DNA Sequencing Analysis Software

Version 3.4

User's Manual

Applied Biosystems

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

Notice to Purchaser: Limited License

For Software License and Warranty information, please see Appendix G of this manual.

Notice to Purchase: 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, impliedly, or by estoppel.

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.

Contents

4 Working with the Gel File

5 Processing Sample Files

6 The Processing Parameters

7 Viewing and Editing Sample Files

A Command Reference

B Input and Output Files

C Troubleshooting Sequencing Analysis

D Troubleshooting the BioLIMS Database

E Creating Instrument Files

F AppleScripting

About This User's Manual 1

Introduction

In This Chapter This chapter provides a general introduction to the ABI PRISM® DNA Sequencing Analysis Software. It gives information about the organization of this manual and instructions on how to get help from Applied Biosystems.

Note Already familiar with previous versions of Sequencing Analysis software and want to know what is new and different in this version? Turn to "New Features in Sequencing Analysis Since Version 3.0" on page 1-5.

Manual Contents

Overview of Contents

The following table describes the contents of this manual.

Chapter Contents

Chapter	Content
1	"About This User's Manual" gives an overview of Sequencing Analysis and related software and tells how to obtain technical help.
2	"Getting Started" describes ♦ The Sequencing Analysis package contents ♦ System requirements ♦ How to install and set up the Sequencing Analysis software ♦ How to exit from the program
3	"Sequencing Analysis and BioLIMS" describes how to set up and use the Sequencing Analysis software to read and write sequence data to a BioLIMS® database.
4	"Working with the Gel File" explains how to ♦ View and edit a gel file in the Gel File window ♦ Generate sample files from a gel file. ♦ Check fifth-dye codes
5	"Processing Sample Files" explains how to ♦ Open and close the Sample Manager window ♦ Add and remove sample files ♦ Change the processing parameters ♦ Submit the list of sample files for processing.
6	"The Processing Parameters" explains the processing parameters and how to decide which parameter values will yield the best results.
7	"Viewing and Editing Sample Files" describes ♦ The six views of the Sample window ♦ Ways to change the appearance of the window ♦ How to search and edit the sequence ♦ How to save the changes made

Chapter Contents *(continued)*

Chapter	Content
Appendix A	"Command Reference" includes brief descriptions of the Sequencing Analysis main menu commands and cross references to other sections that provide more detail.
Appendix B	"Input and Output Files" describes the files created and used by the Sequencing Analysis software.
Appendix D	"Troubleshooting the BioLIMS Database" gives guidance on troubleshooting client connection to a BioLIMS database.
Appendix C	"Troubleshooting Sequencing Analysis" describes Sequencing Analysis error messages and other problems, and what to do about each.
Appendix E	"Creating Instrument Files" explains how to make and change instrument files.
Appendix F	"AppleScripting" lists the AppleScript® commands available in the Sequencing Analysis program and lists the sample scripts included with the software.
Appendix G	"Software License and Warranty" explains your rights and responsibilities regarding this software.
Glossary	The Glossary explains many terms used in this manual.
Index	The index enables you to find information in this manual.

Related Manuals

Sequencing Analysis software is part of a suite of Applied Biosystems hardware and software products that provides a complete sequencing solution.

If the information sought is not in this manual, it may be in one of the other manuals listed in the table below.

For more information about…	See…
a genetic analysis instrument, including data collection software	the user's manual for the instrument.
specific sequencing chemistry protocols, designing experiments, and preparing samples	the *ABI PRISM Automated DNA Sequencing Chemistry Guide* (P/N 4305080) or the protocols that accompany Applied Biosystems sequencing reagent kits.
using Factura™ software to identify and edit out vector and ambiguous regions of sequences	the *ABI PRISM Factura Feature Identification Software User's Manual*.

For more information about…	See…
using AutoAssembler™ software to assemble sequence fragments into contiguous sequence data	the *ABI PRISM AutoAssembler DNA Sequence Assembly Software User's Manual*.
accessing and Managing a BioLIMS database	the *ABI PRISM BioLIMS Genetic Information Management System, System Administration Manual* (P/N 4304071) and the *ABI PRISM BioLIMS Manager™ Software User's Manual* (P/N 4304070).
uploading existing sample files to a BioLIMS database and downloading database records as sample files	*ABI PRISM BioLIMS Sample2DB Software User's Manual* (P/N 4304072).

New Features in Sequencing Analysis Since Version 3.0

New in Version 3.4

Sequencing Analysis version 3.4 software has had the following enhancements since version 3.3:

- The Neural Net Tracker has been modified to improve lane assignment using a fifth dye (see below).
- The Sequencing Analysis software has been refined to improve the performance of the BioLIMS Genetic Information Management System and the Factura Feature Identification Software (see below).

Features Introduced in Version 3.3

The major new features in Sequencing Analysis from version 3.0 to 3.3 are:

- Support for a fifth dye in gel display and extraction
- Support for the BioLIMS Genetic Information Management System
- Enhanced BioLIMS Collection Browser
- New improved Neural Net Tracker
- New improved manual tracking user interface
- Basecaller consolidation
- New Basecaller algorithm for analysis of data from the ABI PRISM® 310 Genetic Analyzer
- Ninety-six lane gel capability
- Basecaller threshold removed
- Maximum number of analyzed scans is increased
- Selectable area around electropherogram base letters increased
- New weighted channel averaging (See “Use Weighted Averaging” on page 6-25)
- The DyeSet/Primer and Instrument file columns in the gel file Sample Sheet can be edited

Fifth-Dye Support for Gel Files

A fifth dye can be used in sequencing setups to facilitate gel tracking. The Neural Net Tracker uses the fifth dye to improve lane assignment.

Sequencing Analysis v. 3.4 can:

- Open and display gel files that contain five colors (fifth dye is orange)
- Extract five-dye data from the gel file into sample files or into the BioLIMS database

Note Since the fifth dye is not used for base calling, the fifth dye color is not displayed in the sample window (electropherogram) within the Sequencing Analysis program.

For more information about the fifth dye, see "When a Fifth Dye Is Used" on page 4-48.

BioLIMS and Factura Software

The Sequencing Analysis v. 3.4 executable is identical to that for v. 3.3.1. The BioLIMS v. 2.0.1 update contained a patch that would bring Sequencing Analysis v. 3.3 software up to v. 3.3.1. Sequencing Analysis v. 3.4 software eliminates the need for the BioLIMS v. 2.0.1 patch. The same is true of Factura v. 2.2 and v. 2.2.2 software.

BioLIMS Support

Sequencing Analysis v. 3.4 software supports the BioLIMS Genetic Information Management System v. 2.0. In BioLIMS mode, Sequencing Analysis extracts sample file data from gel files and writes it to the database. (The contents of the gel file are not saved in the database.) Sequencing Analysis can read and write sequence data to the database just as it writes sample files to the Macintosh® hard disks.

Both Oracle® and Sybase® databases are supported.

For more information about using Sequencing Analysis software in BioLIMS mode, see Chapter 3, "Sequencing Analysis and BioLIMS."

BioLIMS Collection Browser

Sequencing Analysis v. 3.1 was part of the BioLIMS Genetic Information Management System v. 1.0 (Sybase SQL Server™ only).

Sequencing Analysis v. 3.4 has been upgraded to support the new BioLIMS v. 2.0 system. The BioLIMS v. 2.0 system includes support for ABI PRISM® GeneScan Software and for databases on Oracle7™ Workgroup Servers.

In Sequencing Analysis v. 3.1, BioLIMS database sequence records were opened via the Sequence Chooser. In Sequencing Analysis v. 3.4, the Sequence Chooser has been replaced by the BioLIMS Collection Browser, and some new search criteria have been added.

For details about the accessing the BioLIMS database through the BioLIMS Collection Browser, see “Using the Collection Browser Window” on page 3-19.

Neural Net Tracker

This was the most significant improvement in the v. 3.2 release of Sequencing Analysis.

- The Neural Net Tracker program uses a neural net-based algorithm to automatically track gel lanes. The Neural Net Tracker has been taught how to recognize bands and how to track curved lanes.
- The Neural Net Tracker program exists as a separate program within the Sequencing Analysis software folder. Also associated with the Neural Net Tracker program are a set of Tracker settings files that have been optimized for number of lanes and comb-types.
- The Neural Net Tracker program is “headless”. This means that although it stands as a separate program file, it does not have a user interface. The Tracker program is opened automatically from within the Sequencing Analysis application.

IMPORTANT The gel file must be multicomponented using the correct instrument file in order to be auto-tracked.

Tracking times depend upon number of lanes, channels, and scans in the gel file. Consult the table below to estimate gel tracking times for your sequencing system.

Number of Lanes	Number of Channels	Run Time (hr)	Number of Scans	Gel Size (MB)	Auto Tracking Time (min) for CPU/Speed			
					7200/90[a]	4400/200[a]	9500/200[b]	G3/266[a]
36	194	3.5 or 7[c]	8400	25	27	15	11	6
48	388	6	7200	34	36	21	15	8
48	388	10	12,000	54	58	34	25	13
64	388	3.5 or 7[c]	8400	45	54	32	24	13
64	388	12	14,400	65	76	45	32	17
96	480	10	12,000	67	93	57	40	21

a. 7200, 4400, G3: 32MB + 10MB VM

b. 9500: 64MB + 10MB VM

c. 2 x (1200 scan/hr) x 7 hr = 8400 or 4x (2400 scan/hr) x 3.5 hr = 8400

New Manual Tracking UI

The new user interface (UI) for manual tracking is very easy to use. Tracking lines are adjusted by moving and adding control points. A new interpolation mode makes it possible to adjust many lanes at once.

For more information about the new manual tracking user interface, see Chapter 4, “Working with the Gel File.”

