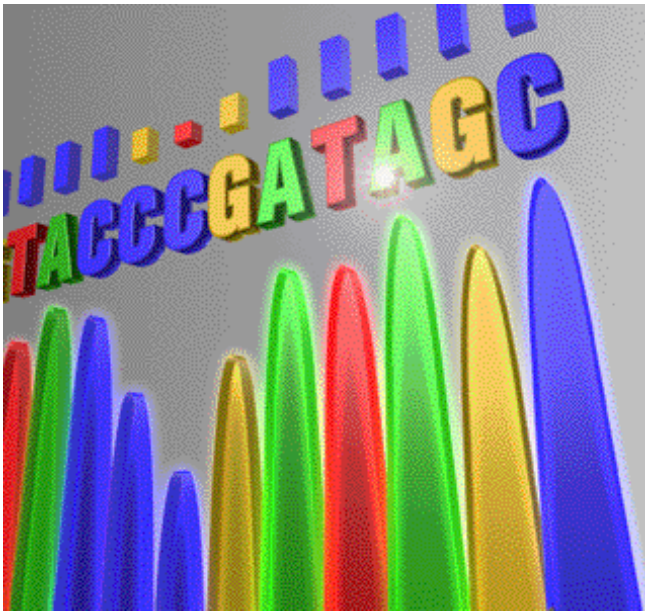


Applied Biosystems DNA Sequencing Analysis Software

Version 5.1 for Windows® XP and 2000 Platforms

User Guide

Applied Biosystems

Applied Biosystems
DNA Sequencing Analysis Software
Version 5.1 for Windows® XP and 2000 Platforms

User Guide

Applied Biosystems

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

Information in this document is subject to change without notice. Applied Biosystems assumes no responsibility for any errors that may appear in this document. This document is believed to be complete and accurate at the time of publication. In no event shall Applied Biosystems be liable for incidental, special, multiple, or consequential damages in connection with or arising from the use of this document.

Notice to Purchaser: License Disclaimer

Purchase of this software product alone does not imply any license under any process, instrument or other apparatus, system, composition, reagent or kit rights under patent claims owned or otherwise controlled by Applera Corporation, either expressly, or by estoppel.

TRADEMARKS:

Applied Biosystems, ABI PRISM, BigDye, Primer Express, and SeqScape are registered trademarks of Applera Corporation or its subsidiaries in the U.S. and/or certain other countries.

AB (Design) and Applera are trademarks of Applera Corporation or its subsidiaries in the U.S. and/or certain other countries.

Microsoft and Windows are registered trademarks of Microsoft Corporation.

All other trademarks are the sole property of their respective owners.

Part Number 4346366 Rev. B
9/2003

Contents

Preface

Safety Information

Chapter 1 Getting Started

Chapter 2 Software Overview

Chapter 3 Sample Files in the Sample Manager

Chapter 4 Viewing and Editing Samples

Chapter 5 Using the Sample Manager

Preface

How to Use This Guide

Purpose of This Guide

The *Applied Biosystems DNA Sequencing Analysis Software v5.1 User Guide* provides information on analyzing and reviewing the sequence data.

Audience

This guide is intended for novice and experienced users who analyze, reanalyze, review, and edit DNA sequence data.

Assumptions

This guide also assumes that you have a working knowledge of:

- Microsoft® Windows® XP or Windows® 2000 operating system
- DNA sequence detection and analysis methods
- DNA and amino acid coding conventions

Text Conventions

This guide uses the following conventions:

- **Bold** indicates user action. For example:

  Type **0**, then press **Enter** for each of the remaining fields.
- *Italic* text indicates new or important words and is also used for emphasis. For example:

  Before analyzing, *always* prepare fresh matrix.
- A right arrow bracket (>) separates successive commands you select from a drop-down or shortcut menu. For example:

  Select **File > Open > Spot Set**.

  Right-click the sample row, then select **View Filter > View All Runs**.
-  indicates a button in the toolbar. For example:

  Click  (Start Analysis).

User Attention Words

Two user attention words appear in Applied Biosystems user documentation. Each word implies a particular level of observation or action as described below:

Note: Provides information that may be of interest or help but is not critical to the use of the product.

IMPORTANT! Provides information that is necessary for proper instrument operation, accurate chemistry kit use, or safe use of a chemical.

Examples of the user attention words appear below:

Note: The size of the column affects the run time.

Note: The Calibrate function is also available in the Control Console.

IMPORTANT! To verify your client connection to the database, you need a valid Oracle user ID and password.

IMPORTANT! You must create a separate Sample Entry Spreadsheet for each 96-well microtiter plate.

How to Obtain More Information

Related Documentation

The following related documents are shipped with the system:

Applied Biosystems Sequencing Analysis Software v5.1 Quick Reference Card– Briefly describes how to use the software to analyze and review the results.

Portable document format (PDF) versions of this guide and the Applied Biosystems document listed above are also available on the *Applied Biosystems Sequencing Analysis Software v5.1* installation CD.

Note: For additional documentation, see "How to Obtain Services and Support" on page xi.

Send Us Your Comments

Applied Biosystems welcomes your comments and suggestions for improving its user documents. You can e-mail your comments to:

techpubs@appliedbiosystems.com

How to Obtain Services and Support

For the latest services and support information for all locations, go to **http://www.appliedbiosystems.com**, then click the link for **Services and Support**.

At the Services and Support page, you can:

- Search through frequently asked questions (FAQs)
- Submit a question directly to Technical Support
- Order Applied Biosystems user documents, MSDSs, certificates of analysis, and other related documents
- Download PDF documents
- Obtain information about customer training
- Download software updates and patches

In addition, the Services and Support page provides access to worldwide telephone and fax numbers to contact Applied Biosystems Technical Support and Sales facilities.

Safety Information

This section includes the following topics:

Safety Conventions Used in This Document

Safety Alert Words

Four safety alert words appear in Applied Biosystems user documentation at points in the document where you need to be aware of relevant hazards. Each alert word implies a particular level of observation or action, as defined below:

Definitions

IMPORTANT! – Indicates information that is necessary for proper instrument operation, accurate chemistry kit use, or safe use of a chemical.

– Indicates a potentially hazardous situation that, if not avoided, may result in minor or moderate injury. It may also be used to alert against unsafe practices.

– Indicates a potentially hazardous situation that, if not avoided, could result in death or serious injury.

– Indicates an imminently hazardous situation that, if not avoided, will result in death or serious injury. This signal word is to be limited to the most extreme situations.

Examples

The following examples show the use of IMPORTANT, CAUTION, and WARNING safety alert words:

IMPORTANT! The sample name, run folder name, and path name, *combined*, can contain no more than 250 characters.

MUSCULOSKELETAL AND REPETITIVE MOTION HAZARD. These hazards are caused by potential risk factors that include but are not limited to repetitive motion, awkward posture, forceful exertion, holding static unhealthy positions, contact pressure, and other workstation environmental factors.

Do not attempt to lift or move the computer or the monitor without the assistance of others. Depending on the weight of the computer and/or the monitor, moving them may require two or more people.

General Instrument Safety

WARNING **PHYSICAL INJURY HAZARD.** Use this product only as specified in this document. Using this instrument in a manner not specified by Applied Biosystems may result in personal injury or damage to the instrument.

Moving and Lifting Stand-Alone Computers and Monitors

WARNING Do not attempt to lift or move the computer or the monitor without the assistance of others. Depending on the weight of the computer and/or the monitor, moving them may require two or more people.

Things to consider before lifting the computer and/or the monitor:

- Make sure that you have a secure, comfortable grip on the computer or the monitor when lifting.
- Make sure that the path from where the object is to where it is being moved is clear of obstructions.
- Do not lift an object and twist your torso at the same time.
- Keep your spine in a good neutral position while lifting with your legs.
- Participants should coordinate lift and move intentions with each other before lifting and carrying.
- Instead of lifting the object from the packing box, carefully tilt the box on its side and hold it stationary while someone slides the contents out of the box.

Workstation Safety

Correct ergonomic configuration of your workstation can reduce or prevent effects such as fatigue, pain, and strain. Minimize or eliminate these effects by configuring your workstation to promote neutral or relaxed working positions.

CAUTION **MUSCULOSKELETAL AND REPETITIVE MOTION HAZARD**. These hazards are caused by potential risk factors that include but are not limited to repetitive motion, awkward posture, forceful exertion, holding static unhealthy positions, contact pressure, and other workstation environmental factors.

To minimize musculoskeletal and repetitive motion risks:

- Use equipment that comfortably supports you in neutral working positions and allows adequate accessibility to the keyboard, monitor, and mouse.
- Position the keyboard, mouse, and monitor to promote relaxed body and head postures.

Getting Started 1

This chapter covers:

Completing the Software Registration Card

License and Warranty

Before you begin, read Appendix F, "Software Warranty Information." This appendix explains your rights and responsibilities regarding the software.

Registering Your Software

To register your copy of the ABI PRISM® Sequencing Analysis Software v5.1, complete the registration card (included in this software package) and return it to Applied Biosystems.

Registering the software enables Applied Biosystems to send you notification of software updates and any other future information that may be specific to Sequencing Analysis Software v5.1 owners.

Your product registration number is located on the registration card. Be sure to record this number here before you return the registration card.

Registration Number:

Hardware and Software Requirements

Introduction

The Sequencing Analysis software can be installed on a:

- Computer connected to your Applied Biosystems 3730/3730*xl* DNA Analzyers or ABI PRISM® 3100/3100-Avant Genetic Analyzers instrument running version 2.0 Data Collection and Microsoft® Windows® 2000
- Computer connected to your ABI PRISM genetic analyzer instrument that is using Microsoft® Windows® XP or Windows® 2000 operating systems.
- Computer dedicated to data analysis only if the minimum requirements stated below are met.

System Requirements

Below are the system requirements for running the Sequencing Analysis software v5.1. for Windows XP or Windows 2000 platforms on your instrument or analysis computer.

Note: These are the minimum requirements. In general, the more memory, the larger the screen size, and the more processing power you have, the better the program performance.

Table 1-1 System Requirements

System Component	Requirements
CPU	733 MHz or faster with a Pentium® III or IV processor. The software does not run on computers with a dual processor or with an Intel Xeon® chipset.
CD-ROM drive	Any
Operating system	Windows XP OS with Service Pack 1 or Windows 2000 OS with Service Pack 3.
RAM	512 MB minimum
Printer	HP® 4600, 8100, 990cxi, or Epson® 980 printer is recommended.
Monitor	17-inch monitor or larger is recommended. 1024 x 768 resolution or greater is recommended.

Table 1-1 System Requirements *(continued)*

System Component	Requirements
Disk space	The hard drive should have 1 GB of space available. 75 MB of free disk space is required to accommodate the Sequencing Analysis software v5.1, and sufficient space for all sample files. Storage requirements depend primarily on the quantity of data to be generated and stored. Because Sequencing Analysis software data files are stored in the area into which the software is installed, install Sequencing Analysis software on a partition with enough space for the files.

Hard Drive Partitions

The installer uses the following default locations for the files:

drive letter:\AppliedBiosystems\SeqA5.1

The drive letter is defined by the following conditions. The installer looks for the Applied Biosystems folder installed by another Applied Biosystems software.

- If the folder exists, the Sequencing Analysis software is installed there.
- If the folder does not exist, the installer uses the default location D:\AppliedBiosystems. If there is no D drive, then the default location C:\AppliedBiosystems is used.

Installing the Sequencing Analysis Software

Before Installation

To prepare for the installation:

1. Check to be sure that your system meets the minimum requirements (see "Hardware and Software Requirements" on page 1-3).
2. Temporarily turn off any virus protection software.
3. Exit all programs, except Applied Biosystems 3730/3730*xl* Data Collection or ABI PRISM® 3100/3100-*Avant* Data Collection software, if applicable.

IMPORTANT! To properly install Sequencing Analysis Software v5.1 on a computer that is connected to Applied Biosystems 3730/3730xl DNA Analyzers or ABI Prism® 3100/3100-Avant Genetic Analyzers, **the data collection software must be running**. If data collection is not running, the Sequencing Analysis software does not register with the Data Service software.

Upgrading from Previous Versions of Software

Sequencing Analysis Software v5.0

The Sequencing Analysis v5.1 installer automatically upgrades the software from v5.0 to 5.1

Sequencing Analysis Software v3.7

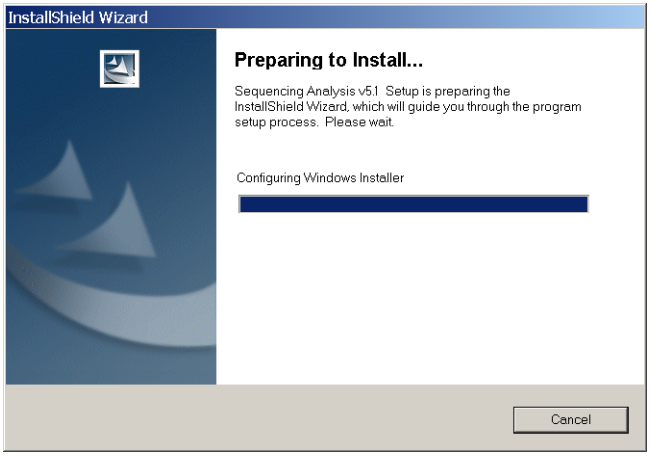
If Sequencing Analysis software v3.7 is installed on your system, uninstall it before you install Sequencing Analysis v5.1.

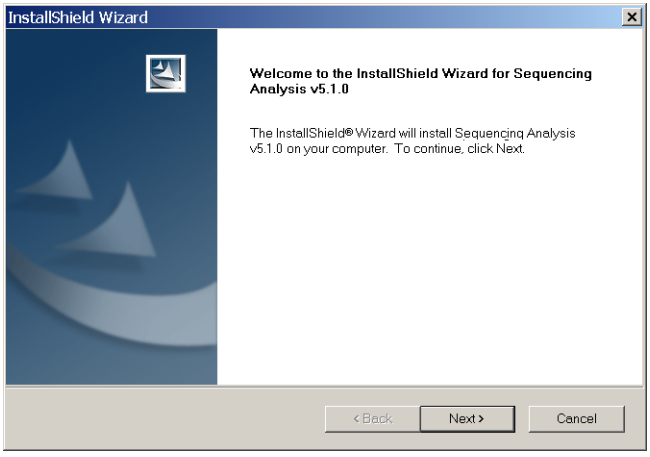
Installing the Sequencing Analysis Software

To install the Sequencing Analysis software from a CD-ROM:

1. Insert the Sequencing Analysis Software v5.1 for Windows XP or Windows 2000 platform CD-ROM into the CD drive.

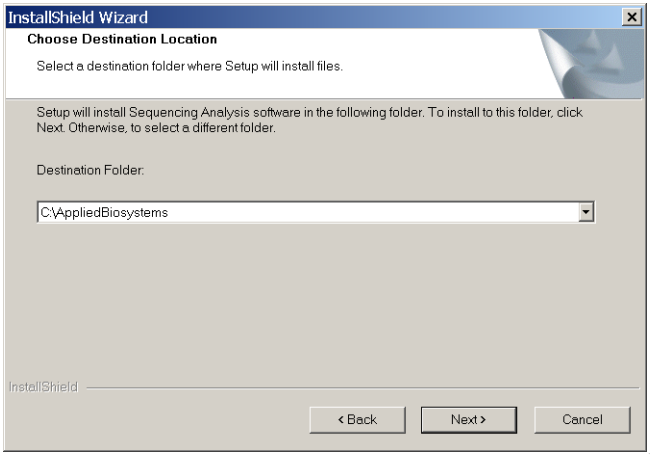
   The installer starts automatically and the following windows open.



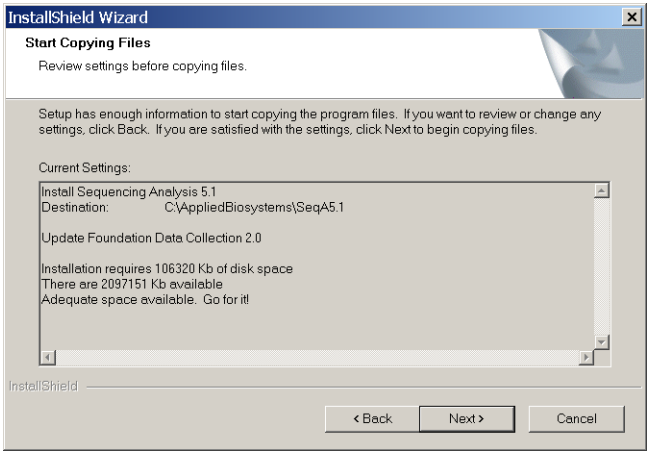


2. Click **Next**.
3. When the License Agreement window opens, read the agreement, then click **Yes**.
4. In Information window, read the ReadMe file, then click **Next**.

5. In the Choose Destination Location window, click **Next**.

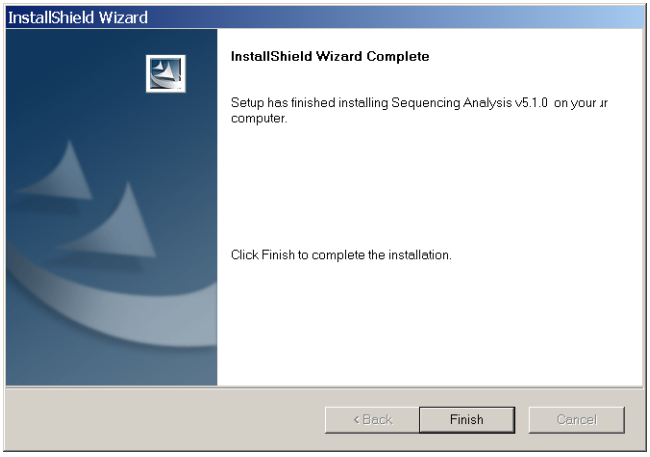


6. Verify the information in the Current Settings pane, then click **Next**.



The files are copied to the hard drive.

7. When the following window opens, click **Finish**.



A Sequencing Analysis 5.1 shortcut is added to the desktop and to the Start menu.

Removing Sequencing Analysis Software

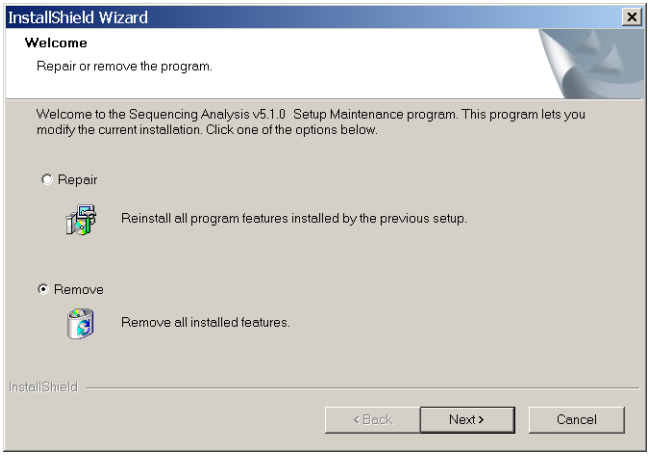
The uninstall process deletes all folders and files installed by the Sequencing Analysis Software v5.1.

To remove installed Sequencing Analysis software:

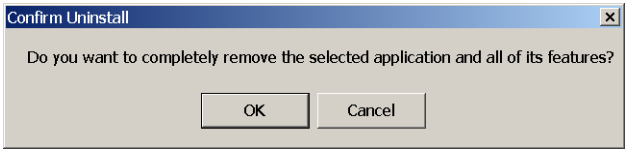
IMPORTANT! If you are uninstalling the software from a computer connected to 3730/3730*xl* DNA Analyzers or 3100/3100-Avant Genetic Analyzers, open the data collection software.

Note: Close all other programs before running the uninstaller.

1. Select **Start > Programs > Applied Biosystems > Sequence Analysis 5.1 > Uninstall Sequencing Analysis 5.1**. The following dialog box opens.



2. Select **Remove**, then click **Next**.
3. In the Confirm Uninstall dialog box, click **OK**.



The files are uninstalled from the hard drive.

4. In the Maintenance Complete dialog box, click **Finish**.

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

Starting the Sequencing Analysis Software for the First Time

Before You Begin

The Sequencing Analysis software is designed with a user login process. When you start the software for the first time, you are prompted with a registration dialog box that creates an administrator account. Log in to the Sequencing Analysis software as Admin and enter the password you created.

To create new users, you must log in as Admin. Logging in with a user name allows Sequencing Analysis software to track each user's interactions with each project.

For information on the privileges for each category of user using the software, refer to Appendix D, "User Privileges."

File-Naming Convention

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

spaces

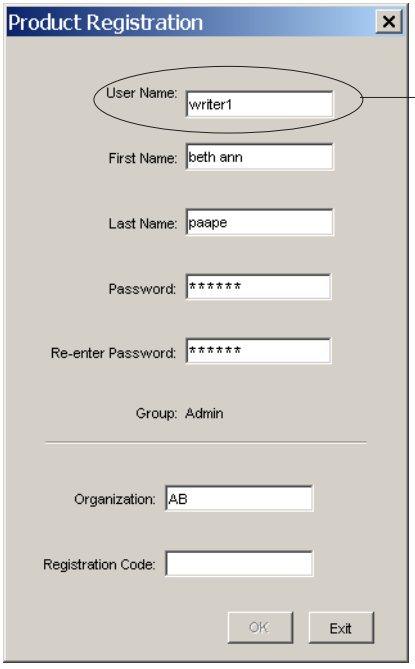
\ / : * ? " < > |

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

Starting Sequencing Analysis Software

To start the software for the first time:

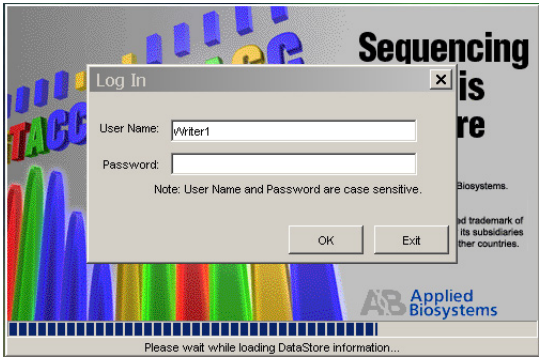
1. Start the Sequencing Analysis Software v5.1 by double-clicking the Sequencing Analysis v5.1 desktop shortcut or select **Start > Programs > Applied Biosystems > Sequencing Analysis 5.1 > Sequencing Analysis 5.1**.
2. In the Product Registration dialog box, enter all the information in the text fields. The User Name and password must be 6 to 15 characters long.



The first user created is automatically assigned Administrator privileges.

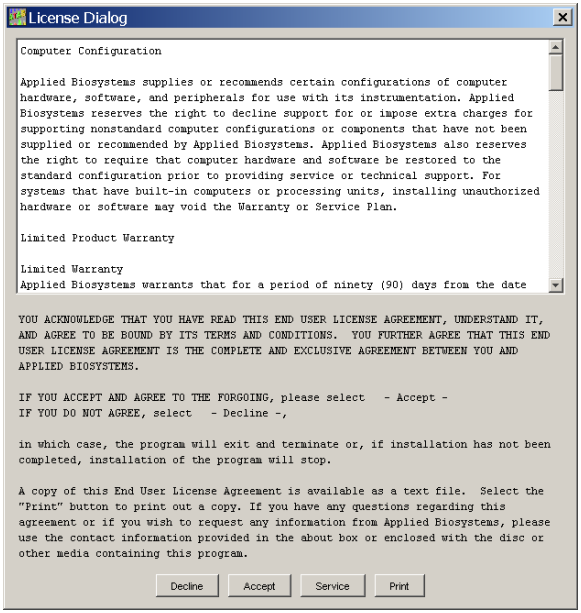
3. Enter the registration code on the registration card you received with your software.
4. Click **OK**.

   While the program is loading, the splash screen opens, then the Log In dialog box opens.
5. Enter your user name and password again.



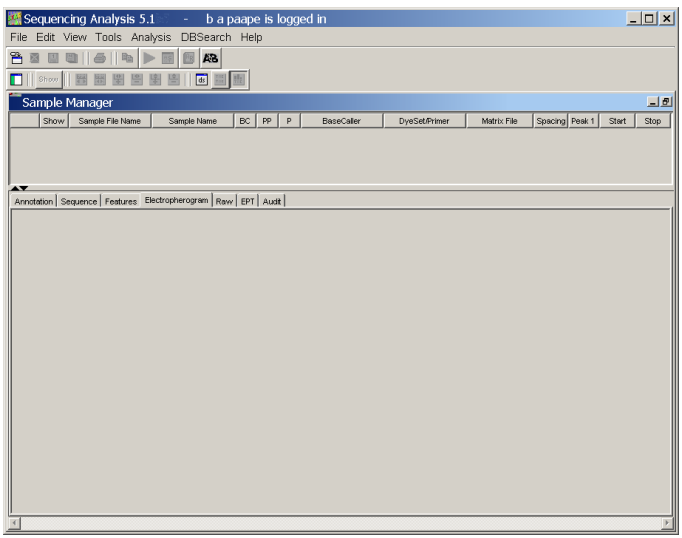
6. Click **OK**.

   The License dialog box opens.



7. Read the license agreement and click **Accept**.

   The main Sequencing Analysis window opens.



User and Audit Trail Set Up

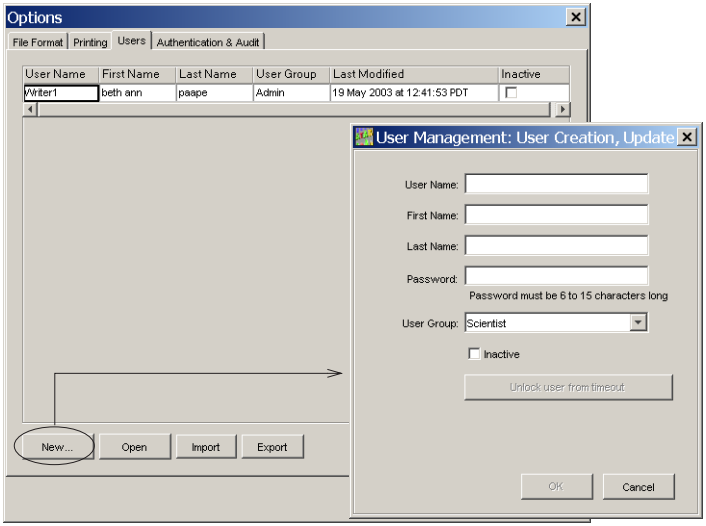
Creating New Users

Because the Sequencing Analysis software tracks the settings for each user, Applied Biosystems recommends that you create users for each individual who uses Sequencing Analysis software on the computer. 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.

To set up new users:

1. Select **Tools** > **Options** to open the Options dialog box.
2. Select the **Users** tab, then click **New**.



3. Fill in the appropriate user name, password, first and last names, then select the level of user from the User Group drop-down list.

   Note: Enter a User Name that contains only alphanumeric characters. This field must not contain any spaces or characters that do not conform with the Microsoft® Windows file system. Refer to "File-Naming Convention" on page 1-11.

   The new user appears in the list in the Users tab.

New users can log in after you exit Sequencing Analysis software and restart the application.

Setting Up Authentication & Audit

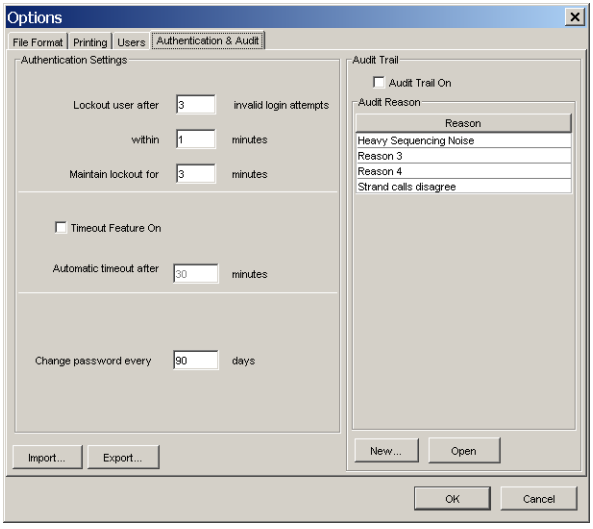
Users belonging to the Administrator group can change the default settings in the Authentication & Audit tab for security features of the application.

Note: 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.

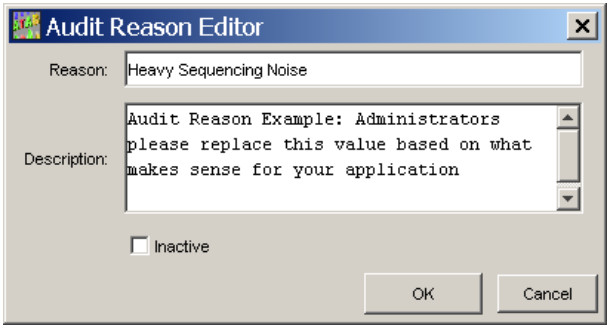
The Authentication & Audit panes provide a way to track the changes such as base change or processes you want to track. You must turn Audit Trail On for tracking to occur.

To set up authentication and auditing:

1. Select the **Authentication & Audit** tab to change the defaults for the Authentication Settings:



2. In the Authentication Settings pane:
   a. Lockout occurs when a user enters an incorrect password or user name the number of times you select for the **Lockout user after invalid login attempts** field. Enter the number or accept the default.
   b. The **within minutes** field indicates that the user will be locked out if the maximum number of attempts occur within the time entered in this field. Enter a number or accept the default.
   c. 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 Sequencing Analysis software. Enter the number of minutes or accept the defaults.
   d. The software times out and the user must log in again to use the software if the **Timeout Feature On** check box is selected. Enter the number of timeout minutes or accept the defaults.
   e. 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. In the Audit Trail pane, select the **Audit Trail On** check box to have a dialog box open whenever an indicated reason occurs.
4. In the Audit Reason pane, enter reasons to provide an audit trail.
   a. Double-click the a field, or highlight it, then click **New**.



   b. In the Reason field, type a reason for a change to the project to identify.
   c. Enter a description of the reason, if desired.

d. Click **OK** in the Audit Reason Editor. The first reason appears in the list in the Options dialog box.

e. Whenever a change is made in any of the data views, the Audit Reason Editor dialog box opens. Select the reason for the change from the drop-down list.

5. Click **OK** in the Options dialog box to save the authentication and audit settings.

Note: It is possible to import or export Authentication & Audit configurations from one computer to another. For example, an administrator may want to set up authentication and audit information for many users, then select all the files and export them to other systems using Sequencing Analysis software.

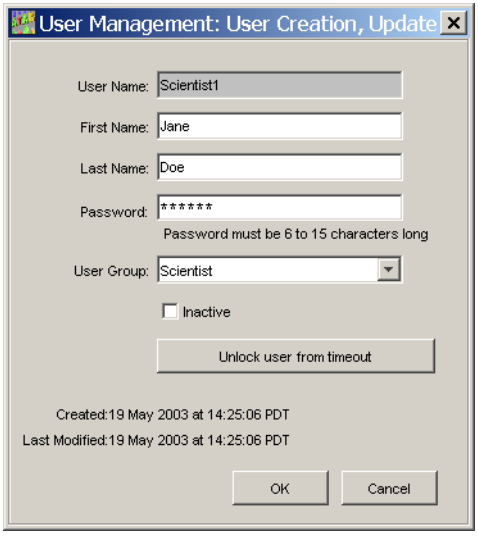
Changing User Information

If desired, change the default settings for all users you are setting up.

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.

To change any of the information for a user:

1. Select **Tools > Options**, then select the **Users** tab.
2. Double-click the name in the list to open the User Management dialog box.



3. Change or correct the user information and click **OK**.
4. If desired, click the **Export** button in the Options dialog box to export the application configuration settings and/or settings for a single user or multiple users in a zipped .ctf format.
5. Enter the path for exporting files in the Export User dialog box, then click **Export**.
6. Click **OK** to close the Options dialog box.

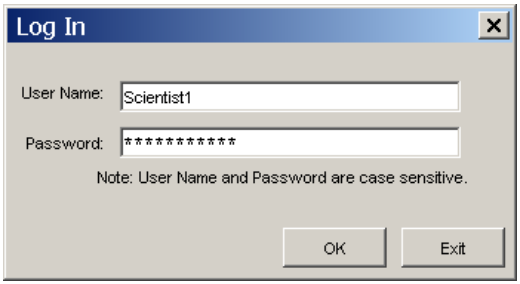
Note: This process can be used by the first administrator to set up additional users or another administrator. It is possible to import or export user settings from one computer to another. For example, an administrator can set up user information for many users, then select all the user files and export them to other systems using Sequencing Analysis software.

When New Users Log In

After the installation and setup are complete, new users can log in to the software.

To log in to the software:

1. Start the Sequencing Analysis software by double-clicking the desktop shortcut  .
2. The Log In dialog box opens, showing the last user's name. Enter your user name and password, then click **OK**.



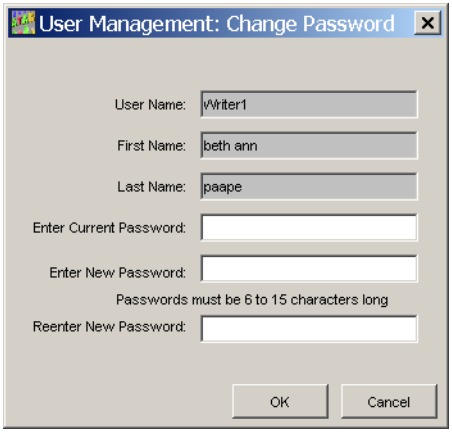
The Sequencing Analysis software is ready for you to use.

Changing the User Password

All user groups (administrator, scientist and analyst) can change their password.

To change the current user's password:

1. Log in to the software.
2. Select **Tools > Change Password**.



3. Enter your current password.
4. Enter your new password, then reenter it again.
5. Click **OK**.

Copying 310 Matrix and DyeSet/Primer Files

Note: Skip this section if you are analyzing sample files not generated on the ABI PRISM® 310 Genetic Analyzer.

Matrix and Mobility Folder Locations

If you are upgrading from Sequencing Analysis software v3.7 to Sequencing Analysis software v5.1, then you will need to copy your matrix and DyeSet/Primer (mobility) files to new folder locations.

Sequencing Analysis Software v3.7 Folder Locations

Sequencing Analysis software v3.7 installation location is D:\AppliedBio, the same installation folder location as the data collection software. Both applications access the Shared folder, which contains the Matrix and Mobility folders. The pathways are:

- D:\AppliedBio\Shared\Analysis\Basecaller\Matrix
- D:\AppliedBio\Shared\Analysis\Basecaller\Mobility

Folder Locations After Sequencing Analysis Software v5.1 Installation

When Sequencing Analysis software v5.1 (SeqA v5.1) is installed, it renames and creates new folders within the Basecaller folder of the D\:AppliedBio folder. See Figure 1-1 on page 1-21.

Reference in Figure 1-1	Within the Basecaller Folder	Contents of the Folder
1	The Matrix folder is renamed to **Old_Matrix** folder	Original 310 matrix files
2	A new **Matrix** folder is created	TestMatrix files installed by SeqA v5.1
3	The Mobility folder is renamed to **Old_Mobility** folder	Original DyeSet/Primer files installed by data collection and SeqA v3.7 software
4	A new **Mobility** folder is created	310 DyeSet/Primer files installed and used by SeqA v5.1

The pathways after Sequencing Analysis software v5.1 are:

- D:\AppliedBio\Shared\Analysis\Basecaller\Old_Matrix
- D:\AppliedBio\Shared\Analysis\Basecaller\Old_Mobility

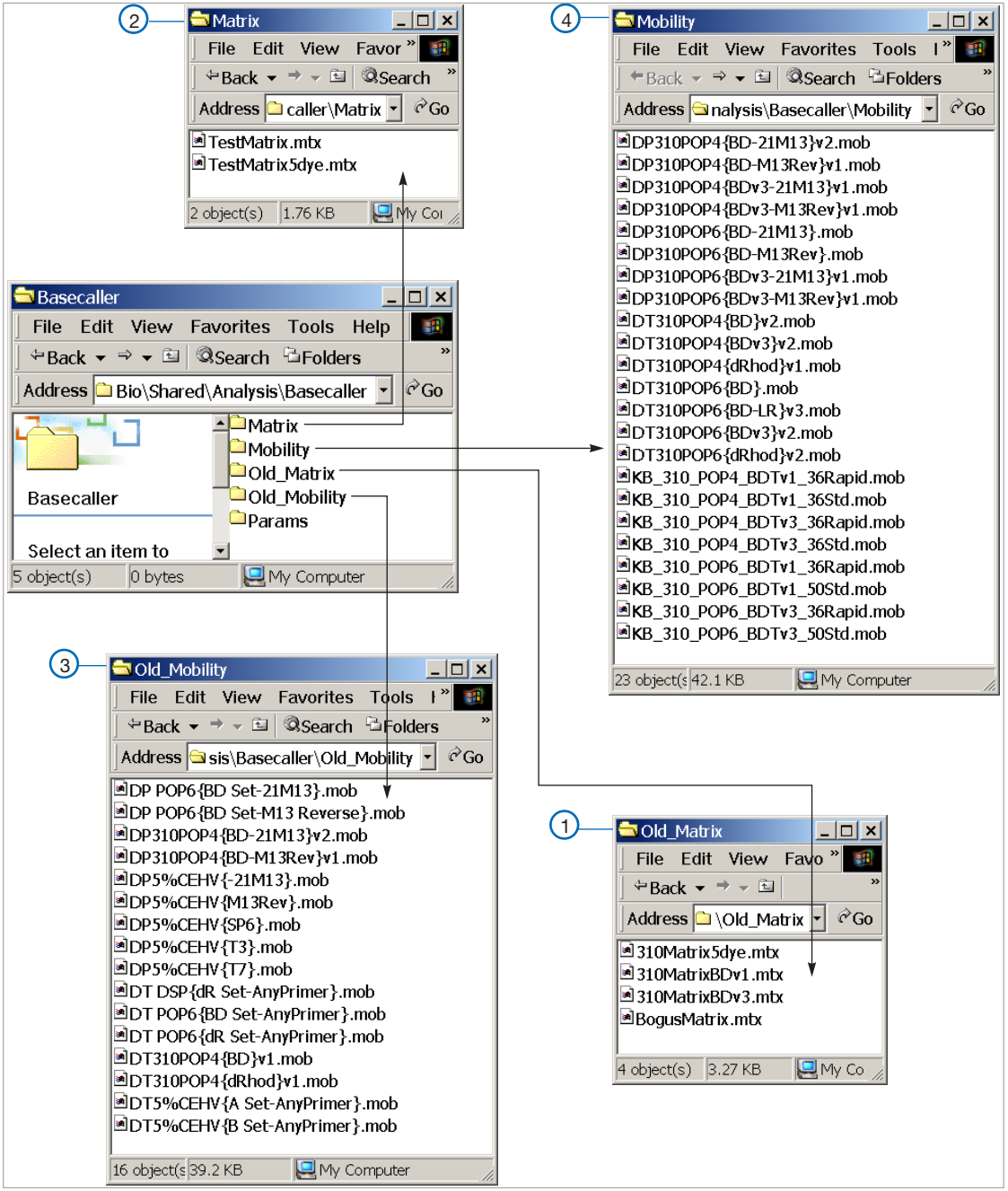


Figure 1-1 Names and contents of folders after the installation of Sequencing Analysis software v5.1

Sequencing Analysis Software v5.1 Folder Locations

Sequencing Analysis software v5.1 installation location is:

- D:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\Basecaller\Matrix
- D:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\Basecaller\Mobility

Because of the location change, the Sequencing Analysis software v5.1 cannot access the Shared folder used by data collection.

Duplicate Copies of Files are Required

To be able to use your matrix files in both the data collection software and Sequencing Analysis software v5.1, you must place a copy of matrix files in the Matrix folder within both the Sequencing Analysis software and data collection software folders (Figure 1-2).

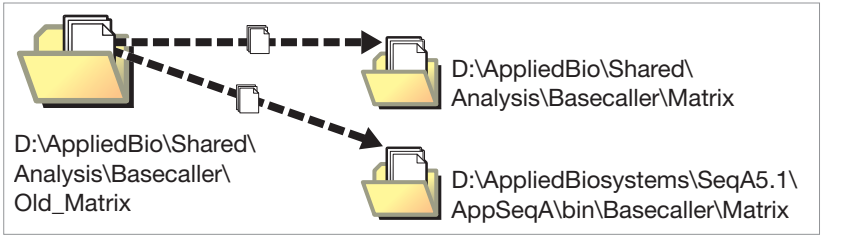


Figure 1-2 Where to copy 310 matrix files

If there are any of the original DyeSet/Primer files you want to continue to use, copies of these files are also required (Figure 1-3).

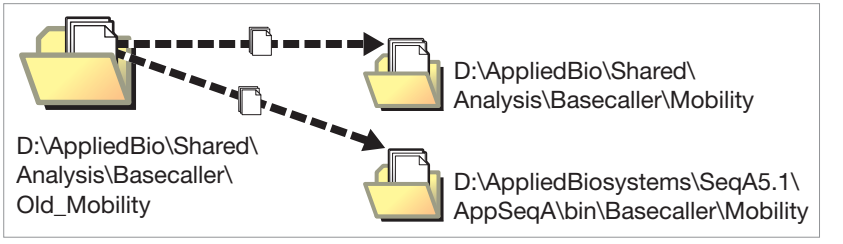


Figure 1-3 Where to copy 310 DyeSet/Primer files

This allows you to:

- Select a matrix or mobility file in a data collection sample sheet – the collection software continues to access the Shared folder
- Analyze a sample file – the analysis software accesses the Matrix or Mobility folder within AppliedBiosystems folder

Sequencing Analysis Software v5.1 Installed on a Separate Computer

If Sequencing Analysis software v5.1 is installed on a separate analysis computer, the installation location is *drive letter*:\AppliedBiosystems (see "Hard Drive Partitions" on page 1-4).

