	The average quality value of the bases in the clear range sequence.
Base Call Start	Raw data start and stop points (scan numbers) used for data analysis.
Base Call End	
Matrix Name	Matrix file used to multicomponent the data. Used for 310 Genetic Analyzer data only.
Clear Range Length	The region of the sequence that remains after excluding the low-quality or error-prone sequence at both the 5' and the 3' ends.
Peak 1 Location	Peak 1 is the raw data point (scan number) where the analyzed data starts.
Ave Signal Intensity	Average fluorescent intensity (signal) for each dye used to label all of the 'A', 'C', 'G', or 'T' bases in a sample.
Noise	Average background fluorescent intensity for each dye.
Signal:Noise	Signal to noise ratio: The average of the signal intensity of the 'A', 'C', 'G', or 'T' base divided by the average of the noise for that base.
Base Spacing Used	Base spacing as calculated by the basecaller or defined by the user.
Base Spacing Calculated	Base spacing as calculated by the basecaller.
QV20+	The number of bases with QVs greather than or equal to 20.
Run Configuration	
Run Protocol	Chemistry, cappilary array length, and DyeSet/Primer file
Run Protocol Version	Version of the run protocol used.
Run Name	Name of the run protocol used.
Run Start Date & Time	Date and time the run started and finished.

(continued)

Field	Description
Run Stop Date & Time	Date and time the run started and finished.
Data Collection Start Date & Time	Date and time the data collection started and finished.
Data Collection Stop Date & Time	
Sampling Rate (Hz)	Data sampling rate.
# Channels Averaged	Average number of dye channels.
Run Module	Module file used to run the sample.
Plate Creator Name	Name of the pate creator.
Plate Name	Name of plate used in run.
Plate ID	Assigned plate identification name.
Plate Type	Plate format (number of wells).
Number of Wells	Number of wells in the plate.
Results Group Name	Assigned group name.
Reinjection	Number of reinjections performed.
Analysis Protocol Name	Analysis protocol last used to analyze the data.
Analysis Protocol Settings Version	Version number of the analysis protocol last used to analyze the data.
Injection Time Setting (Sec)	Injection time setting used on genetic analyzer.
Injection Voltage (Volts)	Injection voltage setting used on genetic analyzer.
Electrophoresis Voltage Settings (Volts)	Electrophoresis voltage setting used on genetic analyzer.
Laser Power Settings (micro Watts)	Laser power setting used on genetic analyzer.
Oven Temperature Settings (deg C)	Oven temperature setting used on genetic analyzer.
Consumables	
Length to Detector	Length of capillary arrays from sample plate to the detector.
Polymer/Gel Type	Polymer/gel type used.
Polymer Lot number	Polymer lot number.
Polymer Expiration date	Expiration date of polymer.

(continued)

Field	Description
Dye Set Name	Name of DyeSet.
Number of Dyes	Number of dyes.

About the Sequence tab

The **Sequence** tab displays the nucleotide sequence called for the selected data file.

- The wide center column contains the sequence data.
- The left and right columns show the base positions at the beginning and end of each row.
- The gray text is data that is outside the clear range.

Note: If the data has not been basecalled, the **Sequence** window is empty.

To edit bases in the **Sequence** tab, see “Edit base calls” on page 77.

Viewing

To display the sequence view, select and show sample(s) in the **Sample Manager**, then select the **Sequence** tab.

Note: If the data has not been basecalled, the Sequence window is empty.

The information in the window can be edited. If you edit a base, then switch to the Electropherogram view, the software displays the same base location.

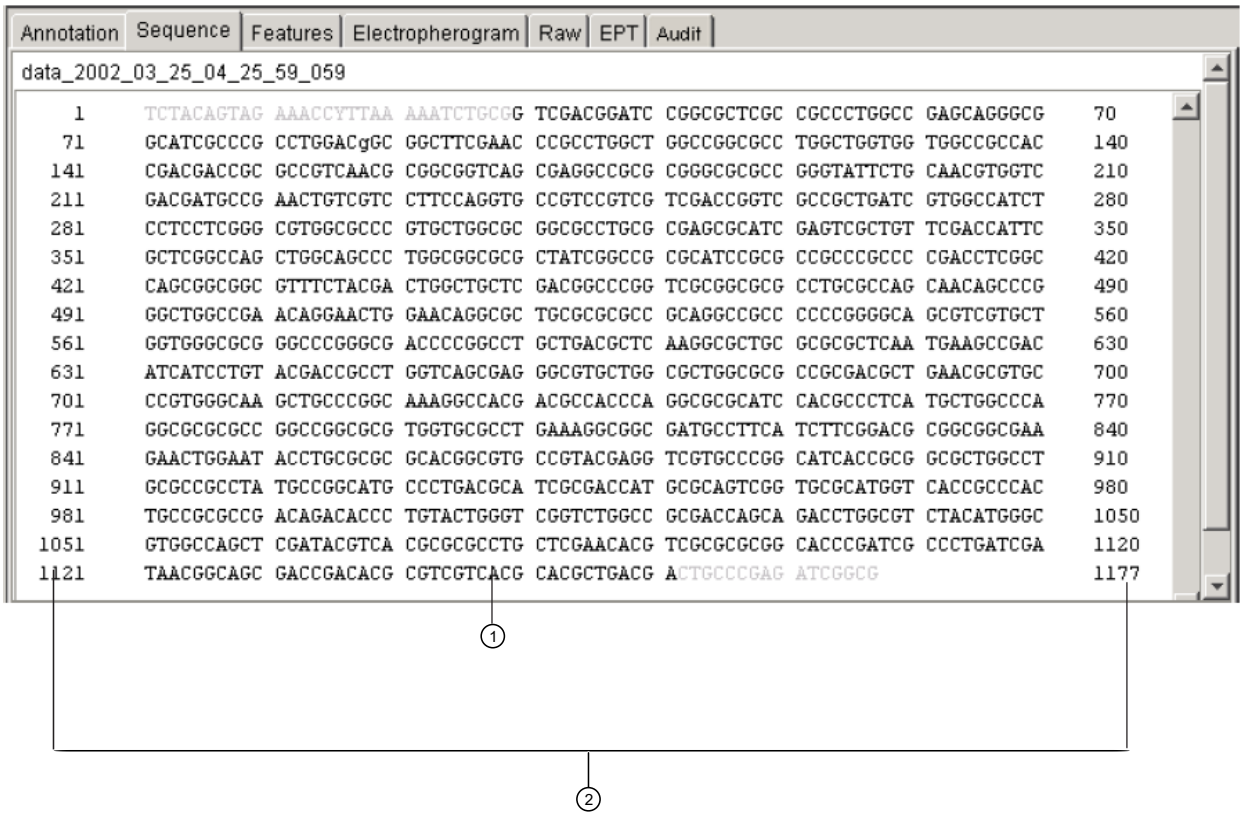


Figure 14 Single sample in the Sequence view

1. Center columns contain sequence data
2. Left and right columns show the base positions at the beginning and end of each row

① Scroll bar to view a single sample
② Scroll bar to view other samples

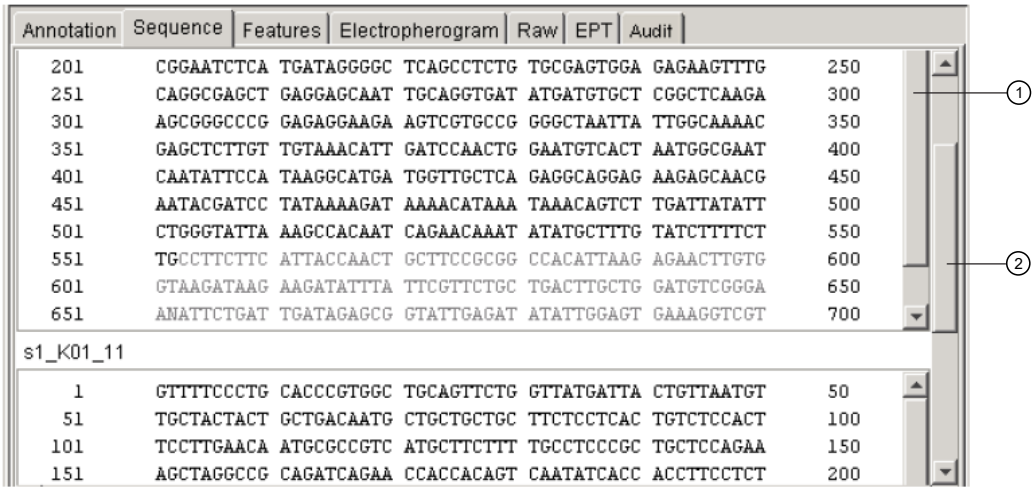


Figure 15 Multiple samples in the Sequence view

Finding, editing, displaying, and printing

To …	See …
Search for a base character, a range of bases, or a specified base pattern using the Find command (Ctrl+F) and Find Again (Ctrl+G)	“Search for a pattern in a sequence” on page 80.
Edit the sequence using any of the standard Windows™-based computer commands	“Edit base calls” on page 77.
Display reverse complement	“Show the reverse complement data” on page 81.
Display quality values and numbers	“Display sample Quality Values” on page 75.
Print the contents of the window	“Manually print the sample file” on page 50.

About the Features tab

The **Features** tab shows features that were added to the analyzed sequence data by the post processing.

Note: If the sequence data has not been post processed, the **Features** tab is empty.

The information in the **Features** tab can be viewed but not edited.

About the Electropherogram tab

The electropherogram is a four-color display of the analyzed sample data, with peaks representing the bases called for the sample. This is the default view for all samples and is editable.

To display the **Electropherogram** view, select and show sample(s) in the **Sample Manager**, then select the **Electropherogram** tab.