Basecaller Consolidation

In Sequencing Analysis v. 3.0, each of the Basecallers existed as a separate program contained in the Basecallers folder.

In Sequencing Analysis v. 3.1, base-calling speed was slightly improved by consolidating the Basecallers into a single base-calling program. The same base-calling algorithms are available in the consolidated Basecaller and are selected in the Sample Manager window as in previous versions.

ABI-CE2 Basecaller for 310 Analysis

New in Sequencing Analysis v. 3.1 was the ABI-CE2 Basecaller for analysis of sequencing data collected on the ABI PRISM 310 Genetic Analyzer. Use this Basecaller to analyze sample data obtained using rhodamine dye terminator chemistries and the POP-6™ polymer.

For more information about when to use the ABI-CE2 basecaller, see “Choosing a Basecaller” and “The ABI Basecallers” on page 6-46.

Note Many ABI PRISM 310 Genetic Analyzer users have already received the ABI-CE2 basecaller separately from the Sequencing Analysis application.

96 Lane Capability

In Sequencing Analysis v. 3.4, 96-lane gel files can be opened, viewed, tracked, and extracted like any other gel.

An optimized Tracker settings file for 96 lane gels is included with Sequencing Analysis v. 3.4. For more information, see the *ABI PRISM 377 DNA Sequencer 96-Lane Upgrade User's Manual* (P/N 4305423).

Basecaller Threshold Removed

The Basecaller in Sequencing Analysis v. 3.4 can read signals of lower intensity.

Before version 3.2, the Sequencing Analysis Basecallers contained a fixed signal cutoff value. Any signal intensity below this preset value caused the Basecaller program to fail. This threshold cutoff was removed from the Basecaller in Sequencing Analysis v. 3.2.

How Will the New Basecaller Effect Data Processing?

The error “signal too weak” no longer appears. This error was most problematic for those using the dRhodamine Terminator DNA sequencing chemistry, which has weaker signal intensity than the other chemistries. With the new Basecaller program, all data is analyzed, regardless of the signal intensity. Due to this, you may want to take more care setting the analysis endpoints. (For more information, see pages 6-16 and 6-28.)

Maximum Number of Scans Increased

Before Sequencing Analysis v. 3.2, the Basecaller maximum number of scans for analyzed data was set to 20,000.

In the new Basecaller, the maximum number of scans for analyzed data has been increased to 32,000. (The raw scan limit remains at 20,000.)

How Will the New Scan Maximum Effect Data Processing?

The error “data too long” should appear less frequently. This problem was most likely to occur with long read formulation gels where the number of analyzed scans could often exceed the 20,000 limit when the base spacing estimate was low, or if the run time was too long.

How Can Analyzed Scans Exceed Raw Scans?

The reason there may be more than 20,000 scans in the analyzed data when there are less than 20,000 raw scans is due to respacing. As part of the base-calling algorithm, raw data is respaced so that the analyzed data will have an average spacing of 12 points, peak-to-peak, throughout the run. If the raw data spacing is less than 12, the basecaller will interpolate, adding more points between peaks as necessary.

Easier Selection of Bases for Editing

In the electropherogram view, the hot spot (selectable area) around the base letters is larger. This makes it easier than before to select bases by clicking on them.

More Gel File Sample Sheet Columns Editable

The DyeSet/Primer and Instrument file columns in the gel file Sample Sheet can be edited. If the wrong file was chosen at data collection, correct it by choosing a new instrument or DyeSet/Primer file in the gel file Sample Sheet.

Sequencing Analysis Software Applies to Three Instruments

Three Sequencing Instruments

The Sequencing Analysis software described in this manual can be used to analyze raw sequencing data collected from the three instruments described in the table below.

Name	This instrument…
ABI PRISM 310 Genetic Analyzer	analyzes one sample at a time using capillary electrophoresis technology. This instrument provides high resolution for short fragments and uses a minimal amount of sample.
ABI 373™ DNA Sequencer (including XL)	performs slab gel electrophoresis, allowing the user to analyze multiple samples on a gel.
ABI PRISM® 377 DNA Sequencer (including XL and 96-Lane Upgrades)	is a high throughput slab-gel electrophoresis instrument, created to meet the needs of high volume DNA sequencing or genetic analysis laboratories. Throughput is more than four times that of the ABI 373.

Margin Notation

Although most of the information in this manual applies to all three instruments, certain parts apply to only one or two of the instruments. Throughout the manual, a notation appears in the left margin when the text applies to only one or two instruments. The instrument or instruments to which the text does not apply are crossed out (⋈). The notation that appears here to the left would indicate that the text applies to the ABI 373 and ABI PRISM 377 instruments, but not to the ABI PRISM 310 instrument.

What Does Sequencing Analysis Software Do?

Introduction

The Sequencing Analysis software can carry out several analysis steps on the data from genetic analysis instruments. These steps can be done manually in the Sample File Manager, or they can be done as part of a fully automated operation. The automated analysis begins with the start of data collection and ends when the data has been analyzed by the Sequencing Analysis software. Additional automatic processing can be carried out using the Factura Feature Identification software. Also, as part of the automatic operation, results can be printed.

For 373 and 377 Instrument Data

~~310~~
373
377

For samples run on ABI 373 and ABI PRISM 377 instruments, the Sequencing Analysis program tracks the gel file:

- Finds the starting position of each lane in the gel file.
- Creates a sample file for each lane marked as used, then transfers the basic sample information (name, run date, etc.) from the sample sheet in the gel file to the sample files.
- Tracks the lanes and transfers the raw data for each lane to the appropriate sample file or BioLIMS database record.

IMPORTANT Do not use Sequencing Analysis v. 3.4 to track a gel file during instrument data collection. Wait until data collection is finished before tracking any gel. Alternatively, load and run Sequencing Analysis v. 3.4 on a separate computer that does not run Data Collection software.

Retracking Gel Data

After analyzing ABI 373 or ABI PRISM 377 data, you can inspect the gel and retrack a lane manually, or specify processing parameters and reanalyze the data for a given lane. This may allow you to salvage an otherwise unusable sample in the case of a chemistry or gel problem.

For 310 Instrument Data

Because each sample on an ABI PRISM 310 instrument is run individually, the data collection program creates the sample files automatically when the samples are run.

310
~~373~~
~~377~~

If data from the ABI PRISM 310 instrument is to be added to a BioLIMS database, this should be done with the BioLIMS Sample2DB program.

Analysis and Printing

Once the sample files are available, the Sequencing Analysis program can:

- Create analyzed data (based on the raw data) in which the bases in the sequence are identified.
- Pass the analyzed sample files to the ABI PRISM® Factura Feature Identification Software for further processing. The Factura software package is included with each copy of Sequencing Analysis software. For example, Factura is used to identify and remove vector sequence and ambiguous regions of sequence. For more information about Factura, see *ABI PRISM Factura Feature Identification Software User's Manual*.
- Print the electropherogram data for each sample file after all requested processing is finished.

Reanalysis of Base Calling

For files and database records from *all* ABI analysis instruments, the software allows re-base-calling of sample files with different settings from those used for the automatic analysis. You can also edit the base sequence in the analyzed data.

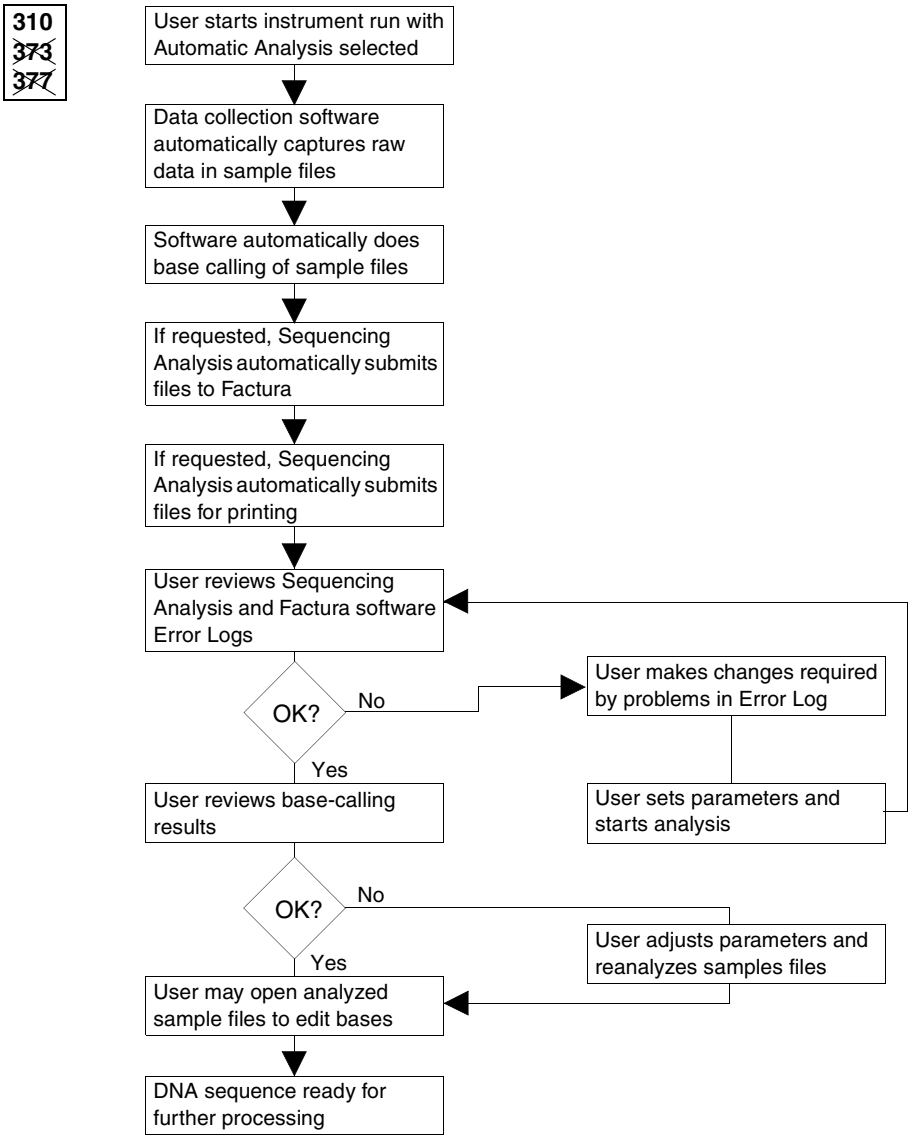
Process Flowcharts

The basic steps used in DNA sample analysis are outlined in the following two flowcharts:

- "Flowchart for ABI PRISM 310" on page 1-14
- "Flowchart for ABI 373 and ABI PRISM 377" on page 1-15

Flowchart for ABI PRISM 310

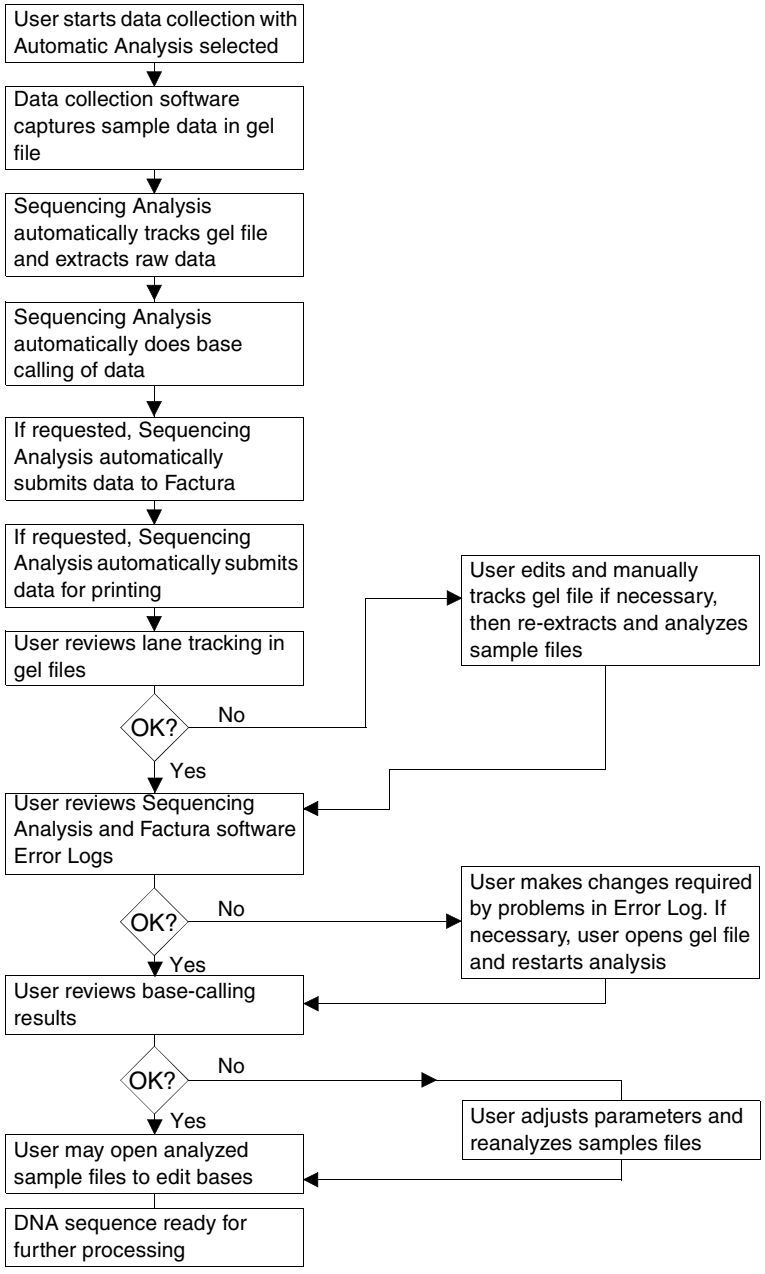
This flowchart shows the analysis procedure for samples from the ABI PRISM 310 Genetic Analyzer.



Flowchart for ABI 373 and ABI PRISM 377

~~310~~
373
377

This flowchart shows the analysis procedure for samples from ABI 373 and ABI Prism 377 Sequencers.



Data Formats

The data that results from the Sequencing Analysis process can be in formats that can be used with commercially available or user-generated programs on the Macintosh® computer or on other compatible computers.

Format	Description
Sample files	These files are written in a Applied Biosystems proprietary format. They contain complete information about the sequence: raw sequence data, basecalls, peak locations, sample information, etc.
BioLIMS database records	These records contain the same information as the sample files, but the data is written directly from the Sequencing Analysis software to a BioLIMS database. For information about using Sequencing Analysis in BioLIMS mode, see Chapter 3, “Sequencing Analysis and BioLIMS.”
Text files	Each time a sample file is created or modified, a text file is created automatically in the same folder as the sample file. By default, text file names have the extension, .Seq. In the Preferences panel (page 6-39), the format of the text file can be specified: ♦ ABI ♦ Intelligenetics ♦ Staden ♦ Wisconsin Regardless of the format chosen, text files are given the default extension “.Seq”. These text files are created automatically whenever sample files are created or updated.

Other ABI PRISM Software

Introduction

This section describes other programs you should know about:

- Utility programs included with the Sequencing Analysis software
- Programs for further processing of sequence data
- Programs for DNA fragment analysis

Other Programs in the Sequencing Analysis Package

This software package includes the following utility programs in addition to the main Sequencing Analysis program:

Program	Description
Neural Net Tracker	Performs the tracking, determining the center of the gel lanes. The Tracker program is called by the Sequencing Analysis software to process gel files from an ABI 373 or ABI PRISM 377 instrument.
Basecaller	Performs the actual base-calling operation. Once you select the Basecaller to be used, the Sequencing Analysis program automatically opens and applies that Basecaller at the appropriate analysis step.
DataUtility	Used to make matrices for instrument files, which are used with the data collection and analysis software, and to monitor noise levels during troubleshooting by Applied Biosystems technical specialists.

Other Programs for Analysis of Sequence Data

After you analyze the raw sample data with the Sequencing Analysis software, that analyzed data can be further processed in any of the following software programs:

Program	Description
Factura Feature Identification Software	Identifies specified vector and ambiguity ranges and a specified confidence range. It also identifies multiple base positions with codes described by the International Union of Biochemists (IUB codes) based on a user-defined threshold. This program is used to prepare the sequence for further analysis using only the target DNA.

Program	Description
Sequence Navigator® DNA and Protein Sequence Comparison Software	Incorporates powerful algorithms for pairwise or multiple alignment of DNA and protein sequences. (Sequence Navigator software is not compatible with the BioLIMS system.)
AutoAssembler DNA Sequence Assembly Software	Allows assembly of small pieces of DNA into larger contiguous segments of DNA, using ABI PRISM genetic analysis instrument data as well as other sequence text files.
EditView	A free, DNA sequence viewer that allows viewing, editing, and printing of sequence data from an ABI PRISM 373, 377, or 310 Genetic Analyzer. Using EditView on a Macintosh computer, you can open an analyzed sample file and view the sequence data either as an electropherogram (traces), or in text format. You can then edit individual bases, export the data to a text file, or print it. EditView is available on the Applied Biosystems web site at www.appliedbiosystems.com/techsupport. (EditView for the Macintosh computer is not compatible with the BioLIMS system.)
Primer Express®	A primer design program with an easy-to-use interface. The software is applications oriented, taking into consideration the most updated criteria for primer design.

Programs for BioLIMS Database Access

If using the BioLIMS database for storage and retrieval of sequence data, these two programs are of use to you:

Program	Description
BioLIMS Manager	A program for searching the BioLIMS database, grouping database records into relevant collections, and deleting sequence records from the database.
Sample2DB	A utility program for uploading data from sample files into the database and for downloading sequence records as sample files.

Programs for Fragment Analysis

In addition to sequencing applications, you can perform sizing and quantifying applications with the genetic analysis instruments. To do so, you must use the GeneScan® Analysis Software instead of the Sequencing Analysis software.

For further analysis of GeneScan data, use the Genotyper® Fragment Analysis Software, which converts data from GeneScan results files into the format required by downstream applications such as linkage analysis programs, databases, or spreadsheets.

Technical Support

Contacting Technical Support

You can contact Applied Biosystems for technical support by telephone or fax, by e-mail, or through the Internet. You can order Applied Biosystems user documents, MSDSs, certificates of analysis, and other related documents 24 hours a day. In addition, you can download documents in PDF format from the Applied Biosystems Web site (please see the section “To Obtain Documents on Demand” following the telephone information below).

To Contact Technical Support by E-Mail

Contact technical support by e-mail for help in the following product areas:

Product Area	E-mail address
Genetic Analysis (DNA Sequencing)	galab@appliedbiosystems.com
Sequence Detection Systems and PCR	pcrlab@appliedbiosystems.com
Protein Sequencing, Peptide and DNA Synthesis	corelab@appliedbiosystems.com
Biochromatography, PerSeptive DNA, PNA and Peptide Synthesis systems, CytoFluor®, FMAT™, Voyager™, and Mariner™ Mass Spectrometers	tsupport@appliedbiosystems.com
Applied Biosystems/MDS Sciex	api3-support@sciex.com
Chemiluminescence (Tropix)	tropix@appliedbiosystems.com

Hours for Telephone Technical Support

In the United States and Canada, technical support is available at the following times:

Product	Hours
Chemiluminescence	8:30 a.m. to 5:30 p.m. Eastern Time
Framingham support	8:00 a.m. to 6:00 p.m. Eastern Time
All Other Products	5:30 a.m. to 5:00 p.m. Pacific Time

To Contact Technical Support by Telephone or Fax

In North America

To contact Applied Biosystems Technical Support, use the telephone or fax numbers given below. (To open a service call for other support needs, or in case of an emergency, dial **1-800-831-6844** and press **1**.)