You must place a copy of the matrix and DyeSet/Primer files in the appropriate Matrix folder and Mobility folder within the Sequencing Analysis folder. You also must copy the files from the primary data collection computer to a floppy disk or CD, then copy them into the appropriate Matrix or Mobility folder within the Sequencing Analysis folder on the secondary analysis computer.

Copying and Pasting Matrix and DyeSet/Primer Files

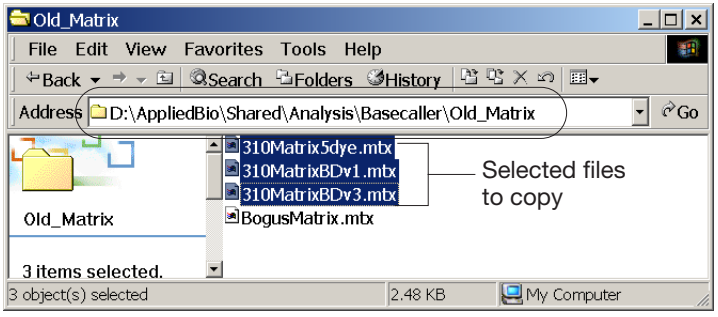
Copying and Pasting Matrix Files

To copy and paste matrix files:

1. Navigate to the location of the matrix files in data collection:

   D:\AppliedBio\Shared\Analysis\Basecaller\Old_Matrix

2. Open the Old_Matrix folder, then select the matrix files to be copied.

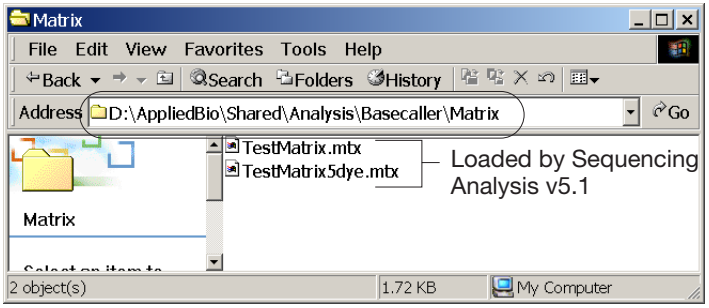


3. Press **Ctrl+C** to copy the files to the clipboard (or copy them onto a floppy disk or CD, if you are using a secondary analysis computer).

4. Copy the matrix files from the Old_Matrix folder to the Matrix folder:

   a. Navigate to the location of the Matrix folder used by the data collection software:

      D:\AppliedBio\Shared\Analysis\Basecaller\Matrix



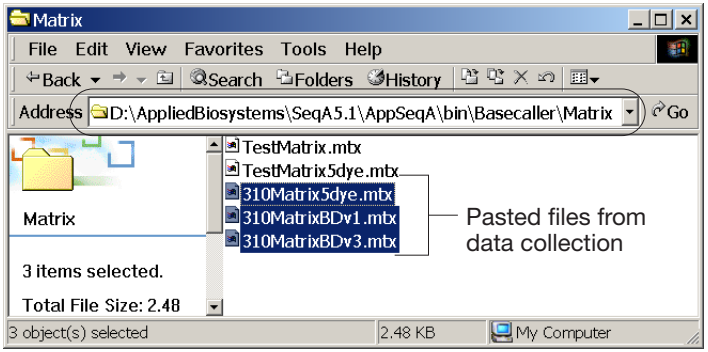
   b. Press **Ctrl+V** to paste the files into the Matrix folder.

5. Paste the matrix files from the Old_Matrix folder in data collection to the Matrix folder used by the analysis software:

   a. Navigate to the location of the Matrix folder used by the analysis software:

      D:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\Basecaller\Matrix

   b. Press **Ctrl+V** to paste the files into the folder or copy them from your floppy disk or CD.



Copying and Pasting Mobility Files

To copy and paste mobility files:

1. Navigate to the location of the DyeSet/Primer files in data collection:

   D:\AppliedBio\Shared\Analysis\Basecaller\Old_Mobility

2. Open the Old_Mobility folder, then select the DyeSet/Primer files to be copied.
3. Press **Ctrl+C** to copy the files to the clipboard (or copy them onto a floppy disk or CD, if you are using a secondary analysis computer).
4. Copy the DyeSet/Primer files from the Old_Mobility folder to the Mobility folder:
   a. Navigate to the location of the Mobility folder used by the data collection software:

      D:\AppliedBio\Shared\Analysis\Basecaller\Mobility

   b. Press **Ctrl+V** to paste the files into the Mobility folder.
5. Paste the DyeSet/Primer files from the Old_Mobility folder in data collection to the Mobility folder used by the analysis software:
   a. Navigate to the location of the Mobility folder used by the analysis software:

      D:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\Basecaller\Mobility

   b. Press **Ctrl+V** to paste the files into the folder or copy them from your floppy disk or CD.

Using the Matrix and Mobility Files

If the data collection and/or analysis software was running while the matrix and DyeSet/Primer files were copied into different folder locations, you need to close the applications and launch them again. Both the collection and analysis software read the contents of the Matrix, Mobility or any other folder at start up only.

Copying 377 Matrix Files

Note: Skip this section if you are analyzing sample files not generated on the ABI PRISM® 377 DNA Sequencer.

Matrix Folder Locations

If you are upgrading from Sequencing Analysis software v3.7 to Sequencing Analysis software v5.1, then you must copy your matrix files to a new folder location.

Sequencing Analysis Software v3.7 Matrix Folder Location

Sequencing Analysis software v3.7 installation location is D:\AppliedBio, the same installation folder location as the data collection software. Both applications access the Shared folder which contains the Matrix folder. The pathway to the Matrix folder is:

D:\AppliedBio\Shared\Analysis\Basecaller\Matrix

Sequencing Analysis Software v5.1 Matrix Folder Location

Sequencing Analysis software v5.1 installation location is:

D:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\Basecaller\Matrix

Because of the location change, the Sequencing Analysis software v5.1 cannot access the Shared folder used by data collection.

Duplicate Copies of the Matrix Files are Required

To be able to use your matrix files in both the data collection software and Sequencing Analysis software v5.1, you must place a copy of the matrix files in the Matrix folder within the Sequencing Analysis folder (Figure 1-4).

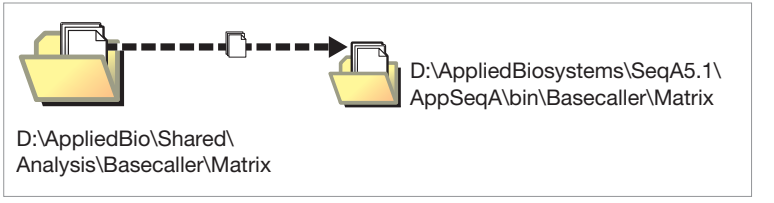


Figure 1-4 Where to copy 377 matrix files

This allows you to:

- Select a matrix file in a data collection sample sheet – the collection software continues to access the Shared folder
- Analyze a sample file – the analysis software accesses the Matrix folder within AppliedBiosystems folder

Sequencing Analysis Software v5.1 Installed on a Separate Computer

If Sequencing Analysis software v5.1 is installed on a separate analysis computer, the installation location is *drive letter*:\AppliedBiosystems (see "Hard Drive Partitions" on page 1-4).

You must place a copy of the matrix files in the Matrix folder within the Sequencing Analysis folder. You also must copy the files from the primary data collection computer to a floppy disk or CD, then copy them into the Matrix folder within the Sequencing Analysis folder on the secondary analysis computer.

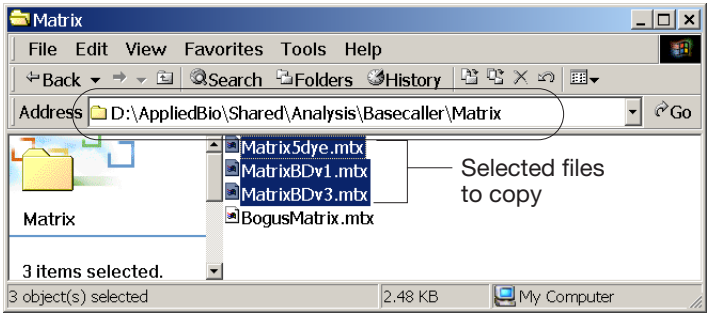
Coping and Pasting Matrix Files

To copy and paste matrix files:

1. Navigate to the location of the matrix files in data collection:

   D:\AppliedBio\Shared\Analysis\Basecaller\Matrix

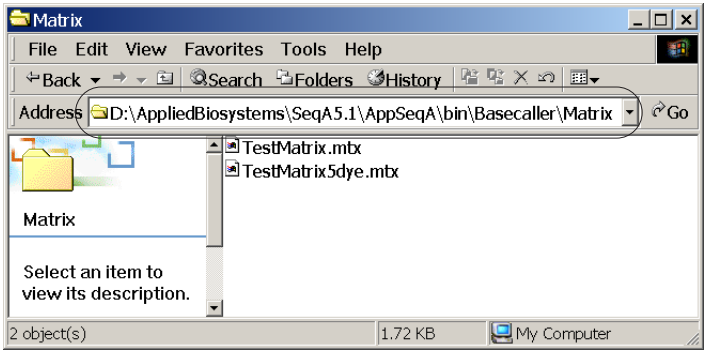
2. Open the Matrix folder, then select the matrix files to be copied.



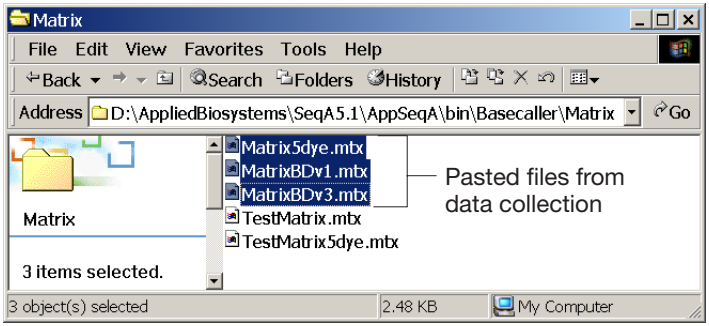
3. Press **Ctrl+C** to copy the files to the clipboard (or copy them onto a floppy disk or CD, if you are using a secondary analysis computer).

4. Navigate to the location of the Matrix folder used by the analysis software:

   D:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\Basecaller\Matrix



5. Press **Ctrl+V** to paste the files into the folder or copy them from your floppy disk or CD.



Using the Matrix Files

If the analysis software was running while the matrix files were copied into the Matrix folder, you need to close the application and launch it again. The analysis software reads the contents of the Matrix or any other folder at start up only.

Automated Sample Analysis

If the Sequencing Analysis software v5.1 is installed on a computer connected to an instrument, you need to set up the data collection software for automatic analysis.

For the:

- 3730/3730xl DNA Analyzer

  Refer to the *Applied Biosystems 3730/3730xl DNA Analyzers User Guide* (PN 4331468).
- 3100/3100-Avant Genetic Analyzer

  Refer to the *ABI PRISM® 3100/3100-Avant Genetic Analyzer User Guide for DNA Sequencing* (PN 4347102).
- 310 Genetic Analyzer

  Refer to Chapter 11, "Autoanalysis of 310 Sample Files."

Note: Automated sample analysis is not supported for the ABI PRISM® 3700 DNA Analyzer or ABI PRISM® 377 DNA Sequencer.

Sequencing Analysis and Primer Express Software

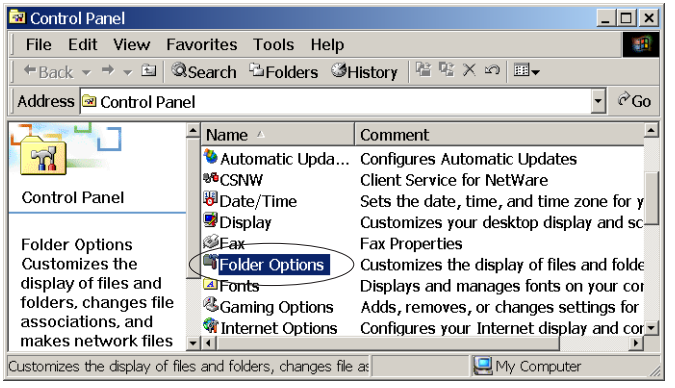
Note: If you do not have Primer Express® software loaded on your computer, then skip this section.

If you have only Sequencing Analysis on your system, then AB1 file type is associated with Sequencing Analysis software. If you double-click a sample file name or its icon ( ), Sequencing Analysis software automatically opens, as well as the sample file.

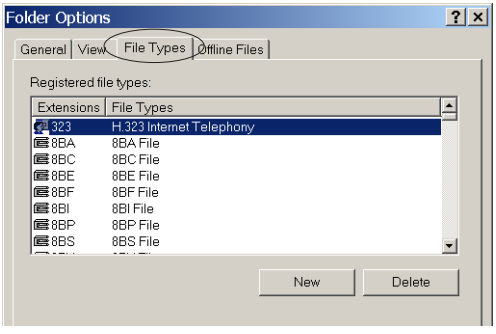
If you have BOTH Sequencing Analysis and Primer Express software, then the AB1 file may be associated with the Sequencing Analysis software or Primer Express software, depending on the order of the installation.

To change the software that is associated with the AB1 file type:

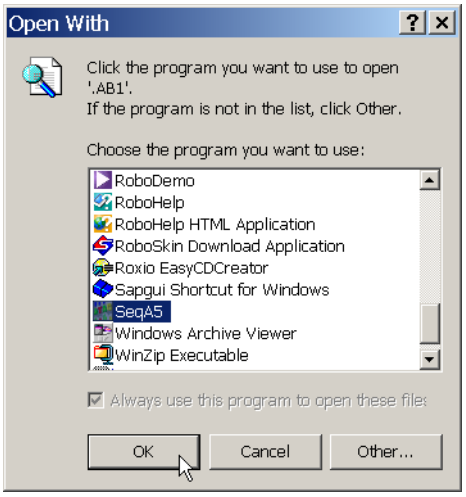
1. Select **Start > Settings > Control Panel**.



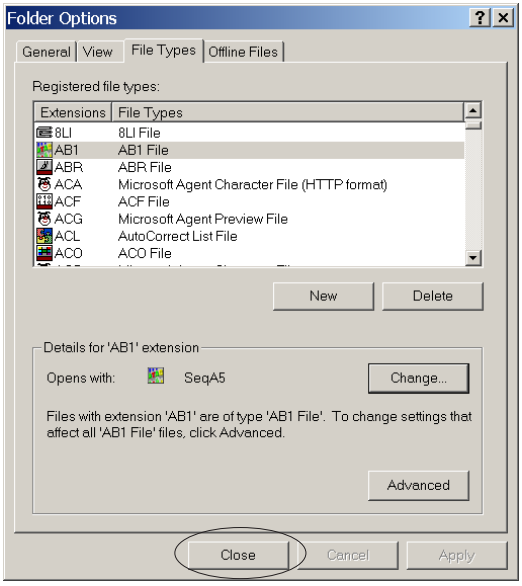
2. Double-click Folder Options, then select the **File Types** tab.



3. Select the File Type that you want to modify, then click **Change**. The Open With dialog box opens.



4. Highlight the program you want to use, then click **OK**. The Folder Option dialog box opens.



5. Click **Close**.

6. Close the Control Panel window.

Software Overview 2

This chapter covers:

About the Sequencing Analysis Software

Applied Biosystems Sequencing Analysis Software v5.1 performs the following tasks:

- Calls the bases
- Defines and displays mixed bases, if the KB basecaller is used
- Calculates and displays quality values, if the KB basecaller is used
- Calculates and display the clear range
- Calculates sample score
- Creates output files in ABI (.seq), FASTA (.seq), Phred (.phd.1), and standard chromatogram format (.scf) formats
- Generates an analysis report containing sample analysis statistics
- Prints data for each sample file
- Prints analysis report
- Creates an audit trail to track all changes to bases and analysis settings, if enabled.

New Features

New in Sequencing Analysis Software v5.1

Sequencing Analysis software v5.1 replaces ABI PRISM® Sequencing Analysis software v3.7 and v5.0. The following key features have been integrated into the application.

- ABI PRISM® 310 Genetic Analyzer automation support
- ABI PRISM® 3100/3100-Avant Genetic Analyzer (running v2.0 Data Collection) automation support
- Analysis of sample files generated on the ABI PRISM® 3700 DNA Analyzer, ABI PRISM® 377 DNA Sequencer and ABI PRISM® 310 Genetic Analyzer
- Make Matrix utility built into the application
- Printing enhancements
- KB basecaller v1.1
  - Option to call Ns when the quality threshold is not met
  - Option to process data with true or flat profile
- KB basecaller support for data generated on 310, 3100 and 3100-Avant systems

- File sharing

  Master analysis protocols are shared between Sequencing Analysis, and 3100/3100-Avant Data Collection software.
- Audit trail

  Optional feature to generate an audit trail of base changes. The audit trail information is created whenever a user:
  - Changes a base
  - Deletes or inserts a base
  - Changes the analysis settings
  - Changes the sample name
  - Analyzes the data

New Features in Sequencing Analysis Software v5.0

The following key features were integrated into Sequencing Analysis software v5.0.

- Single window interface that contains the Sample Manager and sample data windows
- KB basecaller v1.0

  The KB basecaller is a new algorithm for base calling which has the following improvements and features:
  - Calculation of mixed bases

    Mixed bases are one-base positions that contain two bases. These bases are assigned the appropriate IUB code.
  - Calculation and display of quality values (QVs) for pure and mixed bases.

    The QV is a per-base estimate of the basecaller accuracy.
- Generation of additional file formats

  The file formats are Phred (.phd.1) and standard chromatogram format (.scf) formats.

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

- Scrolling and zooming in the horizontal and vertical directions for both raw and analyzed data view
- Resizable Sample File Name column in the Sample Manager
- Multiple files can be viewed from the Sample Manager and Sample Navigator panes
- Scrolling multiple samples horizontally

- Ability to toggle between the Sample Manager and Sample Navigator view
- Analysis Protocol

  An analysis protocol contains all the settings necessary for analysis and post processing. They replace preference settings used in previous versions of Sequencing Analysis software. A protocol is stored in the sample file.

  There are two types of analysis protocols:

  – Per-sample analysis protocol – is the protocol stored within a sample file. This protocol can be edited. The change affects the protocol for the selected sample only.

  – Master analysis 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.
- New Post Processing option: 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.
- Calculation of the length of read (LOR)

  Measurement of the length of quality bases. The LOR is user definable in the Display Settings dialog box and is displayed in the Analysis report.
- Calculation of 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.This information is displayed in the analysis report.
- Generation of an Analysis Report

  Analysis report show the status of the data analysis. The report can be used to help troubleshoot and provide easy assessment of data quality.
- File sharing

  Master analysis protocols are shared between Sequencing Analysis, and 3730/3730xl Data Collection software.

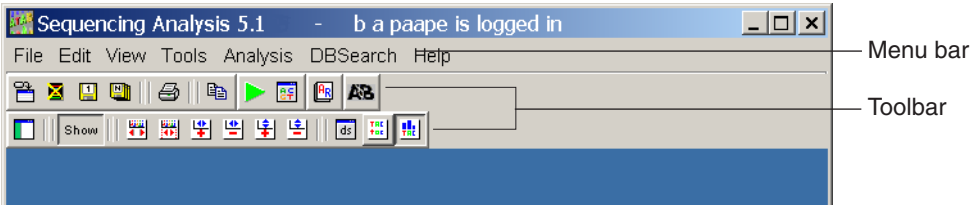
Software Tour

Introduction

The following section is a brief tour of the common software application windows.

Sequencing Analysis Main Window

The Sequencing Analysis Main Window opens when you launch the software. It contains a menu bar and a toolbar that contains buttons for frequently used functions. All operations are performed within this window.



For information on menu commands and toolbar icons, see Appendix B, “Menu Commands and Toolbar Buttons.”

Sample Manager Window

There are two panes within the Sample Manager window.

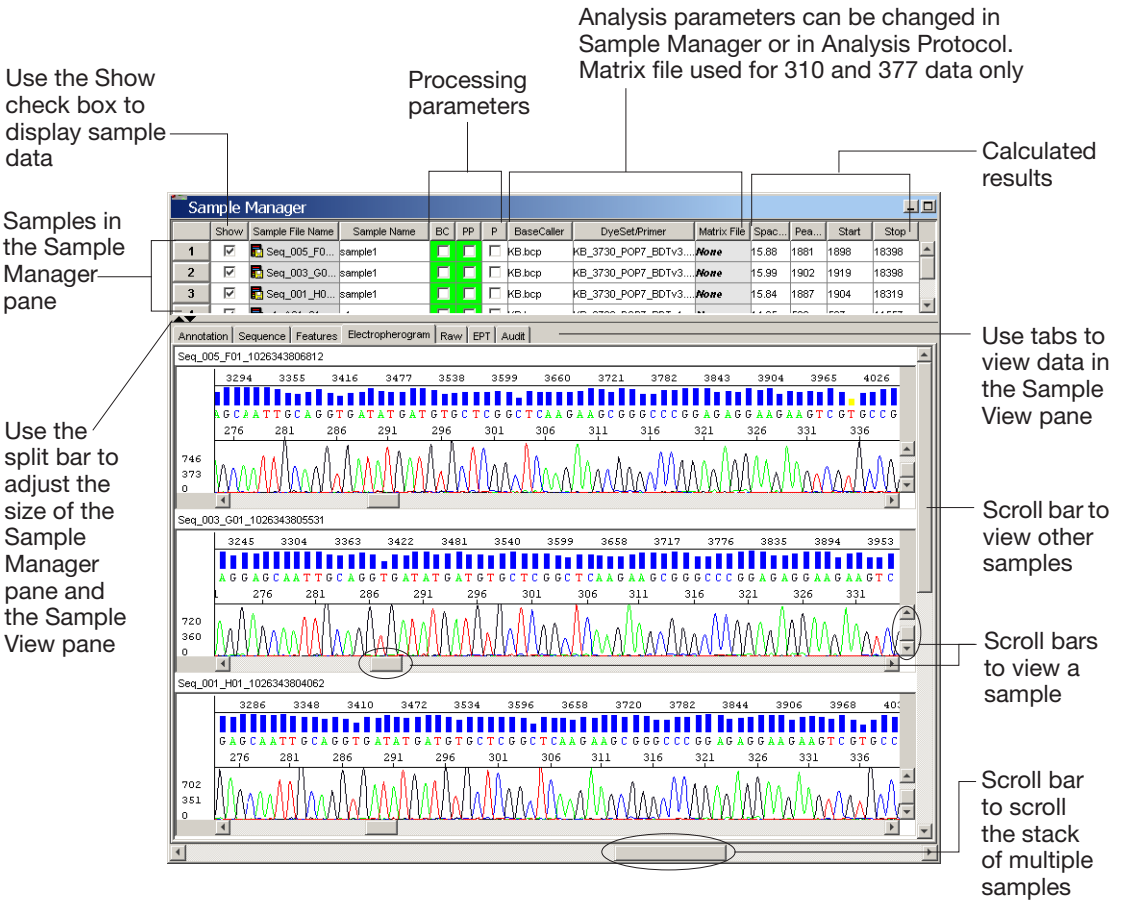
- Sample Manager
- Sample Navigator

To switch between these panes, click , or select **View > Sample Manager** or **Sample Navigator**.

Sample Manager Pane

The Sample Manager view is the default view when the software is launched.

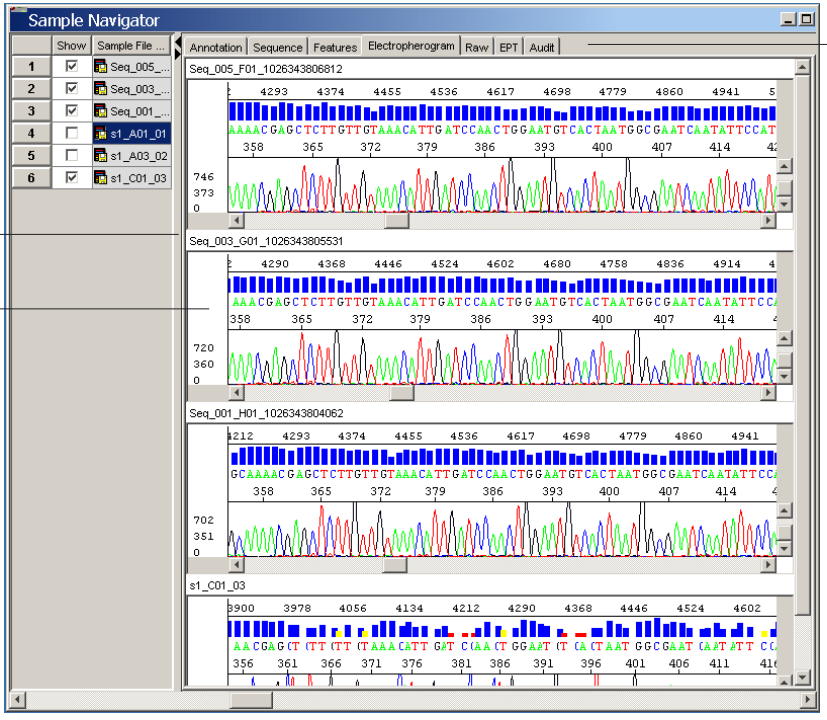
This pane contains samples and the various analysis parameter values. Samples are viewed, analyzed, edited, printed and save in this window pane.



Sample Navigator Pane

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

This pane displays the sample names and sample views. Samples can be edited and saved in this pane.



For more information on samples, refer to Chapter 3, “Sample Files in the Sample Manager.”

Sample View Pane

The sample view is used for viewing all the data characteristics of the sample. Each sample view tab displays information below.

Table 2-1 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.
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.
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 in the Authentication and Audit tab of the Options dialog box.

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

Examples of the seven tab views are displayed in Table 2-2, “Sample Views,” on page 2-9.

Table 2-2 Sample Views

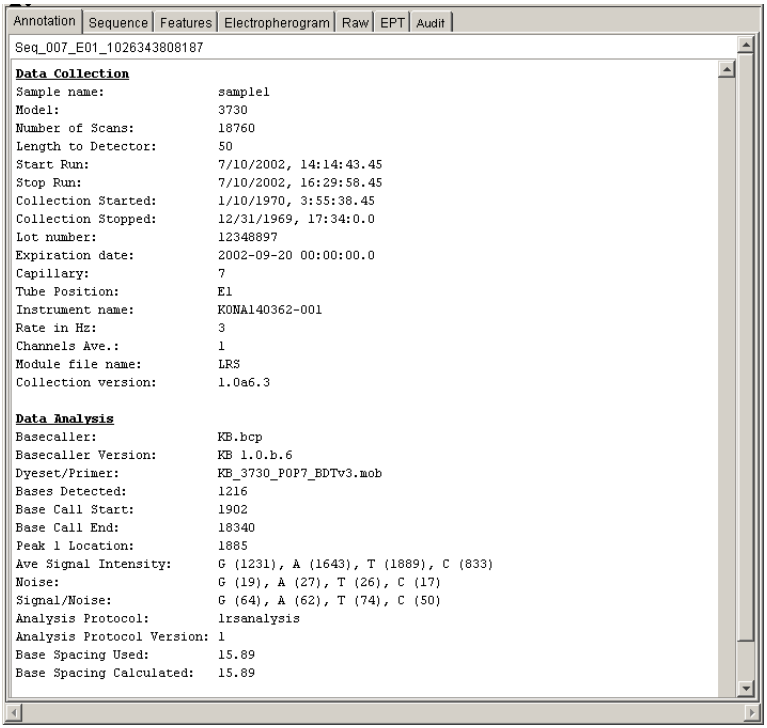
Sample View Tab	Example Data
Annotation	

Table 2-2 Sample Views *(continued)*

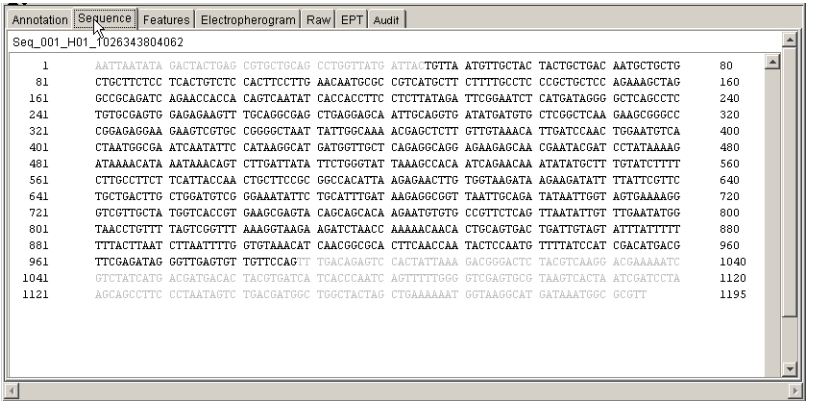
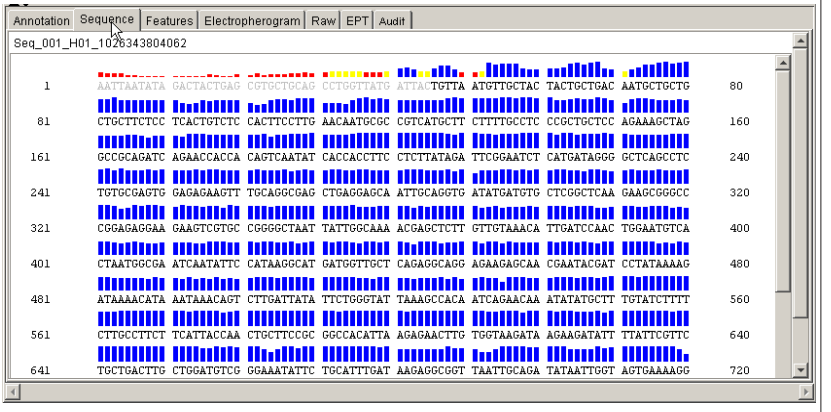
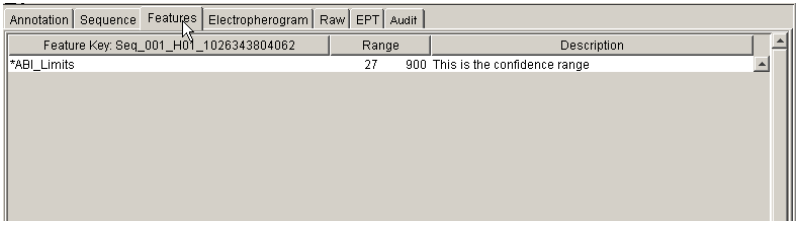
Sample View Tab	Example Data
Sequence	 Use the Show/Hide QV button  to display/hide QVs. 
Feature	

Table 2-2 Sample Views *(continued)*

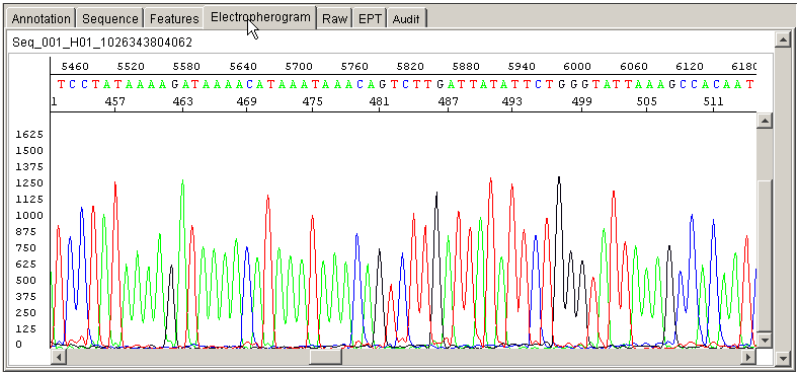
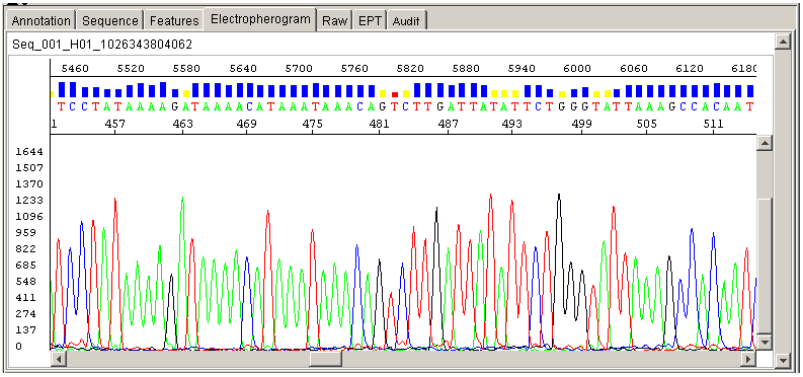
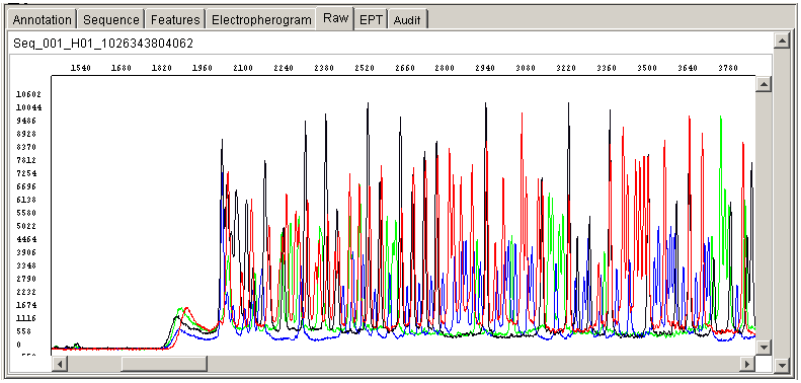
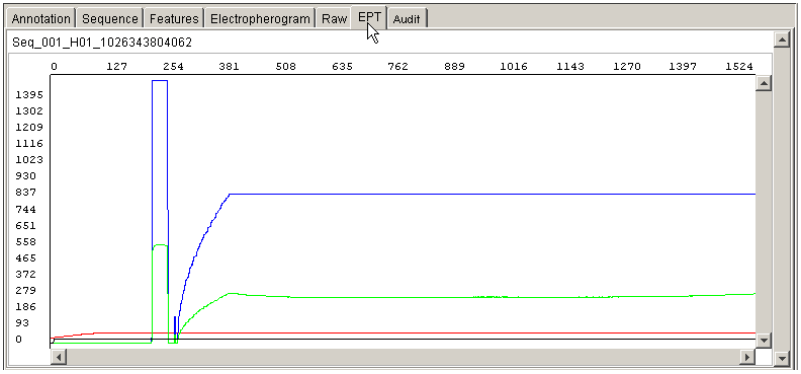
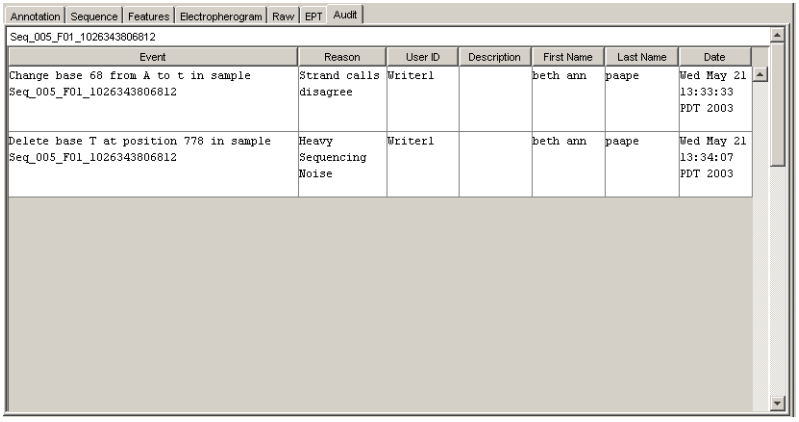
Sample View Tab	Example Data
Electropherogram	 Use the Show/Hide QV button  to display/hide QVs. 
Raw	

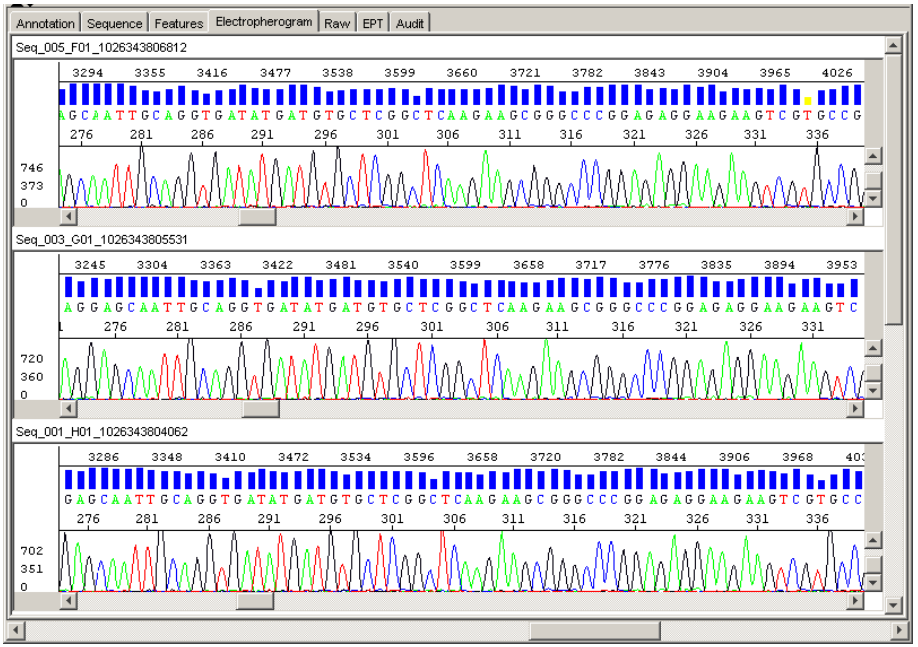
Table 2-2 Sample Views *(continued)*

Sample View Tab	Example Data
EPT	
Audit	

For more information on samples, refer to “About Sample Files” on page 3-4

Viewing Multiple Samples

The same view tabs can be used to view multiple samples simultaneously. An example of multiple samples in the Electropherogram view is shown below.



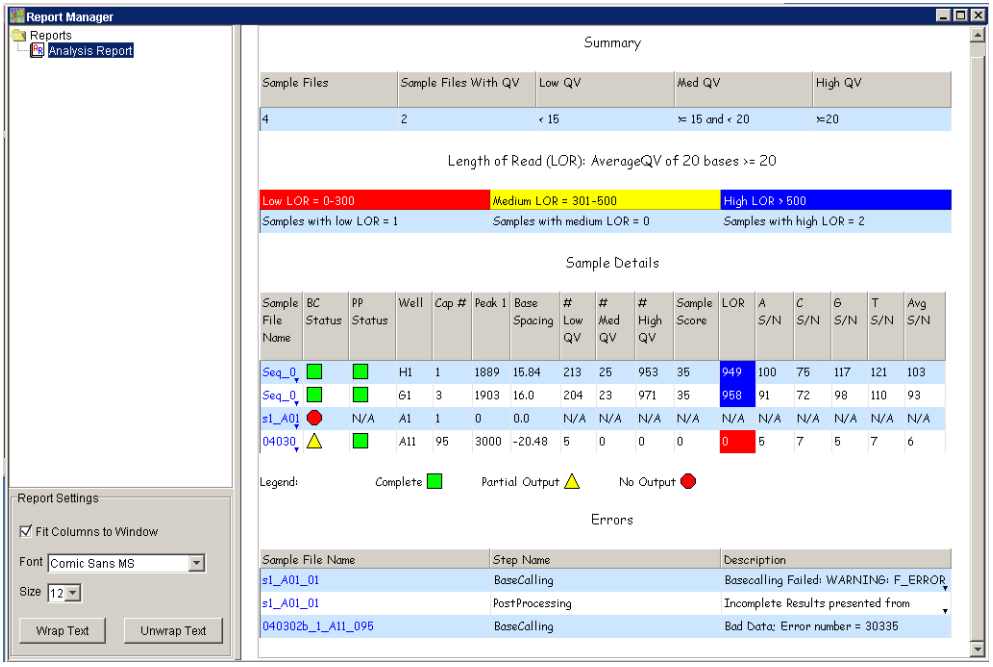
Analysis Reports

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  .

An Analysis Report is displayed below. 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 a partial of failed analysis. The 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



For more information, refer to Chapter 7, “Analysis Report.”

Overview of Operations

Analysis Outline

All analysis, post processing, and printing occur in the Sample Manager. Analyze and review your sample files in the following steps:

1. Add sample(s) to the Sample Manager.
2. Show the sample data.
3. Modify an analysis protocol (optional).
4. Analyze the data.
5. Review and edit your results and generate an analysis report.
6. Save the sample files.

Starting the Software

Double-click  (Sequencing Analysis 5.1 desktop shortcut).

Adding Samples to the Sample Manager

Use the Add Samples function to add samples to the Sample Manager for analysis, printing, viewing, editing data, or generating an analysis report.

To add samples to the Sample Manager:

1. Click  (Add Sample(s)) or select **File > Add Samples**.
2. Navigate to the location of the samples.
3. Select the files that you want to add to the Samples To Add pane of the dialog box.

To add …	Do this ...
A single file to the list	Select the file, then click **Add Selected Samples**.
Continuous multiple files	Use the Shift key to select samples, then click **Add Selected Samples**.
Discontinuous multiple files	Use the Ctrl key to select samples, then click **Add Selected Samples**.
All samples in a single folder	Select the folder, then click **Add Selected Samples**.

a. Click **OK** in the Add Samples dialog box.