Note: If the raw data has not been analyzed, the **Electropherogram** window is empty.

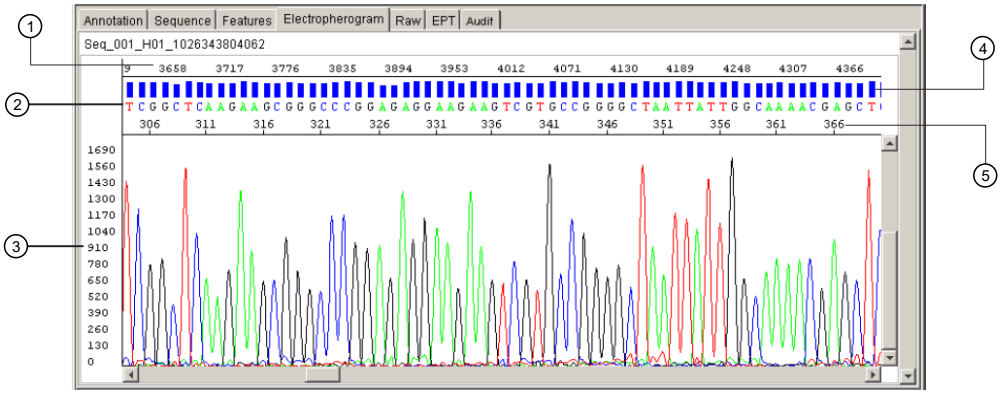


Figure 16 Single sample in the Electropherogram view

1. Spacing-corrected scan number
2. Called bases
3. Normalized fluorescence intensity
4. Quality values
5. Number of the base at this location

Feature	Description
Trace and base colors	The trace lines and the letters above the peaks are colored to represent the four bases. The default color for each base are as follows: • C: Blue • G: Black • A: Green • T: Red The colors the represent the bases can be changed (see “Edit the display settings” on page 87).
Quality Values	The quality value (QV) is a per-base estimate of the basecaller accuracy. The QVs are optionally displayed as bars above each base in the sample. The height and color of the bar indicates its value. The taller the bar, the higher the QV. For more information, see “About Quality Values” on page 95.
Clear range	The clear range is the region of sequence that remains after excluding the low-quality or error-prone sequence at both the 5′ and 3′ ends (see “About the clear range” on page 94). The bases outside the clear range cannot be edited. To edit the clear range, see “Edit the clear range” on page 78.

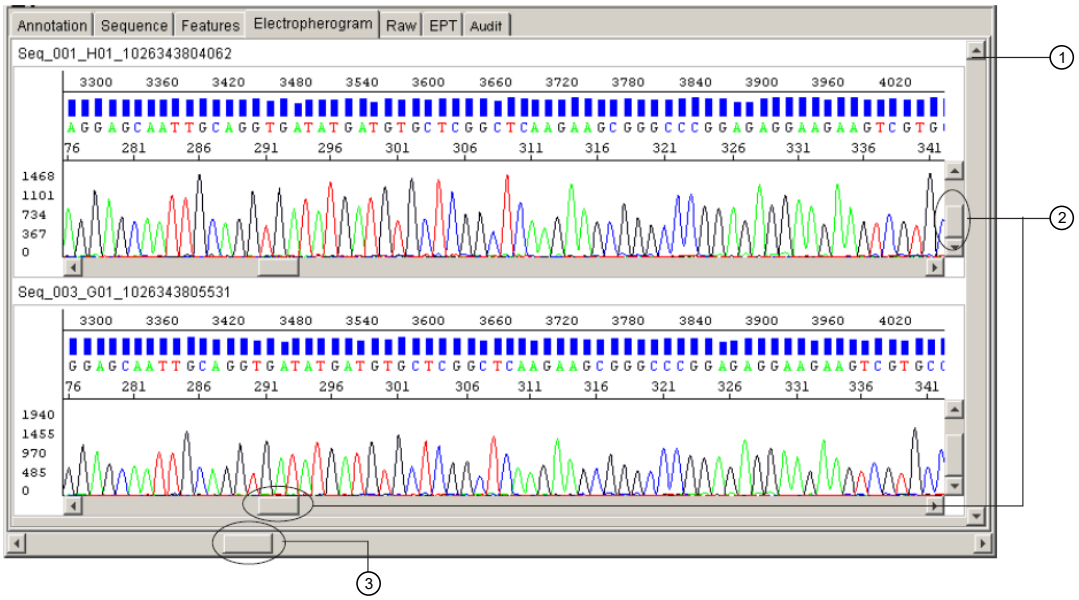


Figure 17 Multiple samples in the Electropherogram view

① Scroll bar to view other samples
② Scroll bar to view a sample
③ Scroll bar to scroll the stack of multiple samples

Note: The pane height of the samples can be lowered so that more samples can be viewed simultaneously. For more information, see Chapter 13, “Sample data displays”.

About mixed bases

Mixed bases are one-base positions that contain two different bases. Mixed base identification is only available if using the Smart Deep™ Basecaller or KB™ Basecaller algorithms.

Mixed base identification is only used if selected in the Analysis Protocol (see “Create or edit an analysis protocol” on page 39):

- If **Use Mixed Base Identification** is selected for mixed base data, then the basecaller assigns A, C, G, T or an IUB code (R, Y, K, M, S, or W) to every base.
- If **Use Mixed Base Identification** is not selected (pure bases), then the basecaller assigns A, C, G, or T to every base.

Note: If the Analysis Protocol is set to assign ’N’ to bases below a designated Quality Value (QV), then N's will be called.

About the Raw tab

The **Raw** tab shows the raw sample data, before any processing is performed. To display the raw data plot, select and show one or more samples in the Sample Manager, then select the **Raw** tab.

The following information can be reviewed in the **Raw** tab:

- Scan numbers used by the software to start and stop basecalling.
- Relative true peak intensities and peak resolution before the smoothing applied by the software.
- Problems or noise in the baseline (for example, electronic spikes in the data or unusual baseline levels) that could result in poor basecalling or could indicate instrument problems.
- Estimate base spacing by measuring the scan points that define two adjacent peaks.

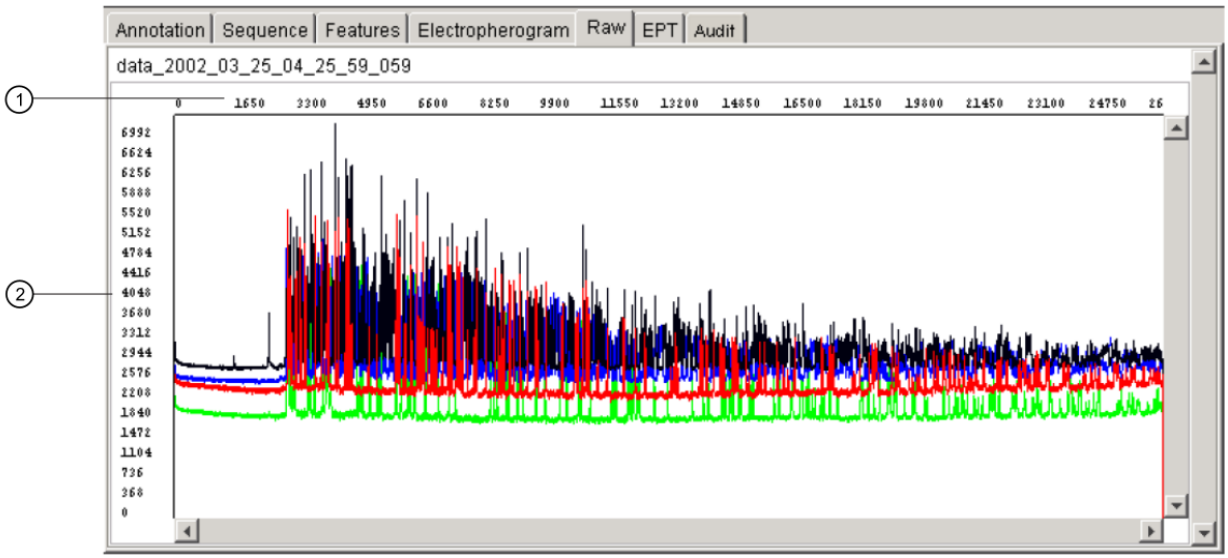


Figure 18 Single sample in the Raw view

① Scan numbers
② Relative peak amplitude (signal intensity)

When raw data is displayed at maximum magnification, four apparent data points (pixels) are displayed for each scan number. The fourth data point is the true intensity value for the scan number. The other three pixels simply fill in the line between the true data points.

The colors shown for raw data in the array view in the data collection software reflect the dyes used to label the bases. The colors represent each base differently for each chemistry.

For *unanalyzed data*, the four colored trace lines represent the fluorescence data from the four fluorescent dyes. The base represented by each color depends on the chemistry.

Table 12 Raw data color display for each chemistry

Color	BigDye primer base	dRhodamine terminator base	BigDye base
Blue	C	G	G
Green	A	A	A
Yellow	G	C	T
Red	T	T	C

After the DyeSet/Primer file has been applied to the sample during analysis, the colors correspond to the bases, as follows.