Product or Product Area	Telephone Dial...	Fax Dial...
ABI PRISM® 3700 DNA Analyzer	**1-800-831-6844,** then press **8**	**1-650-638-5981**
DNA Synthesis	**1-800-831-6844**, then press **21**	**1-650-638-5981**
Fluorescent DNA Sequencing	**1-800-831-6844**, then press **22**	**1-650-638-5981**
Fluorescent Fragment Analysis (includes GeneScan® applications)	**1-800-831-6844**, then press **23**	**1-650-638-5981**
Integrated Thermal Cyclers (ABI PRISM®877 and Catalyst 800 instruments)	**1-800-831-6844**, then press **24**	**1-650-638-5981**
ABI PRISM® 3100 Genetic Analyzer	**1-800-831-6844**, then press **26**	**1-650-638-5981**
BioInformatics (includes BioLIMS™, BioMerge™, and SQL GT™ applications)	**1-800-831-6844**, then press **25**	**1-505-982-7690**
Peptide Synthesis (433 and 43X Systems)	**1-800-831-6844**, then press **31**	**1-650-638-5981**
Protein Sequencing (Procise® Protein Sequencing Systems)	**1-800-831-6844**, then press **32**	**1-650-638-5981**
PCR and Sequence Detection	**1-800-762-4001**, then press **1** for PCR, **2** for the 7700 or 5700, **6** for the 6700 or dial **1-800-831-6844**, then press **5**	**1-240-453-4613**

Product or Product Area	Telephone Dial...	Fax Dial...
Voyager™ MALDI-TOF Biospectrometry and Mariner™ ESI-TOF Mass Spectrometry Workstations	**1-800-899-5858**, then press **13**	**1-508-383-7855**
Biochromatography (BioCAD® Workstations and Poros® Perfusion Chromatography Products)	**1-800-899-5858**, then press **14**	**1-508-383-7855**
Expedite™ Nucleic acid Synthesis Systems	**1-800-899-5858**, then press **15**	**1-508-383-7855**
Peptide Synthesis (Pioneer™ and 9050 Plus Peptide Synthesizers)	**1-800-899-5858**, then press **15**	**1-508-383-7855**
PNA Custom and Synthesis	**1-800-899-5858,** then press **15**	**1-508-383-7855**
FMAT™ 8100 HTS System and Cytofluor® 4000 Fluorescence Plate Reader	**1-800-899-5858**, then press **16**	**1-508-383-7855**
Chemiluminescence (Tropix)	**1-800-542-2369** (U.S. only), or **1-781-271-0045**	**1-781-275-8581**
Applied Biosystems/MDS Sciex	**1-800-952-4716**	**1-650-638-6223**

Outside North America

Region	Telephone Dial...	Fax Dial...
Africa and the Middle East		
Africa (English Speaking) and West Asia (Fairlands, South Africa)	**27 11 478 0411**	**27 11 478 0349**
South Africa (Johannesburg)	**27 11 478 0411**	**27 11 478 0349**
Middle Eastern Countries and North Africa (Monza, Italia)	**39 (0)39 8389 481**	**39 (0)39 8389 493**

Region	Telephone Dial...	Fax Dial...
Eastern Asia, China, Oceania		
Australia (Scoresby, Victoria)	**61 3 9730 8600**	**61 3 9730 8799**
China (Beijing)	**86 10 64106608**	**86 10 64106617**
Hong Kong	**852 2756 6928**	**852 2756 6968**
Korea (Seoul)	**82 2 593 6470/6471**	**82 2 593 6472**
Malaysia (Petaling Jaya)	**60 3 758 8268**	**60 3 754 9043**
Singapore	**65 896 2168**	**65 896 2147**
Taiwan (Taipei Hsien)	**886 2 2358 2838**	**886 2 2358 2839**
Thailand (Bangkok)	**66 2 719 6405**	**66 2 319 9788**
Europe		
Austria (Wien)	**43 (0)1 867 35 75 0**	**43 (0)1 867 35 75 11**
Belgium	**32 (0)2 712 5555**	**32 (0)2 712 5516**
Czech Republic and Slovakia (Praha)	**420 2 61 222 164**	**420 2 61 222 168**
Denmark (Naerum)	**45 45 58 60 00**	**45 45 58 60 01**
Finland (Espoo)	**358 (0)9 251 24 250**	**358 (0)9 251 24 243**
France (Paris)	**33 (0)1 69 59 85 85**	**33 (0)1 69 59 85 00**
Germany (Weiterstadt)	**49 (0) 6150 101 0**	**49 (0) 6150 101 101**
Hungary (Budapest)	**36 (0)1 270 8398**	**36 (0)1 270 8288**
Italy (Milano)	**39 (0)39 83891**	**39 (0)39 838 9492**
Norway (Oslo)	**47 23 12 06 05**	**47 23 12 05 75**
Poland, Lithuania, Latvia, and Estonia (Warszawa)	**48 (22) 866 40 10**	**48 (22) 866 40 20**
Portugal (Lisboa)	**351 (0)22 605 33 14**	**351 (0)22 605 33 15**
Russia (Moskva)	**7 095 935 8888**	**7 095 564 8787**
South East Europe (Zagreb, Croatia)	**385 1 34 91 927**	**385 1 34 91 840**
Spain (Tres Cantos)	**34 (0)91 806 1210**	**34 (0)91 806 1206**
Sweden (Stockholm)	**46 (0)8 619 4400**	**46 (0)8 619 4401**
Switzerland (Rotkreuz)	**41 (0)41 799 7777**	**41 (0)41 790 0676**
The Netherlands (Nieuwerkerk a/d IJssel)	**31 (0)180 331400**	**31 (0)180 331409**

Region	Telephone Dial...	Fax Dial...
United Kingdom (Warrington, Cheshire)	**44 (0)1925 825650**	**44 (0)1925 282502**
All other countries not listed (Warrington, UK)	**44 (0)1925 282481**	**44 (0)1925 282509**
Japan		
Japan (Hacchobori, Chuo-Ku, Tokyo)	**81 3 5566 6230**	**81 3 5566 6507**
Latin America		
Del.A. Obregon, Mexico	**305-670-4350**	**305-670-4349**

To Reach Technical Support Through the Internet

We strongly encourage you to visit our Web site for answers to frequently asked questions and for more information about our products. You can also order technical documents or an index of available documents and have them faxed or e-mailed to you through our site. The Applied Biosystems Web site address is

http://www.appliedbiosystems.com/techsupp

To submit technical questions from North America or Europe:

Step	Action
1	Access the Applied Biosystems Technical Support Web site.
2	Under the **Troubleshooting** heading, click **Support Request Forms**, then select the relevant support region for the product area of interest.
3	Enter the requested information and your question in the displayed form, then click **Ask Us RIGHT NOW** (blue button with yellow text).
4	Enter the required information in the next form (if you have not already done so), then click **Ask Us RIGHT NOW**. You will receive an e-mail reply to your question from one of our technical experts within 24 to 48 hours.

To Obtain Documents on Demand

Free, 24-hour access to Applied Biosystems technical documents, including MSDSs, is available by fax or e-mail or by download from our Web site.

To order documents...	Then...
by index number	a. Access the Applied Biosystems Technical Support Web site at **http://www.appliedbiosystems.com/techsupp** b. Click the **Index** link for the document type you want, then find the document you want and record the index number. c. Use the index number when requesting documents following the procedures below.
by phone for fax delivery	a. From the U.S. or Canada, call **1-800-487-6809,** or from outside the U.S. and Canada, call **1-858-712-0317**. b. Follow the voice instructions to order the documents you want. **Note** There is a limit of five documents per request.
through the Internet for fax or e-mail delivery	a. Access the Applied Biosystems Technical Support Web site at **http://www.appliedbiosystems.com/techsupp** b. Under **Resource Libraries**, click the type of document you want. c. Enter or select the requested information in the displayed form, then click **Search**. d. In the displayed search results, select a check box for the method of delivery for each document that matches your criteria, then click **Deliver Selected Documents Now** (or click the PDF icon for the document to download it immediately). e. Fill in the information form (if you have not previously done so), then click **Deliver Selected Documents Now** to submit your order. **Note** There is a limit of five documents per request for fax delivery but no limit on the number of documents you can order for e-mail delivery.

Getting Started 2

Overview

In This Chapter

This chapter contains information about installing and registering the Sequencing Analysis software. This chapter also describes how to customize the program preference settings.

Registration and Warranty

License and Warranty

Before you begin, please read the “Software License and Warranty” in Appendix G. It explains your rights and responsibilities regarding this software.

Registering Your Software

To register your copy of the Sequencing Analysis software, fill out the registration card (included in this software package) and return it to Applied Biosystems. This enables us to send you notification of software updates and any other future information that may be specific to Sequencing Analysis owners.

IMPORTANT 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 the Macintosh® computer connected to your ABI PRISM instrument or on any other Macintosh computer that meets the minimum requirements stated below. The software can be installed on a computer used for analysis only, as well as on one used for both data collection and analysis.

If Using an ABI 373 Instrument

Sequencing Analysis v. 3.4 only runs on a PowerPC CPU. If you want to use Sequencing Analysis v. 3.4 to analyze gel files collected on a ABI 373 instrument, you must transfer the gel files from the ABI 373 instrument Macintosh to a Power Macintosh® where Sequencing Analysis v. 3.4 software is installed.