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

b. Continue with "Showing the Sample Data."

Showing the Sample Data

To show the data:

1. Use the Show check box to show data for one or more sample files (see "Showing Sample File Data" on page 3-13 for more information.)

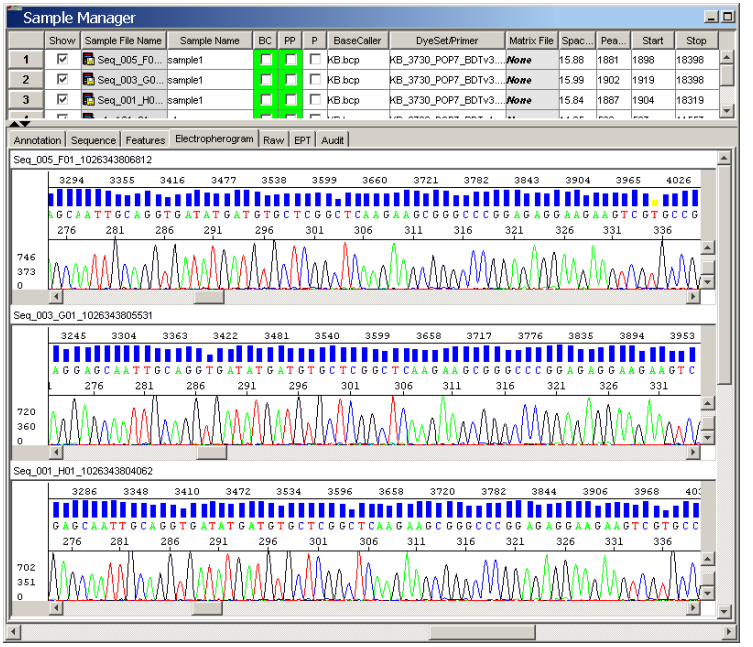


Figure 2-1 Samples in the Sample Manager view

2. To show the data in the Sample Navigator view, click , or select **View > Sample Navigator**.

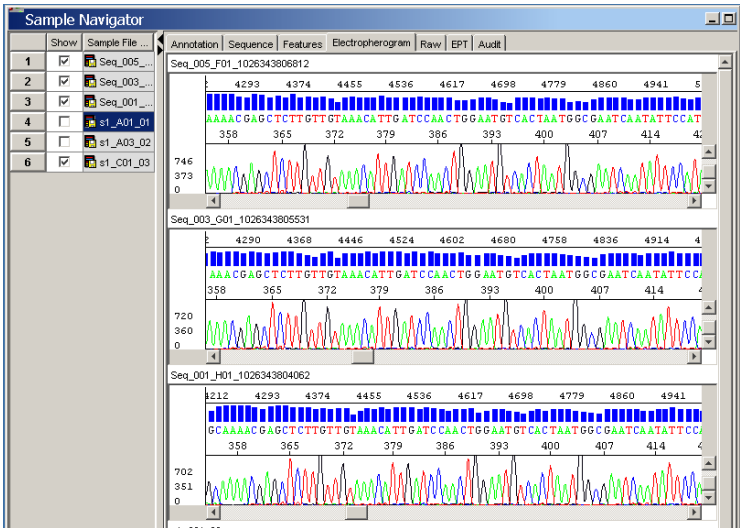


Figure 2-2 Samples in the Sample Navigator view

Editing and Applying an Analysis Protocol

Note: Skip this section, if the existing analysis protocol settings are correct.

Editing a Per-sample Analysis Protocol

Use the Analysis Protocol function to change the analysis protocol for an individual sample in the Sample Manager.

To edit the protocol for a single sample:

1. In the Sample Manager, select a sample row.
2. Click  or select **Analysis > Analysis Protocol**.
3. Edit the protocol, as appropriate.

   Make changes in the General, Basecalling, Mixed Bases, and Clear Range tabs.
4. Click **OK**.

Note: The changes apply only to the protocol for the selected sample.

Editing and Applying a New Analysis Protocol

Use the Analysis Protocol Manager function to change the analysis protocol settings for samples and the master protocol.

To edit and apply a new analysis protocol:

1. Select samples in the Sample Manager.
   - Use the Shift key to select continuous samples.
   - Use the Ctrl key to select discontinuous samples.
2. Select **Analysis > Analysis Protocol Manager**.
3. In the Analysis Protocol column, select the protocol you want to edit.
4. Click the **File** button, then select **Open,** or double-click the protocol name.
5. Make changes in the General, Basecalling, Mixed Bases, and Clear Range tabs, as appropriate.

   Note: For more information on this process, refer to "Creating and Editing Analysis Protocols" on page 8-13.
6. Click **OK** to save the protocol and close the Sequence Analysis Protocol Editor dialog box.

   Note: The version number increments by one.
7. Click **Apply to Selected Samples**.
8. Click **Done** to close the Analysis Protocol Manager dialog box.

Note: The changes apply to the selected sample and the master protocol.

Analyzing the Data

To start the analysis:

Click  (Start Analysis) to start the basecalling, post processing, and printing tasks that you selected.

Basecalling

If the BC parameter (basecalling) is selected, the selected basecaller performs the following:

- KB basecaller
  - Calls mixed bases, if the mixed base option is selected.

    Mixed bases are one base positions that contain two bases. The basecaller assigns A, C, G, T, or an IUB code 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, if the mixed base option is selected.
  - Call Ns when the quality threshold is not met (if selected)
  - Processes data with true or flat profile

Or,

- ABI basecaller

  The basecaller assigns A, C, G, T, or N to every base (no mixed base calling or QV options).

Post Processing

When the PP parameter (post processing) is selected, the clear range is calculated.

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 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.

Printing

When the P parameter (printing) is selected, the sample views are printed automatically after analysis and post processing.

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

Reviewing Overall Results and Generating an Analysis Report

To review overall results:

1. Review your results in the Sample Manager.
   a. Look for green, yellow, or red boxes for the BC parameter. Green indicates the process was successful, yellow indicates poor quality data, and red indicates failure.

      Note: The yellow result applies to samples analyzed with the KB basecaller.

   b. Look for green or red boxes for the PP and/or P parameters. Green indicates the process was successful and red indicates failure.
   c. Review the base spacing, peak 1 location, and, start and stop points. A red value in the Base Spacing column means the spacing could not be calculated and the default value was used for analysis.
2. Review the analysis report.
   a. Click  (Analysis Report) to generate and display the analysis report.
   b. Review the data in the report.

      To export the report, select **File > Export Report**. The file will be exported in a tab-delimited format.

Reviewing Samples and Editing Bases

To review samples and edit base:

1. Select a sample file.
2. Review your results in the sample file:
   a. Review the raw, analyzed, and EPT data.
   b. Review low-quality basecalls (samples analyzed with the KB basecaller) and check for errors.
3. Edit the bases, as needed.

   When you edit bases the QVs change, depending on the change made. If you:
   - Insert a base – No QV is added
   - Delete a base – QV is deleted
   - Change a base – QV will have the same numerical value but be displayed as a gray bar
4. Save the sample file (sample files are not automatically saved after reanalysis or editing).
   - To save selected samples, click  or select **File > Save Sample(s)**.
   - To save all samples, click  or select **File > Save All Samples**.

   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.

Sample Files in the Sample Manager

3

This chapter covers:

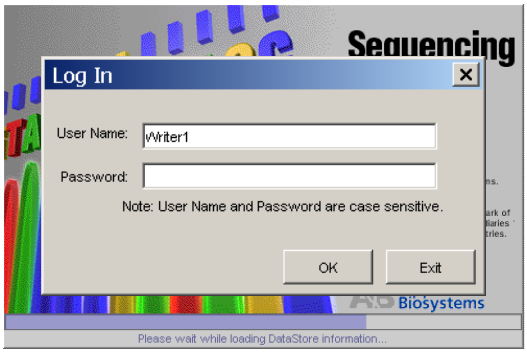
Starting Sequencing Analysis Software

Opening Sequencing Analysis Software

To start the software:

1. Double-click the Sequencing Analysis 5.1 desktop shortcut or select **Start > Programs > Applied Biosystems > Sequencing Analysis 5.1 > Sequencing Analysis 5.1**

   While the program is loading, the splash screen and Log In dialog box opens.



2. Complete the Log In dialog box:

   a. Enter a new user name, if applicable.

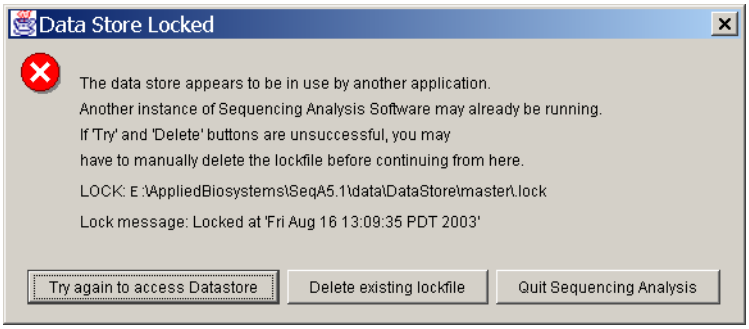
   b. Type in your password.

   c. Click **OK**.

   The Sequencing Analysis main window opens when dialog box has been completed, and the program has completed loading.

If Sequencing Analysis Does Not Open

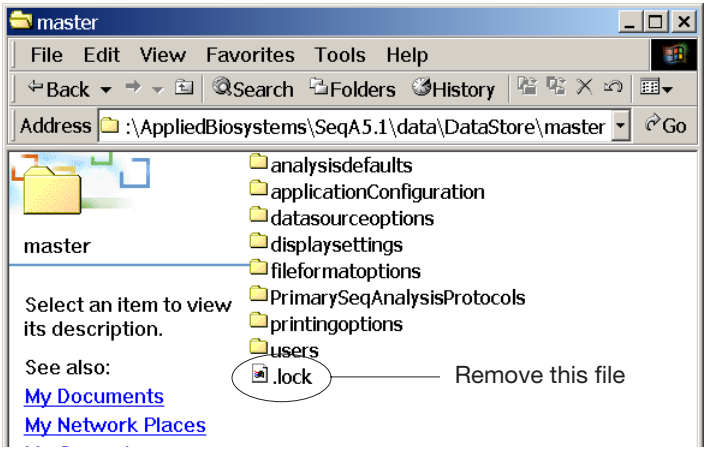
Sometimes when the software has been previously shut down improperly, the software cannot access the DataStore information. When this happens, the following dialog box opens instead of the Sequencing Analysis main window.



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:

  drive letter:\AppliedBiosystems\SeqA5.1\data\DataStore\master\.lock



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

About Sample Files

Sample files in Sequencing analysis v5.1 contain the following information about the DNA sequence:

- Raw data, as captured by the instrument before any post-collection processing
- Analysis settings such as quality values and mixed bases
- Post processing settings 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. Thus, sample files contain the target DNA sequence plus all the historical information about the analysis necessary to interpret the data and processing parameters.

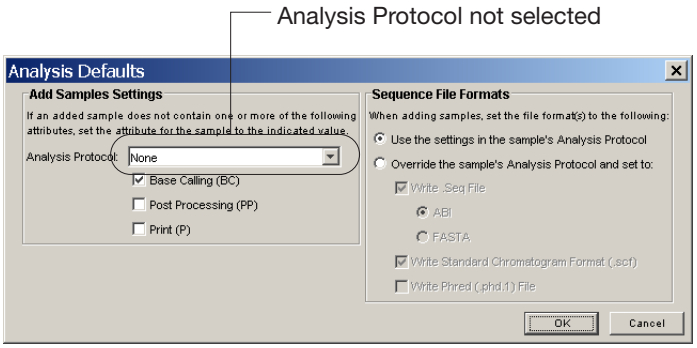
Creating Analysis Defaults for Sample Files

What is an Analysis Default?

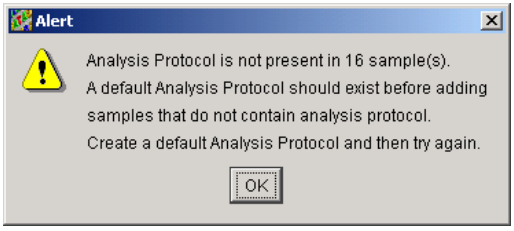
When a sample file is added to the Sample Manager, then it passes through the analysis defaults. It contains processing parameter settings (basecalling, post processing, and printing); sequence file format settings (.seq, .scf and .phd.1); and an analysis protocol. The analysis protocol is assigned to the sample only if it does not already contain one.

Analysis Defaults

An analysis default was included with the installation of your software. The analysis protocol is set to *None*.



If sample files are added to the Sample Manager that do not contain associated analysis protocols, and one is not selected in the Analysis Defaults, then the following alert box opens.



Samples are not added to the Sample Manager until an analysis default with an analysis protocol is created and applied to the sample files.

What is an Analysis Protocol?

An analysis protocol contains all the settings necessary for basecalling and post processing the data. Analysis protocols replace preference settings used in previous versions of Sequencing Analysis software. A protocol is stored in the sample file once it has been applied to the file and saved.

Sample Files That Contain Analysis Protocols

Sample files generated on the following instruments contain analysis protocols:

- Applied Biosystems 3730/3730*xl* DNA Analyzers running versions 1.0 or 2.0 Data Collection software
- ABI PRISM® 3100/3100-*Avant* Genetic Analyzers running version 2.0 Data Collection software.

Sample Files That Do Not Contain Analysis Protocols

Sample files generated on the following instruments *do not* contain analysis protocols:

- ABI PRISM® 3700 DNA Analyzer
- ABI PRISM® 3100 Genetic Analyzer, running version 1.1 (or earlier) Data Collection software
- ABI PRISM® 3100-*Avant* Genetic Analyzer, running version 1.0 Data Collection software
- ABI PRISM® 377 DNA Sequencer
- ABI PRISM® 310 Genetic Analyzer

How the Analysis Defaults are Applied

When sample files are added to the Sample Manager, the analysis protocol is applied. 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) and analysis defaults are applied to the samples. See Figure 3-1 on page 3-7.

IMPORTANT! The DyeSet/Primer file must match the chemistry and basecaller type that you are using in both data collection and the analysis protocol.

The same basecaller related files used in the data collection sample sheet must be used in the Analysis Protocol you create. If in a sample sheet, the selected DyeSet/Primer file is associated with an ABI basecaller but in the Analysis Protocol, the selected basecaller is the KB basecaller, then the analysis fails. To correct these sample files, select the correct basecaller and DyeSet/Primer file from the drop-down lists the Sample Manager, and reanalyze the files.

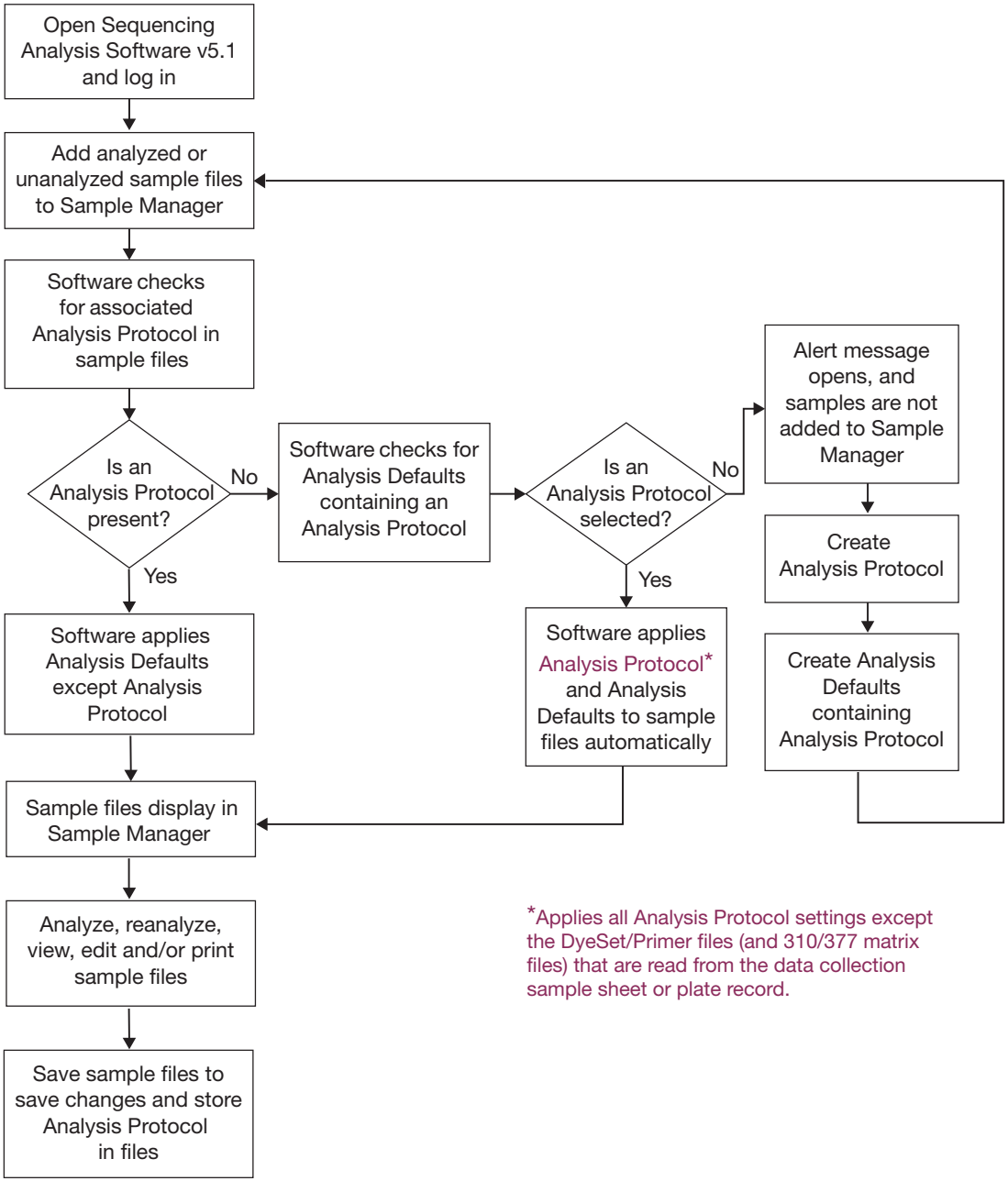


Figure 3-1 How Sequencing Analysis software applies Analysis Defaults to sample files

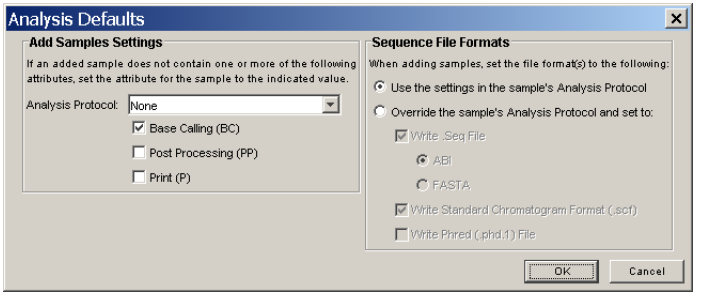
Setting Up Analysis Defaults

Refer to Chapter 8, "Analysis Protocols, Options, and Analysis Defaults," for detailed information on how to create analysis protocols and analysis defaults.

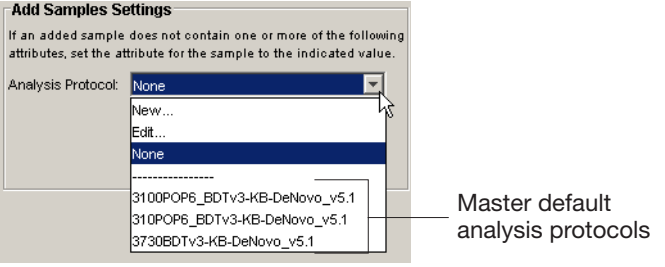
IMPORTANT! When creating or editing an analysis protocol, use Appendix C, Basecallers and DyeSet/Primer Files to select the correct combination of basecaller and DyeSet/Primer for successful analysis.

To set up the Analysis Defaults:

1. Select **Analysis > Analysis Defaults**.



2. In the Add Samples Settings section:
   a. In the Analysis Protocol drop-down list, do one of the following:
      – Select one of the default master analysis protocol, if the basecaller, DyeSet/Primer file, matrix file, and other settings are correct for your run conditions. See "Default Master Analysis Protocol Settings" on page 8-4.



– Create a new analysis protocol or edit an existing one.

Note: To create or edit an analysis protocol, use the Analysis Protocol drop-down list to select New or Edit, then refer to page 8-3 to define your settings.

b. Select/deselect the Basecalling (BC), Post Processing (PP), and Print (P) options as applicable.

3. In the Sequence File Formats section, select to use the current settings from the Analysis Protocol or override them.
4. Click **OK**.

Adding Sample Files to the Sample Manager

Introduction

You can add sample files of the file type .ab1 to the Sample Manager window in any of the following ways:

- Double-click the icons for each file you want to add to the Sample Manager.
- With the Sample Manager active, click either the **Add Samples** button or select **File > Add Sample(s)**.
- Drag selected files to the shortcut icon.
- Select the icon for each file you want to add to the Sample Manager. Right-click, then select **Open with SeqA5App**.

Files are listed in the order in which they were added to the list.

Note: If a Missing Analysis Defaults alert box displays while adding samples to the Sample Manager, see "Creating Analysis Defaults for Sample Files" on page 3-5 to add analysis defaults to the samples.

Adding Files by Double-Clicking on the File Icons

To add sample files to the Sample Manager:

1. Double-click the icon for the file you want to add to the Sample Manager.

   Note: If Sequencing Analysis is not already open, double-clicking the file also opens the software.

2. Repeat step 1 to add additional sample file(s).

Adding Sample File(s) Using the Open with Command

To add multiple files using the Open with command:

1. Select the sample file(s) you want to add to the Sample Manager.

   Note: Limit the number of samples to 15.

2. Right-click, then select **Open with SeqA5App**.

   The files are added to the Sample Manager.

Adding Sample File(s) by Dragging Them onto the Program Icon

To drag samples onto the program icon:

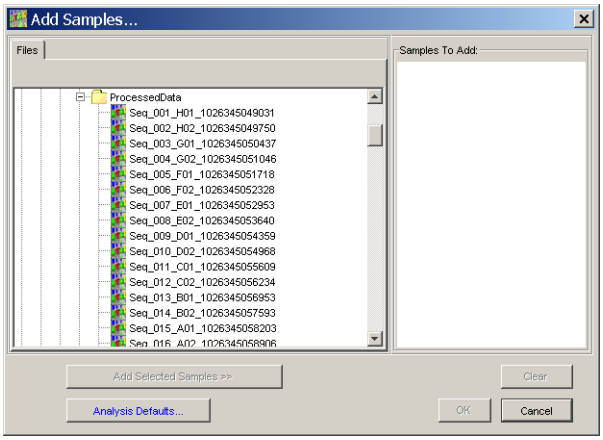
1. Select the sample file(s) you want to add to the Sample Manager.
2. Drag the sample file(s) onto the Seq 5.1.exe or shortcut icon.

   The files will be added to the Sample Manager.

Adding Files from Within Sequencing Analysis Main Window

To add sample files from within the window:

1. Click  or select **File > Add Samples**.
2. In the Add Samples dialog box, locate, then open the folder that contains the files you want to add to the Sample Manager window.



3. In the Samples To Add pane of the dialog box, add the files that you want in the Sample Manager.

To add ...	Do this ...
A single file to the list	Select the file, then click **Add Selected Samples**.
Some or all of the files to the list	• Add the files individually, or • Use the Ctrl key to select discontinuous samples, then click **Add Selected Samples**.
A folder containing sample files	Select the folder, then click **Add Selected Samples**. **Note:** Sample files located in subfolders are not added to the Sample Manager.

4. Click **OK** in the Add Samples dialog box.

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

Removing Samples from the Sample Manager

If you do not want a sample processed, it can be removed from the Sample Manager window at any time except when the program is currently processing that sample. However, you do not need to remove a file from the list to avoid processing it. The Sequencing Analysis software decides whether or not to process files based on the current information in the BC (Basecalling), PP (Post Processing), and P (Printing) parameter check boxes. If the check box for a processing option is empty, that process for the sample is skipped.

Removing a Sample

To remove a single sample from the Sample Manager window:

1. Select a sample row number.
2. Do one of the following:
   - Click  or select **File > Remove Samples.**
   - Use the delete key.

Removing Multiple Samples

To remove …	Then ...
All the samples	Select **File > Remove All Samples**.
Multiple samples that are not next to each other	1. Press the Ctrl key while clicking the row number of each file to be removed. 2. Click  or select **File > Remove Sample(s)**.
Multiple files that are next to each other	1. Click the row number of the top-most file to be removed, then press the Shift key while clicking the row number of the bottom-most file to be removed. or, Drag the mouse down the row number column. 2. Click  or select **File > Remove Samples**.

Showing Sample File Data

Sample files added to the Sample Manager do not automatically display in the Sample View pane. Use the table below to display data for one or more sample files:

To show the data for …	Do this ...
A single sample	Double-click the sample file name or select the corresponding **Show** check box.
Multiple continuous samples	Shift-click or Shift-drag the sample row numbers to select the sample files, then click Show or select **View > Show Data Displays**.
Multiple discontinuous samples	Ctrl-click the sample row numbers to select the sample files, then click Show or select **View > Show Data Displays**.
All samples	Select the empty box above row number 1 or Shift-drag the sample row numbers to select all samples, then click Show or select **View > Show Data Displays**.

Note: Double-clicking a sample file name, deselects the Show check box for all checked samples (if any), and checks the selected sample file only.

Sample Window Views

About the Sample Window

The main portion of the Sample window contains the information pertaining to the sequence. Use this window to view or edit the sequence data.

There are seven views available in the Sample Manager or Navigator panes.

- Annotation
- Sequence
- Feature
- Electropherogram
- Raw (Data)
- EPT
- Audit

To change the view, select the tab for the view you want to see.

Summary of Sample Window Views

The following table provides a description for each view. For more detail on each view, see the sections that follow.

Table 3-1 Parts of the Sample Window

View	Description
Annotation	Summary of the sample information written by the data collection and analysis software (see page 3-16).
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 page 3-19).
Feature	The features that were found in the sequence by the post processing (clear range) (see page 3-21).
Electrophero-gram	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 page 3-23).

Table 3-1 Parts of the Sample Window *(continued)*

View	Description
Raw	The raw data collected by the instrument (see page 3-27).
EPT	A plot of run voltage, current, power and temperature values (see page 3-30).
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 page 3-32).

Annotation View

The Annotation view shows:

- The sample information you entered in the Data Collection software.
- Additional information entered by the Data Collection and Analysis software (for example, the start and stop times).

Viewing

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

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

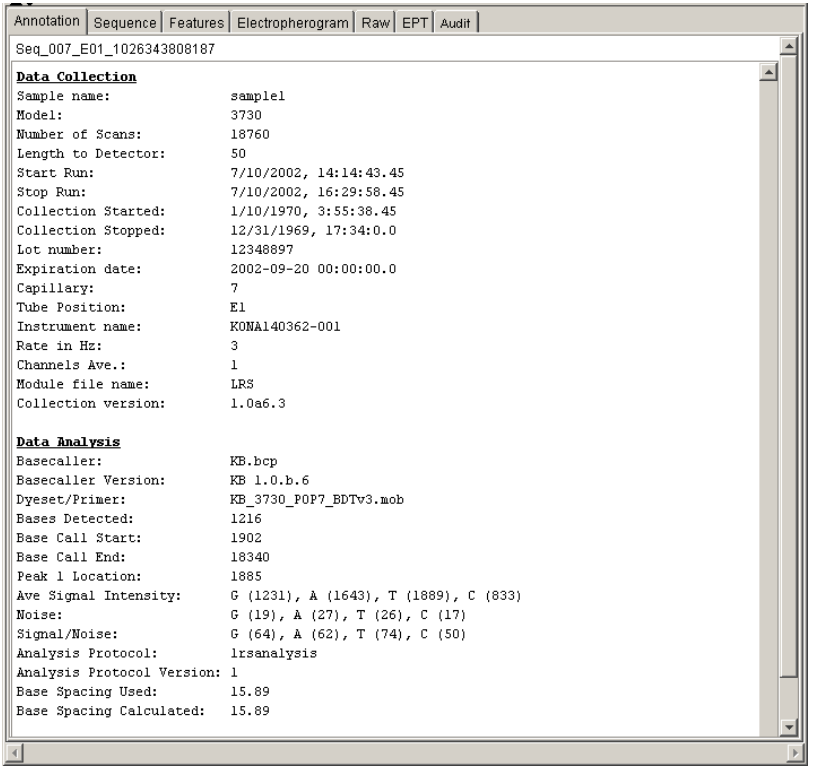


Figure 3-2 Sample in the Annotation view

Annotation View Information

Table 3-2 Annotation View Information

Field	Description
Data Collection	
Sample name	Name entered in the Sample Name column of the Plate Record of data collection.
Model	Instrument model used to collect the data.
Number of Scans	Range of the collected data points that were used to analyze the data.
Start and Stop Run	Date and time the run started and finished.
Collection Started and Stopped	Date and time the data collection started and finished.
Lot number	Lot number of the polymer used for the run.
Expiration date	Expiration date of the polymer used for the run.
Capillary	Capillary number used to electrophorese the sample.
Tube position	Well position of the sample.
Instrument name	Name of the instrument.
Rate in Hz	Data sampling rate.
Module file name	Name of the module file used to run the sample.
Collection version	Version of software used to collect data.
Data Analysis	
Basecaller	Name of the basecaller used to analyze the data.
Basecaller Version	Version of the basecaller used to analyze the data.
DyeSet/Primer file	DyeSet/Primer file used during analysis to adjust for mobility shifts.
Bases detected	Total number of bases in the sample.
Base call start and end	Raw data start and stop points (scan numbers) used for data analysis.
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.

Table 3-2 Annotation View Information *(continued)*

Field	Description
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.
Matrix Name	Matrix file used to multicomponent the data. Used for 310 and 377 data only.
Analysis Protocol	Analysis protocol last used to analyze the data.
Analysis Protocol Version	Version number of the analysis protocol last used to analyze the data.
Spacing Used	Base spacing as calculated by the basecaller or defined by the user.
Spacing Calculated	Base spacing as calculated by the basecaller.

Printing

To print the contents of the Annotation view window, see page 4-21.

Sequence View

The Sequence shows:

- The nucleotide sequence called from the data.
- 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.

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.

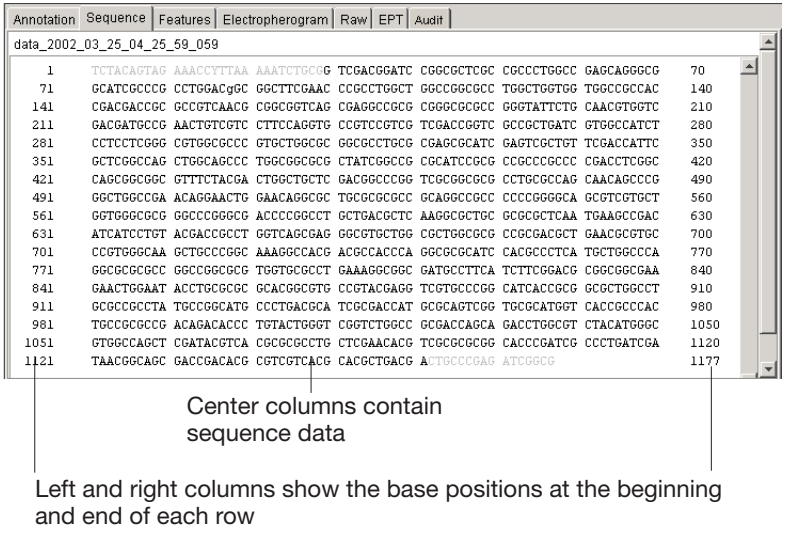


Figure 3-3 Single sample in the Sequence view

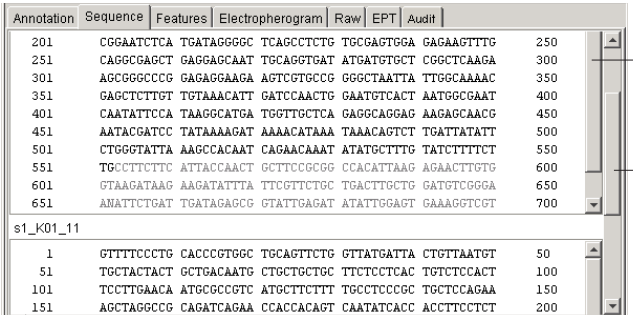


Figure 3-4 Multiple samples in the Sequence view

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)	“Searching for a Pattern in a Sequence” on page 4-12.
Edit the sequence using any of the standard Windows-based computer commands	“Editing Bases in Sequence or Electropherogram View” on page 4-18.
Display reverse complement	“Showing the Reverse Complement Data” on page 4-15.
Display quality values and numbers	“Viewing the Quality Values” on page 4-16.
Print the contents of the window	“Printing the Sample Window Views” on page 4-21.

Feature View

This feature view shows features that were added to the analyzed sequence data by the post processing.

Viewing

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

Note: If the sequence data has not been post processed, the Feature window is empty.

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

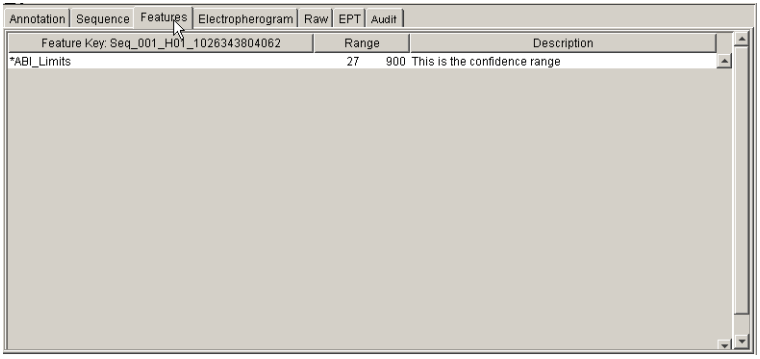


Figure 3-5 Single sample in the Feature view

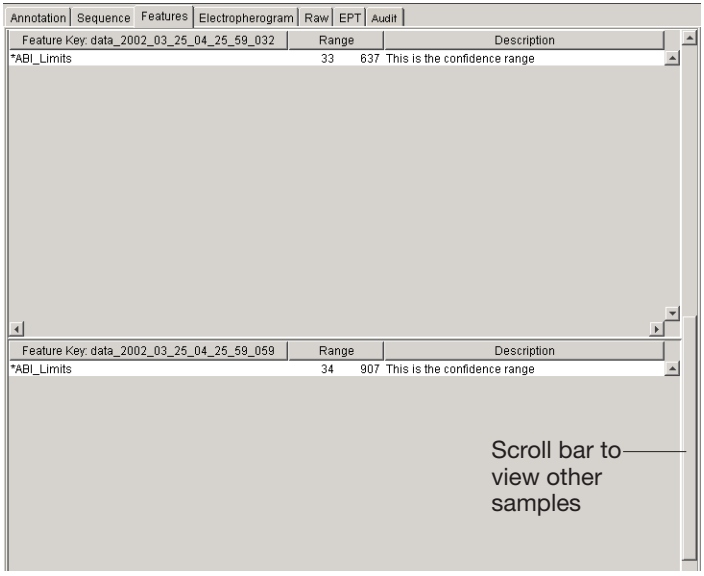


Figure 3-6 Multiple samples in the Feature view

Printing

To print the window contents, see "Printing the Sample Window Views" on page 4-21.

Electropherogram View

The electropherogram view 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.

Viewing

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.

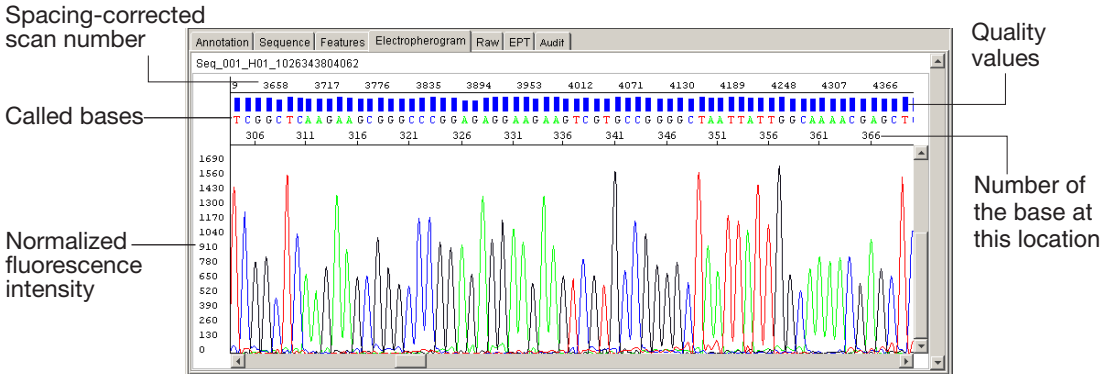


Figure 3-7 Single sample in the Electropherogram view

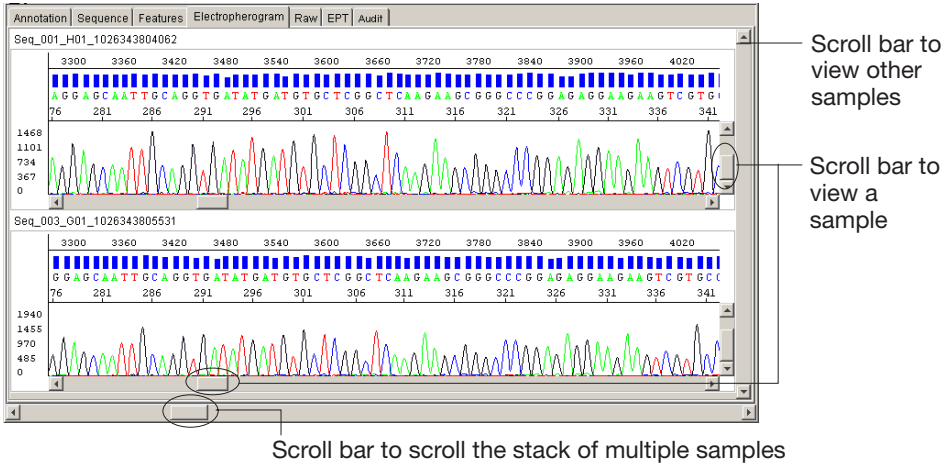


Figure 3-8 Multiple samples in the Electropherogram view

Note: The pane height of the samples can be lowered so that more samples can be viewed simultaneously. For more information, see Chapter 9, "Display Settings."

Trace and Base Colors

The trace lines and the letters above the peaks are colored to represent the four bases. The table below shows the default color for each base.

Base	Color
C	Blue
A	Green
G	Black
T	Red

Note: The colors the represent the bases can be changed. For more information, refer to For more information, see Chapter 9, "Display Settings."

Mixed and Pure Bases

Mixed bases are one-base positions that contain two bases.

KB Basecaller

- If the MixedBases option is selected for mixed base data, then the KB basecaller assigns A, C, G, T or an IUB code to every base.
- If the MixedBases option is not selected (pure bases), then the KB basecaller assigns A, C, G, or T to every base.

Note: Ns will be called, if the Quality Threshold option (in the analysis protocol) is set to Assign 'N' for bases with QV< *X*.

To change the mixed base setting, refer to "Creating and Editing Analysis Protocols" on page 8-13.

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. Refer to "Viewing the Quality Values" on page 4-16.

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. To change the range, refer to “Changing the Clear Range” on page 4-9.

Displaying Edited Bases

Bases that you edit are written in lowercase and the associated QV bar is grayed out as it is no longer applicable to the new base.

Viewing, Editing, and Printing

Table 3-3 Actions Allowed in the Electropherogram View

Action	Procedure or Key Sequence
Zoom in or out to see the data at different magnifications	“Zooming the View” on page 4-5.
Edit the bases	“Editing Bases in Sequence or Electropherogram View” on page 4-18.
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)	“Searching for a Pattern in a Sequence” on page 4-12.
Edit the sequence using any of the standard Windows-based computer commands	“Editing Bases in Sequence or Electropherogram View” on page 4-18.
Display reverse complement	“Showing the Reverse Complement Data” on page 4-15.
Display quality values and numbers	“Viewing the Quality Values” on page 4-16.
Display the original, unedited basecalls while you edit the bases	“Showing the Original Data” on page 4-14.
Display cross hairs and the coordinates for the current cursor location	Click in the data area of the window.
Print the window contents	“Printing the Sample Window Views” on page 4-21.

Table 3-3 Actions Allowed in the Electropherogram View

Action	Procedure or Key Sequence
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

Raw View

The Raw view shows the raw data for the sample, before any processing is performed.

You can use the Raw view to:

- Verify the scan numbers used by the Sequencing Analysis software to start and stop basecalling.
- Measure relative true peak intensities and view peak resolution before the smoothing applied by the Sequencing Analysis software.
- Look for 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.