Color	Base
Blue	C
Green	A
Black	G
Red	T

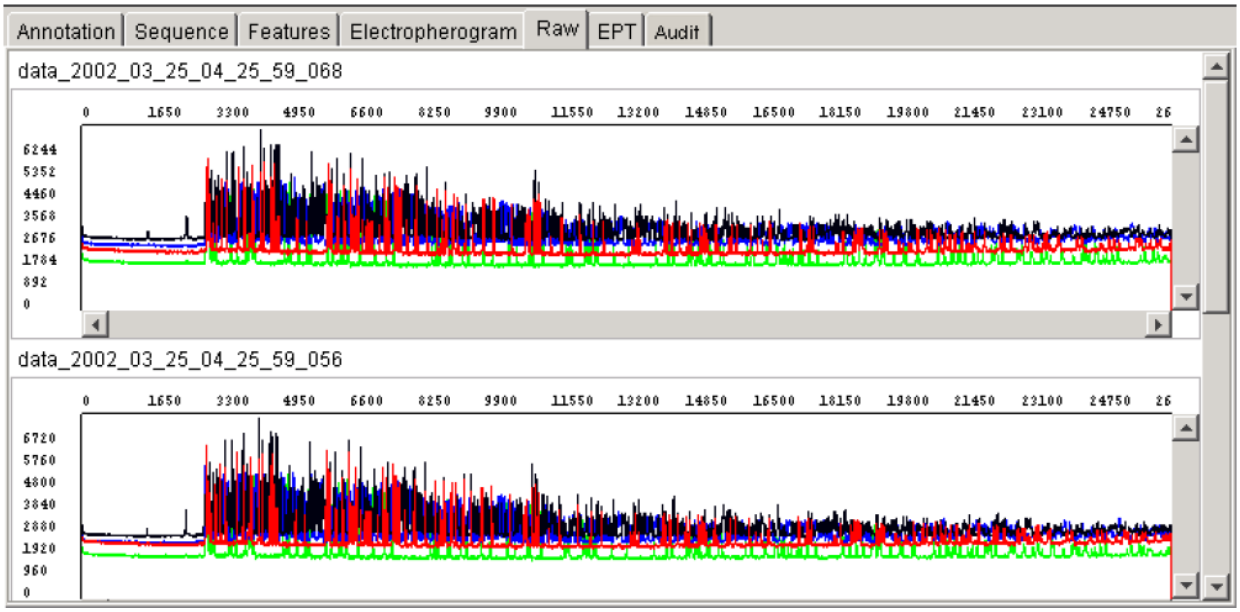


Figure 19 Multiple samples in the Raw view

About the EPT tab

Use the **EPT** tab to review power, temperature, and voltage values after power fluctuations or failures.

The units and default colors used in the EPT view are summarized in the table below.

Measurement plotted	Default color	Unit
Voltage	Blue	V/100
Current	Green	µA
Power	Black	mW × 10
Temperature	Red	℃

About the Audit tab

Review the following data in the **Audit** tab:

- Events (changes) and the reason for the change:
  - Base changes, deletions or insertions
  - Changes made to the analysis settings
  - Changes made to the sample name
  - Analyzing the data
- The description of the change.
- The user ID and name of the person making the change
- The date the change occurred.

About the Electronic Signature tab

Use the **Electronic Signature** tab to electronically sign off when saving, printing, or exporting an analysis report. You can also sign at any other time. Each user can configure their electronic signature for various reasons. The reason is displayed in the **Meaning** column of the **Electronic Signature** tab.

The following information is displayed for each electronic signature:

- **Action Name**
- **User ID**
- **User Name**
- **Date Signed**
- **Meaning**
- **Comment**

See “Enable and set up electronic signature function” on page 26 for information of configuring electronic signature for each user.

14 Copy 310 Genetic Analyzer Matrix and DyeSet/Primer files

Note: Skip this chapter if you are analyzing sample files not generated on the 310 Genetic Analyzer.

If you are upgrading from an older version of the software, then you need to copy the 310 Genetic Analyzer matrix and DyeSet/Primer (mobility) files to new folder locations.

Copy and paste matrix files for 310 Genetic Analyzer

1. Navigate to the location of the matrix files in the older version of the software:
   `C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Matrix`
2. Copy the matrix files (.mtx) for the 310 Genetic Analyzer.
3. Paste the matrix files into the folder used by the current version of the software:
   `C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Matrix`

Copy and paste mobility files for 310 Genetic Analyzer

1. Navigate to the location of the mobility files in the older version of the software:
   `C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Mobility`
2. Copy the mobility files (.mob) for the 310 Genetic Analyzer .
3. Paste the mobility files into the folder used by the current version of the software:
   `C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Mobility`

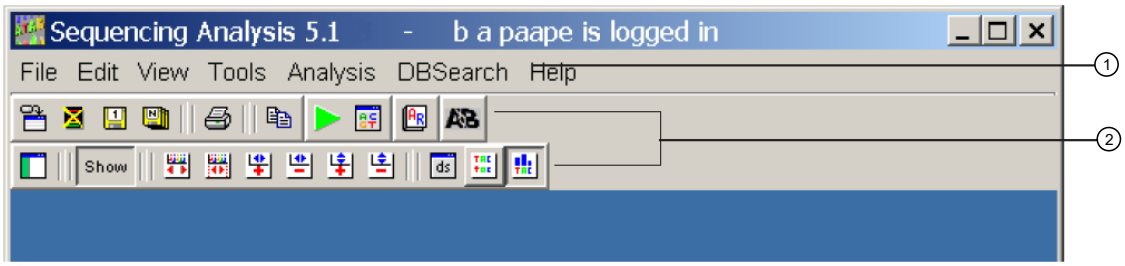
Restart the software

If the software was running while the matrix and mobility files were copied, restart the software. The software read the contents Matrix and Mobility folders at start up only.

A Screen descriptions

Navigating within the software

The software homescreen contains a menu bar and a toolbar. All operations are performed within this screen.



① Menu bar
② Toolbar

For information on menu commands and toolbar icons, see “Menu bar” on page 135 and “Toolbar buttons” on page 139.

Menu bar

File menu

Command	Description
Add Sample(s) (Ctrl+I)	Opens the **Add Sample(s)** dialog box
Remove Sample(s) (Delete)	Removes the selected samples from the **Sample Manager/Navigator**
Remove All Samples	Removes all samples from the **Sample Manager/Navigator** without selecting them
Save Sample(s) (Ctrl+S)	Saves pending changes to selected sample(s)
Save All Sample(s) (Ctrl+Shift+S)	Saves pending changes to all sample(s) in the **Sample Manager**

(continued)

Command	Description
Export Report	Opens the **Export Analysis Report** dialog box to save the report as a tab-delimited file
Page Setup	Opens the **Page Setup** dialog box, which allows you to set the preferences for printing
Print (Ctrl+P)	Opens the **Print** dialog box, which allows you to start printing
Exit (Alt+F4)	Exits the software program

Edit menu

Command	Description
Copy (Ctrl+C)	Copies the selected item in the window to the clipboard
Select All (Ctrl+A)	Selects the entire contents of an active sample window
Find (Ctrl+F)	In the **Electropherogram** or **Sequence** view, searches for a specific base or string of bases
Find Again (Ctrl+G)	In the **Electropherogram** or **Sequence** view, searches for the next occurrence of the string specified in the **Find** dialog box
Fill Down (Ctrl+D)	Copies the value in the top most selected field to all other selected fields in the same column

View menu

Command	Description
Sample Manager/Navigator (Ctrl+N)	Toggles between **Sample Manager** and **Sample Navigator** panes
Show/Hide Data Displays (Ctrl+U)	Shows/hides the selected sample data views
Full View (Ctrl+[)	Displays all the data in a standard size window
Default Size (Ctrl+])	Restores the display to the initial default zoom factor
Zoom In Horizontal (Ctrl+=)	Enlarges the view horizontally so that more detail is visible
Zoom Out Horizontal (Ctrl+Minus)	Reduces the view horizontally so that a larger area is visible

(continued)

Command	Description
Zoom In Vertical (Ctrl+Shift+=)	Enlarges the view vertically so that more detail is visible
Zoom Out Vertical (Ctrl+Shift+Minus)	Reduces the view vertically so that a larger area is visible
Show/Hide Original Sequence (Ctrl+J)	Displays the original basecalls on a separate line above the editable basecalls in the **Electropherogram** view
Show/Hide Quality Bars (Ctrl+K)	Shows/hides the sample quality values

Tools menu

Command	Description
Set Clear Range (Ctrl+Q)	Opens the **Set Clear Range** dialog box which allows you to define the beginning and ending base pair (bp) for the clear range.
Reverse Complement	Displays the complement of a sequence in all sequence views.
Make Matrix	Opens the **Make Matrix** dialog box to create matrix files for the 310 and 377 instruments
Electronic Signature	Opens the **Electronic Signature Verification** dialog box for electronically signing or verifying a sample.
Change Password	Opens the **User Management** dialog box for changing the user's password.

Analysis menu

Command	Description
Start Analysis (Ctrl+R)	Starts the analysis of samples.
Analysis Protocol (Ctrl+T)	Opens the **Analysis Protocol** dialog box which allows you to edit an analysis protocol which defines basecalling, mixed bases and clear range settings for a selected sample in the **Sample Manager**.
Analysis Protocol Manager	Opens the **Analysis Protocol Manager** dialog box which allows you to create, edit, apply, or delete an analysis protocol.
Analysis Defaults	Opens the **Analysis Defaults** dialog box.
Apply Pre-Analysis Settings	Applies original analysis settings to sample.