Computers Connected to ABI Instruments

Each ABI PRISM® genetic analysis instrument is shipped with a Macintosh computer. If you received this software with a newly purchased ABI PRISM genetic analysis instrument, the Sequencing Analysis software is installed by your Applied Biosystems Customer Support Engineer as part of the installation and setup of the instrument. The system requirements for that computer are described in the instrument manual.

If you are replacing an earlier version of the Sequencing Analysis software or if you are using this software for the first time, your computer MUST meet our minimum requirements to run the new Sequencing Analysis software. Our state-of-the-art software requires a PowerPC CPU and at least 32 MB of physical RAM to achieve proper performance. If you are unable to meet these requirements, you will not be able to use this software.

System Requirements and Recommendations

Below are the system requirements and recommendations for running the Sequencing Analysis v. 3.4 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.

System Requirements

System Component	Requirements
CPU	A Power Macintosh (PowerPC CPU) computer. (You will benefit from using the fastest computer available.)
CD-ROM Drive	Any
Operating System	Mac OS version 8.6 or later.
Disk Space	A minimum of 15 MB free disk space. (See also “Disk Space” under Recommendations below.)
Memory (RAM)	The minimum memory requirement is 32 MB total with at least 25 MB of this available to run the Sequencing Analysis program.
Virtual Memory	Virtual memory must be turned on if the physical RAM is less than 48 MB. Set the Memory control panel so that the memory available after restart is 45–55 MB. Using more virtual memory than required can slow software performance.

System Recommendations

System Component	Recommendations
Monitor	A 17-inch monitor or larger is recommended. Although a monitor of 640 x 480 resolution can be used, you will benefit from having a monitor of higher resolution.
Printer	A PostScript-compatible color printer is recommended, *e.g.*, Epson Stylus 900 or HP 2500–series printers.
Disk Space	Storage requirements depend primarily on the quantity of data to be generated and stored. Sample files are approximately 150–250 KB each and gel files are 20–90 MB each. It is common to store many sample files on the analysis computer. Gel files are usually stored only on the computer that is connected to the instrument and are removed or archived frequently.

The table below gives estimates of gel file size according to gel length and number of lanes.

Well-to-Read	36 Lane	48 Lane	64 Lane	96 Lane
36 cm	20 MB	45 MB	40 MB	50 MB
48 cm	30 MB	55 MB	55 MB	65 MB

Installing Sequencing Analysis

Introduction

This section describes the following:

- Before Installing (page 2-6)
- Installing the Sequencing Analysis Software (page 2-7)
- Removing Sequencing Analysis Software (page 2-9)

Before Installing

IMPORTANT This software cannot be installed on a non-PowerPC Macintosh computer.

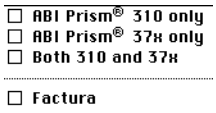
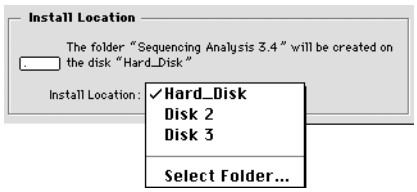
Before you begin installing the Sequencing Analysis software, please do the following:

- Check that you have at least 15 MB of free disk space to accommodate the Sequencing Analysis software.
- Backup the contents of the ABI Folder in the System Folder.
- Quit all open applications.
- Turn off any virus protection software that you may have running.
- Delete any aliases to the previous versions of the Sequencing Analysis program.
- If you are installing Sequencing Analysis v. 3.4 as an update to Sequencing Analysis v. 3.0, you need to find your v. 3.0 registration code. If you cannot find your code: note down the serial number in the splash screen of Sequencing Analysis v. 3.0 (select About Sequencing Analysis from the Apple menu) and call Technical Support (page 1-20).

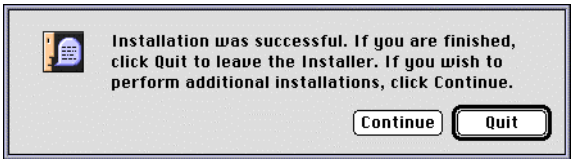
Installing the Sequencing Analysis Software

Follow these steps to install the Sequencing Analysis software.

To install Sequencing Analysis from a CD-ROM:

Step	Action
1	Insert the Sequencing Analysis 3.4 CD-ROM into the computer's CD-ROM drive.
2	Double-click the Sequencing Analysis Installer icon. 
3	When the Installer start-up screen appears, choose Continue. The About Sequencing Analysis 3.4 dialog box appears.
4	This dialog box contains important information that you should read. You may print or save the contents if you want. After you have read About Sequencing Analysis 3.4, click Continue to open the Sequencing Analysis Installer window.
5	Select the check box for the type of data you expect to analyze on this computer. You can also choose to install Factura software.  **Note** The installer loads only the required files, based on the check box you select.
6	The default installation location is the local hard disk, unless there is a BioLIMS 2.0 folder present, in which case the software is installed in the BioLIMS 2.0 folder. If you wish to specify a different installation location, use the pop-up menu. 

To install Sequencing Analysis from a CD-ROM: *(continued)*

Step	Action
7	To begin the installation, choose Install. The following dialog box appears when installation is complete. 
8	Choose Quit.
9	Drag the Sequencing Analysis Installer CD-ROM icon to the trash to eject the CD-ROM.
10	Put the Installation CD-ROM in a safe place. If you ever need to reinstall the application, you will need the Installer CD-ROM. You *cannot* successfully install the Sequencing Analysis software by copying it from one computer to another.
11	The Installer Log File is created by the Installer. The log file is placed in the Sequencing Analysis folder and contains a list of all the files installed. Use this log file if you need to remove Sequencing Analysis from your hard disk. (See “Removing Sequencing Analysis Software” below.)

After Installing the Sequencing Analysis

After installing the Sequencing Analysis software, you should follow the steps in the table below to:

- Rebuild the desktop
- Check your ABI Folder
- Enable any virus protection software you turned off before the installation

To rebuild the desktop, enable virus protection, and check ABI Folder:

Step	Action
1	If you disabled your virus protection before installation, enable it now.

To rebuild the desktop, enable virus protection, and check ABI Folder:

Step	Action
2	While holding down the Command and Option keys, choose Restart from the Special menu in the Finder. Continue to hold down these keys as the computer reboots. Depending on how your system is configured, you may be prompted to respond to various system requests. Respond to each of these prompts appropriately, but do not release pressure on the Command and Option keys until a dialog box appears that asks if you want to rebuild the desktop.
3	Choose OK to rebuild the desktop. When this process is complete, your usual desktop screen will appear.
4	Check the contents of the ABI Folder in the System Folder to be sure that all the DyeSet/Primer and Instrument files that you use are present. If any are missing, copy them from the old ABI Folder that you backed up before installing (page 2-6).

Removing Sequencing Analysis Software

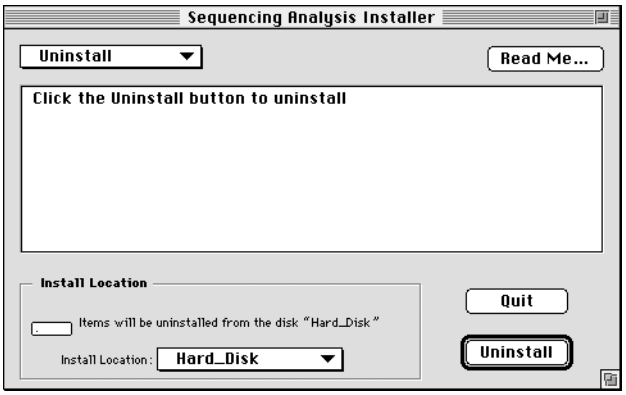
This section describes how to remove the Sequencing Analysis v. 3.4 software from your Macintosh computer. The Uninstall process deletes all folders and files installed by the Sequencing Analysis Installer.

Note If you have moved Sequencing Analysis or Factura files or folders from their original installed locations, they may not be found and deleted by the uninstall operation. Also any files that have been added to the application folders, such as those created when the applications are run, are not deleted by the uninstall operation.

To remove installed Sequencing Analysis software:

Step	Action
1	Follow steps 1–4 in the procedure “Installing the Sequencing Analysis Software” on page 2-7 to open the Installer window.

To remove installed Sequencing Analysis software: *(continued)*

Step	Action
2	Select Uninstall from the pop-up menu at the top left of the window. 
3	Choose the Select Folder item on the Install Location pop-up menu. A Macintosh browser box appears.
4	Use the browser box to locate the folder that contains the Sequencing Analysis and Factura folders.
5	Click the Uninstall button to begin the removal of the files from your disk.
6	At the conclusion of the remove operation, an alert box appears with the message whether or not the remove was successful. **Note** If files have been moved or added to the Sequencing Analysis, Factura, or ABI folders, the remove operation will be reported as unsuccessful; you should then examine and delete the remaining files yourself.

Setting Up the Sequencing Analysis Program

Introduction

After you install the Sequencing Analysis software, you can set up the program for your site. Setting up includes the following operations:

- Starting the Sequencing Analysis program for the first time and entering the requested information. See below.
- Selecting Program Preference settings that are suitable for your site. See “Selecting Processing Preferences” on page 2-17.

Before You Begin

Sequencing Analysis software is easier to use if you make the following adjustment: Choose Control Panels from the Apple menu, then choose General Controls. In the Documents box, select “Last folder used in the Application.”

About This Procedure

Three things are done in the procedure below:

- Page setup and printer defaults are specified. When printing from the Sample Manager window, the Sequencing Analysis software does not ask for page setup and printer information. It uses the values you select in the procedure.
- The program is set for use with the ABI PRISM 310 only, or for use with all instruments.
- The registration code is entered. The registration code is on the Software License and Limited Product Warranty card that comes with the product. If you are upgrading from Sequencing Analysis v. 3.0, 3.1, or 3.2 enter the registration code that you received with your original software.