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

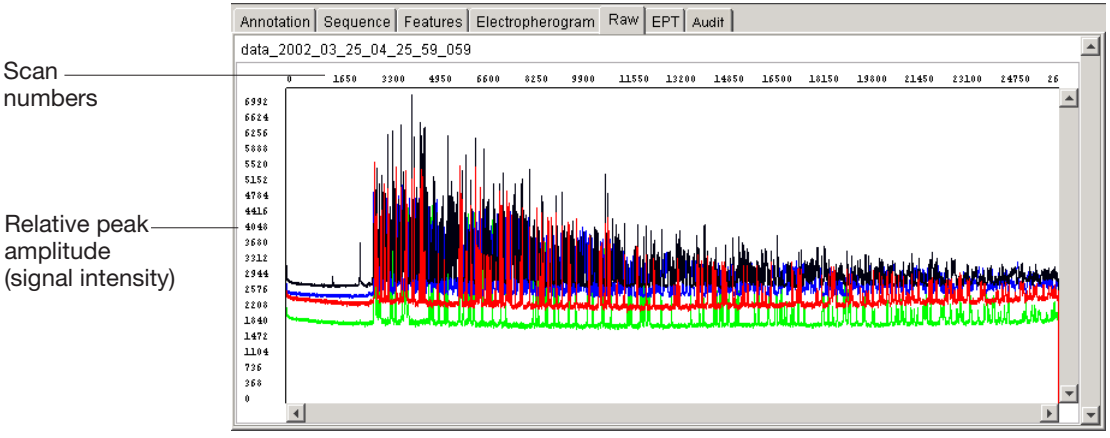


Figure 3-9 Single sample in the Raw view

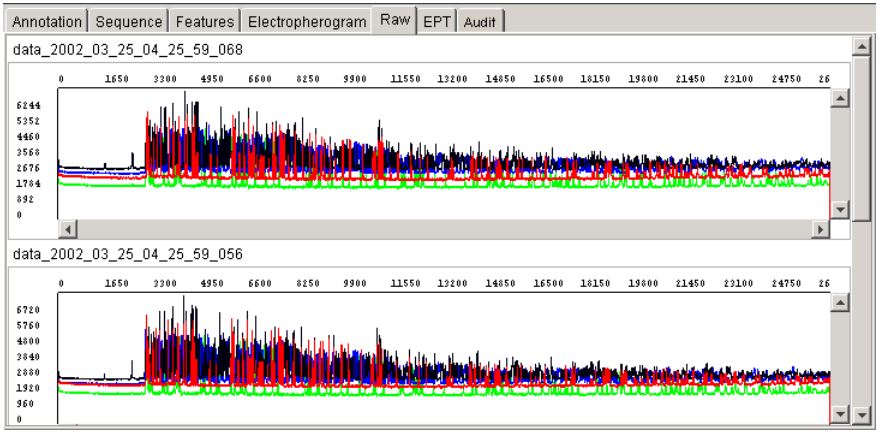


Figure 3-10 Multiple samples in the Raw view

Viewing Raw Data for Analyzed and Unanalyzed Samples

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.

Unanalyzed Data

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

The table below lists the chemistry and the colors that represent each of the four bases.

Raw Data Color Display for Each Chemistry

Color	BigDye® Primers Base	dRhodamine Terminators Base	BigDye® Terminators Base
Blue	C	G	G
Green	A	A	A
Yellow	G	C	T
Red	T	T	C

Analyzed Data

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

Base	Color
C	Blue
A	Green
G	Black
T	Red

Intensity Values in Raw View

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.

Viewing, Editing, and Printing

You can take the following actions in Raw view.

Action	Procedure
Zoom in or out to see the data at different magnifications	See “Zooming the View” on page 4-5.
Change the colors of the trace lines that represent the fluorescent dyes, or hide one or more trace lines	See “Changing the Display Settings” on page 9-10.
Display cross hairs and the coordinates for the current cursor location	Click in the data area of the window.
Print the window contents	See “Printing the Sample Window Views” on page 4-21.

EPT View

The EPT view is useful for reviewing the 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	Units
Voltage	Blue	V/100
Current	Green	µA
Power	Black	mW x 10
Temperature	Red	°C

Viewing

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

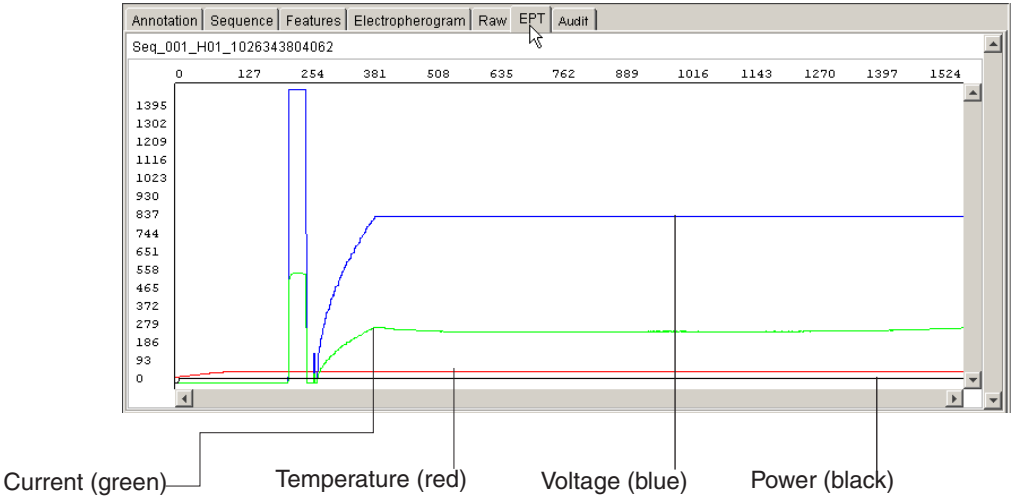


Figure 3-11 Single sample in the EPT view

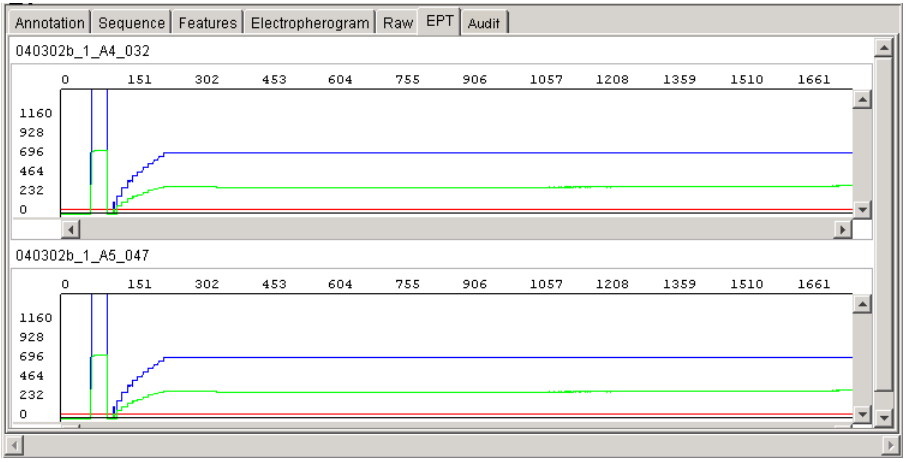


Figure 3-12 Multiple samples in the EPT view

Viewing and Printing

Use the table below to view and print the contents of the EPT window.

Action	Procedure
Display cross hairs and the data values at the current cursor location	Click in the data area of the window.
Open a dialog box and determine the type of information represented by a particular color of line	Select **Analysis > Display Settings**.
Print the window contents	See “Printing the Sample Window Views” on page 4-21.

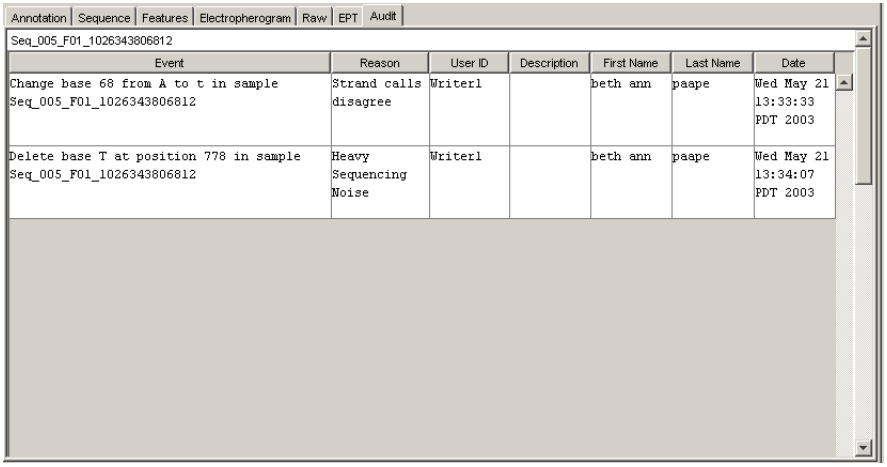
Audit View

The Audit view shows:

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

Viewing

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



Printing

To print the window contents, see “Printing the Sample Window Views” on page 4-21.

Viewing and Editing Samples 4

This chapter covers:

Reviewing the Analysis Results

When sample processing is finished, you should review the results before you begin to work with the analyzed data.

Reviewing the BC, PP, and P Check Boxes

To review the BC (basecalling), PP (post processing), and P (printing) parameter check boxes:

1. Review the BC, PP, and P check boxes in the Sample Manager window.
   a. Look for green, yellow, or red boxes for the BC parameter. Green indicates the process was successful, yellow indicates poor quality data and red indicates failure.

   Note: The yellow result applies to samples analyzed with the KB basecaller.

   b. Look for green or red boxes for the PP and/or P parameters. Green indicates the process was successful and red indicates failure.
   c. Review the base spacing, peak 1 location, and, start and stop points. A red value in the Base Spacing column means the spacing could not be calculated and the default value was used for analysis.
2. If necessary, reanalyze the file.

For more information, see Chapter 5, "Using the Sample Manager."

Reviewing the Analyzed Data

Table 4-1 Steps to Review the Analyzed Data

Step	Description
Review the spacing values.	Review the spacing values in the Sample Manager window. **If ...** / **Then ...** a value is displayed in red text / the basecaller failed while calculating the value, and used a default spacing value.

Table 4-1 Steps to Review the Analyzed Data *(continued)*

<table>
<tr><th>Step</th><th>Description</th></tr>
<tr><td>Review the files used in processing.</td><td>Review the files specified for use during processing.

<table><tr><th>If ...</th><th>Then ...</th></tr><tr><td>the name of a file appears as bold, black, italic text</td><td>the software could not find the file in the expected location.</td></tr></table>
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.

<table><tr><th>To specify a ...</th><th>Use the path ...</th></tr><tr><td>Basecaller</td><td>drive letter:\ AppliedBiosystems\SeqA5.1\ AppSeqA\bin\Basecaller\ Params</td></tr><tr><td>DyeSet/Primer file</td><td>drive letter:\ AppliedBiosystems\SeqA5.1\ AppSeqA\bin\Basecaller\ Mobility</td></tr><tr><td>Matrix file</td><td>drive letter:\ AppliedBiosystems\SeqA5.1\ AppSeqA\bin\Basecaller\ Matrix</td></tr></table></td></tr>
<tr><td>Search for low, medium, and high QVs in the electropherogram (for data analyzed with the KB basecaller).</td><td>You can search for QVs as follows:

<table><tr><th>To move ...</th><th>Press ...</th></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></table></td></tr>
</table>

Table 4-1 Steps to Review the Analyzed Data *(continued)*

<table>
<tr><th>Step</th><th>Description</th></tr>
<tr><td>Scroll through the electropherogram.</td><td>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.

 
Well-resolved peaks

Poorly resolved peaks</td></tr>
<tr><td>Check basecalls in the electropherogram.</td><td>Look at the basecalls in the Electropherogram view.

<table><tr><th>If ...</th><th>Then ...</th></tr><tr><td>two peaks are close together, or the peak is low, or the background noise level is high</td><td>compare each peak to the bases called for that peak.

If necessary, edit incorrect basecalls manually.</td></tr></table></td></tr>
<tr><td>Search for Ns in the electropherogram.</td><td>You can search for Ns as follows:

<table><tr><th>To ...</th><th>Press ...</th></tr><tr><td>move forward</td><td>Tab key</td></tr><tr><td>move backward</td><td>Shift+Tab</td></tr></table>
If you can visually determine the correct basecall at an N location, manually change the N to the correct character.</td></tr>
</table>

Using the Zoom Commands

The View menu has six zoom commands that change the amount of data visible in any of the graphic views.

Zooming the View

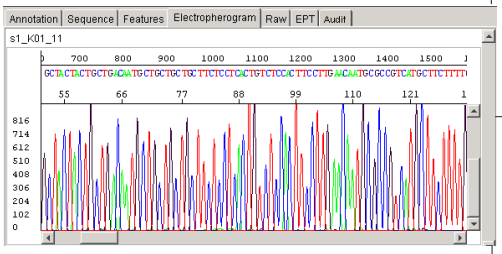
To zoom a view, click the data region that you want to view, then use the procedures described in the following table.

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

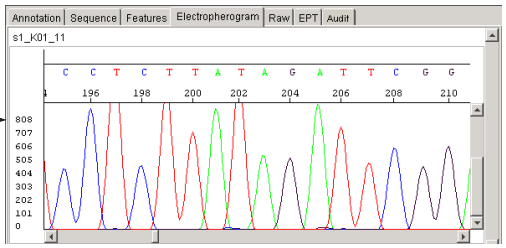
Zoom Commands Illustrated

The examples below of horizontal zoom commands use the Electropherogram view. The commands also work in Raw and EPT views.

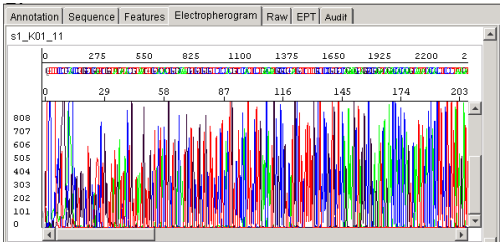
To return to the default view from any other size, select Actual Size.



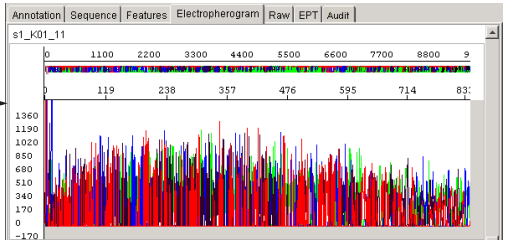
Actual view



Zoom in



Zoom out



Full view

Determining the Value for a Data Point

The values for certain data points are used to set the Peak 1 Location, Start Point, and Stop Point. Using the crosshair feature, you can determine the exact value at any point in an Electropherogram, Raw, or EPT view of the Sample window.

To determine the values for a data point:

1. Click near the point of interest.

   The cursor now includes crosshair locator lines.

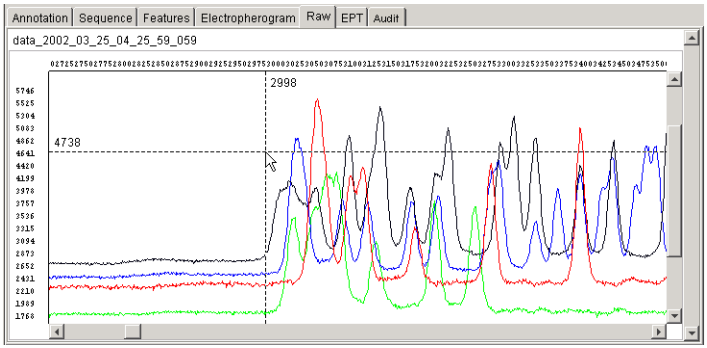


Figure 4-1 Raw data

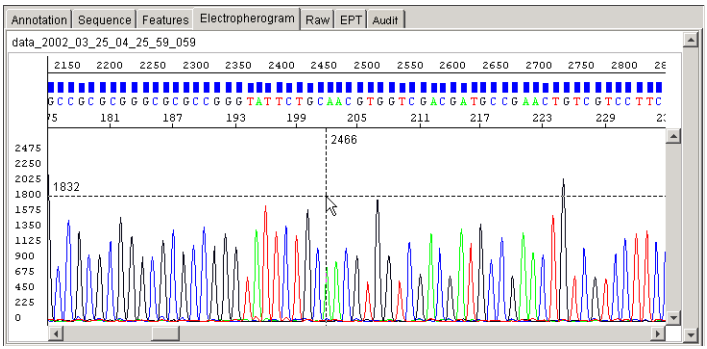


Figure 4-2 Analyzed data

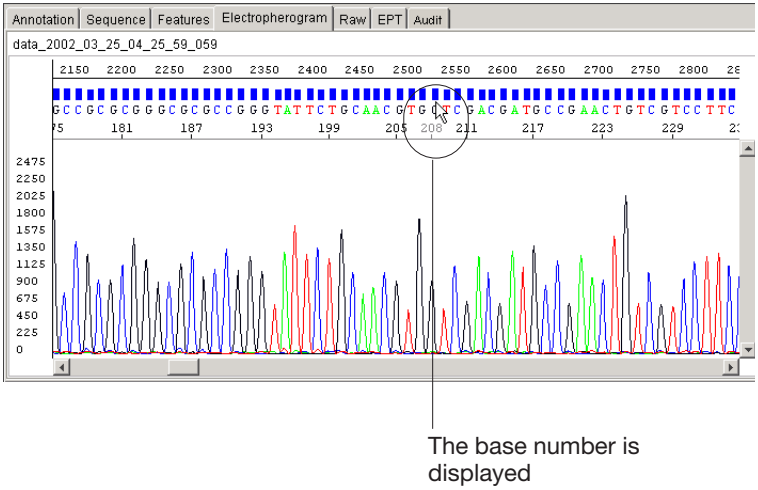
2. Drag the cursor across the window until the locator lines intersect the point of interest.

Displaying the Base Number

To display the base number:

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

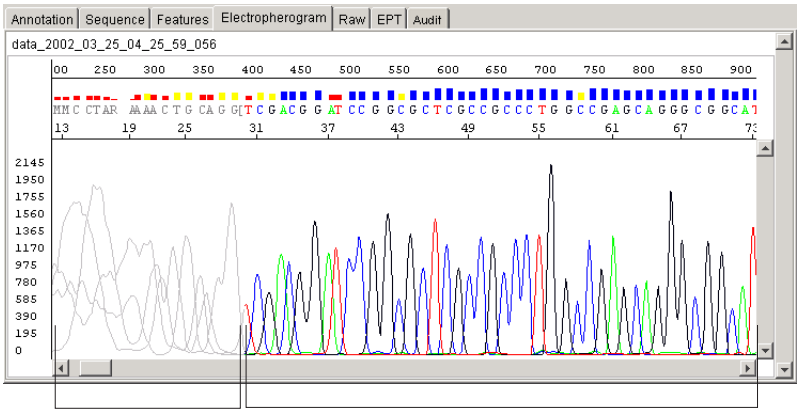
   The base number is displayed. If you drag the cursor across the bases, a base number is displayed for each base.



Changing 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.



Changing the Clear Range

You can change the clear range using the:

- Clear Range widget
- Mouse
- Set Clear Range dialog box

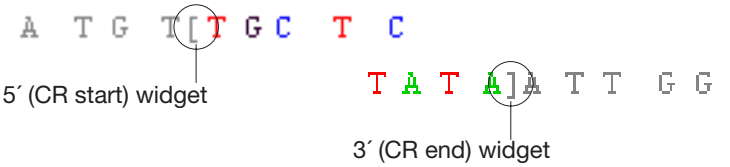
Using the Clear Range Widget

Note: This procedure only applies to samples which already have a clear range,

To change the clear range using the clear range widget:

1. Select the **Electropherogram** tab.
2. Locate and select the 5′ (CR start) or 3′(CR end) widget.

   The widget turns from gray to black when selected.

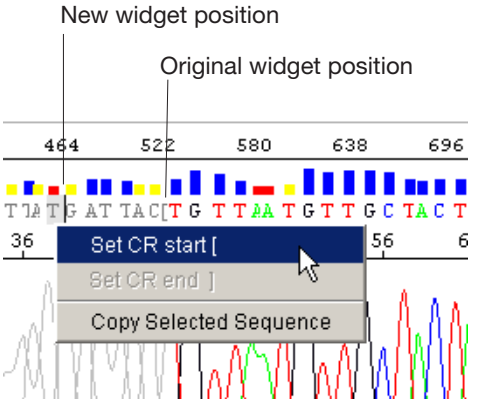


3. Drag the widget along the bases to the right or left, as desired.
4. Release the cursor. The new clear range is displayed.
5. Repeat to define a new clear range for the opposite end.

Using the Mouse

To change the clear range 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.



2. Do one of the following:

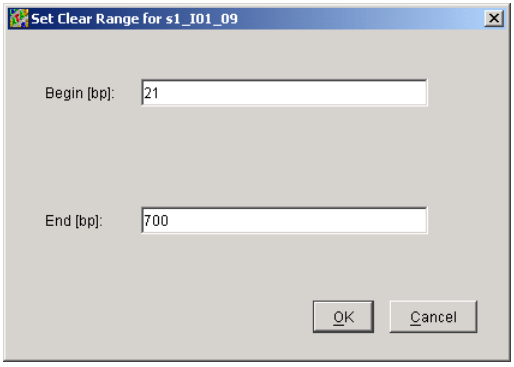
If you are moving the …	Then select …
CR start widget	**Set CR start [**
CR end widget	**Set CR end]**

3. Repeat the process to define a new CR widget position for the opposite end.

Using the Set Clear Range Dialog Box

To change the clear range value using the dialog box:

1. Select a single sample row.
2. In the Electropherogram or Sequence view, determine your new beginning and ending base pair numbers.
3. Select **Tools > Set Clear Range**. The following dialog box opens.



4. Enter the base pair values determined in step 1, then click **OK**.

   The new clear range opens.

Searching for Patterns

Searching Commands

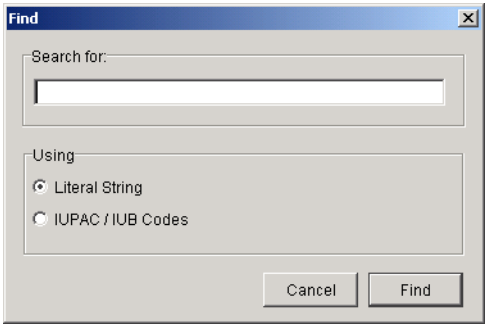
Use the Find (Ctrl+F) and Find Again (Ctrl+G) commands in the Edit menu to search for a particular base or pattern of bases in a sequence. The search operation can be performed in the Sequence or Electropherogram view tabs.

Searching for a Pattern in a Sequence

To find a pattern in a sequence:

1. Select the **Sequence** or **Electropherogram** tab.
2. 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. Select **Edit > Find**. A Find dialog box opens.



4. In the Search for field, enter the search instruction.

   In addition to normal base-character (G, A, T, C) patterns, the search string can include IUPAC/IUB characters (for definitions, see the Glossary).

   For details, refer to “About Search Expressions” on page 4-13.
5. Select the option button that matches the type of instruction entered in the Search for field.

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

   The Sequencing Analysis 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.

7. Select **Edit > Find Again** to find other occurrences of the same pattern.

About Search Expressions

The following table describes the two types of searches you can perform using the Find dialog box:

Search Type	Searches for patterns that ...
Literal String	match exactly what you type in the Find what? field.
IUPAC/IUB Codes	includes an IUB character as part of the pattern. The Find command locates all possible matches. For example, the pattern you enter is TAR, the Find command locates either TAG or TAA. View IUB codes on page "IUPAC/IUB Codes" on page E-2 or select **Help > IUPAC Codes**.

Showing Original Data in Electropherogram View

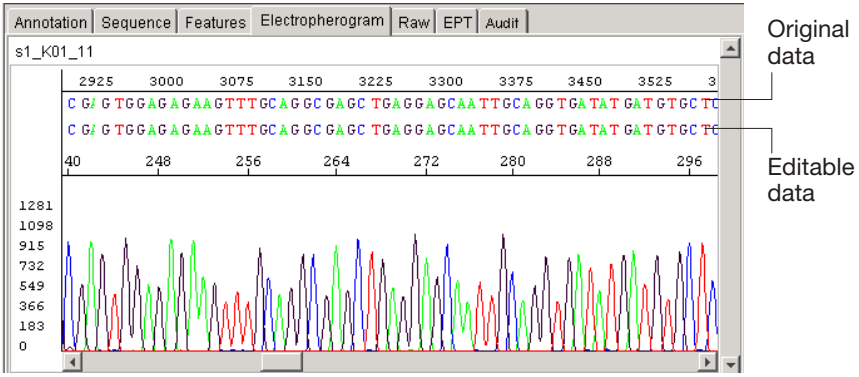
In Electropherogram view, you can display the original sequence data. This is particularly helpful if you are editing bases in this view.

Showing the Original Data

To show the original data in Electropherogram view:

1. Select the **Electropherogram** tab in the Sample Manager window.
2. Select **View > Show Original Sequence** or click  .

Line	Displays the...
Upper	Original, uneditable data
Lower	Editable copy



Hiding Original Data

To switch between showing and hiding the original data, select **View > Show Original Sequence** or click  .

When the original data is visible, Hide Original Sequence is displayed in the menu.

Showing the Reverse Complement of the Data

In Electropherogram and Sequence views, you can display the reverse complement of the data. This alters the display, changing each base to its complement and rewriting it in a 5′ to 3′ direction.

Showing the Reverse Complement Data

To show the reverse complement of the data:

1. Select the appropriate view tab in the Sample Manager window.
2. Select **Tools > Reverse Complement**.

View	Changed Items
Electropherogram	Basecalls and direction
Sequence	
Raw Data	Direction

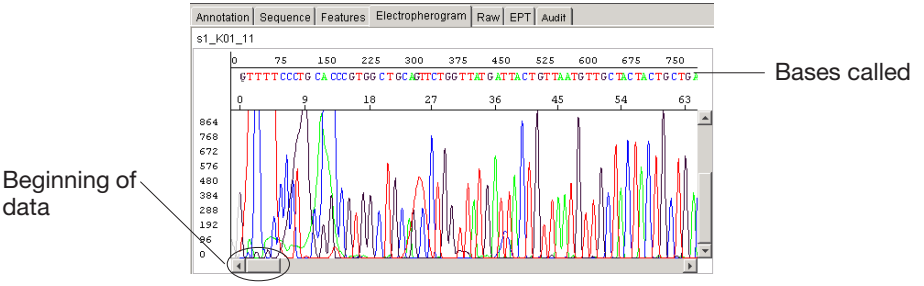


Figure 4-3 Example of 5′ to 3′ direction

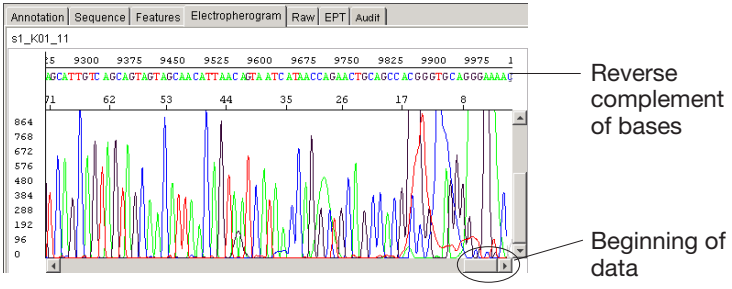


Figure 4-4 Example of reverse complement, 3′ to 5′ direction

Returning to Original Data

To switch between the original and reverse complement data, select **Tools > Reverse Complement**. When the reverse complement data is displayed, a check mark appears next to the command in the menu.

Displaying Quality Values

IMPORTANT! Only samples analyzed with the KB basecaller have QVs.

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 the QV. The taller the bar, the higher the QV. The color that is associated with the value is editable in the display settings.

Mixed base calls yield lower QVs than pure base calls.

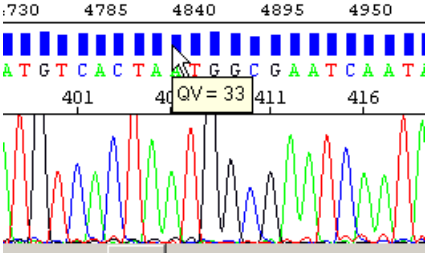
Viewing the Quality Values

If you do not see the QV bars when viewing a sample in the Electropherogram or Sequence view, use one of the following methods.

Using the Show Quality Values Function

To view quality value bars and numbers:

1. Select **View > Show Quality Values** or click  .
2. To obtain a numerical value for a particular bar, place the cursor over the bar for 2 sec. The number is automatically displayed.



Using the Display Setting Function

To view quality value bars:

1. Select **Analysis > Display Settings**, or click  .
2. Select the **Bases** Tab.
3. In the Sample File Display section, select the **QV Bars** check box.
4. Click **OK**.

For more information on QVs and customizing the display, refer to Chapter 6, “Quality Values.”

Editing Analyzed Data

You can edit the bases displayed in the Sequence or Electropherogram view. However, edited data is lost if the sample is reanalyzed or not saved.

Note: The bases outside the clear range cannot be edited or deleted. To edit or delete these bases, adjust the clear range settings (see "Changing the Clear Range" on page 4-9).

Editing Bases in Sequence or Electropherogram View

To ...	Take this 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. This allows you to add any base-identification character that is recognized by the program, including IUPAC/IUB code. See "IUPAC/IUB Codes" on page E-2.
Delete a base from the sequence.	Single-click the base you want to delete, then press the **Delete** key.
Change a base in the sequence.	Single-click the base you want to change, then enter the new character for that position.

Moving from Base to Base

Select the base of interest and use the keyboard shortcuts in Table 3-3 on page 3-25 to move to the next base of interest.

Changing to Views after Adding or Editing Bases

If you edit in the ...	Then the ...
Sequence view	Electropherogram view is immediately updated to match the changed Sequence view data. Edited bases are displayed in lowercase letters.
Electropherogram view	Sequence view is immediately updated to match the changed Electropherogram view data. Edited bases are displayed in lowercase letters.

Saving the Sample Files

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

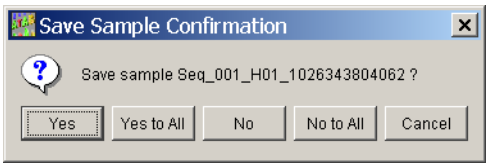
Note: If a .seq file was created when the sample file was analyzed, then both the sample file and .seq file are updated when you save the sample file.

To save selected sample file(s):

1. Select the sample row(s) for the samples you want to save.
2. Click  or select **File > Save Sample(s)**.

To save all sample file(s):

1. Click  or select **File > Save All Samples**. The Save Sample Confirmation dialog box opens.



2. Click **Yes** to save each sample, or **Yes to All** to save all the samples in the sample manager.

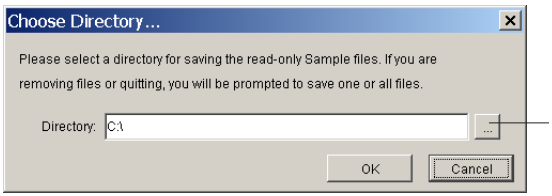
Saving Read-only Sample Files

If you copied sample files from a CD to you hard disk, or imported your samples files from a CD, then these files have read-only attributes. Any changes made to a sample file cannot be saved back to that file. To save read-only sample files follow the procedure on page 4-20.

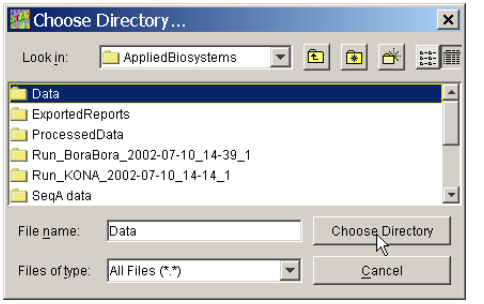
To save read-only sample file(s):

1. *For selected samples:* Select the sample row(s) for the samples you want to save, then click  or select **File > Save Sample(s)**.

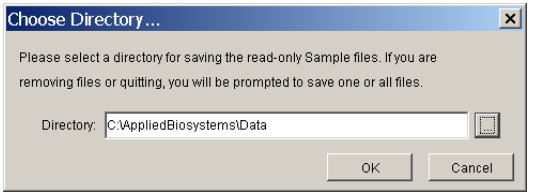
   For all samples: Click  or select **File > Save All Samples**.



2. Click the browse button. A second Choose Directory dialog box opens.



3. Navigate to and select a folder location, then click **Choose Directory**.



4. Click **OK**.
5. *For all samples:* Click **Yes** to save each sample, or **Yes to All** to save all the samples in the sample manager.

Printing the Sample Window Views

Printing can be performed manually or automatically by the Sequencing Analysis software.

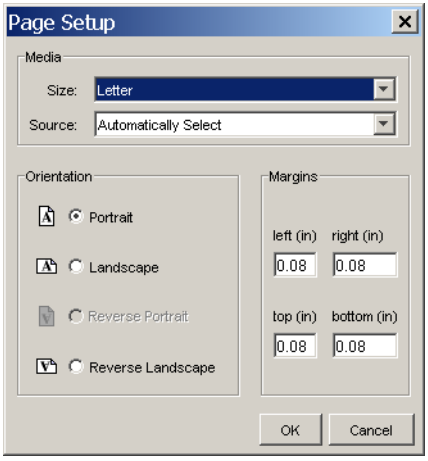
Before you can print:

- Set up and connected the printer to the computer or network.
- Set up the default printer in the Windows® XP or 2000 operating system.

Manually Printing 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 number column, select one or more samples to print.
   - Click to select a sample row
   - Shift-drag to select continuous sample rows
   - Ctrl-click to select discontinuous sample rows
3. Select **File > Page Setup**, then:



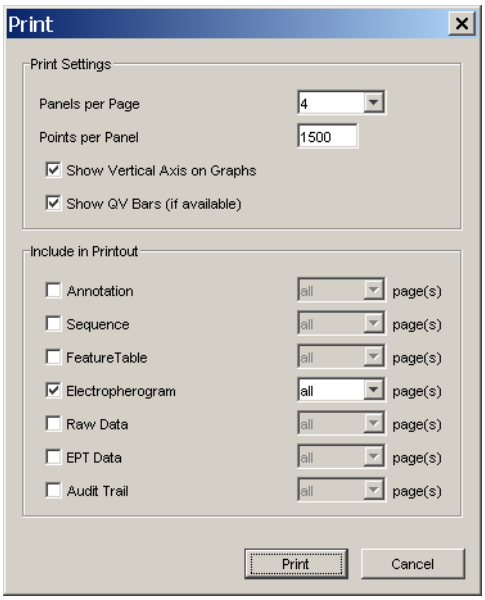
   a. In the Media section, select paper and source from the drop-down lists.

   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**.

4. Select **File > Print** or click  . The Print dialog box opens.



5. In the Print dialog box:

a. Select a value in the Panels Per Page drop-down 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 supressed. This maximum number of panels per page that can be printed with quality values will vary, depending upon paper size, margins and page orientation.

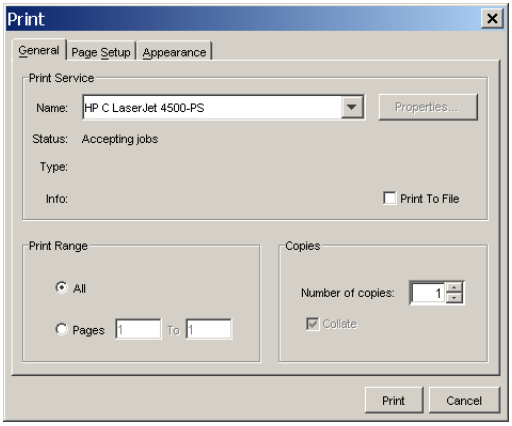
b. Select a value in the Points Per Panel value box. The range is 100 to 12000 and the default is 1500 (about 120 bases).

c. Select to Show/Hide Vertical Axis on graphs when printing electropherogram, raw and EPT data.

d. Select to Show/Hide QV Bars when printing sequence and electropherogram data.

e. Select the view(s) you want to print and the number of pages.

f. Click **Print**. A second Print dialog box opens.

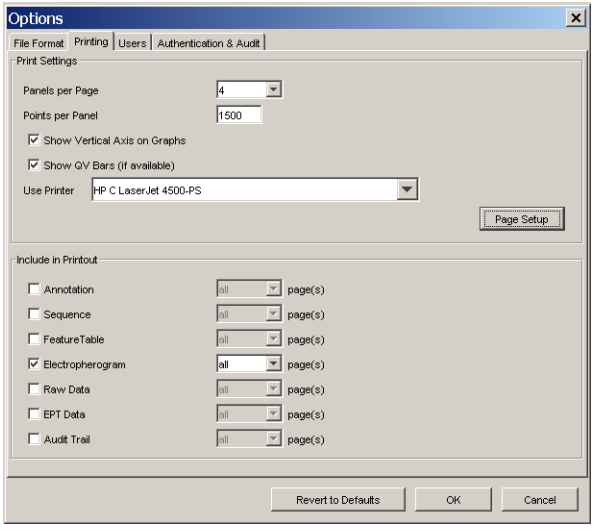


6. Verify and/or change the settings as needed, then click **Print**. The Print dialog box closes, and printing begins.

Automating 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 opens.
3. Select the **Printing** tab, then:



a. Select a value in the Panels Per Page drop-down 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 supressed. 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).

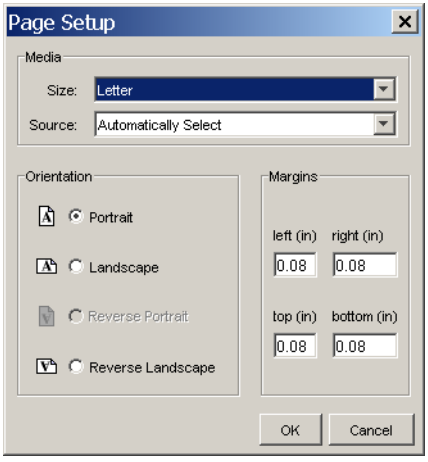
c. Select to Show/Hide Vertical Axis on graphs when printing electropherogram, raw and EPT data.

d. Select to Show/Hide QV Bars when printing sequence and electropherogram data.

e. Select a printer from the Use Printer drop-down list.

f. Select the view(s) you want to print and the number of pages.

4. Click **Page Setup**, then:



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 Options Printing dialog box is 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 check box 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 begins.

After the printing process is completed, a color code status is displayed in the “P” column. Green indicates success and red indicates failure. Also, the checked command in the “P” column is turned off.

Viewing 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, use the Panels Per Page text box in the Printing tab of the Options dialog box or in the Print dialog box from the File menu.

For more information on panels per page, see "Printing Tab" on page 8-28 or "Manually Printing the Sample File" on page 4-21.

Why Print an Electrophero-gram

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 view, 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.

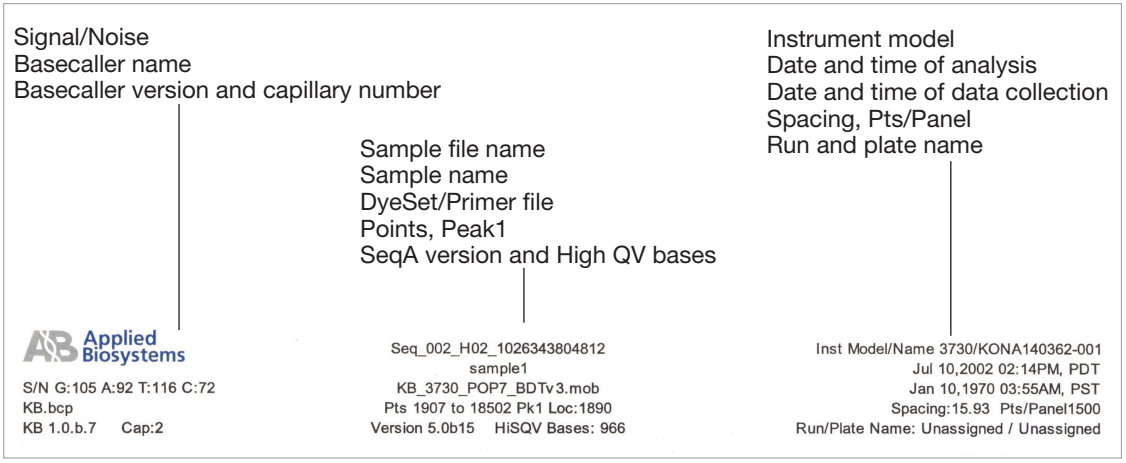


Figure 4-5 Example of printout header information

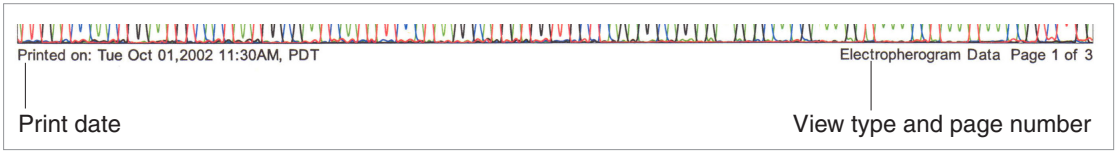


Figure 4-6 Example of printout footer information

Table 4-2 Printout Header Information

Field	Description
First Column	
Signal/Noise	Signal to noise ratio: 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, capillary number	• Version of the basecaller used to analyze the data. • Number of the capillary use to generate the data.

Table 4-2 Printout Header Information *(continued)*

Field	Description
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/Name	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, 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 4-3 Printout Footer Information

Field	Description
Print date	Date and time of printing
View type, 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.

Using the Sample Manager 5

This chapter covers:

About the Sample Manager

Sample Manager Pane

The Sample Manager is a window that can display sample files and their current analysis settings. The settings can be modified and applied to sample(s). You can apply these changes to a single sample, some of the samples, or all of the samples in the window.