(continued)

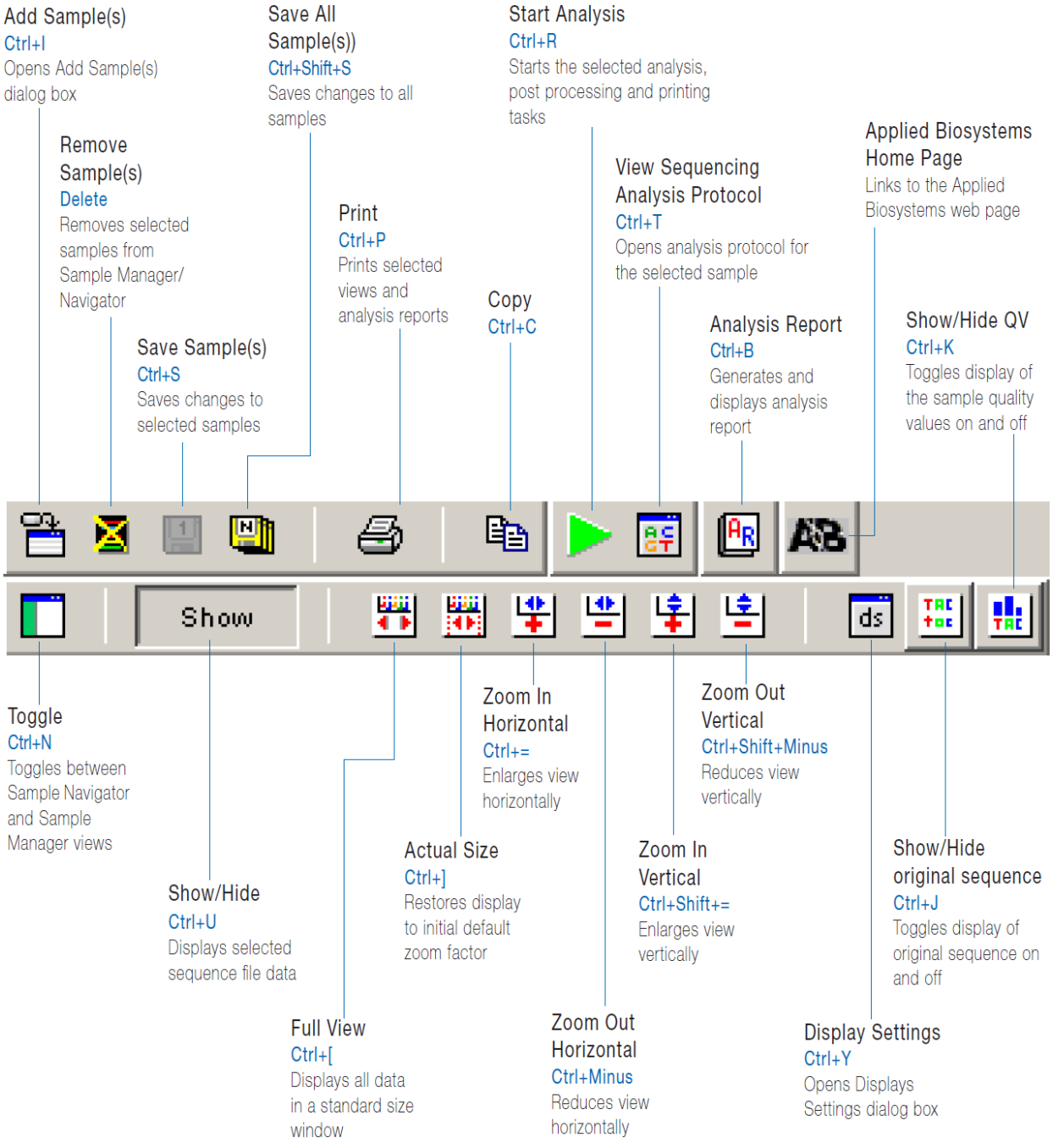
Command	Description
Analysis Report (Ctrl+B)	Opens the analysis report.
Display Settings (Ctrl+Y)	Opens the **Display Settings** dialog box.

Help menu

Command	Description
Sequencing Analysis Software 8 Help (F1)	Opens a PDF file of the user guide.
Sequencing Analysis Software 8 User Bulletin	Opens a PDF file of the user bulletin.
IUPAC Codes	Opens a display box containing a table of the codes for single and multiple bases.
IUPAC Diagram	Opens a display box containing the IUPAC diagram.
Complement Codes	Opens a display box containing a table of the one letter bases and their complements.
Quality Values Chart	Opens a display box containing a table of the quality values.
About Sequencing Analysis	Opens the **About Sequencing Analysis** dialog box.

Toolbar buttons

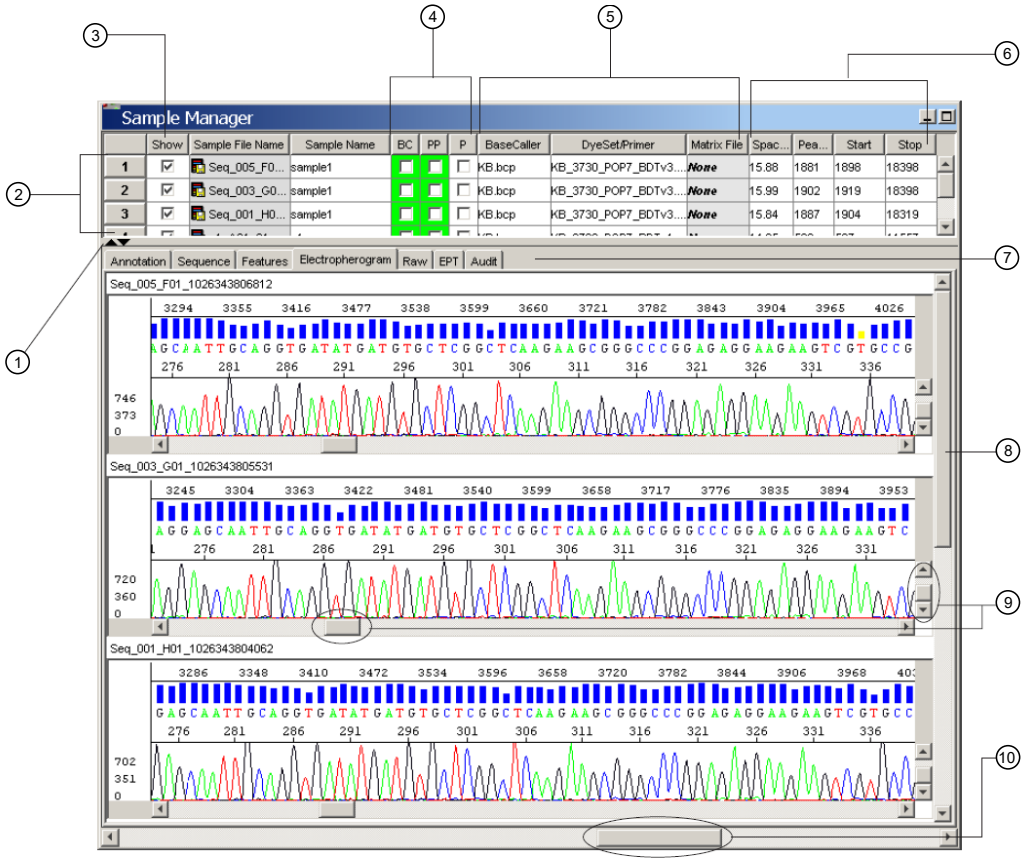
The most frequently used commands are available for quick access in the main window toolbar. They are active only when appropriate.



About the Sample Manager screen

The **Sample Manager** screen is the default view when the software is launched. Samples are viewed, analyzed, edited, printed and save in the **Sample Manager**.

For information on the **Sample Manager** screen, see Table 13.



1. Split bar
2. Samples
3. **Show** check box
4. Processing parameters
5. Analysis parameters
6. Calculated results
7. Display tabs
8. Outer scroll bar
9. Inner scroll bar
10. Bottom scroll bar

Table 13 Sample Manager screen

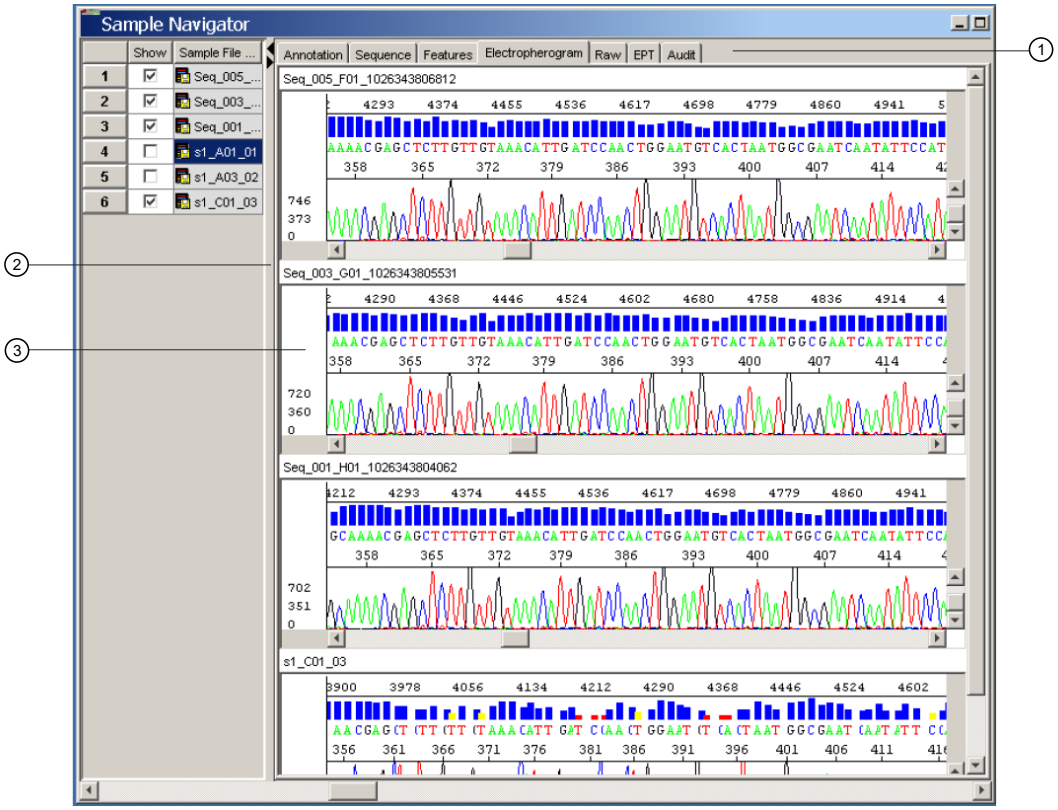
Feature	Description
Split bar	Use the split bar to adjust the size of the **Sample Manager** pane and the **Sample View** pane.
Samples	Displays the added samples.
Show check box	Use the **Show** check box to display sample data.
Processing parameters	Use the check boxes to select or deselect the **BC**, **PP**, and/or **P** parameters, as appropriate. • **BC**—Basecalling • **PP**—Post processing • **P**—Printing
Analysis parameters	Analysis parameters can be changed in **Sample Manager** or in **Analysis Protocol**. **Matrix File** is used for 310 and 377 data only.
Calculated results	Displays the analysis results.
Display tabs	Use tabs to view data in the sample view pane.
Outer scroll bar	Scroll bar to view other displayed samples.
Inner scroll bar	Scroll bar to view within a single sample.
Bottom scroll bar	Scroll bar to scroll the stack of multiple samples.