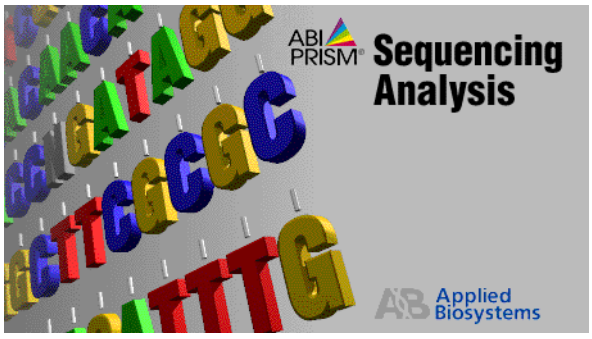
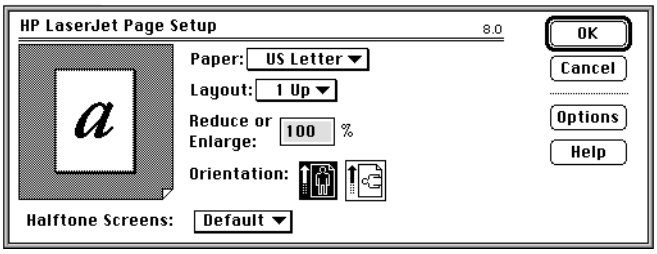
  Each installed Sequencing Analysis program must have a different registration code. Every time the Sequencing Analysis software is started, it searches the network for any other copy with the same registration code. If another copy is found, the program will not start. Because of how the registration code is stored, the Sequencing Analysis software will *not* run from a locked disk, a CD-ROM, or a read-only network volume.

Starting the Program for the First Time

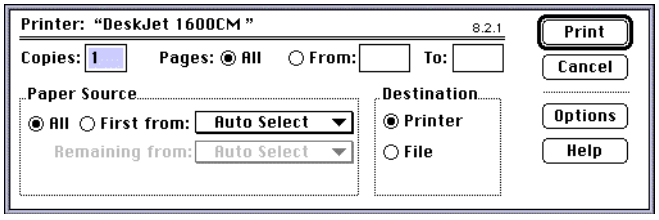
Follow this procedure when you start the Sequencing Analysis program the first time or when you start the program after moving your Seq Analysis v3.4 Prefs file from the System Folder.

Note When printing from the Sample Manager window, the Sequencing Analysis software does not ask for page setup and printer information. It uses the values you select in this procedure.

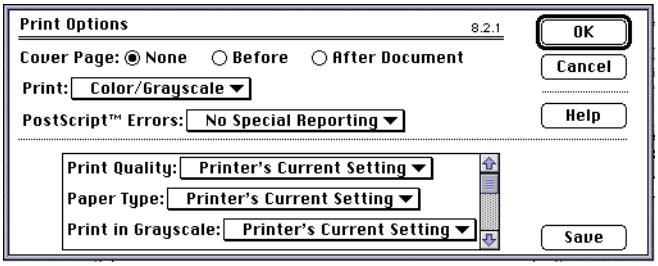
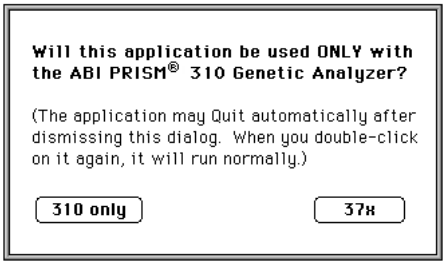
To start Sequencing Analysis for the first time:

Step	Action
1	Before opening the Sequencing Analysis application for the first time: a. Open the Chooser. b. Select the printer you expect to use for Sequencing Analysis data. c. Close the Chooser. The dialog boxes that appear when you first start the Sequencing Analysis program are specific to the selected printer.
2	Double-click on the Sequencing Analysis icon to open the program. The Sequencing Analysis start-up screen appears.  The Page Setup dialog box appears in front of the start-up screen. 

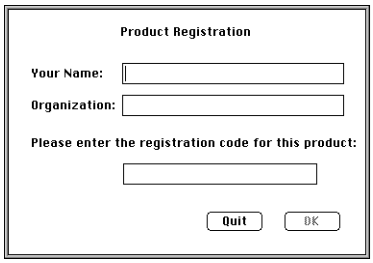
To start Sequencing Analysis for the first time: *(continued)*

Step	Action
3	Select the page setup that you want to use when printing is done as a part of automated sample file processing. The exact contents of the Page Setup dialog box depend on your printer. Normally, you should select your standard paper size and Landscape orientation. If necessary, you can change these settings later in the Preferences dialog box.
4	Choose OK to save the selected page setup to the Seq Analysis v3.4 Prefs file and close the Page Setup dialog box. **IMPORTANT** Each time the printer selected in the Apple Chooser is changed, you must open the Page Setup dialog box to reestablish the default selection. When you close the Page Setup dialog box, the Sequencing Analysis software opens the Printer dialog box. 

To start Sequencing Analysis for the first time: *(continued)*

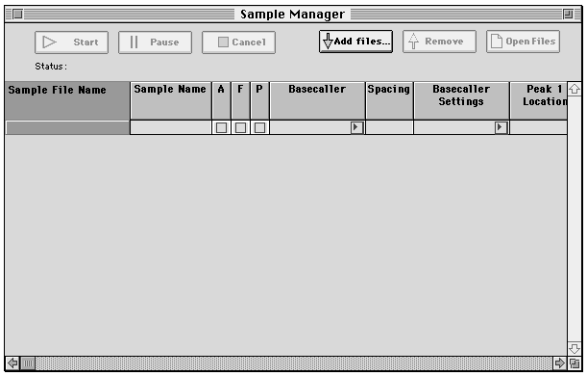
Step	Action
5	Select the printer settings that you want to use when printing is done as a part of automated sample file processing. The exact contents of this dialog box depend on your printer. To print multiple copies of each page, enter the number of copies desired in the Copies field. To ensure base letters on the electropherogram print in color: ♦ Click the Options button to open the Printer Options dialog box.  ♦ Select Color/Grayscale for the Print option, then choose OK to close the Print Options dialog box.
6	Click on the Print button to close the dialog box and save the selected values to the Seq Analysis v3.4 Prefs file. When you choose Print, the settings are only saved to the Pref file. *No printing occurs at this time*. When you close the Printer dialog box, the "310 Only?" dialog box appears. 
7	Consult the If/Then table below to decide which option to choose. The program may quit when the dialog box closes.

To start Sequencing Analysis for the first time: *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td></td><td>
<table>
<tr><th>If…</th><th>Then…</th><th>Result…</th></tr>
<tr><td>on step 4 on page 2-7, you selected to install “ABI Prism® 310 Only”</td><td>choose the 310 Only button in the dialog box above.</td><td>The menu commands and dialog boxes for ABI 373 and ABI PRISM® 377 gel files are removed, and the amount of memory allotted for the program is reduced.</td></tr>
<tr><td>on step 4 on page 2-7, you selected to install either “ABI Prism® 37X only” or “Both 310 and 37x”</td><td>choose the 37x button in the dialog box above.</td><td>All parts of the Sequencing Analysis program are kept.</td></tr>
</table>
</td></tr>
<tr><td>8</td><td>If the program quits, double-click the Sequencing Analysis icon to restart the program and proceed to the next step.</td></tr>
<tr><td>9</td><td>When the following Product Registration dialog box appears, type your registration information into the three fields.
 

This dialog box appears the first time you start the Sequencing Analysis program, and any time you move the program to a different hard disk or partition.</td></tr>
</table>

To start Sequencing Analysis for the first time: *(continued)*

Step	Action
10	Choose OK to save the registration information and close the dialog box. When the Registration dialog box closes, the Sample Manager window appears.  The Sample Manager window is used to specify which sample files to process, and the parameter values to be used during processing. For now, you can ignore this window while you view and adjust the program preference settings, as described in the Selecting Processing Preferences section.

Selecting Processing Preferences

Introduction

Preferred values can be selected for most of the processing parameters used by the Sequencing Analysis software. These values will be used for all fully automated operations (for example, when processing a batch of samples overnight). When performing manual operations (for example, re-analyzing selected files with the Sequencing Analysis software), you can either use the existing preference values, select new preference values, or select temporary values for individual files.

The Preferences

These preferences are grouped into categories on the pages of the Preferences dialog box:

- Gel Preferences (377 and 373 only)
- Basecaller Settings
- Sample Manager Defaults
- Printing Preferences
- Sequence File Formats
- Factura Preferences
- Base Letters Style

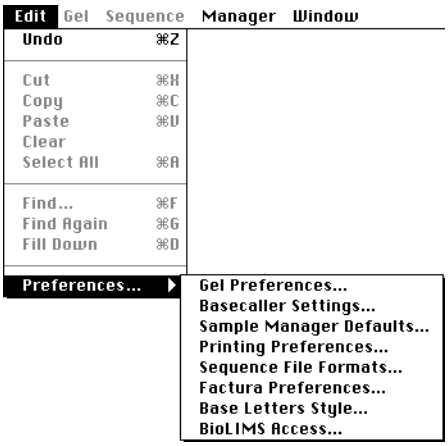
When in Doubt, Use Default Values

The default preference values are those most commonly used by Applied Biosystems customers. If you are a new user or if you are uncertain about how to set a preference, use the default setting.

The preferences are stored in the Seq Analysis v3.4 Prefs file, which is located in the Preferences folder in the System Folder. To return all the preferences to the original installation defaults, delete the Seq Analysis v3.4 Prefs file.