Table 5-1 Parts of the Sample Manager Window

Item	Description
Show	A check box to display sample(s) in the Sample View pane.
Sample File Name	Information from the plate record. It can not be change in the Sample Manager.
Sample Name	Name of the sample, taken from the plate record. It can be changed.
Processing Parameters	
BC (Basecalling)	Basecalls (analyzes) the selected files.
PP (Post processing)	Defines the clear range for the selected files.
P (Print)	Prints the selected files.
Analysis Parameters	
Basecaller	Algorithm used to call the bases.
DyeSet/Primer	A DyeSet/Primer file corrects for mobility shifts and color code changes depending on which chemistry was used.
Matrix File	A file that corrects for the overlap in dye emissions spectra for a set of 4 or 5 dyes. This column is: • used for 310 and 377 data because the matrix is applied 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.
Calculated Results	
Spacing	Spacing is defined by the number of scan points from the crest of one peak to the crest of the next peak.

Table 5-1 Parts of the Sample Manager Window *(continued)*

Item	Description
Peak 1 Location	The first raw data point that is from the sample not including primer peaks in dye primer chemistries. It is the reference point for the spacing and mobility corrections performed by the basecalling software.
Start Point	The Start Point is 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.
Stop Point	The Stop Point specifies the last raw data point to be included in the basecalling.

Show Check Box

The Show check box is used to display sample file data. One or more sample files can be viewed in the Sample Manager or Sample Navigator view.

To show the data for …	Do this ...
A single sample	Double-click the sample file name or select the corresponding **Show** check box.
Multiple continuous samples	Shift-click or Shift-drag the sample row numbers to select the sample files, then click Show or select **View > Show Data Displays**.
Multiple discontinuous samples	Ctrl-click the sample row numbers to select the sample files, then click Show or select **View > Show Data Displays**.
All samples	Select the empty box above row number 1 or Shift-drag the sample row numbers to select all samples, then click Show or select **View > Show Data Displays**.

Note: Double-clicking a sample file name, deselects the Show check box for all checked samples (if any), and checks the selected sample file only.

Sample File Name

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.

Viewing a Sample File

Samples can only be viewed from within the Sample Manager or Navigator pane by selecting the Show check box. Refer to "Adding Sample Files to the Sample Manager" on page 3-10 for information on how to add samples.

Changing a Sample File Name

You cannot change the sample file name from within the Sample Manager window. To rename the file, right-click the file, then enter a name (as with any other file in the Microsoft® Windows® operating system).

Note: For more information on changing the sample file name, refer to the applicable instrument user guide.

Sample Name

The Sample Name is the name of the sample. Set the sample name using the Data Collection software.

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

Changing a Sample Name

You can edit the sample name in the Sample Manager window by entering a name of up to 255 characters. The new name is recorded in the sample when you save the file. The icon next to the sample file name changes from  to  .

Note: The information in the sample is normally connected to the Plate Record information through the sample name. You break this connection when you change the sample name. although, you can use the run time and lane number to find the source of the sample information, it is simpler to keep the original sample name until you no longer need the connection.

When printing the Sample Manager in:

- Portrait mode – 40 characters of the sample name are printed.
- Landscape mode – 43 characters of the sample name are printed.

Processing Parameters

BC (Basecalling) Parameter

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 check box in the analysis defaults.

Changing the BC Parameter Setting

To change the BC parameter setting in the Sample Manager window, select/deselect the check box.

Check Box Color Status

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

Check Box Color	Indicates
Green	Successful analysis, data output
Yellow*	Poor quality data, partial data output
Red	Failed analysis, no data output
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

*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.

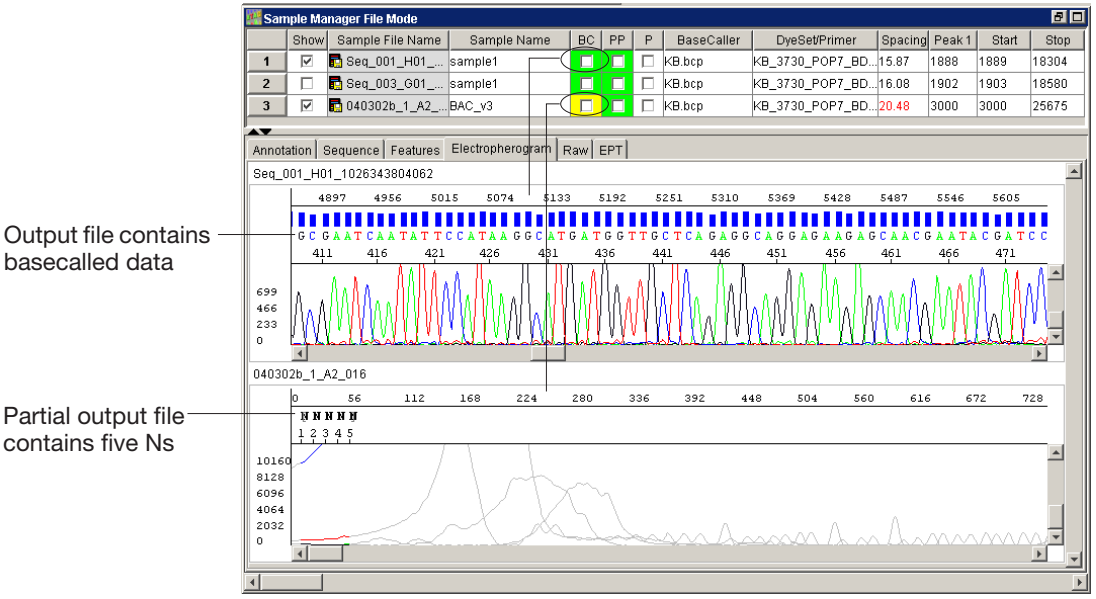


Figure 5-1 Examples of output and partial output files

PP (Post Processing) Parameter

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 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.

Changing the Setting

To change the setting in the Sample Manager window, click the check box.

If ...	Then ...
The PP check box is selected.	Post processing occur. **Note:** Unanalyzed data cannot be post processed until the data is basecalled.
Both BC and PP check boxes are selected.	Basecalling occurs first and then post processing occurs.

Check Box Color Status

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

If the check box color is …	Then Post Processing …
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

The P parameter is a check box option in the Sample Manager window that prints the selected information in Printing options for samples after all processing is complete.

Changing the Setting

Change the setting in the Sample Manager window by clicking the check box. If you also select the BC and/or PP check boxes the printing is done after all other processing of that sample is complete.

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

If the check box color is …	Then printing …
Green	Succeeded.
Red	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

Basecaller File

The basecaller parameter is used to identify bases during the most recent sample analysis.

There are two types of basecallers:

- KB – new algorithm that calculates mixed or pure bases, and quality values.
- ABI – an algorithm used in previous versions of ABI PRISM® Sequencing Analysis software (v3.7 and earlier).

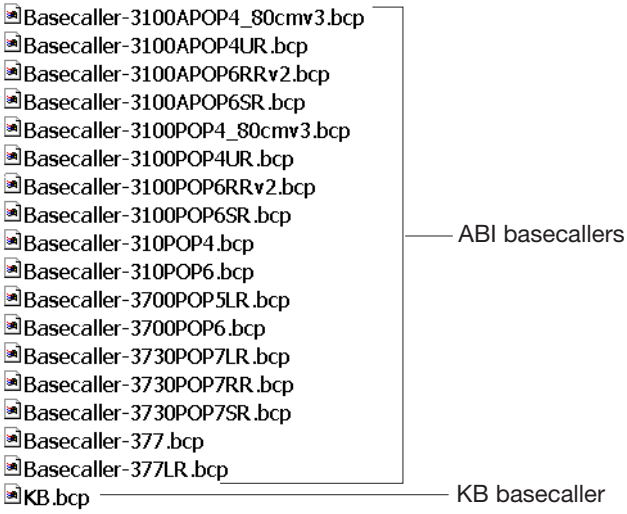


Figure 5-2 List of Basecallers files supplied with the software

Recommended Basecaller

Applied Biosystems recommends that you use the KB basecaller for analysis.

- The KB basecaller was introduced in Sequencing Analysis software v5.0 and ABI PRISM® SeqScape software v2.0. This algorithm will continue to be improved and developed.
- The ABI basecaller is an older algorithm which will be removed in future revisions of the software.

What Happens During Basecalling

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

- Calls the bases with the KB 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 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

Or

- Calls the bases with the ABI basecaller. The basecaller assigns A, C, G, T, or N to every base.

Selecting 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
- Whether you want QVs and mixed bases called (KB basecaller) or not (ABI basecaller)

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

Table 5-2 List of Basecaller files

Instrument	See page …
310	C-2
377	C-5
3100	C-7
3100-Avant	C-10
3700	C-12
3730/3730*xl*	C-14

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.

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 chemistry and basecaller type that you are using.

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

Figure 5-3 List of DyeSet/Primer files supplied with the software

DyeSet/Primer File-Naming Conventions

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

Table 5-3 DyeSet/Primer File Names

Abbreviation	For Runs Using …
Basecaller	
KB	KB basecaller
DP	Dye primer chemistry and ABI basecaller
DT	Dye terminator chemistry and ABI basecaller
Type of Polymer or Gel	
5%LR	% Long Ranger in the gel (377 instrument only)
POP4	ABI PRISM® POP-4™ polymer
POP5	ABI PRISM® POP-5™ polymer
POP6	ABI PRISM® POP-6™ polymer
POP7	ABI PRISM® POP-7™ polymer
Chemistry	
BDTv3	ABI PRISM® BigDye® v3.0 and 3.1 Terminator chemistry
{BDv3}	
BDv1	ABI PRISM® BigDye® v1.0 and 1.1 Terminator chemistry
{BD}	
{BDv1}	
{dRhod}	dRhodamine Terminator chemistry
{-21M13}	Dye primer chemistry – the -21M13 primer is labeled
{M13Rev}	Dye primer chemistry – the M13Rev primer is labeled

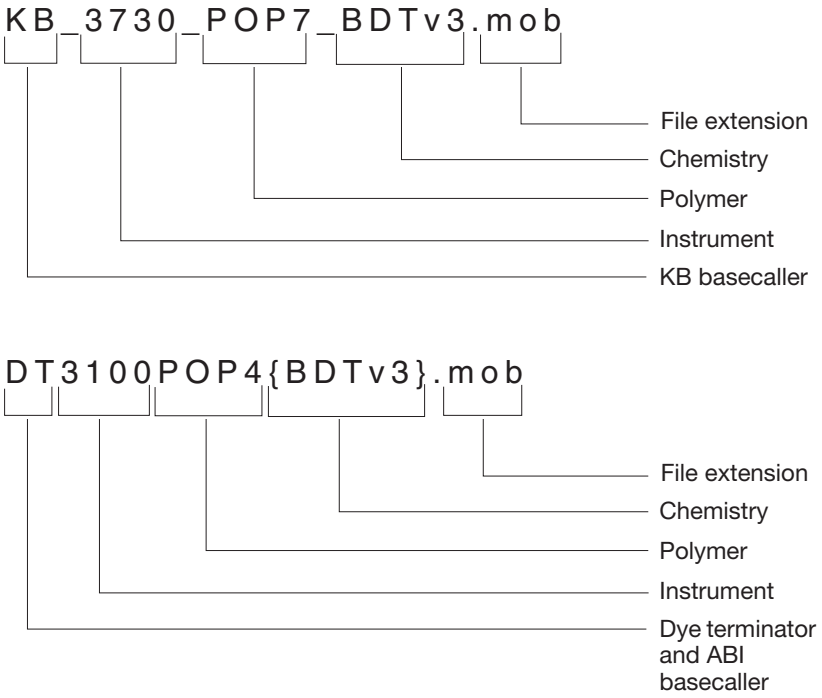


Figure 5-4 Examples of DyeSet/Primer file naming convention

Selecting the Correct DyeSet/Primer File

Basecallers are selected based on a number of factors:

- Basecaller selected
- Instrument model you are using
- Polymer or gel type

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

Table 5-4 List of DyeSet/Primer files

Instrument	See page …
310	C-2
377	C-5

Table 5-4 List of DyeSet/Primer files *(continued)*

Instrument	See page …
3100	C-7
3100-Avant	C-10
3700	C-12
3730/3730*xl*	C-14

Mobility Shift Correction

The basecaller algorithm needs the DyeSet/Primer information to be able to apply the proper mobility shift corrections.

If ...	Or ...	Then ...
You specified the wrong DyeSet/Primer (mobility) file in the Data Collection software	Used a different chemistry from the one for the selected DyeSet/Primer file	You can change this setting for each affected sample file by choosing from the drop-down list and reanalyzing the files.

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 "Copying 310 Matrix and DyeSet/Primer Files" on page 1-20.

The path to the Mobility folder is:
drive letter:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\basecaller\Mobility

Matrix File

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.

The matrix column is:

- Used for 310 and 377 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.

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 "Copying 310 Matrix and DyeSet/Primer Files" on page 1-20 and/or "Copying 377 Matrix Files" on page 1-26 to copy the files in the Matrix folder of the analysis folder.

The path to the matrix folder is:
drive letter:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\Basecaller\Matrix

Calculated Results

Spacing Parameter

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.

The color and text type indicate the status of the calculated base spacing.

If text is …	Then a spacing value …
Normal	Was determined.
Red	Could not be determined and the default value in the basecaller file was used.

Recalculating

To have the software recalculate the spacing parameter, set the value to 0, then reanalyze the sample.

Peak 1 Location Parameter

Introduction

The Peak 1 Location value is defined as the first data point that is from the sample not including primer peaks in dye primer chemistries. The value is initially calculated by the Sequencing Analysis software. It is the reference point for the spacing and mobility corrections performed by the basecalling software.

Changing the Peak 1 Location

To change the Peak 1 Location parameter in the Sample Manager window, enter a new value in the Peak 1 Location field, then reanalyze the sample.

Changing the location value affects the way the DyeSet/Primer file is applied to correct for mobility shifts.This is because the Peak 1 Location is linked to the mobility correction.

Note: Your base spacing may change from the base spacing calculated using the previous Peak 1 Location value.

Not Using the First Base Peak

To start the analysis further along than the actual location of the first base peak, change the value for the Start Point, not for the Peak 1 Location.

If the Peak 1 Location Value Is Wrong

The starting point for data analysis (the Start Point) is normally determined from the Peak 1 Location value.

If ...	Then ...
The Peak 1 Location value is wrong due to low signal or any other aberration	Your data can show bad spacing or strange mobility shifts.
To ...	**See**
Find the beginning of the first base peak for a sample	"Peak 1 Location" on page 5-24.

Recalculating

To recalculate the Peak 1 Location value, reset the value to 0. The software recalculates the Peak 1 Location, Start Point, Stop Point, and Spacing the next time the sample is analyzed.

Start Point Parameter

The Start Point parameter is 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 (refer to "Peak 1 Location Parameter" on page 5-19).

Changing the Start Point Parameter

To start the analysis further along than the location of the first base peak, change the Start Point value, not the Peak 1 Location value. Changing the Peak 1 Location value affects the way the DyeSet/Primer file is applied to correct for mobility shifts.

To change the Start Point parameter in the Sample Manager window, enter a new value in the Start Point field and reanalyze the sample.

The number entered must be equal to, or greater than, the Peak 1 Location value. The number must be the scan number, not the base number, for the point where analysis is to start.

Calling Bases Later in the Data

The following table explains why to start calling bases later in the data.

If ...	Then ...
Any of 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 start calling bases later in the raw data. In such a case, the Start Point value is greater than that of the Peak 1 Location value. The Start Point value can never be less than the Peak 1 Location value.

Recalculating the Start Point Parameter

To have the Sequencing Analysis software recalculate the Peak 1 Location, Start Point, or Stop Point after you have changed a setting, enter a zero in the Peak 1 Location field and reanalyze the data.

Stop Point Parameter

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.

Changing the Stop Point Value

To change the Stop Point parameter in the Sample Manager window, enter a new value in the Stop Point field.

The number must be the scan number, not the base number, for the point where analysis is to stop.

Setting the Stop Basecalling Early

It is possible to stop 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.

Setting the Stop Point Early

Set an earlier Stop Point by entering an earlier Stop Point in the Sample Manager window.

Setting the Stop Point for PCR Products Shorter than the Run

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.

Look at the raw data and choose a scan number after the last peak. Since the Basecaller software calculates spacing and signal strength based on the whole data range, setting an accurate stop point results in better data analysis.

Changing the Analysis Parameters

There are analysis parameters (basecaller and DyeSet/Primer files) associated with every sample file, and are used when the sample file is analyzed.

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

Changing the Analysis Parameters From the Sample Manager

Changing Basecaller, DyeSet/Primer and Matrix Files

If you change the basecaller or DyeSet/Primer file and then reanalyze the sample(s), the basecaller recalculates the Peak 1 Location, Start Point, Stop Point, and Spacing.

Any user-entered values for these parameters are overwritten during the reanalysis process.

To change the basecaller and/or DyeSet/Primer file:

1. In the Sample Manager, select the sample(s).
2. In the basecaller drop-down list, select a new basecaller (see Appendix C, “Basecallers and DyeSet/Primer Files.”.)

3. In the DyeSet/Primer drop-down list, select a new DyeSet/Primer file (see Appendix C, "Basecallers and DyeSet/Primer Files.").

   IMPORTANT! Make sure that the basecaller and the DyeSet/Primer files types match.

4. For 310 or 377 data only: If the incorrect matrix was used to analyze your data, then select the correct file from the drop-down list.
5. Optional: If you want to make other changes, proceed to the next procedure.
6. Select the **BC** check box (and the **PP** and/or **P** check boxes, if desired).
7. Click  (Start Analysis).

Changing the Spacing, Peak 1, Start and Stop Parameters

To change the sample analysis settings:

1. In the Sample Manager, select the sample(s) you want to change.
2. Select a sample, then click the **Raw** tab.
3. Use the Zoom In Horizontal and Zoom In Vertical buttons on the toolbar to expand the raw data view.
4. Use the instructions in Table 5-5, "How to Change Analysis Settings," on page 5-24 to change the settings of interest.

   Note: To recalculate any of the above values, reset the value to 0. The software recalculates the Peak 1 Location, Start Point, Stop Point, and Spacing the next time the sample is analyzed.

5. Select the **BC** check box (and the **PP** and/or **P** check boxes, if desired.)
6. Click  .

Table 5-5 How to Change Analysis Settings

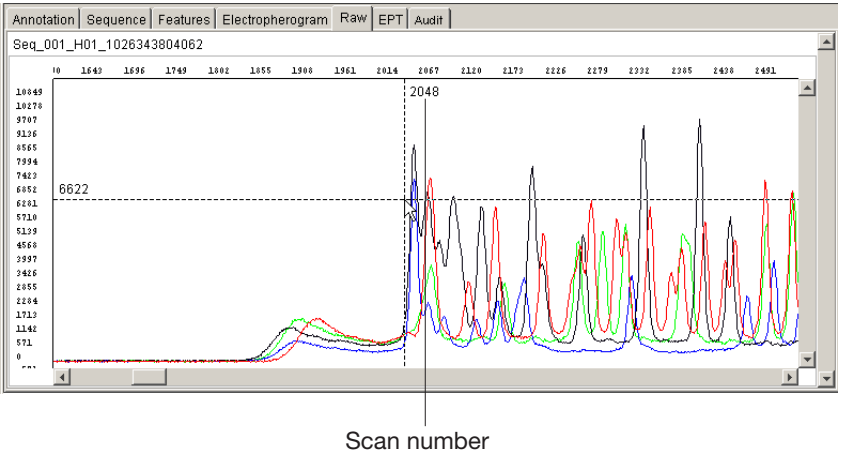
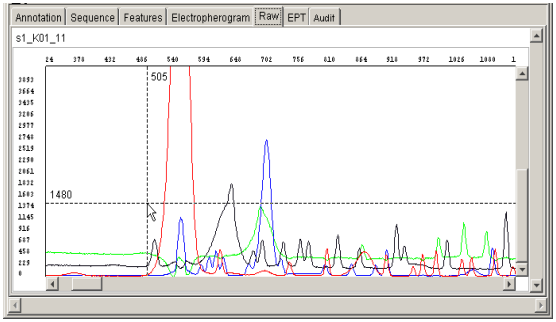
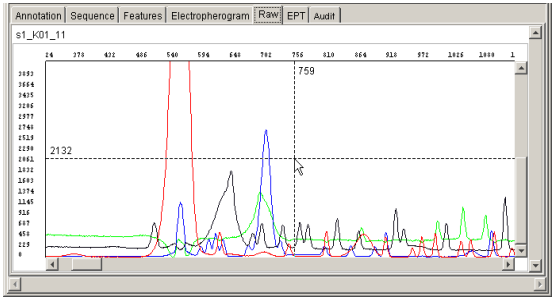
To change…	Procedure
Spacing	1. Use the cross-hair cursor to determine the scan numbers at the tops of two adjacent peaks, then subtract the smaller number from the larger number to determine the spacing. **Note:** Do not use the first 100 bases or the last 200 bases. 2. Record the value. 3. Enter the new value in the Sample Manager. **Note:** To have the software recalculate the spacing parameter, set this value to 0.
Peak 1 Location	1. Find the beginning of the first base peak (the Peak 1 Location value). Point to the beginning of the peak and press the mouse button to display locator lines. 2. Record the cursor position on the X axis. 3. Enter the new value in the Sample Manager. The following figure shows the correct Peak 1 Location value (at scan 2048) for a sample prepared with BigDye® Terminator v3 chemistry:  **Note:** Changing the location value affects the way the DyeSet/Primer file is applied to correct for mobility shifts.

Table 5-5 How to Change Analysis Settings *(continued)*

To change…	Procedure
Start Point	1. Use the cursor to point to the beginning of the peak and press the mouse button to display locator lines. 2. Write down the cursor position on the X axis. This number is the new Start Point value to use for analysis. **Note:** Do not use this value for the Peak 1 Location. The following example shows the difference between a Peak 1 Location and a Start Point.  In the raw data view, the cursor is positioned at the Peak 1 location. The scan number is 505 and represents where the data starts.  In the raw data view, the cursor is positioned at the Start Point. The scan number is 759 and represents where you want the basecalling to start.
Stop Point	1. Use the cursor to point to the beginning of the peak and press the mouse button to display locator lines. 2. Write down the cursor position on the X axis. This number is the new Stop Point value to use for analysis. **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.

Saving the Sample Files

The data in a sample file is not automatically saved after basecalling and/or post processing. See "Saving the Sample Files" on page 4-19.

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.

Changing the Analysis Parameters in the Analysis Protocol

Changing the Per-sample Analysis Protocol

To change a protocol for a single sample:

1. Select a sample row in the Sample Manager.
2. Select **Analysis > Analysis Protocol**.
3. Select the **Basecalling** tab, then use the drop-down lists to select the basecaller and/or the DyeSet/Primer file (see Appendix C, "Basecallers and DyeSet/Primer Files.").

   IMPORTANT! Make sure that the basecaller and the dye set types match.
4. For 310 or 377 data only: If the incorrect matrix was used to analyze your data, then select the correct file from the drop-down list.
5. Click **OK**.
6. Select the **BC** check box (and the **PP** and/or **P** check boxes, if desired).
7. Click  (Start Analysis).
8. Save the samples (see "Saving the Sample Files" on page 4-19).

To change a protocol for multiple samples:

1. Select the sample rows in the Sample Manager.
   - Use the Shift key to select continuous samples.
   - Use the Ctrl key to select discontinuous samples.
2. Select **Analysis > Analysis Protocol Manager**.
3. Select the protocol you want to edit.
4. Click the **File** button, then select **Open,** or double-click the protocol name.
5. Select the **Basecalling** tab, then use the drop-down lists to select the basecaller and/or the DyeSet/Primer file (see Appendix C, "Basecallers and DyeSet/Primer Files.").

   IMPORTANT! Make sure that the basecaller and the dye set types match.

6. For 310 or 377 data only: If the incorrect matrix was used to analyze your data, then select the correct file from the drop-down list.
7. Click **OK** to save the protocol and close the Sequence Analysis Protocol Editor dialog box.

   Note: The version number will increment by one.
8. Click:
   - **Apply to Selected Samples** to apply the protocol to the samples selected in step 1, or
   - **Apply to All Samples** to apply the protocol to all the samples in the Sample Manager.
9. Click **Done** to close the Analysis Protocol Manager dialog box.
10. Select the **BC** check box (and the **PP** and/or **P** check boxes, if desired).
11. Click  (Start Analysis).
12. Save the samples (see "Saving the Sample Files" on page 4-19).

For more information, see "Creating and Editing Analysis Protocols" on page 8-13.

Quality Values 6

This chapter covers:

About Quality Values

One of the features introduced in ABI PRISM® Sequencing Analysis Software v5.0, is the KB basecaller which assigns quality values (QVs) for each base, including pure and mixed bases. The QV is a per-base estimate of the basecaller accuracy.

Interpreting the Per-Base Quality Values

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

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

where *Pe* is the probability of error.

The KB basecaller generates QVs from 1 to 99, with 1 being low confidence and 99 being high confidence. See Table 6-1, "Quality Value and Probability of Error," on page 6-3 for the probability of basecall error for QVs ranging from 1 to 99.

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

To Use QVs

To use QVs for data quality review, determine what is the lowest QV that your lab accepts as good data. Refer to "Customizing the Quality Value Display" on page 6-6.

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.

Table 6-1 Quality Value and Probability of Error

Quality Value = $-10\log_{10}(Pe)$

***Pe* is the probability of error**

KB basecaller generates QVs from 1 to 99

Typical high quality pure bases will have QV 20 to 50

Typical high quality mixed bases will have QV 10 to 50

Size and color of QVs bars are identical for QVs 50 to 99

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%
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*	10%	**30***	0.10%	**50***	0.0010%
11	7.9%	**31**	0.079%	**60**	0.0001%
12	6.3%	**32**	0.063%	**70**	0.00001%
13*	5.0%	**33**	0.050%	**80**	0.000001%
14*	4.0%	**34**	0.040%	**90**	0.0000001%
15*	3.2%	**35**	0.320%	**99**	0.00000012%
16	2.5%	**36**	0.025%		
17*	2.0%	**37**	0.020%		
18	1.6%	**38**	0.016%		
19	1.3%	**39**	0.013%		
20*	1.0%	**40***	0.010%		

*Commonly used cut-off values for quality values

Viewing the Quality Values

IMPORTANT! Only samples analyzed with the KB basecaller have QVs.

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. The color, associated with the value, is editable in the Display Settings.

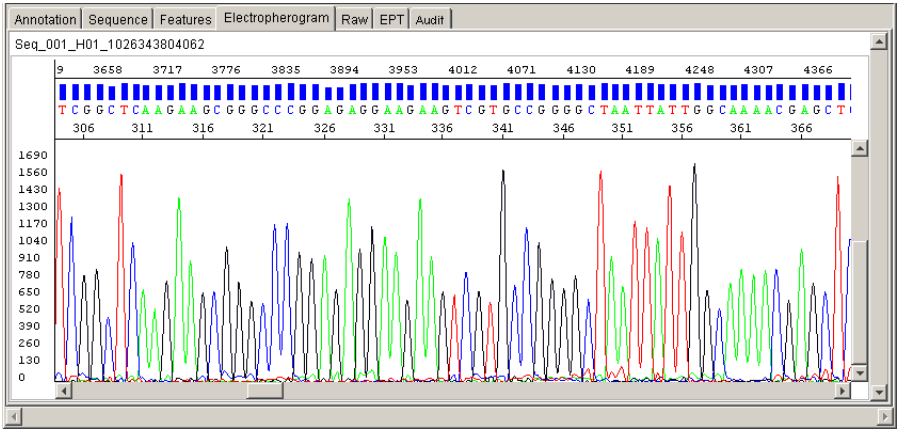


Figure 6-1 Example of analyzed data with QV values

QVs for Mixed Base Data

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 8-11.

Viewing the Quality Value Bars

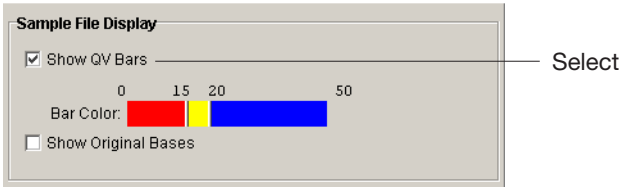
If you do not see the QV bars when viewing a sample in the Electropherogram or Sequence view, use one of the following methods.

Using the Display Setting Function

To view quality value bars:

1. Select **Analysis > Display Settings** or click ds.
2. Select the **Bases** tab.

3. In the Sample File Display section, select the **Show QV Bars** check box.

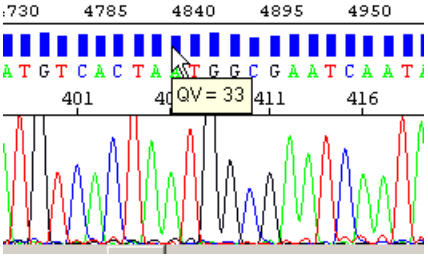


4. Click **OK**.

Using the Show Quality Values Function

To view quality bars and values:

1. Select **View > Show Quality Values** or click  .
2. To display the numerical value for a particular bar, place the cursor over the bar for 2 sec. The number automatically displays.



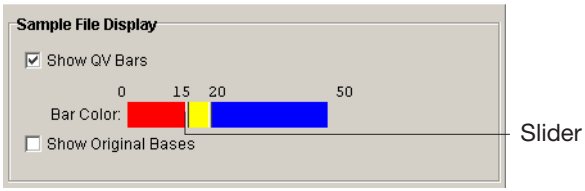
Customizing the Quality Value Display

The low, medium, and high ranges and the color associated with a QV can be modified.

To modify the QV display:

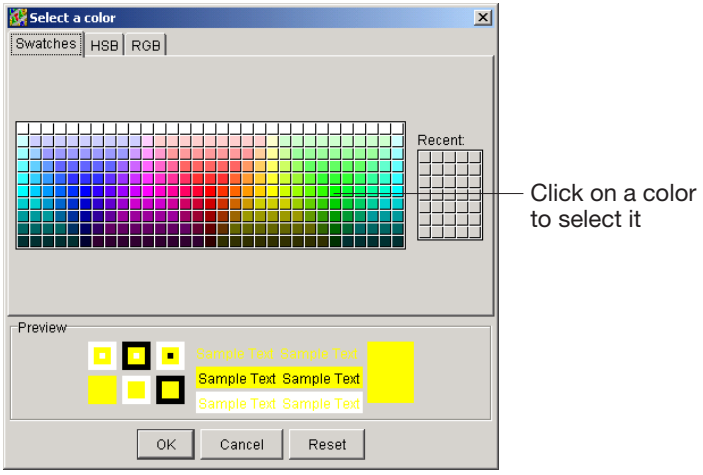
1. Select **Analysis > Display Settings** or click  .
2. In the Sample File Display section, click the slider on the color bar and drag it to left or right to the desired value.

   This will change the limits for low, medium, and high quality values.



QV Bar	Default Color and Range	Set the range to identify data that is …
Low	Red 0 to 14	Not acceptable
Medium	Yellow 15 to 19	Needs manual review
High	Blue 20 or higher	Acceptable

3. To change the color of the QV colorbar:
   a. Click on the QV color you want to change. The Select a color dialog box opens.



   b. Click on a new color in the Swatches tab, or use the HSB or RGB tabs to define a new color.
   c. Click **OK**.
4. Click **OK** to close the Display Settings dialog box.

Editing Bases with Quality Values

About Editing with QVs

Changing, deleting, and inserting a base will affect the QVs displayed.

If you ...	Then the ...
Change a base	New base is in lowercase and the QV has the same value but be displayed as a gray bar, as the bar is no longer applicable to the new base.
Change a base back to the original call	Base displays in uppercase and the quality value bar color is restored.
Insert a base	Inserted base displays in lowercase and it has no QV.
Delete a base	Quality value for that base disappears.
Reinsert a deleted base	Reinserted base displays in lowercase and it has no QV.

Analysis Report 7

This chapter covers:

About the Analysis Report

An analysis report shows the success and/or failure of data analysis. An analysis report can be generated for any samples added to the Sample Manager. If the data has been analyzed, the report displays a summary of QVs and LORs, as well as individual sample information and errors. If the data is unanalyzed, the report displays status information. 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.

Displaying the Analysis Report

To display the report, select **Analysis** > **Analysis Report** or click  . Applied Biosystems recommend that after analysis, you review the report before examining each sample file.

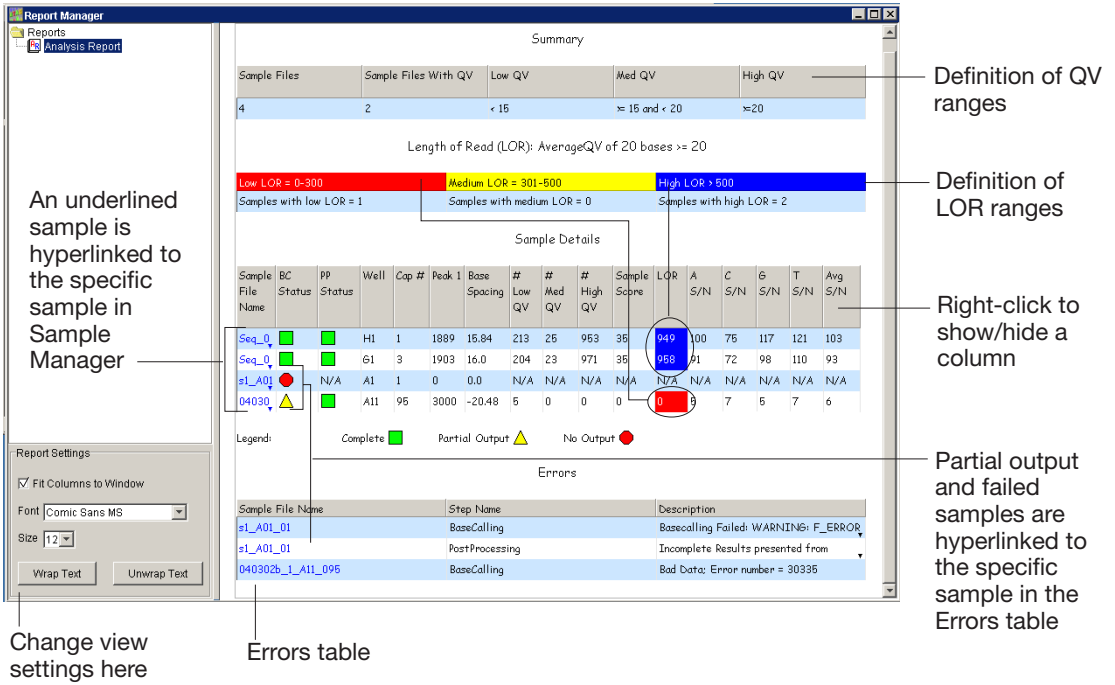


Figure 7-1 Analysis report example

Parts of the Analysis Report Window

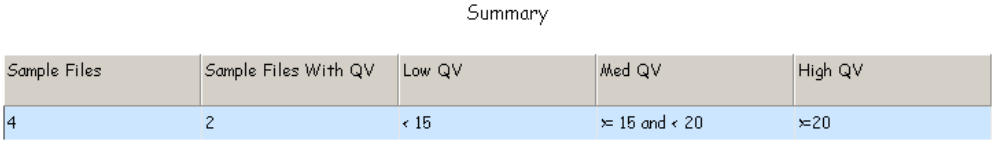
The report contains four separate tables of information.

Table 7-1 Parts of the Analysis Report

Table	Description
Summary	Displays the total number of samples in the report (depends on the number of samples in the Sample Manager), number of samples that contain QVs, and the definition of the QV ranges.
Length of Read (LOR)	Displays the definition of the LOR ranges and colors, and the number of samples in each range.
Sample Details	Displays sample statics. Each sample has a hyperlink to the specific sample in the Sample Manager. Partial output and failed samples are hyperlinked to the specific sample in the Errors table.
Errors	Displays the errors that occurred during analysis and post processing; each sample has a hyperlink to the specific sample in the Sample Manager.

Summary Table

This part of the report contains a summary of the number of samples and the quality values in the report.



Summary

Sample Files	Sample Files With QV	Low QV	Med QV	High QV
4	2	< 15	>= 15 and < 20	>=20

Table 7-2 Parts of the Summary Table

Column	Description
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	The total number of samples analyzed with the KB basecaller that contain QVs.
Low, Medium, or High QV	The low, medium, and high ranges for the QVs. **Note:** The ranges are defined in the Display Settings. For more information, see "Sample File Display Section" on page 9-4.

Length of Read (LOR) Table

This part of the report contains a summary of the low, medium, and high LOR information.

The LOR is the usable range of high-quality or high-accuracy bases, as determined by quality values. The LOR information is calculated when the analysis report opens.

Note: LOR information is displayed for samples analyzed with the KB basecaller only.

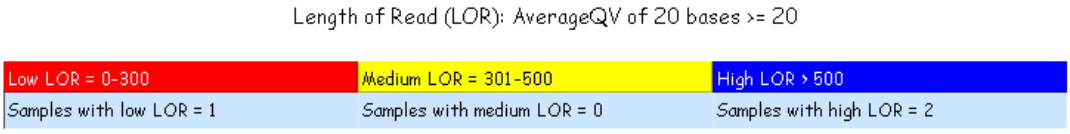


Table 7-3 Parts of the LOR Table

Column	Description
Low LOR	The range and color for low LOR results*. The default color is red.
Medium LOR	The range and color for medium LOR results. The default color is yellow.
High LOR	The range and color for high LOR results. The default color is blue.
Samples with Low LOR	The number of samples in the low LOR range.
Samples with Medium LOR	The number of samples in the medium LOR range.
Samples with High LOR	The number of samples in the high LOR range.

*The ranges and colors for the LORs are defined in the Display Settings. For more information, refer to “Sample File Display Section” on page 9-4.

Sample Details Table

This part of the report contains a list of each sample and its associated analysis information.

Note: QVs, sample score, LOR, and signal to noise information is displayed for samples analyzed with the KB basecaller only.

In the Sample Details table, 'N/A' means no analysis or not available in this session.

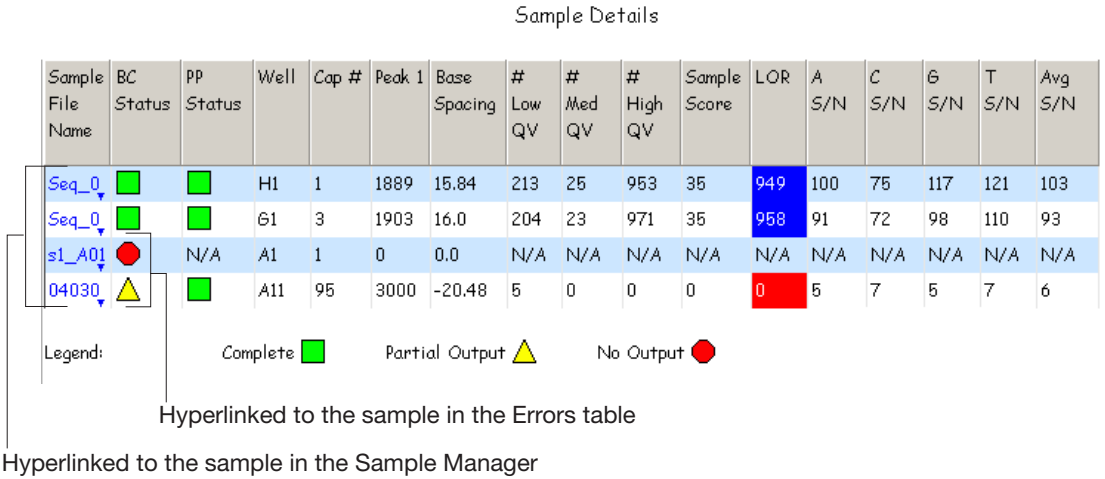


Table 7-4 Parts of the Sample Details Table

Column	Description
Sample File Name	The name of the sample. The files are hyperlinked to the specific sample in Sample Manager.
BC Status	Pass, poor quality data, fail, and not analyzed indicator for basecalling • Green icon = passed • Yellow icon = poor quality data, partial output (for data analyzed with the KB basecaller only) • Red icon = failed • N/A = no analysis or not available in this session Samples with red or yellow icons are hyperlinked to the specific sample in the Error table.

Table 7-4 Parts of the Sample Details Table *(continued)*

Column	Description
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 ample 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.
Bases with Low, Medium, or High QVs	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’, ‘G’, ‘C’, 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.

Errors Table

This part of the report displays the errors that occurred during analysis and post processing; each failed sample has a hyperlink to the specific sample in the Sample Manager.

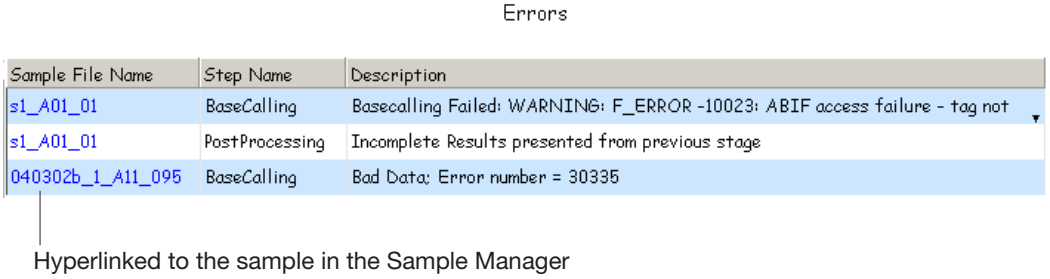


Table 7-5 Parts of the Errors Table

Column	Description
Sample File Name	The name of the sample.
Step Name	Where in basecalling or post processing process the error occurred.
Description	The description of the error.

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

Hyperlinks

Hyperlinks are used in the analysis report to aid in sample evaluation and troubleshooting.