To switch to the **Sample Navigator** view, click  , or select **View** ▸ **Sample Manager** or **Sample Navigator**.

About the Sample Navigator view

The **Sample Navigator** view displays the sample names data. Samples can be edited and saved in this view.

To switch to the **Sample Navigator** view, click , or select **View ▸ Sample Navigator**.



① Sample view tabs
② Split bar
③ **Sample Navigator** pane

Sample view pane

The sample view is used for viewing all the data characteristics of the sample. Each sample view tab displays information below. The same view tabs can be used to view multiple samples simultaneously.

Table 14 Sample Views Tabs

Tab	Displayed Information
Annotation	Information about the data and its analysis.
Sequence	Sequence of the sample. Quality bars and values, and original data displays are optional. For readability the display clusters the sequences into substrings of 10 characters each, separated by blanks. Use the **Show/Hide QV Bars** button to display/hide QVs.
Features	Calculated clear range.
Electropherogram	Electropherogram and basecall data for the sample. Quality bars and values, and original data displays are optional. The data excluded from the clear range is shown in gray. Use the **Show/Hide QV** button to display/hide QVs.
Raw	Raw data collected by the instrument.
EPT	Volts, watt, current and temperature data collected by the instrument.
Audit	Information about modifications to the data (base change, deletion, insertion, change in analysis settings, sample name change). This window contains data only if the Audit Trail feature is activated. Click **Tools** ▸ **Options**. In the **Audit** pane, select the **Audit Trail On** checkbox.

Note: For unanalyzed samples, only the **Annotation**, **Raw**, and **EPT** tabs contain information.

For more information on samples, refer to “About sample files” on page 11

Analysis report screen description

Whenever sample files are in the Sample Manager, an Analysis Report can be generated to show the status of the samples in the Sample Manager.

To display the report, click or click **Analysis ▸ Analysis Report**.

An **Analysis Report** is displayed. It shows the success and/or failure of the data analysis. There are multiple parts to an analysis report:

- Summary – Contains a summary of all samples in the report
- Length of Read (LOR) – Contains LOR for each sample
- Sample Details – Contains list of each sample, basecalling status, and its associated quality values and sample score. A yellow triangle or red stop sign in the BC Status column indicates failed analysis. A blue diamond in the BC Status column indicates successful basecalling with some anomalies. The blue, yellow, and red icons are hyperlinked to a specific error message in the error table, as well a sample file name hyperlinked to the Sample Manager.
- Errors – Lists the errors that occurred during analysis of sample files

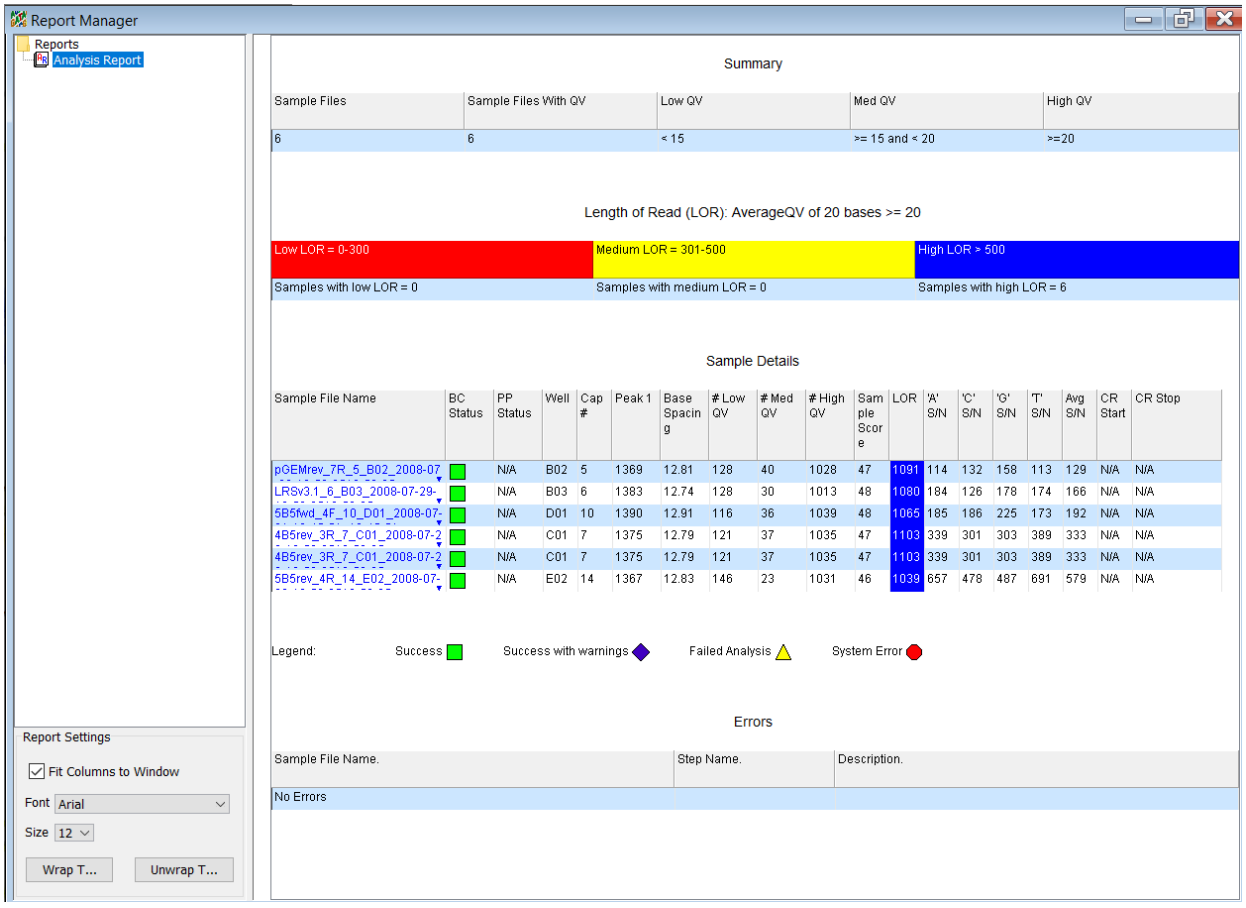


Figure 20 Analysis report example

For more information, refer to Chapter 8, “Review the Analysis Report”.

B

Frequently asked questions

This appendix provides an answer to the most commonly asked questions regarding the ABI PRISM™ Sequencing Analysis Software v5.1.

General questions and answers

Table 15 General Questions and Answers

Question	Answer
What is Sequencing Analysis software?	Sequencing Analysis software is a multi-purpose software used to view, display, analyze, edit, save, and print sequencing files.
How does Sequencing Analysis Software 8 differ from previous versions?	Novel Smart Deep™ Basecaller, in addition to KB™ Basecaller and ABI™ Basecaller. See “Smart Deep™ Basecaller overview” on page 10. A 90-day Smart Deep™ Basecaller demonstration license is included with the Sequencing Analysis Software 8.

Table 15 General Questions and Answers *(continued)*

Question	Answer
What is Smart Deep™ Basecaller?	A new basecalling algorithm in Sequencing Analysis Software responsible for basecalling, identifying mixed bases, and generating per-base quality values (QVs). For more information on its improvements relative to KB™ Basecaller, see "Smart Deep™ Basecaller overview" on page 10. For more information on purchasing and activating the Smart Deep™ Basecaller license, see Chapter 4, "Activate the Smart Deep™ Basecaller license". For more information on the **SDB Trace Visualizaztion** feature, see "Basecalling tab" on page 105.
What is the KB™ Basecaller?	A basecalling algorithm in Sequencing Analysis Software responsible for basecalling, identifying mixed bases, and generating per-base quality values (QVs).
What is an ABI™ Basecaller?	A basecaller that uses the algorithms used in previous versions Sequencing Analysis Software (v3.7 and earlier).
What is FASTA format? How can I convert non-FASTA files into the correct format?	A sequence in FASTA format begins with a single-line description, followed by lines of sequence data. The description line is distinguished from the sequence data by a greater-than (>) symbol in the first column. When creating a file in Microsoft™ Word™, be sure to save it in text only format (line breaks are OK but spaces are not OK). >HumMitoCamb from 15871 to 450 (hard return) aatactcaaatgggcctgtccttgtagtataaactaatacac cagtcttgtaaaccggagatgaaaacctttccaaggacaa atcagagaaaaagtctttaactccaccattagcacccaaagct (hard return)
What is a .phd.1 file and how can it be opened?	A Phred file contains a header with a data description, revised base calls, assigned quality values, and peak location. The file can be opened with any text editor.