Before Using Sequencing Analysis

Before using Sequencing Analysis software, review and edit (if necessary) the processing parameters preferences shown in the Preferences dialog box, which is accessed from the Preferences… submenu of the Edit menu.



Note If you do not have BioLIMS® software installed, you will not see the BioLIMS Access… item at the bottom of the menu.

Please read pages 6-21 to 6-44 for detailed information about the processing parameters and how to change them.

When Do Preferences Take Effect?

You can open the Preferences dialog box and change these processing preference values at any time. The new values take effect when you close the Preferences dialog box.

Sequencing Analysis and BioLIMS Database 3

Overview

In This Chapter This chapter explains how to set up and use Sequencing Analysis software in conjunction with the BioLIMS® database.

Note If you do not have the BioLIMS Genetic Information Management System v. 2.0, you should skip this chapter entirely.

About Sequencing Analysis and BioLIMS

What is BioLIMS?

The BioLIMS Genetic Information Management System provides a relational database for storage and retrieval of DNA sequence and fragment data. In addition to the database itself, BioLIMS contains a set of software applications and tools for:

- Analyzing and processing DNA sequence and fragment data
- Maintaining and interacting with the database

The BioLIMS database resides on a UNIX workstation. The client applications run on Power Macintosh® computers and/or on UNIX workstations.

Modes in Sequencing Analysis

If you have obtained Sequencing Analysis software as part of the BioLIMS Client Package, you can use Sequencing Analysis in either of two modes:

- Sample File mode
- BioLIMS mode

Sample File Mode

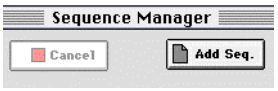
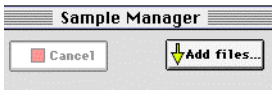
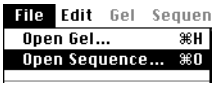
In Sample File mode, sequence data extracted from gel files is written out to individual sample files. Extracted sequence data that is viewed and processed within Sequencing Analysis software is read from and saved to sample files.

BioLIMS Mode

In BioLIMS mode, sequence data extracted from gel files is written directly to a BioLIMS database that resides on an Oracle® or Sybase SQL Server™ running on a SUN workstation. Extracted sequence data viewed and processed within Sequencing Analysis software is read from and written back to the same BioLIMS database.

Sample files from the ABI PRISM 310 Genetic Analyzer can be uploaded to the BioLIMS database using the Sample2DB utility program.

Comparing Modes

Feature	BioLIMS Mode	Sample File Mode
Data extracted from a gel file is written to…	the BioLIMS database.	sample files.
Sequence data is opened for viewing and for analysis from…	the BioLIMS database.	sample files.
The main window is called the…	Sequence Manager window. 	Sample Manager window. 
In menus, windows, and dialog boxes the word…	"Sequence" is used to indicate that the program is in BioLIMS mode. 	"Sample" is used to indicate that the program is in Sample File mode. 

Examples in This Manual Assume Sample File Mode

Most of the instructions and screen shots in chapters 4–7 of this manual assume that the Sequencing Analysis program is running in Sample File mode. In general, it is left to the reader to substitute "sequence" for "sample" and "database record" for "sample file" in the text as necessary.

In cases where confusion between the modes is likely to arise or the behavior of the program is significantly different in Sample File mode, the features of the program in BioLIMS mode are described explicitly.

The most important BioLIMS-specific concepts and features of the Sequencing Analysis program are explained in this chapter.

About Server Names

Sybase or Oracle?

The BioLIMS Session Manager decides whether you are connecting to a Sybase SQL Server or an Oracle Server database by looking at the name in the Server field in the Session Manager dialog box. The table below summarizes how server names are recognized.

If the Session Manager sees a Server name…	It assumes a…	Example
All in uppercase letters	Sybase SQL Server database connection	MOZART
Suffixed by ":s" or ":S"	Sybase SQL Server database connection	Offenbach:S
Containing any lowercase letters	Oracle Server database	Oramozart
Suffixed by ":o" or ":O"	Oracle Server database	SIBELIUS:O

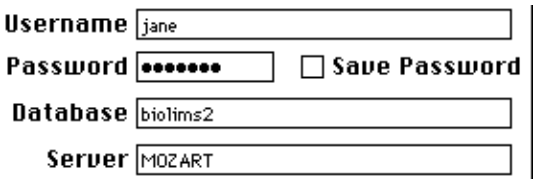
Sybase SQL Server Examples

Example 1

If the interfaces file contains this:

```
MOZART
query MacTCP mac_ether mozart.apldbio.com 2500
```

The Session Manager would look like this:



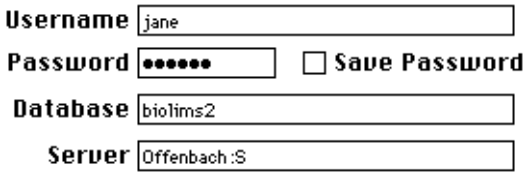
MOZART is recognized as a Sybase SQL Server because the server name is in all uppercase letters.

Example 2

If the interfaces file contains this:

```
Offenbach
query MacTCP mac_ether mozart.apldbio.com 2500
```

The Session Manager would look like this:



In order for Offenbach to be recognized as a Sybase SQL Server, the name in the Server field is suffixed with ":S".

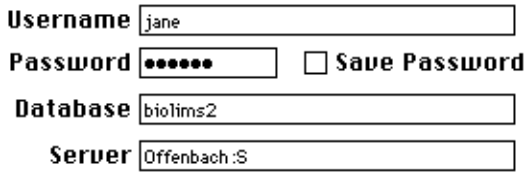
Oracle Server Examples

Example 1

If the tnsnames.ora file contains this:

```
Oramozart=(DESCRIPTION=
             (ADDRESS=
              (PROTOCOL=TCP)(host=mozart)(port=1521))
                   (CONNECT_DATA=(SID=WG733)
                  )
           )
```

The Session Manager would look like this:



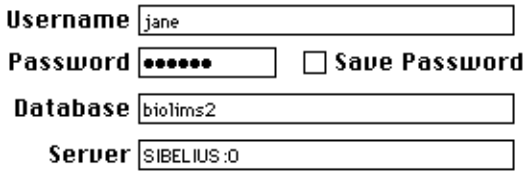
Oramozart is recognized as an Oracle Server because the server name contains lowercase letters.

Example 2

If the tnsnames.ora file contains this:

```
SIBELIUS = (DESCRIPTION=
            (ADDRESS=
            (PROTOCOL=TCP)(host=SIBELIUS)(port=1521))
                   (CONNECT_DATA=(SID=WG733)
                  )
           )
```

The Session Manager would look like this:



In order for SIBELIUS to be recognized as an Oracle Server, the name in the Server field is suffixed with ":O".

Configuring for Connection to the BioLIMS Database Server

Sybase or Oracle?

This section gives instructions on how to configure the client computer (that runs the Sequencing Analysis program) for database access.

The BioLIMS database resides on either a Sybase SQL Server or an Oracle Server.

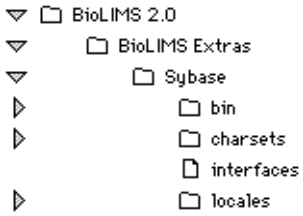
To configure for the…	See…
Sybase SQL Server	"Configuring for Sybase SQL Server Connection" on page 3-7
Oracle Server	"Configuring for Oracle Server Connection" on page 3-10

Configuring for Sybase SQL Server Connection

Follow the steps below to configure a Macintosh computer for connection to the Sybase SQL Server.

IMPORTANT Any time the name, port number, IP address, or host and domain name of the BioLIMS database server is changed, you need to repeat this procedure.

To configure for Sybase SQL Server connection:

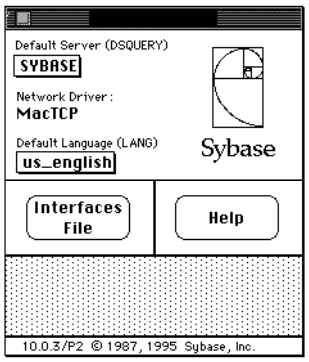
Step	Action
1	Find the interfaces file in the Sybase folder in the BioLIMS Extras folder. 
2	Open the file with SimpleText or a similar text editing application.

To configure for Sybase SQL Server connection: *(continued)*

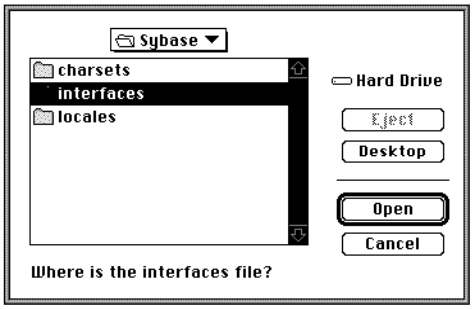
Step	Action

3	Find the lines:
	`SYBASE`
	`query MacTCP mac_ether neuron.apldbio.com 2500`
	and edit them:
	♦ Replace SYBASE by an alias name for the database server (see “About Server Names” on page 3-4)
	♦ Replace neuron.apldbio.com with the IP address or host and domain name of the server machine
	♦ 2500 is the default port number. If necessary replace 2500 with the port number recommended by your BioLIMS database administrator.
	You can find this information in the interfaces file on the Sybase® server, or your BioLIMS database administrator can provide this information.

4	If you have access to more than one server, duplicate the two lines and edit them for the other server(s). For example, for two servers, one called SYBASE and one called SERVER2, the interfaces file might look like this:
	`SYBASE`
	`query MacTCP mac_ether neuron.apldbio.com 2500`
	`SERVER2`
	`query MacTCP mac_ether 192.135.191.128 2025`

5	Save and close the interfaces file.
6	Open the SybaseConfig control panel. This control panel is found in the Control Panels folder in the System folder. 