- The samples in the Sample Details table and Error table of the analysis report are hyperlinked to the specific sample file in the Sample Manager. When you click a sample in the Sample File Name column the corresponding file is displayed in the Sample Manager.
- The samples that have partial output (display a yellow icon) or failed basecalling or post processing (display a red icon) are hyperlinked to the Error table. When you click the yellow or red icon for a file, the corresponding file in the Error table is highlighted.
- Hyperlinks move only from the analysis report to the Sample Manager, but not vice versa. To return to the analysis report, click  (Analysis Report).

Viewing the Analysis Report

IMPORTANT! Any action that renders the report to be invalid such as adding more samples forces the report view to close.

To view the Analysis Report:

1. Select **Analysis > Analysis Report** or click  .

   The analysis report opens.

2. To view the data in the report, perform the following as needed:

   a. Use the scroll bars to see all of the report.

   b. Use the hyperlinks to move from the analysis report to the Sample Manager.

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

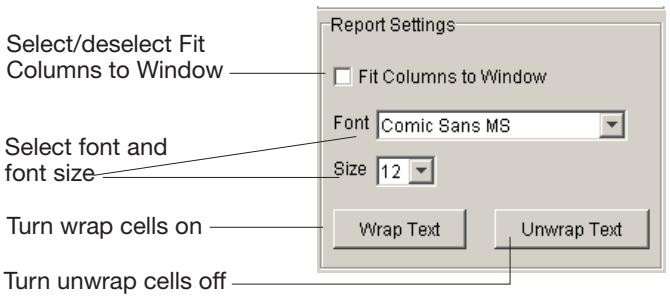
3. Click  to return to the analysis report.

Customizing the Display

Customizing the Font and Text

To customize the font and text in the cells:

To customize the view, use the Report Settings section which is located in the bottom left corner of the Report Manager window.



1. Select or deselect **Fit Columns to Window** check box. The default is off.

Figure 7-2 Examples of Fit Columns to Window check box on and off

2. Select your font type and font size in the appropriate drop-down lists. The default is font and size is Arial 12.
3. Click **Wrap Text** or **Unwrap Text**.

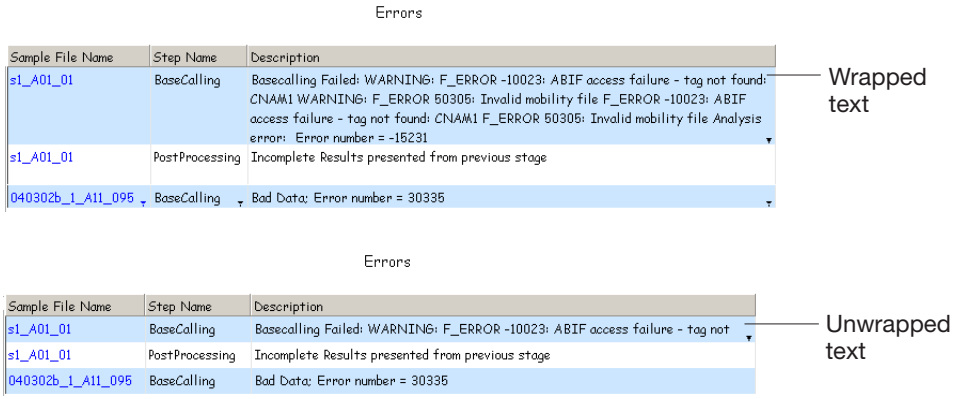
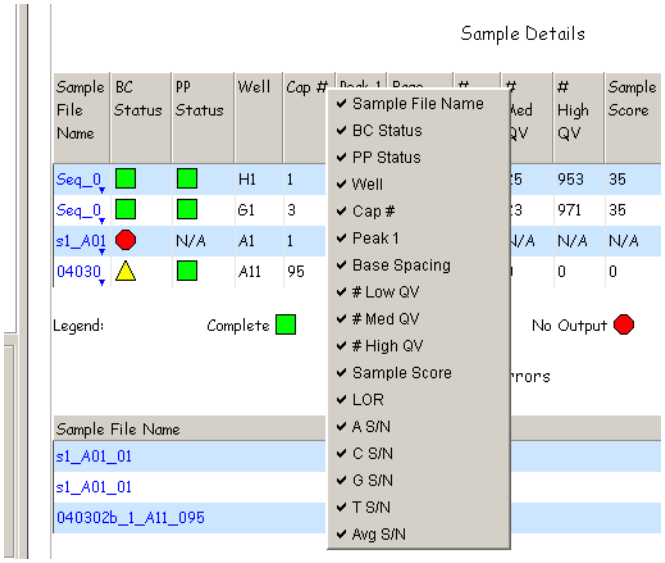


Figure 7-3 Examples of wrapped and unwrapped text

Customizing the Data View

To customize the information displayed in the report:

1. Right-click in any column heading of a table. A list of the column headings in the table displays.



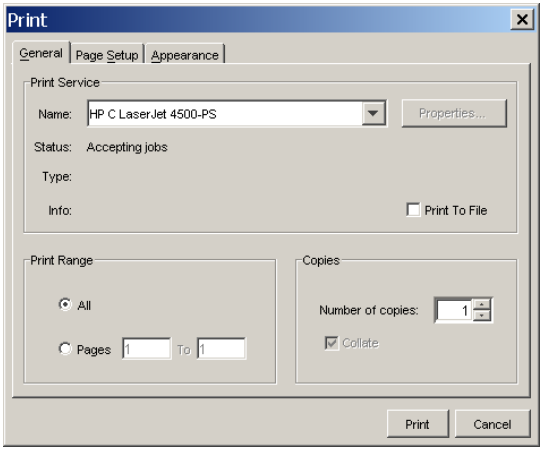
2. To hide a column, deselect the column heading.
3. Repeat steps 1 and 2 to deselect additional headings.
4. To redisplay a column, right-click in any column heading, then select the column heading.
5. To sort the data A to Z or Z to A in a Sample Details or Errors table column, double-click in the 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.

Printing and Exporting the Analysis Report

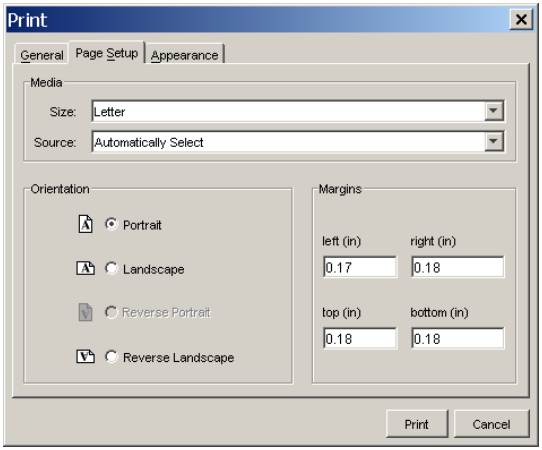
Printing the Analysis Report

To print the analysis report:

1. Open the analysis report.
2. Select **File > Print** or click  . The Print dialog box opens.
3. In the General tab:



   a. In the Print Services section, select a printer from the Name drop-down list.
   b. Select to print all or selected pages.
   c. Change the number of copies, if needed.
4. Select the **Page Setup** tab, then:

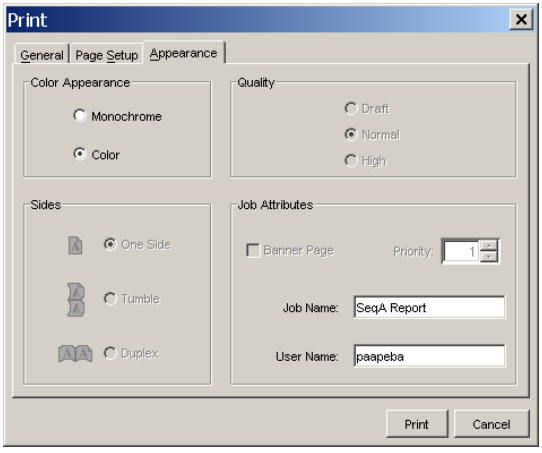


a. In the Media section, select paper and source from the drop-down lists.

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.

5. Select the **Page Setup** tab, then:



a. In the Color appearance section, use the radio button to select the monochrome or color option.

b. In Quality section, use the radio button to select your print quality.

c. In Sides section, use the radio button to select your preference.

Note: If you are not using a double-sided printer, the options are not available and One Page is the default setting.

d. In the Job Attributes section, change the job name and/or user name, if desired.

6. Click **Print**.

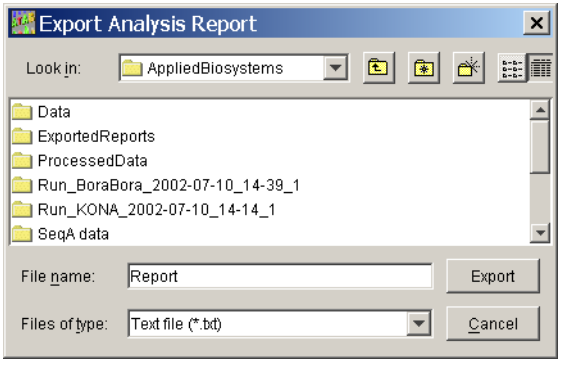
   The report prints.

Exporting the Analysis Report

The analysis report can be exported as a tab-delimited file. It can be opened in Microsoft® Excel software (or any application that reads this type of format) for trend analysis.

To export the analysis report:

1. Open the analysis report.
2. Select **File > Export Report**.
3. In the Export Analysis Report dialog box:

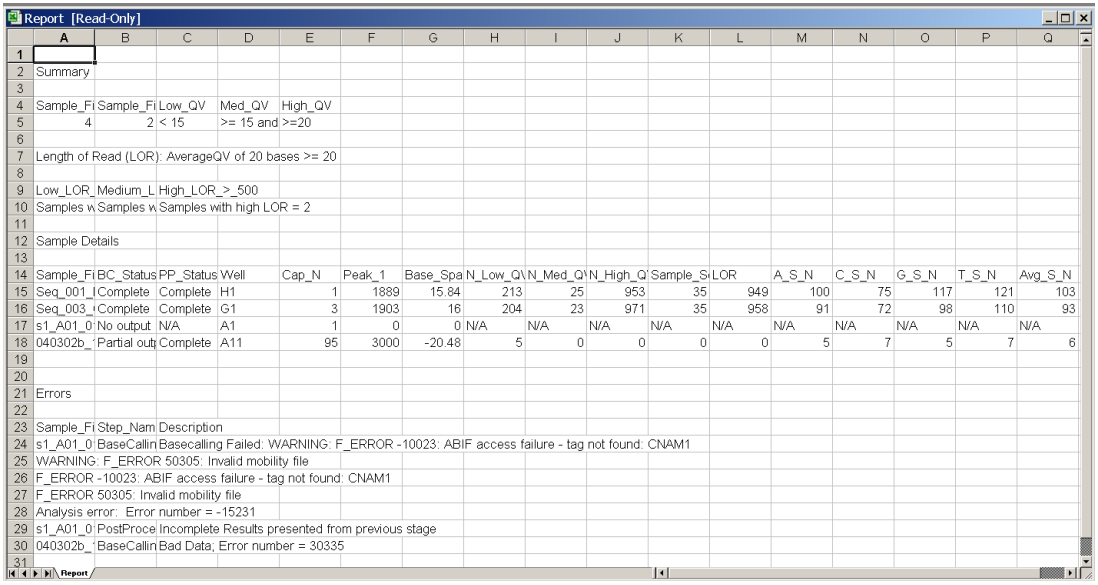


a. Define a folder location to store the file.

b. Enter a file name.

c. Click **Export**.

The file exports in a tab-delimited format.



Report [Read-Only]

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1																	
2	Summary																
3																	
4	Sample_Fi	Sample_Fi	Low_QV	Med_QV	High_QV												
5	4	2	< 15	>= 15 and	>=20												
6																	
7	Length of Read (LOR): Average QV of 20 bases >= 20																
8																	
9	Low_LOR_	Medium_L	High_LOR_>_500														
10	Samples w	Samples w	Samples with high LOR = 2														
11																	
12	Sample Details																
13																	
14	Sample_Fi	BC_Status	PP_Status	Well	Cap_N	Peak_1	Base_Spa	N_Low_Q	N_Med_Q	N_High_Q	Sample_S	LOR	A_S_N	C_S_N	G_S_N	T_S_N	Avg_S_N
15	Seq_001_I	Complete	Complete	H1	1	1889	15.84	213	25	953	35	949	100	75	117	121	103
16	Seq_003_	Complete	Complete	G1	3	1903	16	204	23	971	35	958	91	72	98	110	93
17	s1_A01_0	No output	N/A	A1	1	0	0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
18	040302b_	Partial out	Complete	A11	95	3000	-20.48	5	0	0	0	0	5	7	5	7	6
19																	
20																	
21	Errors																
22																	
23	Sample_Fi	Step_Nam	Description														
24	s1_A01_0	BaseCallin	Basecalling Failed: WARNING: F_ERROR -10023: ABIF access failure - tag not found: CNAM1														
25	WARNING: F_ERROR 50305: Invalid mobility file																
26	F_ERROR -10023: ABIF access failure - tag not found: CNAM1																
27	F_ERROR 50305: Invalid mobility file																
28	Analysis error: Error number = -15231																
29	s1_A01_0	PostProce	Incomplete Results presented from previous stage														
30	040302b_	BaseCallin	Bad Data, Error number = 30335														
31																	

Report

Figure 7-4 Example of an analysis report in Excel software

Analysis Protocols, Options, and Analysis Defaults

8

This chapter covers:

About Analysis Protocols

Analysis protocols were introduced in ABI PRISM® Sequencing Analysis Software v5.0. They replace preference settings used in previous versions of Sequencing Analysis software. An analysis protocol contains all the settings necessary for basecalling and post processing. A protocol is stored in the sample file.

The following settings are contained in an analysis protocol:

- Protocol name
- Basecalling and file format settings
- Mixed bases settings
- Clear range settings

Types of Analysis Protocols

There are two types of analysis protocols, per-sample and master.

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.

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 described under "Applying Analysis Protocols to Data" on page 8-18, or the analysis default, if the sample does not have a protocol.

The following default master analysis protocols are provided with the software:

- 3730BDTv3-KB-DeNovo_v5.1
- 3100POP6_BDTv3-KB-DeNovo_v5.1
- 310POP6_BDTv3-KB-DeNovo_v5.1

Analysis Protocol Manager

The Analysis Protocol Manager is an interface for managing the creation, modification, application, and deletion of master analysis protocols.

Parts of an Analysis Protocol

An analysis protocol is defined in the Sequence Analysis Protocol Editor window of the Analysis Protocol Manager, which has four tabs described in Table 8-1.

Note: For information on creating and editing analysis protocols, see "Creating and Editing Analysis Protocols" on page 8-13.

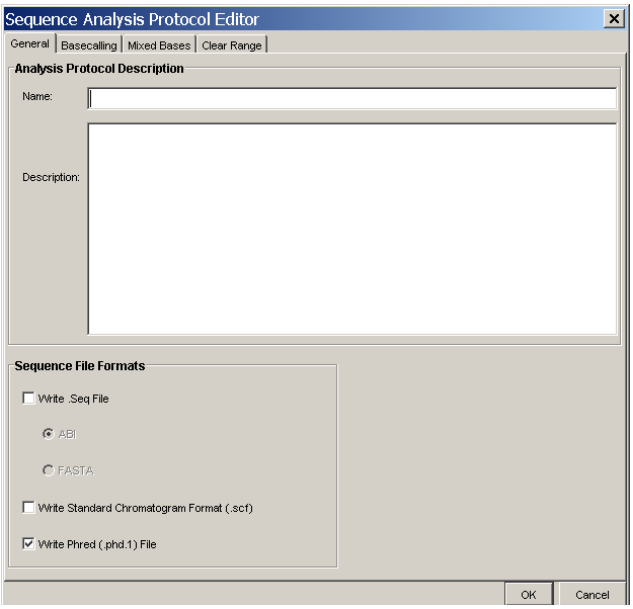


Table 8-1 Tabs in an Analysis Protocol

Tab	In this view you can …
General	Name and describe the analysis protocol, and define the sequence file formats to be used.
Basecalling	Select the basecaller, DyeSet/Primer file, matrix, and analysis stop point.
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.
Clear Range	Specify the clear range based on base positions, quality values, or number of ambiguities (Ns).

Default Master Analysis Protocol Settings

The default settings for the master analysis protocols, are listed in the tables below.

Table 8-2 3730BDTv3-KB-DeNovo_v5.1 Protocol Settings

Tab	Default Settings
General	Name: **3730BDTv3-KB-DeNovo_v5.1** Sequence file formats: **Write Phred (phd.1) File**
Basecalling	Basecaller: **KB.bcp** DyeSet/Primer file: **KB_3730_POP7_BDTv3.mob** Matrix File: **None** Ending base options: **All deselected** Processed Data: **True Profile** Quality Threshold: **Do not assign Ns to Basecalls**
Mixed Bases	Mixed base identification: **Deselected**
Clear Range	Use clear range: **Deselected** Use quality values: **Selected** Use identification of N calls: **Deselected**

Table 8-3 3100POP6_BDTv3-KB-DeNovo_v5.1 Protocol Settings

Tab	Default Settings
General	Name: **3100POP6_BDTv3-KB-DeNovo_v5.1** Sequence file formats: **Write Phred (phd.1) File**
Basecalling	Basecaller: **KB.bcp** DyeSet/Primer file: **KB_3100_POP6_BDTv3.mob** Matrix File: **None** Ending base options: **All deselected** Processed Data: **True Profile** Quality Threshold: **Do not assign Ns to Basecalls**
Mixed Bases	Mixed base identification: **Deselected**

Table 8-3 3100POP6_BDTv3-KB-DeNovo_v5.1 Protocol Settings *(continued)*

Tab	Default Settings
Clear Range	Use clear range: **Deselected** Use quality values: **Selected** Use identification of N calls: **Deselected**

Table 8-4 310POP6_BDTv3-KB-DeNovo_v5.1 Protocol Settings

Tab	Default Settings
General	Name: **310POP6_BDTv3-KB-DeNovo_v5.1** Sequence file formats: **Write Phred (phd.1) File**
Basecalling	Basecaller: **KB.bcp** DyeSet/Primer file: **KB_310_POP6_BDTv3_50Std.mob** Matrix File: **TestMatrix.mtx** Ending base options: **All deselected** Processed Data: **True Profile** Quality Threshold: **Do not assign Ns to Basecalls**
Mixed Bases	Mixed base identification: **Deselected**
Clear Range	Use clear range: **Deselected** Use quality values: **Selected** Use identification of N calls: **Deselected**

General Tab

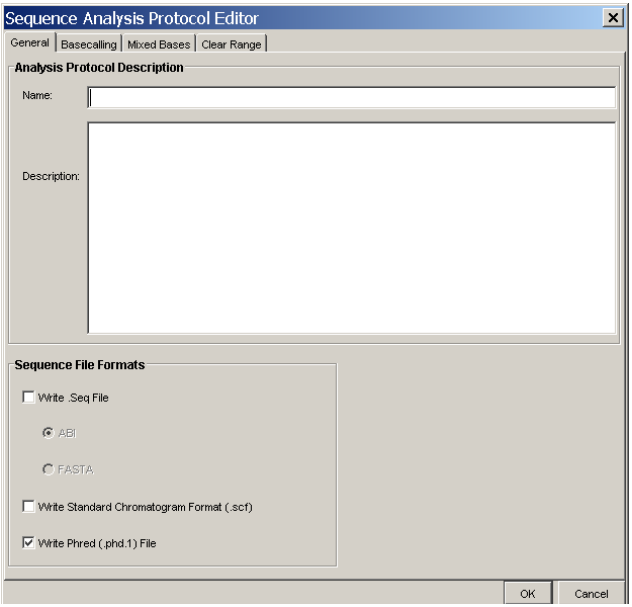


Table 8-5 Parts of the General Tab

Item	Function
Analysis Protocol Description Section	
Name	Name of the analysis protocol
Description	Description of the protocol
Sequence File Formats Section	
Write .Seq File check box	When selected, the software creates 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
Write Standard Chromatogram Format file (.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 the KB basecaller is used, the software creates a .phd.1 file that can be used with other software.

Basecalling Tab

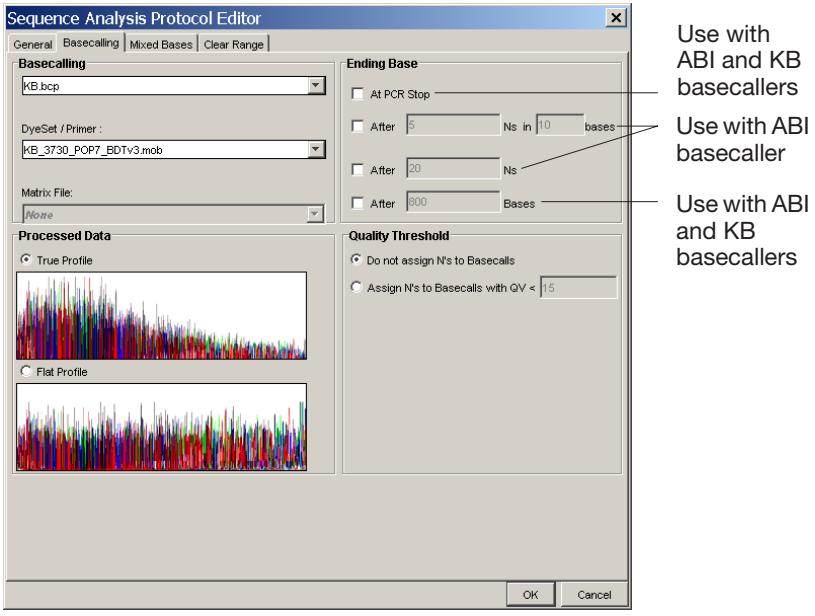


Table 8-6 Parts of the Basecalling Tab

Item	Function
Basecalling Section	
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. • ABI basecaller – Algorithms used in previous versions of ABI PRISM Sequencing Analysis software. Refer to Appendix C for a list of basecallers sorted by instrument.

Table 8-6 Parts of the Basecalling Tab *(continued)*

Item	Function
DyeSet/Primer drop-down 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 for a list of DyeSet/Primer files sorted by instrument and basecaller. **IMPORTANT!** The DyeSet/Primer file type must match the basecaller type. • 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 drop-down list	Used to select a matrix file for 310 or 377 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 and 377 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 8-6 Parts of the Basecalling Tab *(continued)*

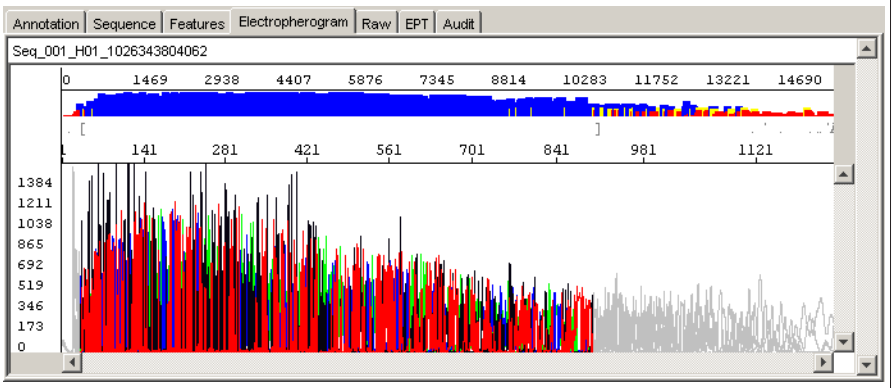
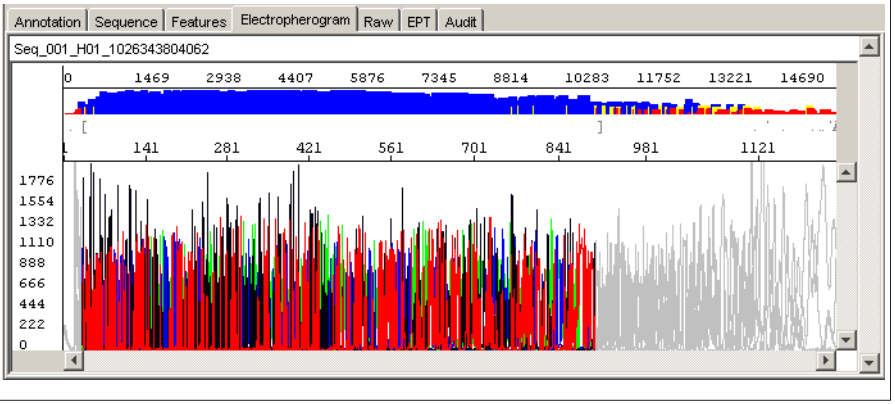
Item	Function
Processed Data	
True Profile	Used to display data as processed traces scaled uniformly so that the average height of peaks in the region of strongest signal is about equal to a fixed value. The profile of the processed traces will be very similar to that of the raw traces.
	
Flat Profile	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 only. If you use the ABI basecaller the profile option reverts to True Profile.
	

Table 8-6 Parts of the Basecalling Tab *(continued)*

Item	Function
Ending Base Section	
At PCR Stop check box	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 check box	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 check box	Sets the analysis endpoint after a certain number of Ns occur (for example, after 20 Ns are detected).
After __ Bases check box	Sets the analysis endpoint after a certain number of bases (for example, after 800 bases are detected).
Quality Threshold	
Call all bases and assign QV	When using the KB basecaller, use this setting assign a base to every position, as well as the QV.
Assign ‘N’ for bases with QV>	When using the KB basecaller, use this setting assign Ns to base with QVs less than the set point. The QV will still be displayed.

Mixed Bases Tab

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

Note: This option should be used with the KB basecaller only.

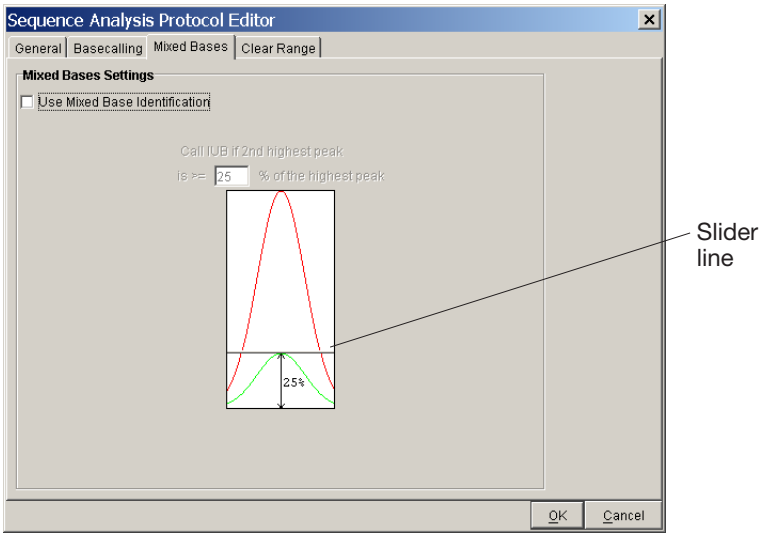


Table 8-7 Parts of the Mixed Bases Settings Section

Item	Description
Use Mixed Base Identification	• Select the Mixed Bases option for mixed base data. The KB basecaller assigns A, C, G, T, or an IUB code to every base. • Deselect the Mixed Bases option for pure base data. The KB basecaller assigns A, C, G, or T to every base. **Note:** The QVs indicate the quality of the basecalls.
Call IUB if 2nd highest peak is ≥ % of the highest peak	Set the % limit by typing in a value or moving the slider line on the graphic up or down. The default is 25%.

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.

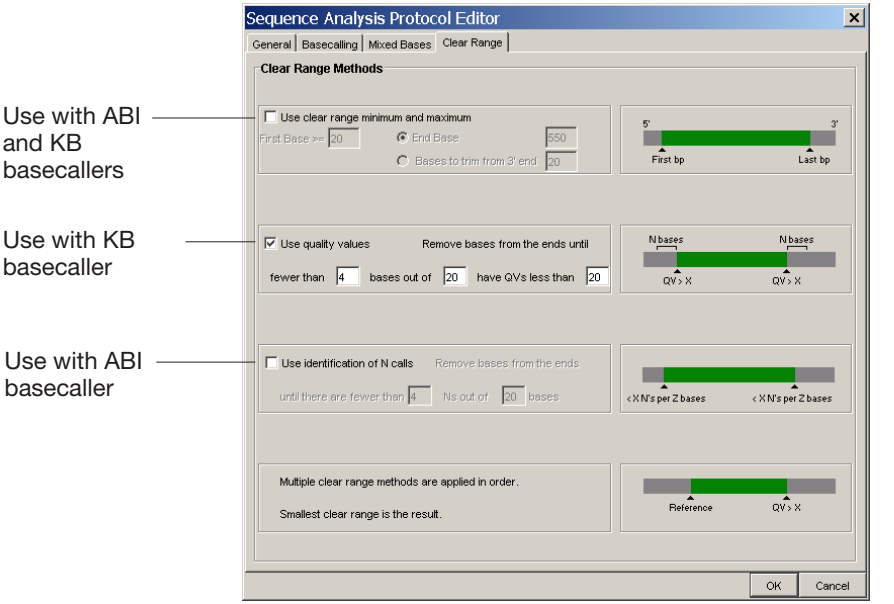


Table 8-8 Parts of the Clear Range Methods Section

Item	Function
Use clear range minimum and maximum check box	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 check box	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 check box	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).

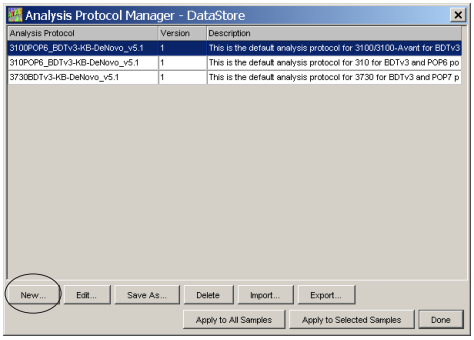
Note: More than one method can be used at a time.The clear range methods are applied in order. The smallest clear range is the result.

Creating and Editing Analysis Protocols

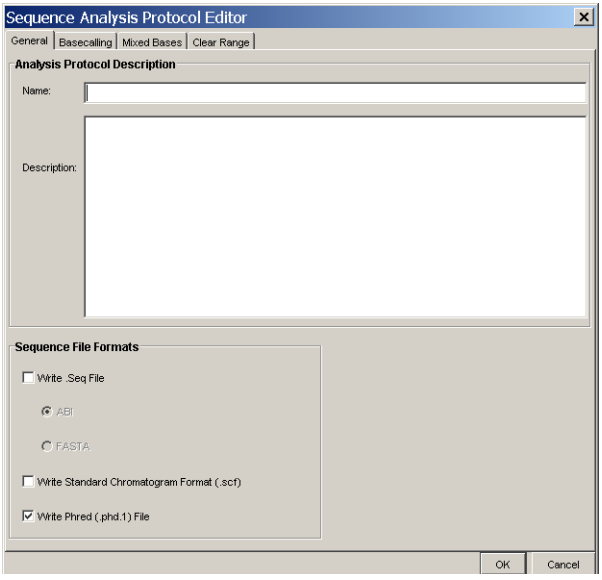
Creating a Master Analysis Protocol

To create a master analysis protocol:

1. Select **Analysis > Analysis Protocol Manager**. The Analysis Protocol Manager opens.

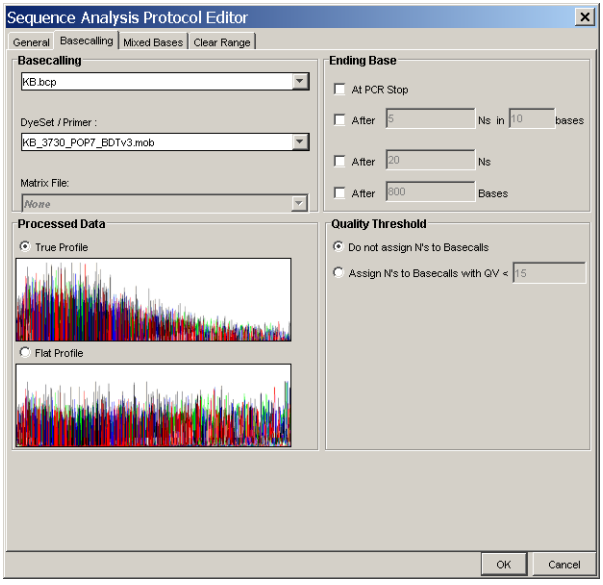


2. Click **New**.
3. In the **General** tab:



   a. Enter a unique name and description for the new protocol.
   b. Select the appropriate Sequence File formats settings.

4. Select the **Basecalling** tab, then:



a. Select the correct Basecaller and DyeSet/Primer file from the drop-down lists. Refer to Appendix C, “Basecallers and DyeSet/Primer Files.”

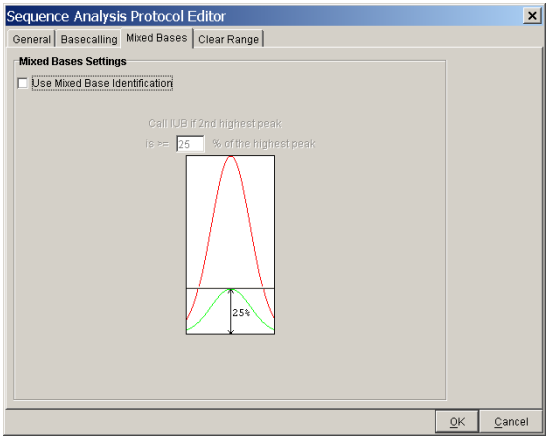
b. For 310 and 377 data select a matrix from the drop-down list.

c. In the Processed data pane, select True or Flat profile (for KB basecaller only).

d. If desired, select one or more stop points for data analysis.

e. Select your Threshold Quality option (for KB basecaller only).

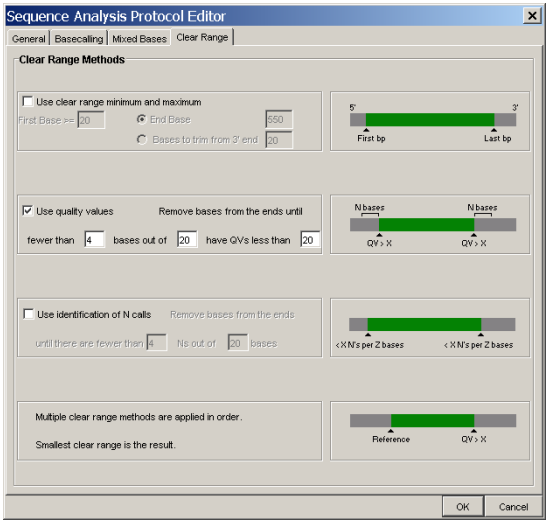
5. Select the **Mixed Bases** tab, then:



   a. For mixed base data, select **Use Mixed Base Identification** check box.

   b. Use the default setting of **25%** or change the detection level by either entering a new value or by dragging the % line up or down.

6. Select the **Clear Range** tab, then:



   If desired, select one or more stop points for data analysis.

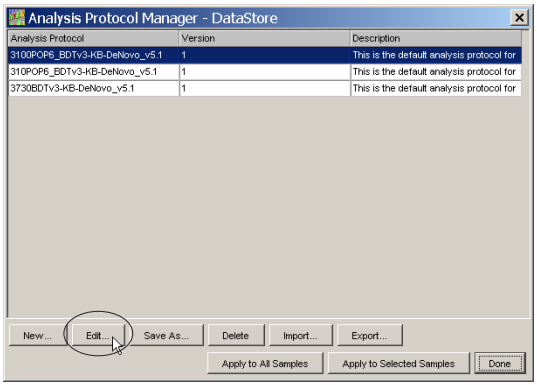
7. Click **OK** to save the protocol and close the Sequence Analysis Protocol Editor dialog box.
8. Click **Done** to close the Analysis Protocol Manager.

To apply the Analysis Protocol to samples, refer to page 8-18.

Editing a Master Analysis Protocol

To edit a master analysis protocol:

1. Select **Analysis > Analysis Protocol Manager**. The Analysis Protocol Manager opens.



2. In the Analysis Protocol column, select the protocol you want to edit.
3. Click **Edit**, or double-click the protocol name.
4. Make changes in the General, Basecalling, Mixed Bases, and Clear Range tabs, as appropriate.
5. Click **OK** to save the protocol and close the Sequence Analysis Protocol Editor dialog box.

   Note: The version number increments by one every time you press OK.
6. Click **Done** to close the Analysis Protocol Manager dialog box.

To apply the analysis protocol to samples, refer to page 8-18.

Deleting a Master Analysis Protocol

To delete a master analysis protocol:

1. Select **Analysis > Analysis Protocol Manager**.
2. In the Analysis Protocol column, select the protocol you want to delete, then select **Delete**.
3. In the Deletion Confirmation dialog box, click **Yes**.
4. Click **Done** to close the Analysis Protocol Manager.

Editing a Per-sample Analysis Protocol

Use the Analysis Protocol function to change the analysis protocol for an individual sample in the Sample Manager.

To edit the per-sample analysis protocol:

1. Add sample(s) to the Sample Manager (see "Adding Sample Files to the Sample Manager" on page 3-10).
2. Select a sample row in the Sample Manager.
3. Select **Analysis > Analysis Protocol**.
4. Optional: Make changes in the General, Basecalling, Mixed Bases, and Clear Range tabs, as appropriate.
5. Click **OK**.
6. Reanalyze the sample.
7. If you want to save the changes, save the sample file (see "Saving the Sample Files" on page 4-19).

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

Applying Analysis Protocols to Data

Note: After applying a master protocol to sample(s), you must reanalyze the sample(s) for the protocol settings to take affect.

Applying a Master Protocol to Multiple Samples

Use the Analysis Protocol Manager function to change the analysis protocol for multiple samples.

To apply a master protocol to multiple samples:

1. Add sample(s) to the Sample Manager (see "Adding Sample Files to the Sample Manager" on page 3-10).
2. Select the sample rows in the Sample Manager.
   - Use the Shift key to select continuous samples.
   - Use the Ctrl key to select discontinuous samples.
3. Select **Analysis > Analysis Protocol Manager**. The Analysis Protocol Manager opens.
4. Select the analysis protocol you want to apply, then click:
   - **Apply to Selected Samples** to apply the protocol to the sample files selected in step 2, or
   - **Apply to Selected Samples** to apply the protocol to all the sample files in the Sample Manager
5. Click **Done** to close the Analysis Protocol Manager.
6. Reanalyze the samples.
7. If you want to save the changes, save the sample files (see "Saving the Sample Files" on page 4-19).

Applying Original Analysis Settings to Samples

Use the Apply Pre-Analysis Settings function to apply original analysis settings to sample(s).

To apply the pre-analysis settings:

1. Add sample(s) to the Sample Manager.
2. Select the sample rows in the Sample Manager.
   - Use the Shift key to select continuous samples.
   - Use the Ctrl key to select discontinuous samples.
3. Select **Analysis > Apply Pre-Analysis Settings**.

4. Reanalyze the sample(s).
5. If you want to save the changes, save the sample files (see “Saving the Sample Files” on page 4-19).

Analysis Protocol Sharing Between Data Collection and Sequencing Analysis Software

For 3730/3730*xl* *and* 3100/3100-Avant Data Collection Only

IMPORTANT! For proper installation of the Sequencing Analysis Software v5.1 on a computer that is connected to a Applied Biosystems 3730/3730*xl* DNA Analyzer or ABI PRISM® 3100/3100-Avant Genetic Analyzer, the data collection software version 2.0 must be running. See Chapter 1, "Installing the Sequencing Analysis Software," for information on properly installing the software.

A master analysis protocol can also be created in 3730/3730*xl* Data Collection software v2.0 and 3100/3100-Avant Data Collection software v2.0.

In the following table, the conditions for file sharing are defined. The term "MAP" refers to master analysis protocol.

Conditions	Result	Fix
Sequencing Analysis software installed while data collection was open (proper installation)		
• MAP created in Sequencing Analysis • Data collection open	MAP is registered in both applications, and is available for use in the data collection collection software.	—
• MAP created in Data Collection • Sequencing Analysis open	MAP is registered in both applications, and is available for use in the analysis software.	—
Sequencing Analysis software installed while data collection was closed (improper installation)		
• MAP created in Sequencing Analysis or in data collection • Other software open or closed	Sequencing Analysis was never registered in the Data Service—no communication between the software.	1. Uninstall the Sequencing Analysis software. 2. Open the data collection software. 3. Reinstall the Sequencing Analysis software. 4. Register the software.

Analysis Defaults

When a sample file is added to the Sample Manager, then it passes through the 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.

Parts of the Analysis Defaults

An example of the Analysis Defaults window (below) contains the Add Samples Settings and Sequence File Formats sections.

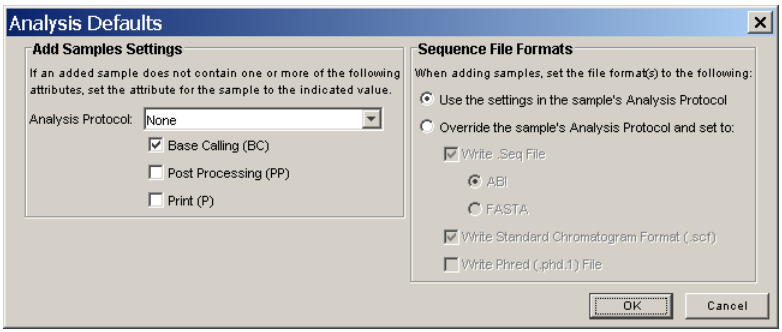


Table 8-9 Parts of the Analysis Defaults Dialog Box

Item	Function
Add Samples Settings Section	
Analysis Protocol	Displays a drop-down list of analysis protocols to apply to sample files that have no protocols. You can create a new master analysis protocol or edit an existing one.
Basecalling check box	When selected, the BC parameter check box is selected for each sample you add to the Sample Manager.
Post Processing check box	When selected, the PP parameter check box is selected for each sample you add to the Sample Manager.
Print check box	When selected, the P parameter check box is selected for each sample you add to the Sample Manager.