Table 15 General Questions and Answers *(continued)*

Question	Answer
What is a .scf file and how can it be opened?	A standard chromatogram format (.scf) file format that is compatible with Staden package. **Note:** When standard chromatogram file format is created, the .scf extension is not appended to the file name. However, the file format is correct.
What ABI instruments can I use to generate data for Sequencing analysis?	Sequencing Analysis software will analyze and post process sequence files generated from the 3730/3730*x*/ DNA Analyzers, 3100/3100-Avant™ Genetic Analyzers, 3130/3130*x*/ Genetic Analyzers, 3500/3500xL Genetic Analyzers, and 310 Genetic Analyzers.
What are the minimum computer requirements for Sequencing Analysis software?	Recommended system requirements for Windows™ 10: • CPU:Intel™ Core™ 2 Duo Processor, 3.0 GHz • Memory: 4 GB • Disk space: 1 GB for application on the drive for the software Minimum system requirements for Windows™ 10: • CPU: 2 GHz or faster Intel™ processor • Memory: 2 GB • Disk space: 1 GB for application on the drive for the software
What monitor resolution I should have?	• Minimum: 1280 x 1024 pixels • Maximum: 2048 x 1152 pixels • Display scale should be set at 100%. **Note:** Any resolution higher than the recommended configuration may result in display of smaller icons and fonts.
What kind of performance can I expect from my Sequencing Analysis software?	Performance depends on the computer specifications.
Do I need ABI PRISM™ SeqScape™ software if I have Sequencing Analysis software?	Sequencing Analysis software is a multi-purpose software used to view, display, analyze, edit, and print sequencing files. SeqScape™ software is designed specifically for sequence comparison. Sequencing Analysis software should be used in every laboratory for general troubleshooting and viewing of data.

Table 15 General Questions and Answers *(continued)*

Question	Answer
Can I BLAST® against a database?	To search a database using a sequence generated with Sequencing Analysis software, create a FASTA file from the data. Open this file in a text viewer, then cut and paste the sequence you would like to search for in your BLAST® query.
Can the Sequencing Analysis software perform just the basecalling for samples?	Yes. Samples can be basecalled only.

Sample manager questions and answers

Table 16 Sample Manager Questions and Answers

Question	Answer
What is the Sample Manager?	Sample Manager allows you to analyze, view, and edit data.
How do I add samples?	Add sample files from any local or networked directory using the Add Samples dialog box.
How can I remove samples?	Select the item to be deleted and click the **Remove Samples** button in the toolbar.
Can I add samples from a CD?	Yes, but the files are read-only. To change the read-only attribute: **1.** Copy the files to the hard disk. **2.** Select the sample names, then right-click and select **Properties**. **3.** Deselect the Read-only check box, then click **OK**.
What can I export from Sequencing Analysis software?	Analysis reports. Reanalysis .ab1 files for samples that have been reanalyzed are saved and available only if you select a sample file in the **Sample Manager**, then click  .
What are Display Settings?	Display settings control the font styles and colors for bases, electropherogram display, axis scale, and length of read on the analysis report.
How do I begin analysis?	Select **Analysis** ▸ **Start Analysis**, or click .
How does reverse complement or show original bases work?	Reverse complement or show original bases apply to selected samples in the manager.

Sample file questions and answers

Table 17 Sample File Questions and Answers

Question	Answer
How does editing affect my data? What gets updated?	Editing does not affect raw data, only the results of basecalling. See “Save sample files” on page 61.
How can I distinguish between edited and non-edited data?	When a base is edited, it displays in lower case while the unedited bases are displayed in upper case letters.
What will happen to my edited sequence when I start analysis?	Once basecalling begins, all current edits will be overwritten.
How do I remove unwanted spaces in my samples?	To remove unwanted spaces in the sample, double-click on the space and press the delete or backspace key.
What can I do if I deleted too many bases?	Start the analysis over or close the file without saving it.

Analysis protocol questions and answers

Table 18 Analysis protocol questions and answers

Question	Answer
What is an analysis protocol?	An analysis protocol contains all the settings necessary for analysis, and is used to perform basecalling and post processing.
What are mixed bases?	A single base position that contains two bases.
What is clear range?	The region of sequence that remains after excluding the low-quality or error prone sequence at both the 5′ and 3′ ends.
What is the LOR?	The length of read (LOR) is the usable range of high quality or high accuracy bases, as determined by the sample quality values. The range is user defined.
Can an analysis protocol in use by data collection be deleted?	Yes.
How can I select the Enhanced View and identify mixed bases?	When you select the Enhanced View, Mixed Base Identification is disabled. You can select either the Classic View and Mixed Base Identification, or Enhanced View and Pure Basecalling only.

Quality values (QVs) questions and answers

Table 19 Quality Values Questions and Answers

Question	Answer
What are QVs?	A quality value is a per-base estimate of the base calling accuracy.
What is the QV equation?	Per-base QVs are calibrated on a scale corresponding to: $QV = -10\log_{10}(Pe)$ where *Pe* is the probability of error. See "Interpretation of Quality Values" on page 96.
What happens to the QVs when I edit bases?	They change depending on what you do: • 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
How can I change the display settings for the QVs?	See "Adjust Quality Value ranges" on page 76.
What is the Sample Score?	A sample score is generated from QVs. It is the average quality value of the bases in the clear range sequence for that sample.

Analysis report questions and answers

Table 20 Analysis Report Questions and Answers

Question	Answer
What is the analysis report?	The analysis report shows the success and/or failure of the data analysis. The report can be used to help troubleshoot and provide easy assessment of data quality.
How can I access my analysis report?	To open the analysis report, click  , or select **Analysis ▸ Analysis Report**.
How can I export my analysis report?	1. Select **File ▸ Export**. 2. Enter a name and storage location for the report. 3. Click **Save**.

Printing questions and answers

Table 21 Printing Questions and Answers

Question	Answer
What can I print in Sequencing Analysis software?	You can print sample views (annotative, sequence, feature, electropherogram, raw, and EPT) and analysis reports.
What printers are recommended for use with Sequencing Analysis software?	An HP™ 8100, 4600, 990cxi, or an Epson™ 980 color printer is recommended.
Why are my printouts are chopped off?	Use the paper size that matches your settings and make sure that your printer is configured to use that size paper.
Why does not automated printing work?	A default printer must be set in Windows™. Refer to your Windows™ operating system documentation.
How many bases does 1500 pts/panel equal?	~120 bases.

C Basecallers and DyeSet/Primer files

SeqStudio™ Flex Series Genetic Analyzer files

Table 22 SeqStudio™ Flex Series Genetic Analyzer Basecaller and DyeSet/Primer files used for dye terminator chemistry

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
SDB basecalling			
BigDye™ Terminator v1.1	36 cm 50 cm	SDB.bcp	SDB_3600_SeqStudioFlex_POP7_BDTv1.mob
	36 cm 50 cm	SDB.bcp	SDB_3600_SeqStudioFlex_POP6_BDTv1.mob
	36 cm 50 cm	SDB.bcp	SDB_3600_SeqStudioFlex_POP4_BDTv1.mob
BigDye™ Terminator v3.1	36 cm 50 cm	SDB.bcp	SDB_3600_SeqStudioFlex_POP7_BDTv3.mob
	36 cm 50 cm	SDB.bcp	SDB_3600_SeqStudioFlex_POP6_BDTv3.mob
	36 cm 50 cm	SDB.bcp	SDB_3600_SeqStudioFlex_POP4_BDTv3.mob
BigDye™ Direct	36 cm 50 cm	SDB.bcp	SDB_3600_SeqStudioFlex_POP7_BDTv3direct.mob

Table 22 SeqStudio Flex Series Genetic Analyzer Basecaller and DyeSet/Primer files used for dye terminator chemistry *(continued)*

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
KB basecalling			
BigDye™ Terminator v1.1	36 cm 50 cm	KB.bcp	KB_3600_SeqStudioFlex_POP7_BDTv1.mob
	36 cm 50 cm	KB.bcp	KB_3600_SeqStudioFlex_POP6_BDTv1.mob
	36 cm 50 cm	KB.bcp	KB_3600_SeqStudioFlex_POP4_BDTv1.mob
BigDye™ Terminator v3.1	36 cm 50 cm	KB.bcp	KB_3600_SeqStudioFlex_POP7_BDTv3.mob
	36 cm 50 cm	KB.bcp	KB_3600_SeqStudioFlex_POP6_BDTv3.mob
	36 cm 50 cm	KB.bcp	KB_3600_SeqStudioFlex_POP4_BDTv3.mob
BigDye™ Direct	36 cm 50 cm	KB.bcp	KB_3600_SeqStudioFlex_POP7_BDTv3direct.mob

SeqStudio™ Genetic Analyzer files

Table 23 SeqStudio™ Genetic Analyzer Basecaller and DyeSet/Primer files used for dye terminator chemistry

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
SDB basecalling			
BigDye™ Terminator v1.1	28 cm	SDB.bcp	SDB_3200_SeqStudio_POP1_BDTv1.mob
BigDye™ Terminator v3.1	28 cm	SDB.bcp	SDB_3200_SeqStudio_POP1_BDTv3.mob
BigDye™ Direct	28 cm	SDB.bcp	SDB_3200_SeqStudio_POP1_BDTv3direct.mob

Table 23 SeqStudio Genetic Analyzer Basecaller and DyeSet/Primer files used for dye terminator chemistry *(continued)*

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
KB basecalling			
BigDye™ Terminator v1.1	28 cm	KB.bcp	KB_3200_SeqStudio_POP1_BDTv1.mob
BigDye™ Terminator v3.1	28 cm	KB.bcp	KB_3200_SeqStudio_POP1_BDTv3.mob
BigDye™ Direct	28 cm	KB.bcp	KB_3200_SeqStudio_POP1_BDTv3direct.mob