Table 8-9 Parts of the Analysis Defaults Dialog Box *(continued)*

Item	Function
Sequence File Formats Parameters	
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 settings in 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

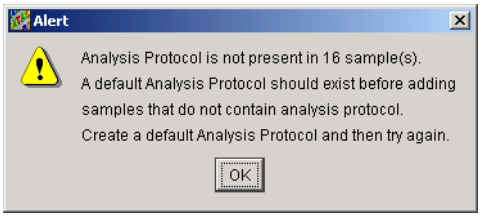
Table 8-10 When the Settings are Applied to the Samples

Setting	Settings are applied to samples …
Analysis protocol	Without an analysis protocol.
Processing parameters	When they are added to the Sample Manager.
Sequence file formats parameters	When they are added to the Sample Manager and analyzed.

IMPORTANT! If a sample has an analysis protocol associated with it, then an edited protocol or a new protocol is not applied to the sample. Refer to "Creating and Editing Analysis Protocols" on page 8-13 for information regarding changing the analysis protocol.

Missing Analysis Defaults

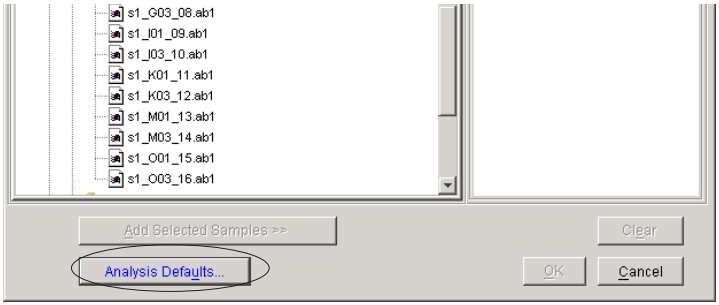
A default analysis default was included with the installation of your software. If the analysis protocol is set to None, *and* the samples added to the Sample Manager do not contain associated analysis protocols, then the following alert box displays. Samples will not be added to the Sample Manager until an analysis default with an analysis protocol is created.



Ways to Access Analysis Defaults

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

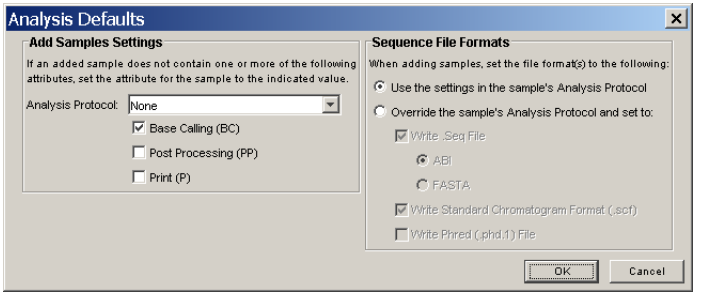
- Select **Analysis > Analysis Defaults**, or
- Click **Analysis Defaults** in the Add Samples window.



Editing and Applying the Analysis Defaults

To edit and apply the Analysis Defaults:

1. Select **Analysis > Analysis Defaults**.



2. In the Add Samples Settings section:

   a. Select an analysis protocol from the drop down list.

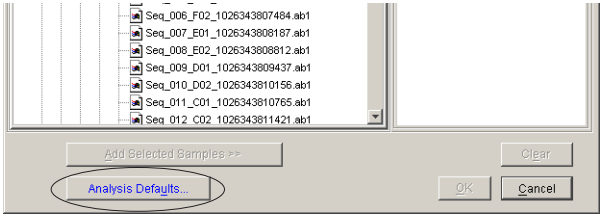
   Note: To create or edit an analysis protocol, use the Analysis Protocol drop-down list to select New or Edit, then refer to page 8-3 to define your settings.

   b. Select/deselect the Basecalling (BC), Post Processing (PP), and Print (P) options as applicable.

3. In the Sequence File Formats section, select to use the current settings or override them, then select/deselect the option to write Phred files.
4. Click **OK**.
5. Add samples to the Sample Manager.
6. If you made changes to the sequence file formats settings, then analyze the samples.

To edit and apply the Analysis Defaults in the Add Samples dialog box:

1. Click (Add Sample(s)) or select **File > Add Samples**.
2. In the Add Samples dialog box, locate, then open the folder that contains the files you want to add to the Sample Manager window.



3. In the Samples To Add pane of the dialog box, select the files that you want in the Sample Manager.
4. Click **Analysis Defaults**.
5. In the Add Samples Settings section:
   a. Select an analysis protocol from the drop down list.

   Note: To create or edit an analysis protocol, use the Analysis Protocol drop-down list to select New or Edit, then refer to page 8-3 to define your settings.

   b. Select/deselect the Basecalling (BC), Post Processing (PP), and Print (P) options as applicable.
6. In the Sequence File Formats section, select to use the current settings or override them, then select/deselect the option to write Phred files.
7. Click **OK** to close the Analysis Defaults dialog box.
8. Click **Add Selected Samples**.
9. Click **OK** to add the samples and close the dialog box.
10. If you made changes to the sequence file formats settings, then analyze the samples.

Options

The Options dialog box allows you to set the options for the file format, automated printing, create users, and audit trail.

To open the Options dialog box, select **Tools > Options**.

Parts of the Options Dialog Box

The four tabs in the Options dialog box are described in the table below.

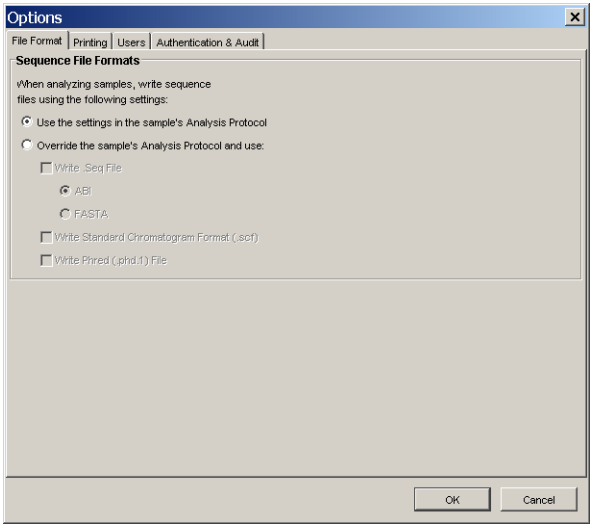


Table 8-11 Parts of the Options Dialog Box

Tab	In this view you can …
File Format	Choose to override the analysis protocol settings for the file format.
Printing	Change the default printing parameters for automatic printing.
Users	Create new users and edit current user information. Only the Administrator can use this tab.
Authentication & Audit	Set user lockout, password change, active the audit trail and define reasons for changes in the data.

File Format Tab

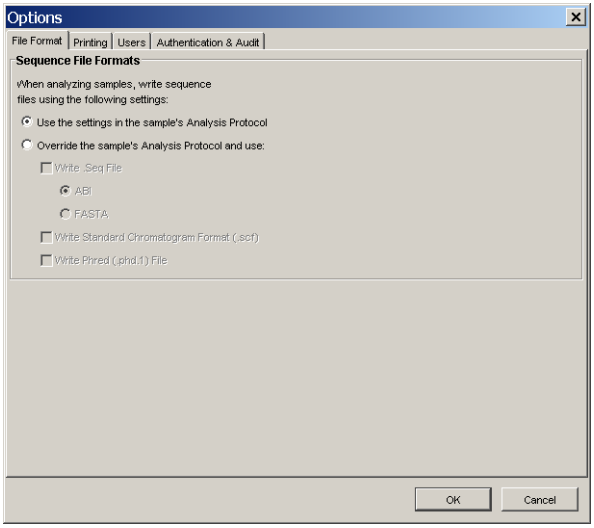


Table 8-12 Parts of the Sequence File Formats Section

Item	Function
Use the settings option button 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, during analysis the software creates: • 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

Printing Tab

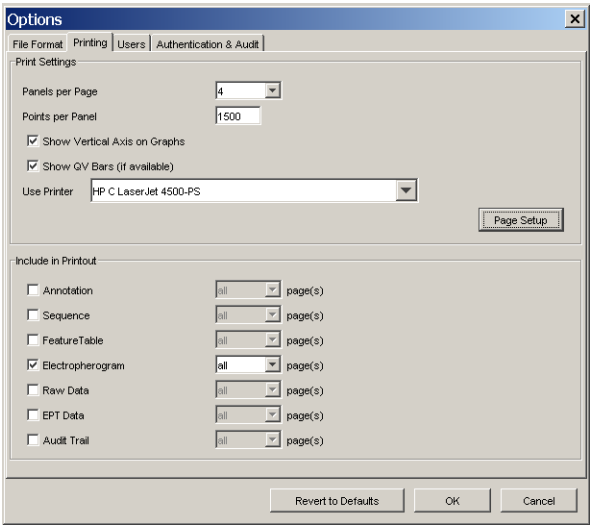
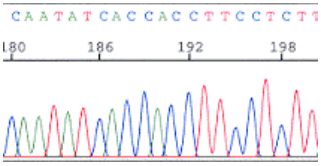
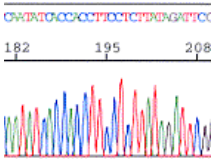


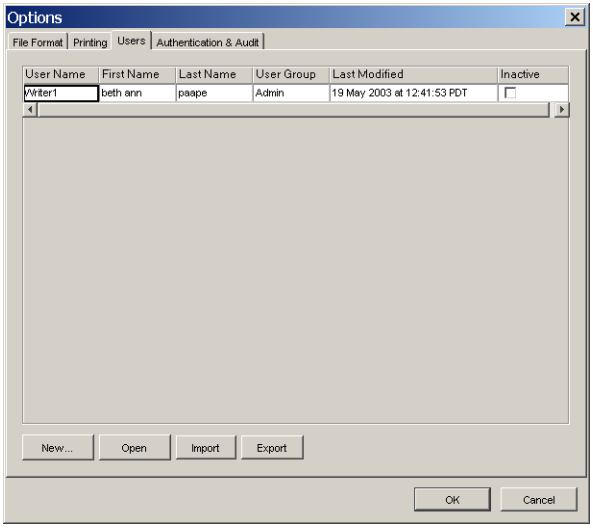
Table 8-13 Parts of the Printing Tab

Item	Function
Print Settings Section	
Panels Per Page drop-down list	Used to 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.

Table 8-13 Parts of the Printing Tab *(continued)*

Item	Function
Points Per Panel value box	Used to display the number of data points 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. 700 points per panel  1500 points per panel 
Show Vertical Axis on Graphs	Used to show /hide the vertical axis on the graphs
Show QV Bars (if available)	Used to show/hide the QV bars in the electropherogram and sequence views
Use Printer	Used to select a printer
Include in Printout Section	
View and Pages	Used to select the view to be printed and the number of pages of that view. The default is Electropherogram and all pages. The page range is 1-5.

Users Tab

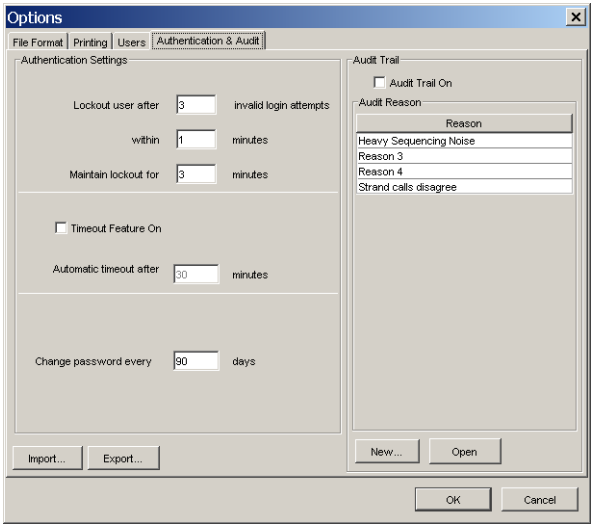


IMPORTANT! The administrator is the only person who can set up and change the information in the Users tab. The selections in this tab are inactive for all other users.

Table 8-14 Parts of the Users Tab

Item	Function
User Name list	List of all users names, user group, and activity status.
New	To create a new user and assign a user group.
Open	To open and change or verify user information.
Import	To import user profiles from other computers.
Export	To export user profiles from other computers.

Authentication & 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 8-15 Parts of the Authentication and Audit Tab

Item	Function
Authentication Settings Section	
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	The amount of time the user is locked out of the software.
Timeout Feature On check box	Used to turn off or on the timeout feature. The default is off.
Automatic timeout after __ minutes	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.

Table 8-15 Parts of the Authentication and Audit Tab *(continued)*

Item	Function
Audit Trail Section	
Audit Trail On check box	Used to turn off or on the audit trail feature. The default is off. If the Audit trail is on, is information 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 take with the data.
New button	Used to create addition reasons.
Open button	Used to modify or verify a reason.

Control Buttons

Available in the Printing tab only

Table 8-16 Control Button Function

Button	Function
Revert to Defaults (Printing tab only)	Returns all settings within the tab to their default settings.
OK	Accepts any pending changes and closes the dialog box.
Cancel	Closes the dialog box without accepting changes.

Ways to Change the Sequence File Formats

There are multiple ways to change the sequence file formats.

In the Analysis Defaults Dialog Box

- Select **Analysis > Analysis Defaults** (see page 8-24), or
- Click **Analysis Defaults** in the Add Samples window (see page 8-25).

In the Options Dialog Box

To change the Options dialog box settings:

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.
5. Reanalyze the samples to create the new file formats.
6. Save the samples.

In the Per-sample Analysis Protocol

To change the per-sample analysis protocol:

1. Add sample(s) to the Sample Manager.
2. Select a sample row in the Sample Manager.
3. Select **Analysis > Analysis Protocol**.
4. In the General tab, select the desired file formats.
5. Click **OK**.
6. Reanalyze the sample to create the new file formats.
7. Save the samples.

Ways to Change the Processing Parameters

There are multiple ways to change the processing parameters (BC, PP, and P).

In the Analysis Defaults Dialog Box

- Select **Analysis > Analysis Defaults** (see page 8-24), or
- Click **Analysis Defaults** in the Add Samples window (see page 8-25).

In the Sample Manager

To change the parameters in the Sample Manager:

1. Add sample(s) to the Sample Manager.
2. Select/deselect the **BC, PP** and/or **P** check boxes as appropriate.

Ways to Change the Analysis Parameters

There are multiple ways to change the analysis parameters (basecaller and DyeSet/Primer file).

In the Per-sample Analysis Protocol

To change the per-sample analysis protocol:

1. Add sample(s) to the Sample Manager.
2. Select a sample row in the Sample Manager.
3. Select **Analysis > Analysis Protocol**.
4. Select the Basecalling tab, then change the basecaller and DyeSet/Primer file, as needed.

   IMPORTANT! Make sure that the basecaller and the DyeSet/Primer files types match.
5. For 310 and 377 data select a matrix from the drop-down list.
6. Click **OK**.
7. Reanalyze the sample.
8. Save the samples.

In the Sample Manager

To change the basecaller and/or DyeSet/Primer file:

1. In the Sample Manager, select the sample(s).
2. In the basecaller drop-down list, select a new basecaller.
3. In the DyeSet/Primer drop-down list, select a new DyeSet/Primer file.

   IMPORTANT! Make sure that the basecaller and the DyeSet/Primer files types match.
4. For 310 and 377 data select a matrix from the drop-down list
5. Select the **BC** check box (and the **PP** and/or **P** check boxes, if desired).
6. Reanalyze the sample.
7. Save the samples.

Ways to Change the Analysis Protocol Settings

There are multiple ways to change the analysis protocol settings.

In the Per-sample Analysis Protocol

To change the per-sample analysis protocol:

1. Add sample(s) to the Sample Manager.
2. Select a sample row in the Sample Manager.
3. Select **Analysis > Analysis Protocol**.
4. Make changes in the General, Basecalling, Mixed Bases, and Clear Range tabs, as needed.
5. Click **OK**.
6. Reanalyze, then save the sample.

In the Analysis Protocol Manager

In the Analysis Protocol Manager

1. Add sample(s) to the Sample Manager.
2. Select the sample rows in the Sample Manager.
   - Use the Shift key to select continuous samples.
   - Use the Ctrl key to select discontinuous samples.
3. Select **Analysis > Analysis Protocol Manager**. The Analysis Protocol Manager opens.
4. Open the Analysis protocol you want to change, then:
   a. Make changes in the General, Basecalling, Mixed Bases, and Clear Range tabs, as needed.
   b. Click **OK**.
5. Select the analysis protocol you want to apply, then click:
   - **Apply to Selected Samples** to apply the protocol to the sample files selected in step 2, or
   - **Apply to Selected Samples** to apply the protocol to all the sample files in the Sample Manager
6. Click **Done** to close the Analysis Protocol Manager.
7. Reanalyze the samples.
8. Save the samples.

Display Settings 9

This chapter covers:

About Display Settings

For the Electropherogram view, Raw view, and EPT view of the sample window, you can use the Display Settings dialog box to:

- Determine which color is used to represent each kind of data.
- Change the colors of the trace lines to make them easier to see on screen.
- Turn off one or more trace line(s) selectively.
- Change the type of scaling used for the display.
- Turn on the display of quality values (QVs) and original bases, selectively.
- Select colors to represent QV and Length of Read (LOR) ranges.

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  .

Parts of the Display Settings Dialog Box

There are two tabs in the Display Settings dialog box, Bases and Data, and are described in Table 9-1 on page 9-3.

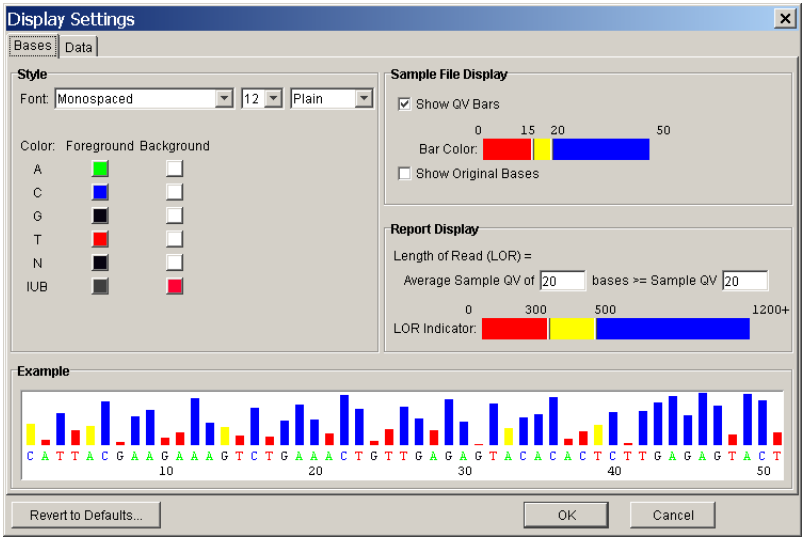


Table 9-1 Parts of the Display Settings Dialog Box

Tabs	In this tab you can…
Bases	• 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
Data	• Change the viewing parameters of electropherogram, raw, and EPT data • Selectively turn off one or more trace lines

Bases Tab

Parts of the Bases Tab

There are four sections within the Bases tab of the Display Settings dialog box. The sections are:

- Style
- Sample File Display
- Report Display
- Example

Style Section

This section allows you to change the font type, size, and style, and along with the color for both screen and printed data.

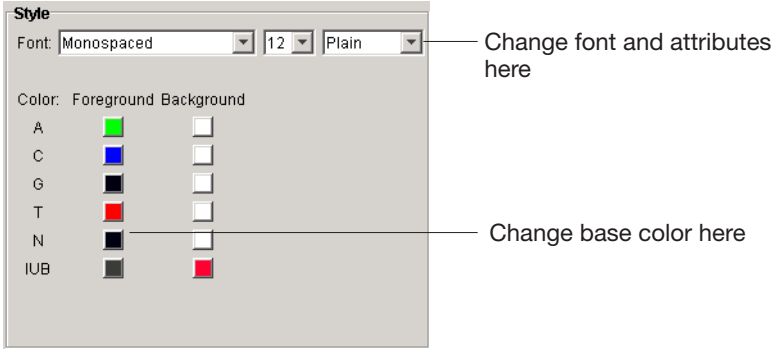
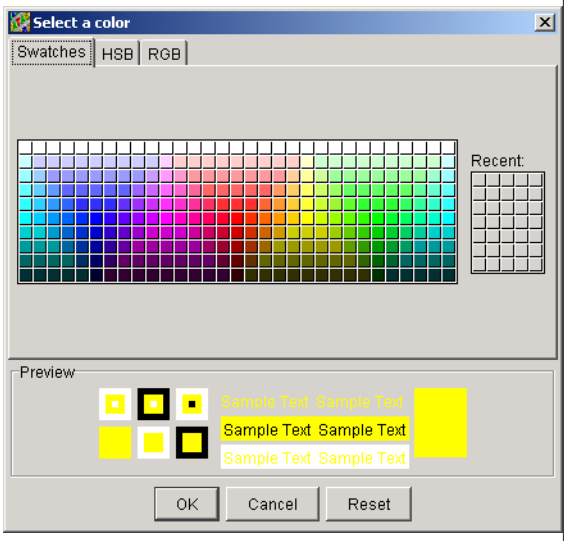


Table 9-2 Parts in the Style Section

Item	Function
Font	Use the drop-down list to select a different font type, size, and/or style.
Color	Use this option to change the color of the foreground (bases) and background. To change the color: 1. Select a color to change. The Select a color dialog box opens.  2. Select a new color. 3. Click **OK**.

Sample File Display Section

This section allows you to display/hide QV bars, and/or original bases for both screen and printed data.

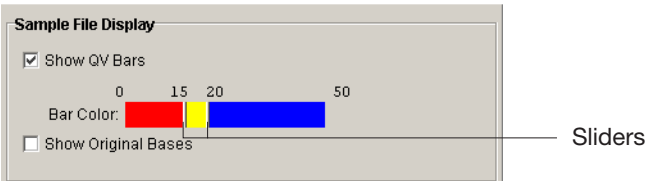


Table 9-3 Parts in the Sample File Display Section

<table>
<tr><th>Item</th><th>Function</th></tr>
<tr><td>Show QV Bars check box</td><td>Displays a QV for each base for samples analyzed with the KB basecaller. A selected check box is the default setting.

For more information, see Chapter 6, “Quality Values.”</td></tr>
<tr><td>Bar Color</td><td>The low, medium, and high ranges and the color associated with a QV can be modified using the Bar Color.

To change the ranges use the two sliders to define the low, medium, and high range.

<table>
<tr><th>QV Bar</th><th>Default Color and Range</th><th>Set the range to identify data that is …</th></tr>
<tr><td>Low</td><td>Red
0 to 14</td><td>Not acceptable</td></tr>
<tr><td>Medium</td><td>Yellow
15 to 19</td><td>Needs manual review</td></tr>
<tr><td>High</td><td>Blue
20 or higher</td><td>Acceptable</td></tr>
</table>
To change the color:

1. Select a color to change. The Select a color dialog box opens.
2. Select a new color.
3. Click OK.</td></tr>
<tr><td>Show Original Bases</td><td>If selected, two lines of data will be displayed. The top line is the original data and the bottom line is editable sequence. A clear check box is the default setting.

For more information, refer to “Showing the Original Data” on page 4-14.</td></tr>
</table>

Report Display Section

This section allows you to set LOR and display information for the analysis report.

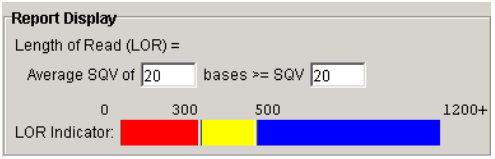
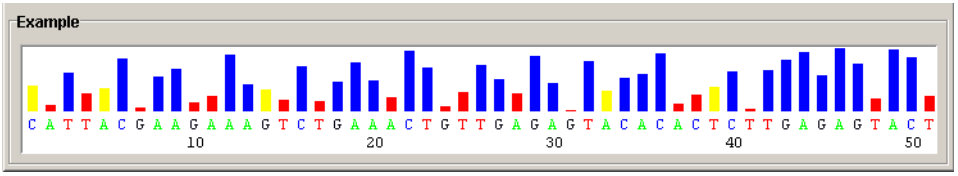


Table 9-4 Parts in the Report Display Section

Item	Function
Length Of Read (LOR)	The usable range of high quality or high accuracy bases, as determined by quality values. You can view it in the analysis report instead of opening the sample file.
LOR Indicator bar	The short, medium, and long ranges and the color associated with a LOR can be modified using the LOR Indicator bar. To change the range: Use the two sliders to define the short, medium, and long ranges. The sliders move in increments of 5 bases. To change the color: 1. Select a color to change. The Select a color dialog box opens. 2. Select a new color. 3. Click **OK**.

Example Section

This section allows you to view an example sequence with the various settings applied, including color and QVs.



Data Tab

The Data tab contains the Display Setting controls for Electropherogram, Raw Data, and EPT views.

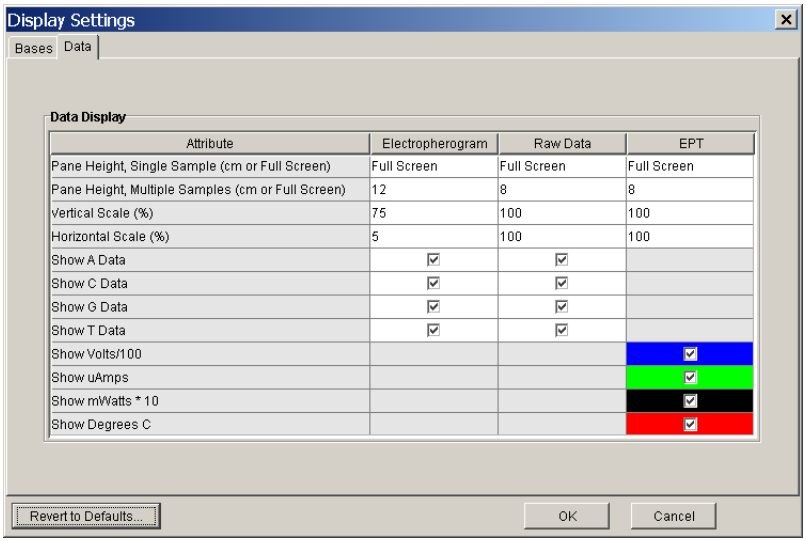


Table 9-5 Parts in the Data Display Section

Item	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.

Table 9-5 Parts in the Data Display Section *(continued)*

Item	Function
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 to100. 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 to100.
Show A, C, G, and T Data	When selected the electropherogram and raw data traces for the A, C, G, and T are displayed. All are selected by default. The default colors are: • A = Green • C = Blue • G = Yellow • T= Red
Show Volts/100, Show µAmps, Show mWatts x 10 and Show Degrees C	When selected, the data traces for the volts, µAmps, mWatts, and temperature are displayed. All are selected by default and the colors represent: • Volts/100 = Blue • µAmps = Green • mWatts = Black • Degrees C = Red

Control Buttons

Table 9-6 Control Button Function

Button	Function
Revert to Defaults	Returns all settings within a tab to their default settings
OK	Accepts any pending changes and closes the dialog box
Cancel	Closes the dialog box without accepting changes

Changing the Display Setting

Changing the Display Settings

To change the Display Settings:

1. Select **Analysis > Display Settings** or click  . The Display Settings dialog box opens.
2. Select the **Bases** tab and make changes, as needed.
3. Select the **Data** tab and make changes, as needed.
4. Click **OK**.

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.

Reverting to the Default Settings

To revert to the default settings:

1. Select **Analysis > Display Settings** or click  . The Display Settings dialog box opens.
2. In the Bases tab:
   a. Click **Revert to Defaults**.
   b. In the Revert to Defaults dialog box, click **Yes**.
3. In the Data tab:
   a. Click **Revert to Defaults**.
   b. In the Revert to Defaults dialog box, click **Yes**.
4. Click **OK**.

Matrix Maker 10

This chapter covers:

Making a Matrix File in Sequencing Analysis

This utility is used to make matrix files from data generated on an ABI PRISM® 310 Genetic Analyzer or ABI PRISM® 377 DNA Sequencer.

Matrix files can be create from two different types of data:

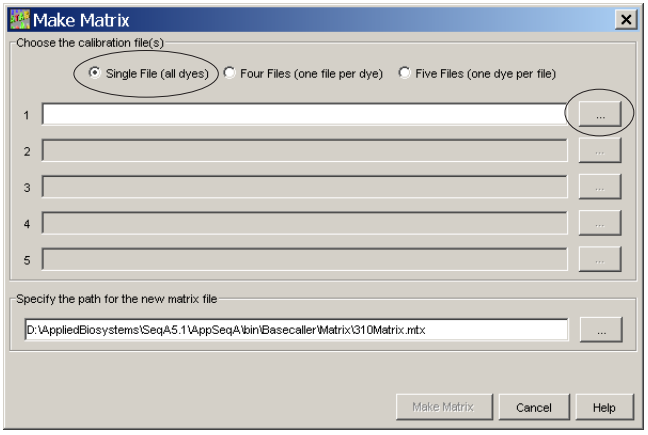
- A single sample file that has a good distribution of As, Gs, Cs, and Ts in the data.
- Matrix standard data files

  These are files contain only one color per sample.

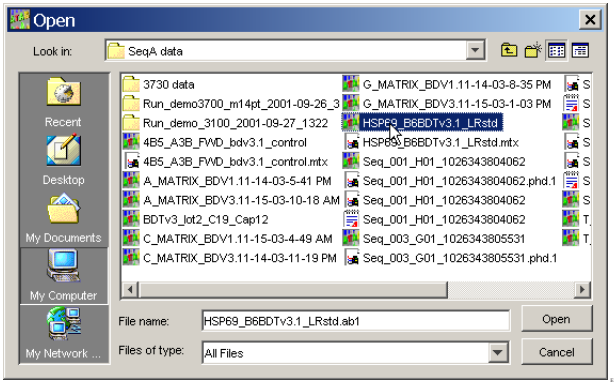
Using a Single Sample File

To create a matrix from a single sample file:

1. Select **Tools > Make Matrix**. The Make Matrix dialog box opens.

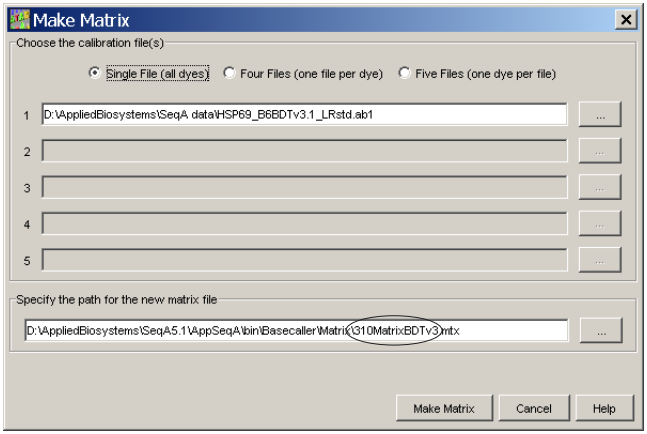


2. Verify the radio button for **Single File (all dyes)** is selected.
3. Click the button next to line 1. An Open dialog box opens.
4. Locate and select the file you want to use for the matrix, then click **Open**.



5. Enter a name for the new matrix file using a .mtx extension.

   IMPORTANT! Do not change the pathway, only change the matrix name. If the file is not stored in the following location ***drive letter*:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\Basecaller\Matrix*matrixname.mtx***, the matrix is not available for selection in the Sample Manager or in an analysis protocol.



6. Click **Make Matrix**.

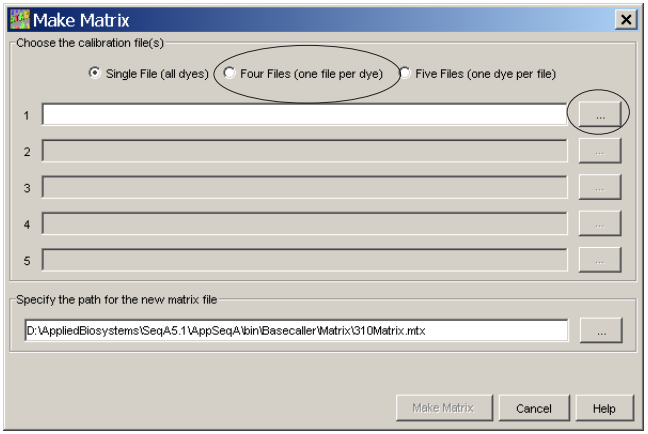
   If the matrix was successfully made, then the following message opens.

   Click **OK** to close the Matrix File created and Make Matrix dialog boxes.

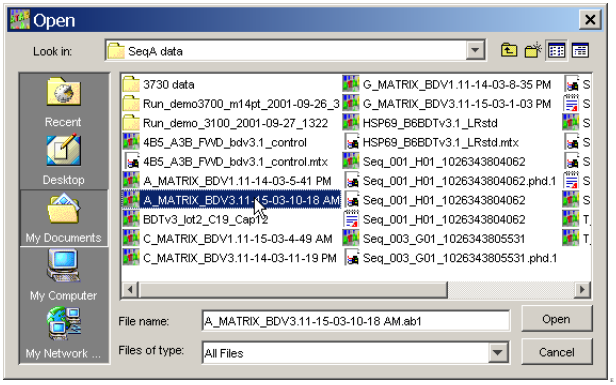
Using Matrix Standard Files

To create a matrix from four matrix standard files:

1. Select **Tools > Make Matrix**. The Make Matrix dialog box opens.



2. Select the radio button **Four Files (one file per dye)**.
3. Click the button next to a line. An Open dialog box opens.
4. Locate and select one of the file you want to use to make the matrix, then click **Open**.

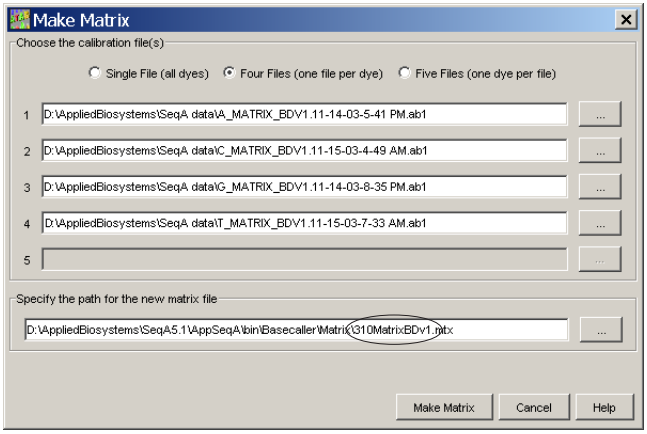


5. Repeat steps 3 and 4 until all four files have been selected.

   IMPORTANT! Verify you have a file that represents green, blue, yellow, and red data. The order of the colors does not matter.

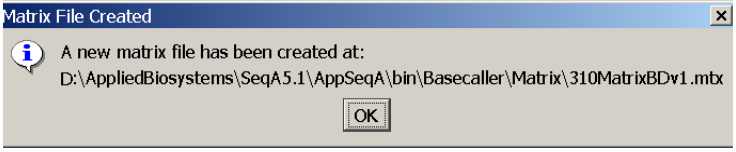
6. Enter a name for the new matrix using a .mtx extension.

   IMPORTANT! Do not change the pathway, only change the matrix name. If the file is not stored in the following location ***drive letter*:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\Basecaller\Matrix*matrixname.mtx***, the matrix will not be available for selection in the Sample Manager or in an analysis protocol.



7. Click **Make Matrix**.

   If the matrix was successfully made, then the following message opens.



   Click **OK** to close the Matrix File created and Make Matrix dialog boxes.

Duplicating Matrix Files for Data Collection

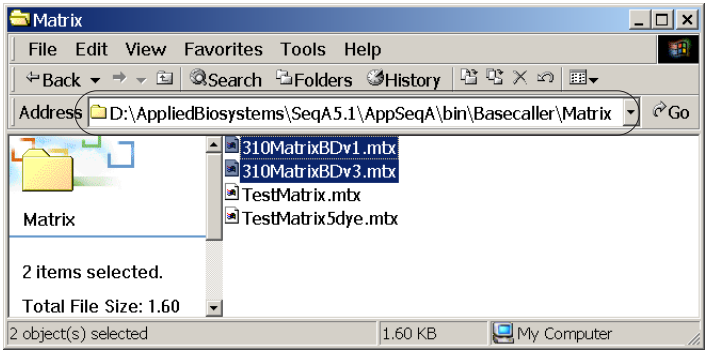
The 310 and 377 data collection software use a different folder location to access and store matrix files than the analysis software. To be able to select your new matrix file in the preferences and a sample sheet, copy a duplicate of the matrix file into the Matrix folder within Data Collection folder.

To copy and paste a matrix file in to the data collection folder:

1. Navigate to the location of the Matrix folder in Sequencing Analysis v5.1:

   D:\AppliedBiosystems\SeqA5.1\AppSeqA\bin\Basecaller\Matrix

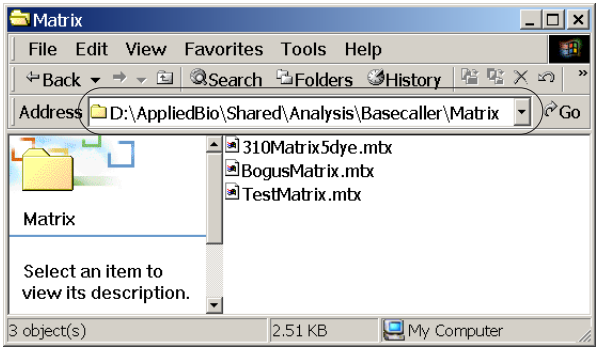
2. Open the Matrix folder, then select the matrix files to be copied.



3. Press **Ctrl+C** to copy the files.

4. Navigate to the location of the Matrix folder used by the data collection:

   D:\AppliedBio\Shared\Analysis\Basecaller\Matrix



5. Press **Ctrl+V** to paste the files into the folder.

Autoanalysis of 310 Sample Files

11

This chapter covers:

Autoanalysis of 310 Sample Files

Requirements

Sequencing data that is generated on the ABI PRISM® 310 Genetic Analyzer can be automatically analyzed by Applied Biosystems Sequencing Analysis Software v5.1. Autoanalysis can be performed only on the same instrument computer that collected the sample files. You can configure the software packages to perform data collection, then data analysis without requiring user interaction.

Autoanalysis requires:

- Microsoft® Windows® 2000, Service Pack 3 operating system
- 310 Data Collection software and Sequencing Analysis software v5.1 loaded on the same computer
- Matrix and DyeSet/Primer files have been copied into the correct folders for use in both data collection and analysis software (see "Copying 310 Matrix and DyeSet/Primer Files" on page 1-20)
- 310 Data Collection software set up to open the automated version of the analysis software (see page 11-3)
- Analysis Defaults be created in the analysis software (page 11-6)

Automated Version of the Analysis Software

The Automated version of the Sequencing Analysis software is named Automation310.exe.

This version of Sequencing Analysis software is identical to the regular version of the software, except that no user interface exists. The data collection opens this version of software to analyze the data.

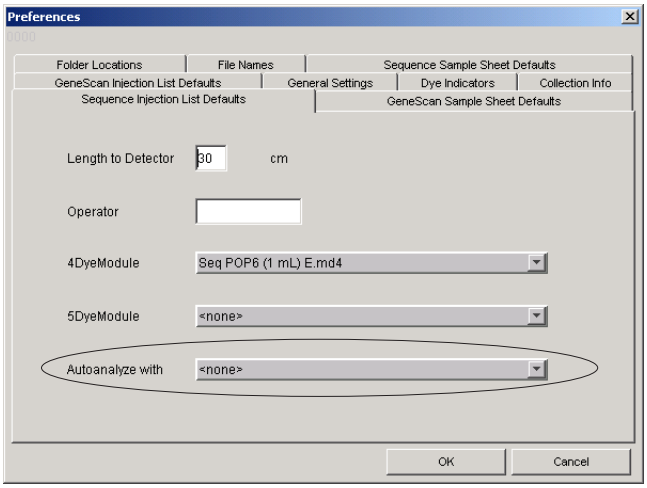
The automated version and the standard version of Sequencing Analysis software are automatically installed from the Sequencing Analysis Software v5.1 installation CD.

Setting Up the Software for Autoanalysis

Data Collection Software Set Up

To set up the data collection software:

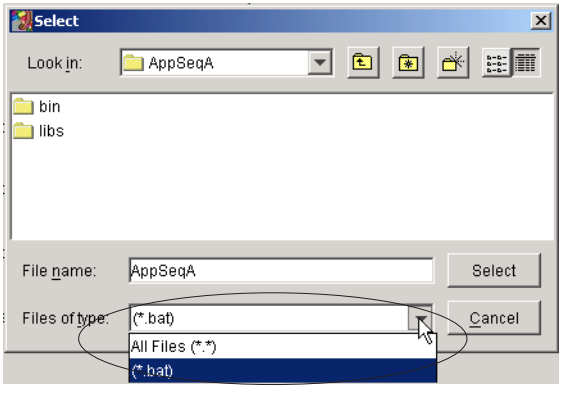
1. Open the 310 Data Collection software.
2. Select **Windows > Preferences > General Settings**.
3. Select the **Sequence Injection List Defaults** tab, then:



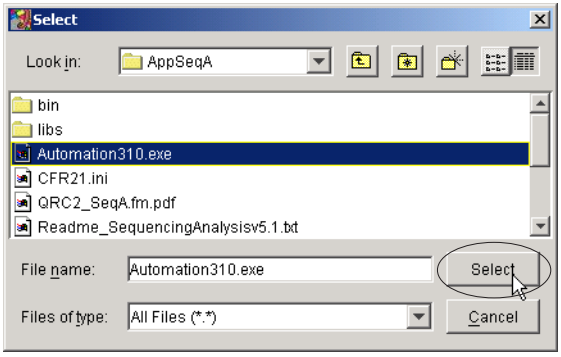
a. In the Autoanalyze with drop-down list, select **Other**.

b. In the Select dialog box, navigate to:
D:\AppliedBiosystems\SeqA5.1\AppSeqA

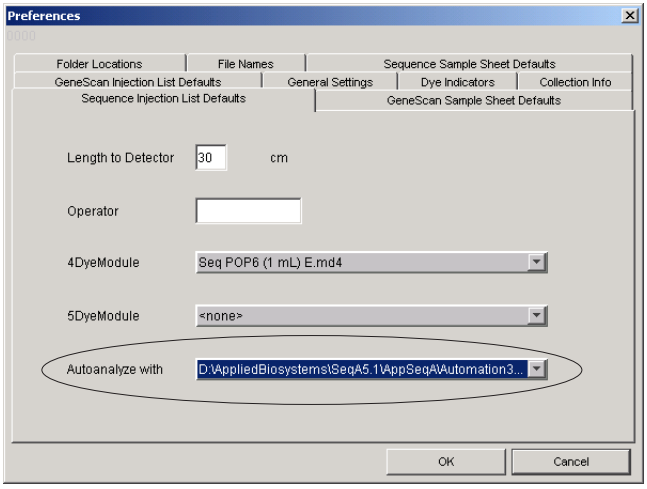
c. In the Files of type drop-down list, select **All Files**.



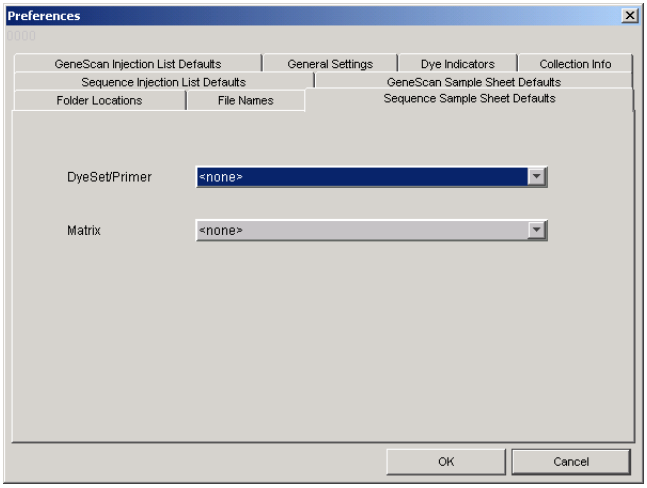
d. Select the Automation310.exe file, then click **Select**.



The Select dialog box closes and the pathway displays in the Sequence Injection List Defaults tab.



4. Select the **Sequence Sample Sheet Defaults** tab, then:



   a. In the DyeSet/Primer drop-down list, select a DyeSet/Primer file. See Appendix C, Basecallers and DyeSet/Primer Files to select the correct DyeSet/Primer and basecaller combination.

   IMPORTANT! The DyeSet/Primer file must match the chemistry and basecaller type that you are using in data collection and in the analysis protocol you create.

   b. In the Matrix drop-down list, select a matrix file.

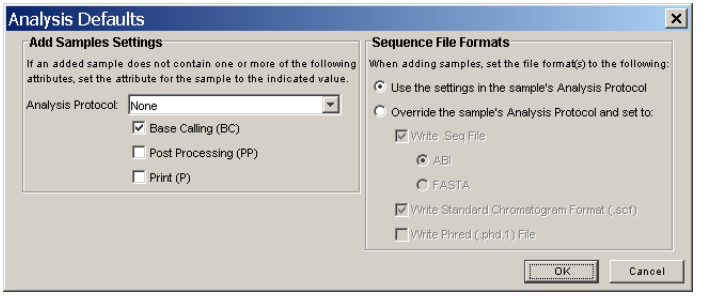
5. Click **OK**.

Sequencing Analysis Software Set Up

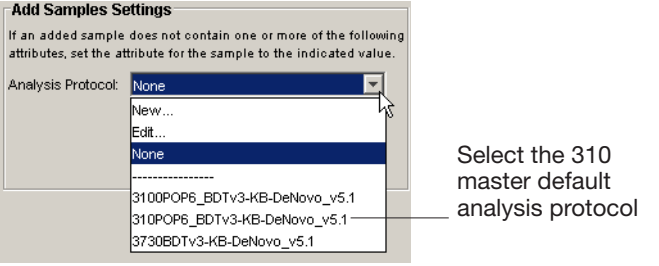
When a sample file is added to the Sample Manager, then it passes through the analysis defaults. Analysis defaults contain processing parameter settings (basecalling, post processing and printing), file formats settings (.seq, .scf and .phd.1) and an analysis protocol. The analysis protocol is assigned to the sample only if it does not already contain one. 310 sample files do not contain analysis protocols.

To set up the Analysis Defaults:

1. Select **Analysis > Analysis Defaults**.



2. In the Add Samples Settings section:
   a. In the Analysis Protocol drop-down list, do one of the following:
      – Select the default master analysis protocol, 310POP6_BDTv3-KB-DeNovo_v5.1, if the basecaller, DyeSet/Primer file, matrix file and other settings are correct for your run conditions. See Table 8-4 on page 8-5.



      – Create a new analysis protocol or edit an existing one.

   Note: To create or edit an analysis protocol, use the Analysis Protocol drop-down list to select New or Edit, then refer to page 8-3 to define your settings.

b. Select/deselect the Basecalling (BC), Post Processing (PP), and Print (P) options as applicable.

3. In the Sequence File Formats section, select to use the current settings or override them, then select/deselect the option to write Phred files.
4. Click **OK**.

How Autoanalysis Works

Data Collection Software

The data collection software runs the instrument, collects fluorescent data, and stores the raw data into sample files. At the completion of the run, the datacollection software launches Automation310.exe.

Automation310 Software

The Automation310 version of the analysis software applies the analysis defaults with an analysis protocol to the data. The software analyzes, stores, and saves the analyzed data and analysis protocol into the sample file. The Automation310 software automatically closes when the analysis is complete.

IMPORTANT! Because this version of analysis software does not have a user interface, the analysis process is not visible.

Sequencing Analysis Software

To work with your sample files, open the Sequencing Analysis software v5.1. Add the files of interest to the Sample Manager, and display, edit, reanalyze, print and/or save the data.

Frequently Asked Questions A

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

This appendix covers:

General Questions and Answers

Table A-1 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 v5.1 differ from v3.7?	New features include: • Novel basecaller algorithm that performs base calling for pure and mixed base calls • Generation of quality values to provide basecall accuracy information for pure and mixed base calls • Analysis report to help troubleshoot and provide easy assessment of data quality • New Sample Manager interface • Calculation of length of read (LOR) • Optional feature to generate an audit trail of base changes
What is the KB basecaller?	A new 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 algorithm 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 atcagagaaaaagtctttaactccaccattagcacccaaag ct (hard return)

Table A-1 General Questions and Answers *(continued)*

Question	Answer
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.
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 Applied Biosystems 3730/3730*x*/ DNA Analyzers, ABI PRISM® 3100/3100-Avant Genetic Analyzers , ABI PRISM® 310 Genetic Analyzers and ABI PRISM® 377 DNA Sequencer.
What are the minimum computer requirements for Sequencing Analysis software?	• CPU – 733 MHz or faster, single processor • Memory – 512 MB RAM • OS – Microsoft® Windows XP with Service pack 1 or Windows® 2000 Service pack 3 • 1 GB hard drive • Intel Pentium® III or IV chip, not Xeon
What is the minimal monitor resolution I should have?	1024 x 768 pixels
What kind of performance can I expect from my Sequencing Analysis software?	Performance depends on the computer specifications.

Table A-1 General Questions and Answers *(continued)*

Question	Answer
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.
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 A-2 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. See "Saving Read-only Sample Files" on page 4-19. 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.
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 A-3 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 “Saving the Sample Files” on page 4-19.
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 A-4 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.

Quality Values (QVs) Questions and Answers

Table A-5 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 "Interpreting the Per-Base Quality Values" on page 6-2.
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 "Customizing the Quality Value Display" on page 6-6.
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 A-6 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 A-7 Printing Questions and Answers

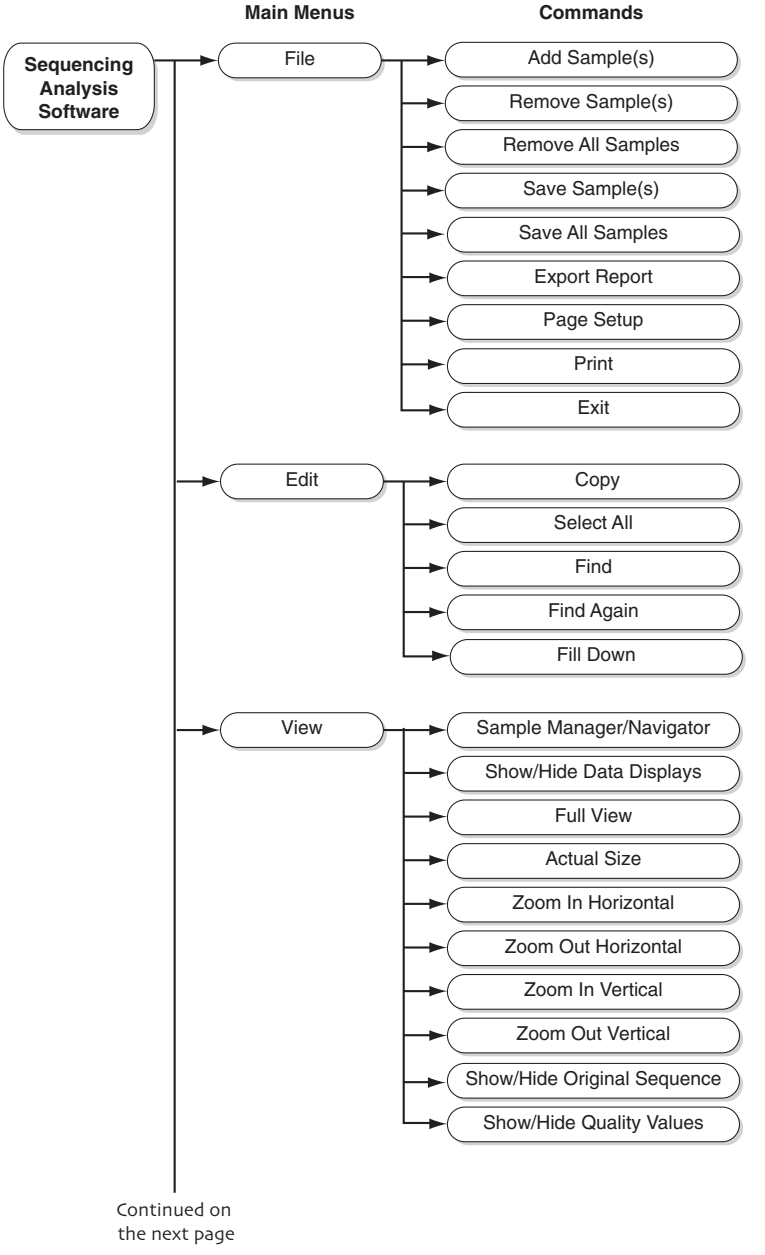
Question	Answer
What can I print in Sequencing Analysis software?	You can print sample views (annotative, sequence, feature, electropherogram, raw, and EPT) and analysis reports.
What printers are recommended for use with Sequencing Analysis software?	An HP® 8100, 4600, 990cxi, or an Epson® 980 color 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.

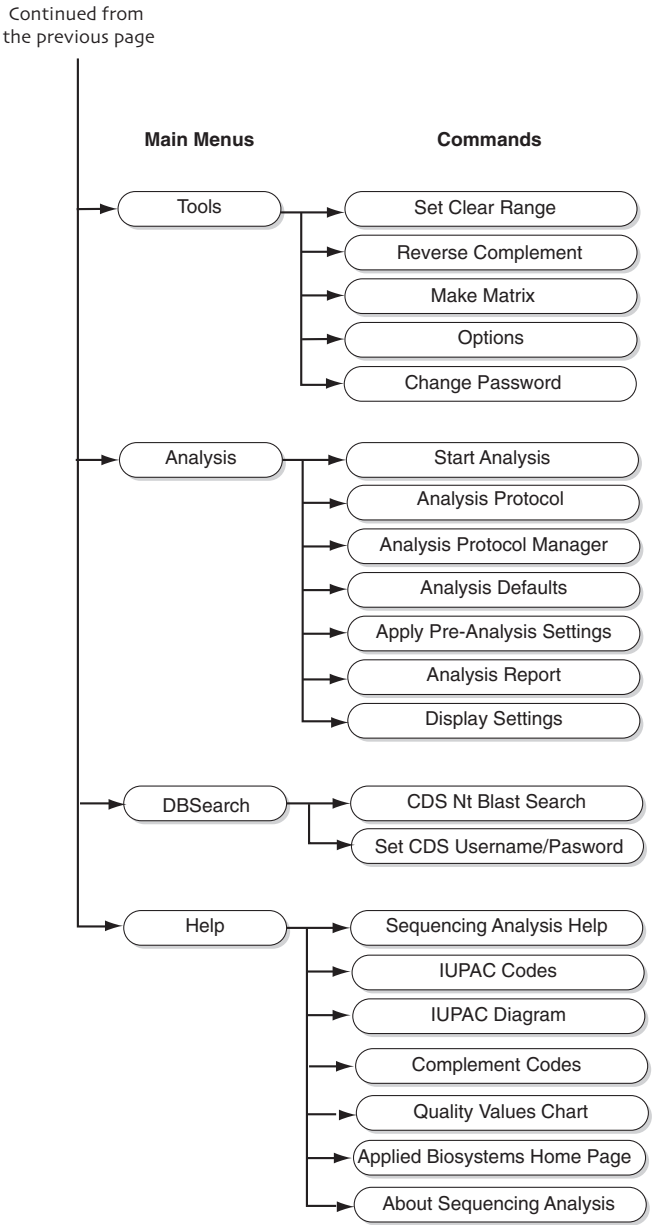
Menu Commands and Toolbar Buttons

B

This appendix covers:

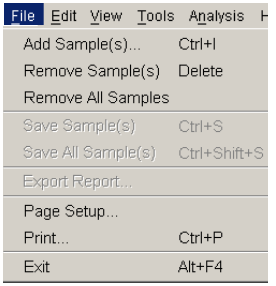
Menus Flowchart




Continued from
the previous page
Main Menus
Commands
Tools
Set Clear Range
Reverse Complement
Make Matrix
Options
Change Password
Analysis
Start Analysis
Analysis Protocol
Analysis Protocol Manager
Analysis Defaults
Apply Pre-Analysis Settings
Analysis Report
Display Settings
DBSearch
CDS Nt Blast Search
Set CDS Username/Pasword
Help
Sequencing Analysis Help
IUPAC Codes
IUPAC Diagram
Complement Codes
Quality Values Chart
Applied Biosystems Home Page
About Sequencing Analysis

Menu Commands

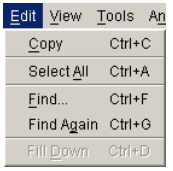
File Menu



The File menu contains the following commands:

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
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

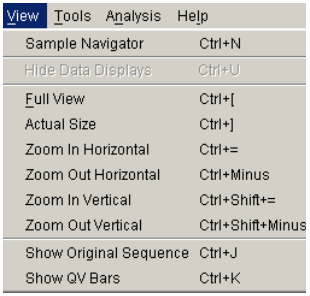
Edit Menu



The Edit menu contains the following commands:

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

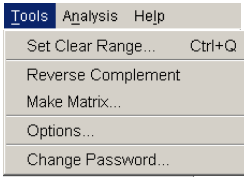
View Menu



The View menu contains the following commands:

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
Actual 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
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

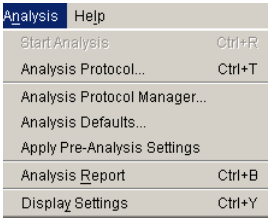
Tools Menu



The Tools menu contains the following commands:

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
Options	Opens the Options dialog box for selecting file formats, printing, users, and audit trail options
Change Password	Opens the User Management dialog box for changing the user's password

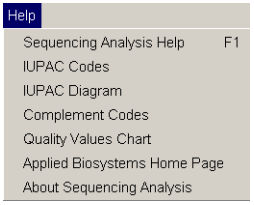
Analysis Menu



The Analysis menu contains the following commands:

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
Analysis Report (Ctrl+B)	Opens the Analysis report
Display Settings (Ctrl+Y)	Opens the Display Settings dialog box

Help Menu

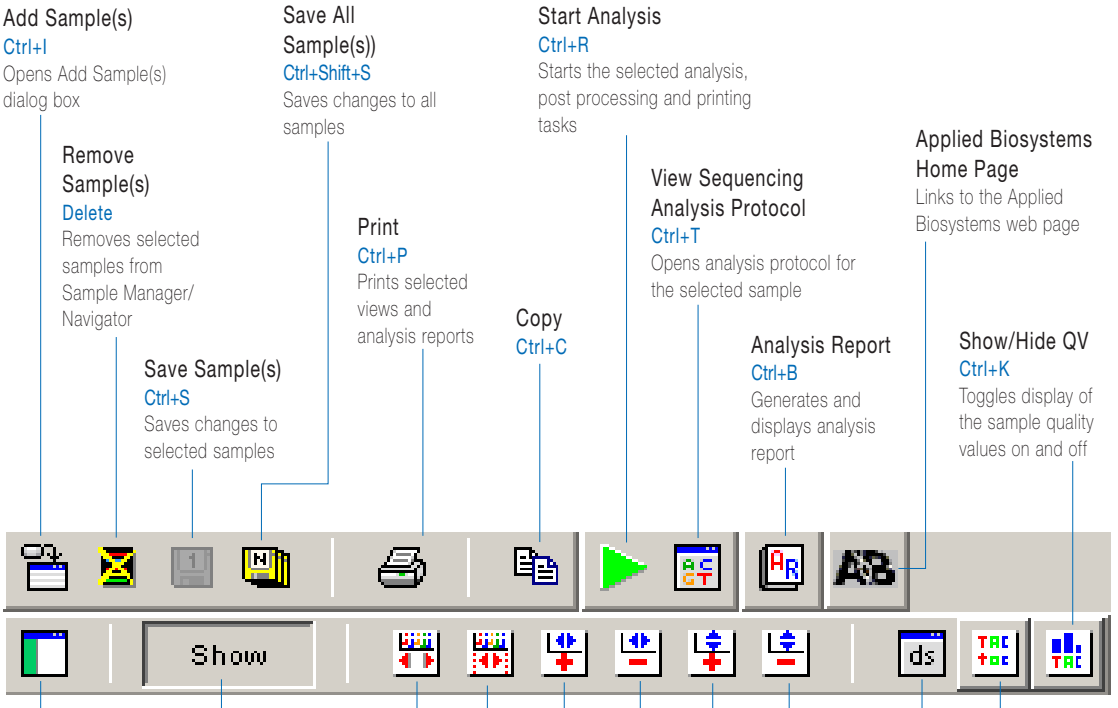


The Help menu contains the following commands:

Command	Description
Sequencing Analysis Help (F1)	Opens a PDF file of the *ABI PRISM® Sequencing Analysis Software v5.1 User Guide*
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
IUPAC Codes	Opens a display box containing a table of the quality values
Applied Biosystems Home Page	Opens a link to the home page of Applied Biosystems web site
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.



Basecallers and DyeSet/Primer Files

C

This appendix contains:

ABI PRISM 310 Genetic Analyzer Files

Note: 47 cm capillary array length = 36 cm read length
61 cm capillary array length = 50 cm read length

Table C-1 310 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
KB Basecalling			
ABI PRISM® BigDye® Terminator v1.0 and v1.1	47	KB.bcp	KB_310_POP4_BDTv1_36Rapid.mob KB_310_POP4_BDTv1_36Std.mob
	47	KB.bcp	KB_310_POP6_BDTv1_36Rapid.mob
	61		KB_310_POP6_BDTv1_50Std.mob
ABI PRISM® BigDye® Terminator v3.0 and v3.1	47	KB.bcp	KB_310_POP4_BDTv3_36Rapid.mob KB_310_POP4_BDTv3_36Std.mob
	47	KB.bcp	KB_310_POP6_BDTv3_36Rapid.mob
	61		KB_310_POP6_BDTv3_50Std.mob
ABI Basecalling			
ABI PRISM BigDye Terminator v1.0 and v1.1	47	Basecaller-310POP4.bcp	DT310POP4{BD}v2.mob
	47	Basecaller-310POP6.bcp	DT310POP6{BD}.mob
	61		DT310POP6{BD-LR}v3.mob

Table C-1 310 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry *(continued)*

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI PRISM® dRhodamine Terminator	47	Basecaller-310POP4.bcp	DT310POP4{dRhod}v1.mob
	47	Basecaller-310POP6.bcp	DT310POP6{dRhod}v2.mob
	61		
ABI PRISM BigDye Terminator v3.0 and v3.1	47	Basecaller-310POP4.bcp	DT310POP4{BDv3}v2.mob
	47	Basecaller-310POP6.bcp	DT310POP6{BDv3}v2.mob
	61		

Table C-2 310 Basecaller and DyeSet/Primer Files Used for Dye Primer Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecaller			
ABI PRISM® BigDye® Primer v1.0 and v1.1	47	Basecaller-310POP4.bcp	DP310POP4{BD-21M13}v1.mob DP310POP4{M13Rev}v1.mob
	47	Basecaller-310POP6.bcp	DP310POP6{BD-21M13}v1.mob DP310POP6{M13Rev}v1.mob
	61		

Table C-2 310 Basecaller and DyeSet/Primer Files Used for Dye Primer Chemistry *(continued)*

ABI PRISM BigDye® Primer v3.0 and v3.1	47	Basecaller-310POP4.bcp	DP310POP4{BDv3-21M13}v1.mob DP310POP4{BDv3-M13Rev}v1.mob
	47	Basecaller-310POP6.bcp	DP310POP6{BDv3-21M13}v1.mob DP310POP6{BDv3-M13Rev}v1.mob
	61		

ABI PRISM 377 DNA Sequencer Files

Table C-3 377 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	WTR (cm)/Scan Rate (scans/hr)	Basecaller	DyeSet/Primer
ABI Basecalling			
• ABI PRISM BigDye Terminator v1.0 and v1.1 • ABI PRISM dGTP BigDye Terminator	36/2400	Basecaller-377.bcp	DT377{BD}.mob
	36 & 48/1200	Basecaller-377LR.bcp	
ABI PRISM dRhodamine Terminator	36/2400	Basecaller-377.bcp	DT377{dRhod}.mob
	36 & 48/1200	Basecaller-377LR.bcp	
• ABI PRISM BigDye Terminator v3.0 and 3.1 • ABI PRISM dGTP BigDye v3.0 Terminator	36/2400	Basecaller-377.bcp	DT377{BDv3}v2.mob
	36 & 48/1200	Basecaller-377LR.bcp	DT377LR{BDv3}v1.mob

Table C-4 377 Basecaller and DyeSet/Primer Files Used for Dye Primer Chemistry

DNA Sequencing Chemistry	WTR (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
ABI PRISM BigDye Primer v1.0 and v1.1	36/2400	Basecaller-377.bcp	DP377-5%LR{BD-21M13}.mob DP377-5%LR{BD-M13Rev}.mob,
	36 & 48/1200	Basecaller-377LR.bcp	
ABI PRISM BigDye Primer v3.0 and 3.1	36/2400	Basecaller-377.bcp	DP377{BDv3-21M13}v1.mob DP377{BDv3-M13Rev}v1.mob
	36 & 48/1200	Basecaller-377LR.bcp	

ABI PRISM 3100 Genetic Analyzer Files

Table C-5 3100 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
KB Basecalling			
ABI PRISM BigDye Terminator v1.0 and v1.1	36: ultra rapid	KB.bcp	KB_3100_POP4_BDTv1.mob
	50: std read		
	80: long read		
	36: rapid read	KB.bcp	KB_3100_POP6_BDTv1.mob
	50: std read		
ABI PRISM BigDye Terminator v3.0 and v3.1	36: ultra rapid	KB.bcp	KB_3100_POP4_BDTv3_.mob
	50: std read		
	80: long read		
	36: rapid read	KB.bcp	KB_3100_POP6_BDTv3.mob
	50: std read		

Table C-5 3100 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry *(continued)*

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
• ABI PRISM BigDye Terminator v1.0 and v1.1 • ABI PRISM dGTP BigDye Terminator	36: ultra rapid	Basecaller-3100POP4UR.bcp	DT3100POP4LR{BD}v1.mob
	80: long read	Basecaller-3100POP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100POP6RRv2.bcp	DT3100POP6{BD}v2.mob
	50: std read	Basecaller-3100POP6SR.bcp	
• ABI PRISM BigDye Terminator v3.0 and 3.1 • ABI PRISM dGTP BigDye v3.0Terminator	36: ultra rapid	Basecaller-3100POP4UR.bcp	DT3100POP4{BDv3}v1.mob
	80: long read	Basecaller-3100POP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100POP6RRv2.bcp	DT3100POP6{BDv3}v1.mob
	50: std read	Basecaller-3100POP6SR.bcp	
ABI PRISM dRhodamine Terminator	36: ultra rapid	Basecaller-3100POP4UR.bcp	DT3100POP4{dRhod}v2.mob
	80: long read	Basecaller-3100POP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100POP6RRv2.bcp	DT3100POP6{dRhod}v2.mob
	50: std read	Basecaller-3100POP6SR.bcp	

Table C-6 3100 Basecaller and DyeSet/Primer Files Used for Dye Primer Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
ABI PRISM BigDye Primer v1.0 and v1.1	36: rapid read	Basecaller-3100POP6RRv2.bcp	DP3100POP6{BD-21M13}v1.mob DP3100POP6{BD-M13Rev}v1.mob
	50: std read	Basecaller-3100POP6SR.bcp	
ABI PRISM BigDye Primer v3.0 and 3.1	36: rapid read	Basecaller-3100POP6RRv2.bcp	DP3100POP6{BDv3-21M13}v1.mob DP3100POP6{BDv3-M13Rev}v1.mob
	50: std read	Basecaller-3100POP6SR.bcp	
ABI PRISM BigDye v3 Primer (All primers)	36: ultra rapid	Basecaller-3100POP4UR.bcp	DP3100POP4{BDv3}v1.mob
	80: long read	Basecaller-3100POP4_80cmv3.bcp	

ABI PRISM 3100-*Avant* Genetic Analyzer Files

Table C-7 3100-*Avant* Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
KB Basecalling			
ABI PRISM BigDye Terminator v1.0 and v1.1	36: ultra rapid	KB.bcp	KB_3100_POP4_BDTv1.mob
	50: std read		
	80: long read		
	36: rapid read	KB.bcp	KB_3100_POP6_BDTv1.mob
	50: std read		
ABI PRISM BigDye Terminator v3.0 and v3.1	36: ultra rapid	KB.bcp	KB_3100_POP4_BDTv3_.mob
	50: std read		
	80: long read		
	36: rapid read	KB.bcp	KB_3100_POP6_BDTv3.mob
	50: std read		

Table C-7 3100-*Avant* Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry *(continued)*

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
ABI PRISM BigDye Terminator v1.0 and v1.1	36: ultra rapid	Basecaller-3100APOP4UR.bcp	DT3100POP4LR{BD}v1.mob
	80: long read	Basecaller-3100APOP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100APOP6RRv2.bcp	DT3100POP6{BD}v2.mob
	50: std run	Basecaller-3100APOP6SR.bcp	
ABI PRISM BigDye Terminator v3.0 and 3.1	36: ultra rapid	Basecaller-3100APOP4UR.bcp	DT3100POP4{BDv3}v1.mob
	80: long read	Basecaller-3100APOP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100APOP6RRv2.bcp	DT3100POP6{BDv3}v1.mob
ABI PRISM dRhodamine Terminator	36: ultra rapid	Basecaller-3100APOP4UR.bcp	DT3100POP4{dRhod}v2.mob
	80: long read	Basecaller-3100APOP4_80cmv3.bcp	
	36: rapid read	Basecaller-3100APOP6RRv2.bcp	DT3100POP6{dRhod}v2.mob
	50: std run	Basecaller-3100APOP6SR.bcp	

ABI PRISM 3700 DNA Analyzer Files

Table C-8 3700 Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
ABI PRISM BigDye Terminator v1.0 and 1.1	50	Basecaller-3700POP6.bcp	DT3700POP6{BD}v5.mob
		Basecaller-3700POP5LR.bcp	DT3700POP5{BD}v3.mob
ABI PRISM BigDye Terminator v3.0 and 3.1	50	Basecaller-3700POP6.bcp	DT3700POP6{BDv3}v1.mob
		Basecaller-3700POP5LR.bcp	DT3700POP5{BDv3}v1.mob
ABI PRISM dRhodamine Terminator	50	Basecaller-3700POP6.bcp	DT3700POP6{dRhod}v3.mob
		Basecaller-3700POP5LR.bcp	DT3700POP5{dRhod}v1.mob

Table C-9 3700 Basecaller and DyeSet/Primer Files Used for Dye Primer Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
ABI Basecalling			
ABI PRISM BigDye Primer v1.0 and v1.1	50	Basecaller-3700POP6.bcp	DP3700POP6{BD-21M13}v3.mob DP3700POP6{BD-M13Rev}v2.mob
		Basecaller-3700POP5LR.bcp	DP3700POP5{BD-21M13}v1.mob DP3700POP5{BD-M13Rev}v1.mob
ABI PRISM BigDye Primer v3.0 and v3.1	50	Basecaller-3700POP6.bcp	DP3700POP6{BDv3-21M13}v1.mob DP3700POP6{BDv3-M13Rev}v1.mob
		Basecaller-3700POP5LR.bcp	DP3700POP5{BDv3-21M13}v1.mob DP3700POP5{BDv3-M13Rev}v1.mob

Applied Biosystems 3730/3730*xl* DNA Analyzers Files

Table C-10 3730/3730*xl* Basecaller and DyeSet/Primer Files Used for Dye Terminator Chemistry

DNA Sequencing Chemistry	Capillary Array Length (cm)	Basecaller	DyeSet/Primer
KB Basecalling			
ABI PRISM BigDye v3.0 Terminator	all lengths	KB.bcp	KB_3730_POP7_BDTv3.mob
ABI PRISM BigDye Terminator	all lengths	KB.bcp	KB_3730_POP7_BDTv1.mob
ABI Basecalling			
ABI PRISM BigDye Terminator v1.0 and v1.1	36: rapid read	Basecaller-3730POP7RR.bcp	DT3730POP7{BD}.mob
	36: std read	Basecaller-3730POP7SR.bcp	
	50: long read	Basecaller-3730POP7LR.bcp	
ABI PRISM BigDye Terminator v3.0 and v3.1	36: rapid read	Basecaller-3730POP7RR.bcp	DT3730POP7{BDv3}.mob
	36: std read	Basecaller-3730POP7SR.bcp	
	50: long read	Basecaller-3730POP7LR.bcp	

User Privileges

D

This appendix covers:

Tables of User Privileges

This appendix contains a list of privileges for users of the three categories, Administrator, Scientist, and Analyst, when they use the ABI PRISM® Sequencing Analysis Software v5.1.

Table D-1 Access for Admin Level

		Description of access for users of Admin level only	Admin	Scientist	Analyst
Admin only access	1	Create User Accounts	Allowed	Not Allowed	Not Allowed
	2	Exporting/Importing User Accounts			
	3	Turning on/off the Timeout feature			
	4	Turning on/off the Audit Trail feature			
	5	Mark an user inactive			

Table D-2 Access for Admin and Scientist Levels

		Description of access for users of Admin and Scientist levels	Admin	Scientist	Analyst
Analysis Protocol & Settings	1	Creating an analysis protocol	Allowed	Allowed	Not Allowed
	2	Editing an existing analysis protocol			
	3	Apply an analysis protocol to a set of samples			
	4	Delete an Analysis protocol			
	5	Set Clear range determination in Analysis settings or analysis defaults			
	6	Edit Display Settings			
	7	Edit an analysis protocol from the Analysis Defaults			
	8	Edit Analysis Protocol per sample			
	9	Use Save as to create a new Analysis protocol			

Table D-3 Access for Admin, Scientist and Analyst Levels

		Description of access for users of Admin, Scientist and Analyst levels	Admin	Scientist	Analyst
Reports	1	View Analysis Report	Allowed	Allowed	Allowed
	2	View Report with enabled links back to primary data			
	3	Customize report			
	4	Export report			
	5	Print report			
Sample Manager	6	Browse/locate data in the file system			
	7	Add samples to the Sample Manager			
	8	Delete samples in the Sample Manager			
	9	Print sample file views			
	10	Change the Basecaller and mobility file in the Sample Manager			
	11	Insert or delete a base in the electropherogram or sequence view			
	12	Change a base in the electropherogram or sequence view			
	13	Search for text in the electropherogram or sequence view			
	14	Edit sample name			
	15	View data in Sample Navigator view			
	16	Search for text strings in any sequence data			
Sequencing Analysis	17	Open the Sequencing Analysis			
	18	Exit the Sequencing Analysis			

Key Codes

E

This appendix provides translations for codes used in the Applied Biosystems Sequencing Analysis Software v5.1.

This appendix contains the following topics:

Translation Tables

Introduction

This section provides the following translation tables:

- IUPAC/IUB Codes
- Complements
- Universal Genetic Code
- Amino Acid Abbreviations

Note: These tables are available from the Help menu in the main window of the software.

IUPAC/IUB Codes

The table below provides translations for IUPAC/IUB codes used in the Sequencing Analysis 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

The table below provides complements for reference.

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 Sequencing Analysis 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
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G
Stop Codes: AMBer, OCHer, OPA					

Amino Acid Abbreviations

The table below provides 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

Software Warranty Information

F

This appendix contains the following:

Computer Configuration

Applied Biosystems supplies or recommends certain configurations of computer hardware, software, and peripherals for use with its instrumentation.
Applied Biosystems reserves the right to decline support for or impose extra charges for supporting nonstandard computer configurations or components that have not been supplied or recommended by Applied Biosystems. Applied Biosystems also reserves the right to require that computer hardware and software be restored to the standard configuration prior to providing service or technical support. For systems that have built-in computers or processing units, installing unauthorized hardware or software may void the Warranty or Service Plan.

Limited Product Warranty

Limited Warranty

Applied Biosystems warrants that for a period of ninety (90) days from the date the warranty period begins, its Applied Biosystems DNA Sequencing Analysis Software will perform substantially in accordance with the functions and features described in its accompanying documentation when properly installed on the instrument system for which it is designated, and that for a period of ninety (90) days from the date the warranty period begins, the tapes, diskettes, or other media bearing the software product will be free of defects in materials and workmanship under normal use. If buyer believes that it has discovered a failure of the software to satisfy the foregoing warranty, and if buyer notifies Applied Biosystems of such failure in writing during the ninety (90) day warranty period, and if Applied Biosystems is able to reliably reproduce such failure, then Applied Biosystems, at its sole option, will either (i) provide any software corrections or “bug-fixes” of the identified failure, if and when they become commercially available, to buyer free of charge, or (ii) notify buyer that Applied Biosystems will accept a return of the software from the buyer and, upon such return and removal of the software from buyer's systems, terminate the license to use the software and refund the buyer's purchase price for the software. If there is a defect in the media covered by the above warranty and the media is returned to Applied Biosystems within the ninety (90) day warranty period, Applied Biosystems will replace the defective media. Applied Biosystems does not warrant that the software will meet buyer's requirements or conform exactly to its documentation, or that operation of the software will be uninterrupted or error free.

Warranty Period Effective Date

Any applicable warranty period under these sections begins on the earlier of the date of installation or ninety (90) days from the date of shipment for software installed by Applied Biosystems personnel. For all software installed by the buyer or anyone other than Applied Biosystems, the applicable warranty period begins the date the software is delivered to the buyer.

Warranty Claims

Warranty claims must be made within the applicable warranty period.

Warranty Exceptions

The above warranties do not apply to defects resulting from misuse, neglect, or accident, including without limitation: operation outside of the environmental or use specifications, or not in conformance with the instructions for the instrument system, software, or accessories; improper or inadequate maintenance by the user; installation of software or interfacing, or use in combination with software or products, not supplied or authorized by Applied Biosystems; and modification or repair of the product not authorized by Applied Biosystems.

THE FOREGOING PROVISIONS SET FORTH APPLIED BIOSYSTEMS' SOLE AND EXCLUSIVE REPRESENTATIONS, WARRANTIES, AND OBLIGATIONS WITH RESPECT TO ITS PRODUCTS, AND APPLIED BIOSYSTEMS MAKES NO OTHER WARRANTY OF ANY KIND WHATSOEVER, EXPRESSED OR IMPLIED, INCLUDING WITHOUT LIMITATION, WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHETHER ARISING FROM A STATUTE OR OTHERWISE IN LAW OR FROM A COURSE OF DEALING OR USAGE OF TRADE, ALL OF WHICH ARE EXPRESSLY DISCLAIMED.

Warranty Limitations

THE REMEDIES PROVIDED HEREIN ARE THE BUYER'S SOLE AND EXCLUSIVE REMEDIES. WITHOUT LIMITING THE GENERALITY OF THE FOREGOING, IN NO EVENT SHALL APPLIED BIOSYSTEMS BE LIABLE, WHETHER IN CONTRACT, TORT, WARRANTY, OR UNDER ANY STATUTE (INCLUDING WITHOUT LIMITATION ANY TRADE PRACTICE, UNFAIR COMPETITION, OR OTHER STATUTE OF SIMILAR IMPORT) OR ON ANY OTHER BASIS, FOR DIRECT, INDIRECT, PUNITIVE, INCIDENTAL, MULTIPLE, CONSEQUENTIAL, OR SPECIAL DAMAGES SUSTAINED BY THE BUYER OR ANY OTHER PERSON OR ENTITY, WHETHER OR NOT FORESEEABLE AND WHETHER OR NOT APPLIED BIOSYSTEMS IS ADVISED OF THE POSSIBILITY OF SUCH DAMAGES, INCLUDING WITHOUT LIMITATION, DAMAGES ARISING FROM OR RELATED TO LOSS OF USE, LOSS OF DATA, FAILURE OR INTERRUPTION IN THE OPERATION OF ANY EQUIPMENT OR SOFTWARE, DELAY IN REPAIR OR REPLACEMENT, OR FOR LOSS OF REVENUE OR PROFITS, LOSS OF GOOD WILL, LOSS OF BUSINESS, OR OTHER FINANCIAL LOSS OR PERSONAL INJURY OR PROPERTY DAMAGE.

NO AGENT, EMPLOYEE, OR REPRESENTATIVE OF APPLIED BIOSYSTEMS HAS ANY AUTHORITY TO MODIFY THE TERMS OF THIS LIMITED WARRANTY STATEMENT OR TO BIND APPLIED BIOSYSTEMS TO ANY AFFIRMATION, REPRESENTATION, OR WARRANTY CONCERNING THE PRODUCT THAT IS NOT CONTAINED IN THIS LIMITED WARRANTY STATEMENT, AND ANY SUCH MODIFICATION, AFFIRMATION, REPRESENTATION, OR WARRANTY MADE BY ANY AGENT, EMPLOYEE, OR REPRESENTATIVE OF APPLIED BIOSYSTEMS WILL NOT BE BINDING ON APPLIED BIOSYSTEMS UNLESS IN A WRITING SIGNED BY AN EXECUTIVE OFFICER OF APPLIED BIOSYSTEMS.

THIS WARRANTY IS LIMITED TO THE BUYER OF THE PRODUCT FROM APPLIED BIOSYSTEMS AND IS NOT TRANSFERABLE.

Some countries or jurisdictions limit the scope of or preclude limitations or exclusion of warranties, of liability, such as liability for gross negligence or willful misconduct, or of remedies or damages, as or to the extent set forth above. In such countries and jurisdictions, the limitation or exclusion of warranties, liability, remedies or damages set forth above shall apply to the fullest extent permitted by law, and shall not apply to the extent prohibited by law.

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 two types of basecallers:

KB – new algorithm that calculates mixed or pure bases, and sample quality values.

ABI – 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.

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 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: *drive letter*:\AppliedBiosystems\SeqA5.1\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” and “original data.”

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 IUB codes (see "IUB code"), 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.

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" and "sample files."

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.

views Various displays provided in the Sample window.

Index

Symbols

A

B

C

D

E

F

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W

Z

Headquarters
850 Lincoln Centre Drive
Foster City, CA 94404 USA
Phone: +1 650.638.5800
Toll Free (In North America): +1 800.345.5224
Fax: +1 650.638.5884

Worldwide Sales and Support
Applied Biosystems vast distribution and service network, composed of highly trained support and applications personnel, reaches 150 countries on six continents. For sales office locations and technical support, please call our local office or refer to our Web site at www.appliedbiosystems.com.

www.appliedbiosystems.com

Applied Biosystems

Applera Corporation is committed to providing the world's leading technology and information for life scientists. Applera Corporation consists of the Applied Biosystems and Celera Genomics businesses.

Printed in USA, 9/2003
Part Number 4346366 Rev. B

an **Applera** business