3500/3500xL Genetic Analyzer files

Table 24 3500/3500xL Genetic Analyzer Basecaller and DyeSet/Primer files used for dye terminator chemistry

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
SDB basecalling			
BigDye™ Terminator v1.1	36 cm 50 cm	SDB.bcp	SDB_3500_POP7_BDTv1.mob
	36 cm 50 cm	SDB.bcp	SDB_3500_POP6_BDTv1.mob
	36 cm 50 cm	SDB.bcp	SDB_3500_POP4_BDTv1.mob
BigDye™ Terminator v3.1	36 cm 50 cm	SDB.bcp	SDB_3500_POP7_BDTv3.mob
	36 cm 50 cm	SDB.bcp	SDB_3500_POP6_BDTv3.mob
	36 cm 50 cm	SDB.bcp	SDB_3500_POP4_BDTv3.mob
BigDye™ Direct	36 cm 50 cm	SDB.bcp	SDB_3500_POP7_BDTv3direct.mob

Table 24 3500/3500xL Genetic AnalyzerRUO Basecaller and DyeSet/Primer files used for dye terminator chemistry *(continued)*

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
KB basecalling			
BigDye™ Terminator v1.1	36 cm 50 cm	KB.bcp	KB_3500_POP7_BDTv1.mob
	36 cm 50 cm	KB.bcp	KB_3500_POP6_BDTv1.mob
	36 cm 50 cm	KB.bcp	KB_3500_POP4_BDTv1.mob
BigDye™ Terminator v3.1	36 cm 50 cm	KB.bcp	KB_3500_POP7_BDTv3.mob
	36 cm 50 cm	KB.bcp	KB_3500_POP6_BDTv3.mob
	36 cm 50 cm	KB.bcp	KB_3500_POP4_BDTv3.mob
BigDye™ Direct	36 cm 50 cm	KB.bcp	KB_3500_POP7_BDTv3direct.mob

3730/3730*xl* DNA Analyzer files

Table 25 3730/3730*xl* DNA Analyzer Basecaller and DyeSet/Primer files used for dye terminator chemistry

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
SDB basecalling			
BigDye™ Terminator v3.1	36 cm 50 cm	SDB.bcp	SDB_3730_POP7_BDTv3.mob SDB_3730_POP6_BDTv3.mob
BigDye™ Terminator v1.1	36 cm 50 cm	SDB.bcp	SDB_3730_POP7_BDTv1.mob SDB_3730_POP6_BDTv1.mob
BigDye™ Direct	36 cm 50 cm	SDB.bcp	SDB_3730_POP7_BDTv3direct.mob

Table 25 3730/3730xl DNA Analyzer Basecaller and DyeSet/Primer files used for dye terminator chemistry *(continued)*

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
KB basecalling			
BigDye™ Terminator v3.1	36 cm 50 cm	KB.bcp	KB_3730_POP7_BDTv3.mob KB_3730_POP6_BDTv3.mob
BigDye™ Terminator v1.1	36 cm 50 cm	KB.bcp	KB_3730_POP7_BDTv1.mob KB_3730_POP6_BDTv1.mob
BigDye™ Direct	36 cm 50 cm	KB.bcp	KB_3730_POP7_BDTv3direct.mob
ABI basecalling			
BigDye™ Terminator v1.1	36 cm (rapid read)	Basecaller-3730POP7RR.bcp	DT3730POP7{BD}.mob
	36 cm (std read)	Basecaller-3730POP7SR.bcp	
	50 cm (long read)	Basecaller-3730POP7LR.bcp	
BigDye™ Terminator v3.1	36 cm (rapid read)	Basecaller-3730POP7RR.bcp	DT3730POP7{BDv3}.mob
	36 cm (std read)	Basecaller-3730POP7SR.bcp	
	50 cm (long read)	Basecaller-3730POP7LR.bcp	

3130/3130*xl* Genetic Analyzer files

Table 26 310 Basecaller and DyeSet/Primer files used for dye terminator chemistry

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
KB basecalling			
BigDye™ Terminator v1.0 and v1.1	36 cm (ultra)	KB.bcp	KB_3130_POP4_BDTv1.mob
	50 cm (std read)		
	80 cm (long)		
	36 cm (rapid read)	KB.bcp	KB_3130_POP6_BDTv1.mob
	50 cm (std read)		
	36 cm (ultra)	KB.bcp	KB_3130_POP7_BDTv1.mob
	50 cm (fast std)		
	80 cm (long)		
BigDye™ Terminator v3.0 and v3.1	36 cm (ultra)	KB.bcp	KB_3130_POP4_BDTv3.mob
	50 cm (std read)		
	80 cm (long)		
	36 cm (rapid read)	KB.bcp	KB_3130_POP6_BDTv3.mob
	50 cm (std read)		
	36 cm (ultra)	KB.bcp	KB_3130_POP7_BDTv3.mob
	50 cm (fast std)		
	80 cm (long)		
BigDye™ Direct	36 cm (ultra)	KB.bcp	KB_3130_POP7_BDTv3direct.mob
	50 cm (fast std)		

310 Genetic Analyzer files

- 47 cm capillary array length = 36 cm read length
- 61 cm capillary array length = 50 cm read length

Table 27 310 Basecaller and DyeSet/Primer files used for dye terminator chemistry

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
KB basecalling			
BigDye™ Terminator v1.0 and v1.1	47 cm	KB.bcp	KB_310_POP4_BDTv1_36Rapid.mob KB_310_POP4_BDTv1_36Std.mob
	47 cm	KB.bcp	KB_310_POP6_BDTv1_36Rapid.mob
	61 cm		KB_310_POP6_BDTv1_50Std.mob
BigDye™ Terminator v3.0 and v3.1	47 cm	KB.bcp	KB_310_POP4_BDTv3_36Rapid.mob KB_310_POP4_BDTv3_36Std.mob
	47 cm	KB.bcp	KB_310_POP6_BDTv3_36Rapid.mob
	61 cm		KB_310_POP6_BDTv3_50Std.mob
ABI basecalling			
BigDye™ Terminator v1.0 and v1.1	47 cm	Basecaller-310POP4.bcp	DT310POP4{BD}v2.mob
	47 cm	Basecaller-310POP6.bcp	DT310POP6{BD}.mob
	61		DT310POP6{BD-LR}v3.mob
dRhodamine Terminator	47 cm	Basecaller-310POP4.bcp	DT310POP4{dRhod}v1.mob
	47 cm	Basecaller-310POP6.bcp	DT310POP6{dRhod}v2.mob
	61 cm		
BigDye™ Terminator v3.0 and v3.1	47 cm	Basecaller-310POP4.bcp	DT310POP4{BDv3}v2.mob
	47 cm	Basecaller-310POP6.bcp	DT310POP6{BDv3}v2.mob
	61 cm		

Table 28 310 Basecaller and DyeSet/Primer files used for dye primer chemistry

DNA sequencing chemistry	Capillary array length	Basecaller	DyeSet/Primer
ABI basecalling			
BigDye™ Primer v1.0 and v1.1	47 cm	Basecaller-310POP4.bcp	DP310POP4{BD-21M13}v1.mob DP310POP4{M13Rev}v1.mob
	47 cm	Basecaller-310POP6.bcp	DP310POP6{BD-21M13}v1.mob DP310POP6{M13Rev}v1.mob
	61 cm		
BigDye™ Primer v3.0 and v3.1	47 cm	Basecaller-310POP4.bcp	DP310POP4{BDv3-21M13}v1.mob DP310POP4{BDv3-M13Rev}v1.mob
	47 cm	Basecaller-310POP6.bcp	DP310POP6{BDv3-21M13}v1.mob DP310POP6{BDv3-M13Rev}v1.mob
	61 cm		

D

User privileges

This following table contains a list of privileges for three categories of software users: Administrator, Scientist, and Analyst.

Privilege	Admin	Scientist	Analyst
User accounts, timeout, and audit trail (administrator only)			
Create User Accounts	Allowed	Not Allowed	Not Allowed
Exporting/Importing User Accounts			
Turning on/off the Timeout feature			
Turning on/off the Audit Trail feature			
Mark an user inactive			
Analysis Protocol and Settings			
Creating an analysis protocol	Allowed	Allowed	Not Allowed
Editing an existing analysis protocol			
Apply an analysis protocol to a set of samples			
Delete an Analysis protocol			
Set Clear range determination in Analysis settings or analysis defaults			
Edit Display Settings			
Edit an analysis protocol from the Analysis Defaults			
Edit Analysis Protocol per sample			
Use Save as to create a new Analysis protocol			

(continued)

Privilege	Admin	Scientist	Analyst
Reports			
View Analysis Report	Allowed	Allowed	Allowed
View Report with enabled links back to primary data			
Customize report			
Export report			
Print report			
Sample Manager			
Browse/locate data in the file system	Allowed	Allowed	Allowed
Add samples to the Sample Manager			
Delete samples in the Sample Manager			
Print sample file views			
Change the Basecaller and mobility file in the Sample Manager			
Insert or delete a base in the electropherogram or sequence view			
Change a base in the electropherogram or sequence view			
Search for text in the electropherogram or sequence view			
Edit sample name			
View data in Sample Navigator view			
Search for text strings in any sequence data			

(continued)

Privilege	Admin	Scientist	Analyst
Sequencing Analysis			
Open the Sequencing Analysis	Allowed	Allowed	Allowed
Exit the Sequencing Analysis			

E Key codes and translation tables

The tables in this appendix are available from the Help menu in the software.

IUPAC/IUB codes

The table below provides translations for IUPAC/IUB codes used in the software.

Code	Translation
A	Adenosine
C	Cytidine
G	Guanine
T	Thymidine
B	C, G, or T
D	A, G, or T
H	A, C, or T
R	A or G (puRine)
Y	C or T (pYrimidine)
K	G or T (Keto)
M	A or C (aMino)
S	G or C (Strong—3 H bonds)
W	A or T (Weak—2 H bonds)
N	aNy base
V	A, C, or G

Complements

A	T	S	S
C	G	W	W
G	C	—	—
T	A	B	V
—	—	D	H
R	Y	H	D
Y	R	V	B
K	M	N	N
M	K	—	—

Universal genetic code

The table below provides Universal Genetic Codes for use with the software.

5′ End	2nd Position				3′ End
T	C	A	G		
T	Phe	Ser	Tyr	Cys	T
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	OCH	OPA	A
	Leu	Ser	AMB	Trp	G
C	Leu	Pro	His	Arg	T
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	T
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	T
	Val	Ala	Asp	Gly	C

(continued)

5′ End	2nd Position				3′ End
G	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G
Stop Codes: AMBer, OCHer, OPA					

Amino acid abbreviations

Amino acid	Three letters	One letter
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Any amino acid		X

F

Glossary

This glossary includes some of the terms used in the *Applied Biosystems™ DNA Sequencing Analysis Software v5.1 User Guide*. If a term is not defined here, check the index to see if it is explained elsewhere in the manual.

analysis defaults

Analysis defaults contain processing parameter settings (basecalling, post processing and printing), file formats settings (.seq, .scf and .phd.1) and an analysis protocol. The analysis protocol is assigned to the sample only if it does not already contain one.

analysis protocol

Analysis protocols contain all the settings necessary for analysis, and it is used to perform basecalling and post processing. A protocol is stored in the sample file. Analysis protocols replace preference settings used in previous versions of Sequencing Analysis software.

base spacing

Base spacing is the number of data points from one peak to the next.

Spacing of a negative or red number indicates a problem with your samples, and/or the analysis parameters.

basecaller

The basecaller is an algorithm that determines the bases of a sequence during analysis.

There are three types of basecallers:

KB™ Basecaller—An algorithm that calculates mixed or pure bases, and sample quality values.

Smart Deep™ Basecaller—New algorithm that calculates mixed or pure bases, sample quality values, and enables enhanced trace visualization.

ABI™ Basecaller—An algorithm used in previous versions of ABI PRISM™ Sequencing Analysis software (v3.7 and earlier).

basecalling

Basecalling is the primary function of the Sequencing Analysis software. Basecalling identifies each base in the sample and the order in which the bases are arranged and marks locations where there is some question about the base identification, such as when two bases seem to occur at the same position, with an N (instead of one of the four bases A, C, G, and T).

chromatogram

See Electropherogram on page 167.

clear range

The clear range is the region of sequence that remains after excluding the low-quality or error-prone sequence at both the 5′ and 3′ ends. If the KB™ Basecaller/Smart Deep™ Basecaller was used for analysis the clear range is calculated from the QVs. If an ABI™ Basecaller was used, the range is calculated from the Ns in the data and/or trim by the number of bases at the start and end of the data.

complement

The opposite strand of double-stranded DNA. For example, if you sequenced the 3′ to 5′ strand, then the 5′ to 3′ strand is the complement.

data point

A sampling of fluorescence.

Each data point is associated with a scan number.

dyeset/primer file

A file used to adjust for varying mobility between the dyes and primers used to label DNA for runs on the ABI PRISM™ Genetic Analysis instruments.

These files are sometimes referred to as mobility files. The Sequencing Analysis installer installs DyeSet/Primer files in the Mobility Folder. The path is `C:\Applied Biosystems\SeqA\AppSeqA\bin\Basecaller\Mobility`.

editable data

Sequencing Analysis basecalled data is saved as “original data” in the sample file.

If you edit bases in the sample files, these are stored as editable data. Parallel copies of the original and edited data are maintained in the sample file.

The data displayed in the Sample window is the editable copy (unless you choose to display both the editable data and original data).

See also “sample files” on page 169 and “original data.” on page 169

ept

A multi-color graph displaying the values for the voltage, power, current and temperature for the entire run.

electropherogram

A multi-color picture of a sequence showing peaks that represent the bases.

feature

The clear range (confidence range) for a sample.

The Sample views pane includes a Feature view that displays feature information if any is present in the file.

heterozygote

A position at which the electropherogram displays more than one base.

IUB code

An alphabetic character representing the occurrence of mixed bases at a given position in a sequence.

A code originally defined by the International Union of Biochemistry.

The table below contains a table of IUB codes, the mixed bases they represent, and a listing of the complements.

Base	IUB Code	Complement
Adenosine	A	T
Cytidine	C	G
Guanosine	G	C
Thymidine	T	A
Adenosine or Guanosine (puRine)	R	Y
Cytidine or Thymidine (pYrimidine)	Y	R
Guanosine or Thymidine (Keto)	K	M
Adenosine or Cytidine (aMino)	M	K
Guanosine or Cytidine (Strong — 3 H bonds)	S	S
Adenosine or Thymidine (Weak — 2 H bonds)	W	W
Cytidine, Guanosine, or Thymidine	B	V
Adenosine, Guanosine, or Thymidine	D	H
Adenosine, Cytidine, or Thymidine	H	D
Adenosine, Cytidine, or Guanosine	V	B
Adenosine, Cytidine, Guanosine, or Thymidine (any base)	N	N

IUPAC

International Union of Pure and Applied Chemistry.

This acronym is also used to refer to “IUPAC/IUB codes” on page 163 (see “IUB code” on page 167), because IUPAC adopted the codes as a standard.

length

The length of a sequence is the number of characters it contains, including gap characters.

For example, GAATTC has a length of 6 while GAA-TTC has a length of 7.

length of read

The usable range of high-quality or high-accuracy bases, as determined by quality values. This information is displayed in the Analysis report.

master analysis protocol

A master protocol is not associated with any sample. They are copied and assigned to a sample by using either the Apply to Selected Samples feature, or the analysis default, if the sample does not have a protocol.

mixed bases

Mixed bases are one base positions that contain 2, 3, or 4 bases. These bases are assigned the appropriate IUB code.

mobility file

See dyeset/primer file on page 167.

noise

Average background fluorescent intensity for each dye.

original data

The sequence data created the last time the Basecaller was run.

This basecalled data is maintained in the sample file. If you edit the bases in the sample file, your edits are saved as editable data.

The original basecalled data is not overwritten by your edits but it is overwritten if the sample is reanalyzed with a different Basecaller or Basecaller settings. See also “editable data” on page 167 and “sample files.” on page 169

per-sample analysis protocol

A per-sample protocol is the protocol stored within a sample file. This protocol can be edited. The change affects the protocol for the selected sample only. You cannot apply this protocol to other samples.

.phd.1 file

An additional file format that can be generated during sample analysis. The file contains base calls and quality values.

raw data

A multi-color graph displaying the fluorescence intensity (signal) collected for each of the four fluorescent dyes.

sample files

A sample file contains raw DNA sequence data (as read by the electrophoresis instrument), and the basecalls, peak locations, and electropherogram created by the Sequencing Analysis software.

For the ABI PRISM™ genetic analysis instruments, raw sample files are created and can be analyzed by the data collection. Raw or previously analyzed sample files are analyzed by Sequencing Analysis.

sample score

A sample score is generated from QVs. It is the average quality value of the bases in the clear range sequence for that sample.

quality values

An estimate (or prediction) of the likelihood that a given basecall is in error. Typically, the quality value is scaled following the convention established by the phred program: QV = -10 log10(*Pe*), where *Pe* stands for the estimated probability that the call is in error.

scan number

On an ABI PRISM™ genetic analysis instrument, one sampling is taken during each scan and the information is stored as a data point.

.scf file

An additional file format that can be generated during sample analysis. The file contains base calls, electropherogram and quality values, but no raw data.

Note: When standard chromatogram file format is created, the .scf extension is not appended to the file name. However, the file format is correct.

.seq files

Text files created by the Sequencing Analysis software.

The .seq files contain only the characters of the sequence, and can be created in several formats (ABI and FASTA) for use with other software.

sequence

A linear series of characters.

The characters are displayed in rows from left to right. More specifically, a sequence is a series of nucleotide base characters that represent a linear DNA sequence, or a series of amino acid characters that represent a protein sequence.

sequencing reactions

The reactions performed to incorporate fluorescent dye labels into DNA extension products.

signal

A number that indicates the intensity of the fluorescence from one of the dyes used to identify bases during a data run.

Signal strength numbers are shown in the Annotation view of the sample file.

signal/noise

The average of the signal intensity of the ‘A’, ‘C’, ‘G’, or ‘T’ base divided by the average noise for that base.

spacing

See base spacing on page 166.

views

Various displays provided in the Sample window.

Documentation and support

Related documentation

Document	Pub. No.
Sequencing Analysis Software 8 Quick Reference	MAN0019596
Sequencing Analysis Software 8 User Bulletin	MAN0019597
ABCall Utility 1.4 User Guide	MAN0028060
User Guide: Applied Biosystems™ 3730/3730xl DNA Analyzer	4331468

Customer and technical support

Visit **thermofisher.com/support** for the latest service and support information.

- Worldwide contact telephone numbers
- Product support information
  - Product FAQs
  - Software, patches, and updates
  - Training for many applications and instruments
- Order and web support
- Product documentation
  - User guides, manuals, and protocols
  - Certificates of Analysis
  - Safety Data Sheets (SDSs; also known as MSDSs)

Note: For SDSs for reagents and chemicals from other manufacturers, contact the manufacturer.

Limited product warranty

Life Technologies Corporation and/or its affiliate(s) warrant their products as set forth in the Life Technologies' General Terms and Conditions of Sale at **www.thermofisher.com/us/en/home/global/terms-and-conditions.html**. If you have any questions, please contact Life Technologies at **www.thermofisher.com/support**.