To configure for Sybase SQL Server connection: *(continued)*

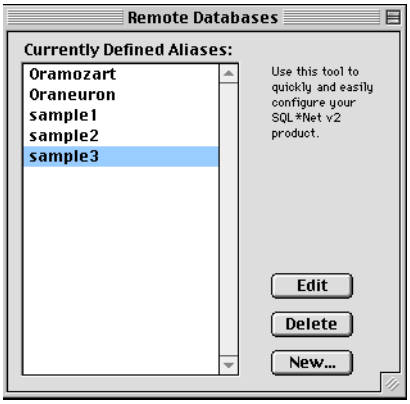
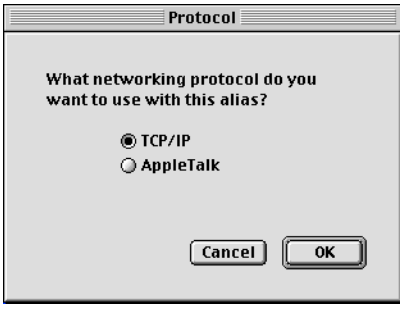
Step	Action
7	The first time the SybaseConfig control panel is opened, a file browser opens automatically. If a file browser does not open immediately, click the Interfaces File button to open one. 
8	Use the file browser to locate and open the interfaces file edited in the steps above.
9	Set the Default Language pop-up menu to us_english.
10	Close the SybaseConfig control panel.

Configuring for Oracle Server Connection

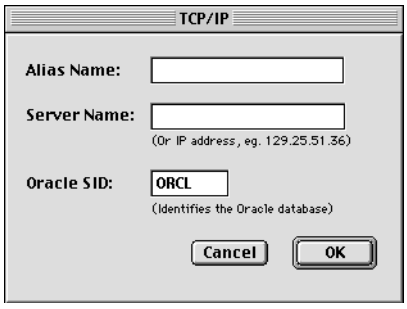
Follow the steps below to configure your Macintosh computer for connection to the Oracle Server.

IMPORTANT Any time you change the BioLIMS database server name, its IP address or host and domain name, or the port number, you need to repeat this procedure.

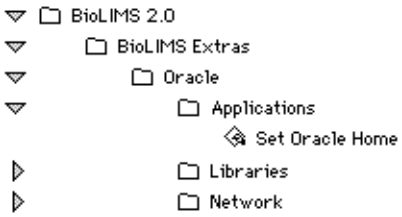
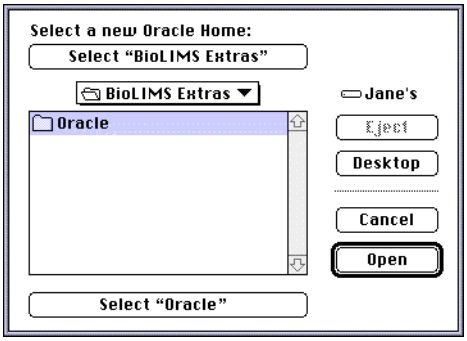
To configure for Oracle Server connection:

Step	Action
1	Use the program Easy Config to configure for the Oracle server connection. At installation, Easy Config is placed into the BioLIMS 2.0:BioLIMS Extras:Oracle:Applications:Networking folder. Open the Easy Config program. The Remote Databases window appears. 
2	Click New. The Protocol dialog box appears. 

To configure for Oracle Server connection: *(continued)*

Step	Action
3	Select TCP/IP and click OK. The TCP/IP dialog box appears. 
4	Enter text in the fields as follows: ♦ Alias Name: Enter an alias name for the database server. (See "About Server Names" on page 3-4.) ♦ Server Name: Enter the server name. This may be an IP address or host (and domain name) of the server machine. **Note** This field does not scroll horizontally for display even though it accepts characters typed past the end of the field. If the server name is longer than 20 characters, you may want to enter the end characters first and go back or just use the IP address. ♦ Oracle SID: Enter the value of the ORACLE_SID environment variable. You can find this information in the tnsnames.ora file on the Oracle server, or your BioLIMS database administrator can provide you with the information.
5	Click OK to close the TCP/IP dialog box.
6	From the File menu, Choose Save Configuration.
7	From the File menu, Choose Quit to exit the Easy Config program.

To configure for Oracle Server connection: *(continued)*

Step	Action
8	Find the application Set Oracle Home.  The application is contained in your BioLIMS 2.0:BioLIMS Extras:Oracle:Applications folder. 
9	Open the application Set Oracle Home. 
10	Use the file browser to find and highlight the Oracle folder in the BioLIMS Extras folder. Press the Select "Oracle" button.

Accessing BioLIMS Database

Introduction

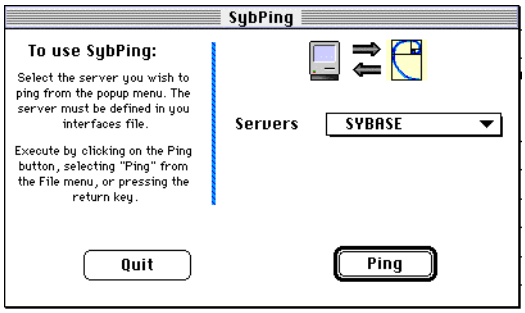
The following procedure describes how to access the BioLIMS database by completing the Preferences dialog box.

Before you can work with sequences in the BioLIMS database using the Sequencing Analysis software, you must establish a connection to the database. This connection is made through the BioLIMS Access page of the Preferences dialog box.

Before Accessing a Sybase-based BioLIMS Database

This section applies to Sybase-based BioLIMS databases only. If you are connecting to an Oracle-based BioLIMS database, skip to “To Access the BioLIMS Database” on page 3-14.

Before accessing a Sybase-based BioLIMS database, you need to ping the database using SybPing to ensure that the database connection is working.

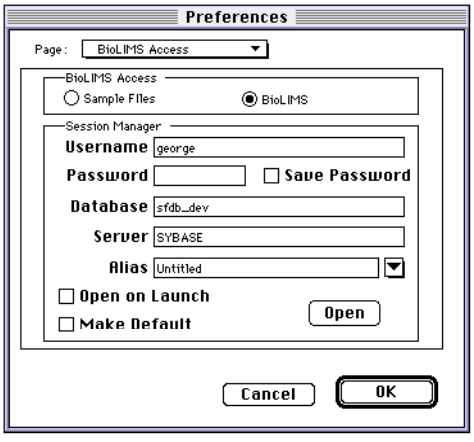
Step	Action
1	Open the SybPing program. At installation this program is placed in the BioLIMS 2.0:BioLIMS Extras:Sybase:bin folder. 
2	Select the name of the database Server that you want to use. These Server names correspond to those listed in the interfaces file. (See “Configuring for Sybase SQL Server Connection” on page 3-7.)
3	Click Ping.

Step	Action
4	<table><tr><th>If the connection is…</th><th>Then…</th></tr><tr><td>working</td><td>the Ping window reports that the “Ping was successful.”</td></tr><tr><td>not working</td><td>the Ping window reports that the “Ping Failed.” or the Ping window gives no response. Refer to Appendix D, “Troubleshooting the BioLIMS Database.”</td></tr></table>
5	Quit the SybPing application.

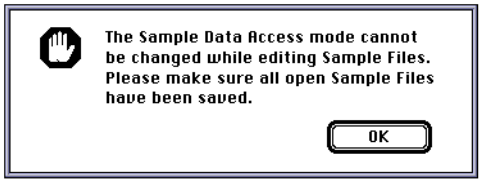
IMPORTANT If you try to connect to a Sybase-based BioLIMS database through the Sequencing Analysis program when the network is down, the computer may hang or freeze. If this happens, you have to restart the Macintosh computer. Force the computer to restart by pressing the key combination, Command–Control–Power key.

To Access the BioLIMS Database

Accessing the database:

Step	Action
1	Choose Preferences from the Edit menu and BioLIMS Access from the submenu. The Preferences dialog box appears. 

Accessing the database: *(continued)*

Step	Action
2	In the BioLIMS Access section of the dialog box, click the BioLIMS button. The fields in the Session Manager subpanel are activated. **Note** The following alert appears if there are unsaved sample files open when you switch to BioLIMS mode. When you click OK, the alert is dismissed and the Preferences dialog box appears with the Sample Files button selected. You need to save your sample files before you switch modes. When you switch modes, the files in the Sample Manager are removed. 
3	In the Session Manager section of the BioLIMS Access page of the Preferences dialog box, enter: ♦ Your user name on the server. ♦ Your password for your server account. ♦ The name of the database on the server (You may have access to more than one database on the server.) ♦ The server name. (The server name is contained in the interfaces file (Sybase) or in the tnsnames.ora file (Oracle). **IMPORTANT** All these text boxes are case sensitive.
4	Click the check box labeled Save Password if you want to: ♦ Save your password so that you do not have to enter it every time you open the connection. ♦ Run AppleScripts that do not contain password information. ♦ Automatically analyze the data after the Data Collection software has collected the data. ☒ Save Password

Accessing the database: *(continued)*

Step	Action
5	If you want the database to open automatically when you launch the Sequencing Analysis application, click the check box labeled Open on Launch. ☒ Open on Launch **Note** You must also click the check box labeled Save Password if you want the database to open automatically, allowing the data to be analyzed automatically.
6	If you intend to use more than one database or user account, enter an alias for this BioLIMS session information. Use the pop-up menu to add, change, or remove aliases. If you have more than one alias, click the Make Default check box to choose which one appears when you first open the Edit Session dialog box. ☒ Make Default **Note** The default alias is the database that opens if you choose to automatically analyze data. **Note** If both the Make Default and the Save Password boxes are checked, no dialog box will appear when a connection to the server is requested. Since all the information required of the user has been saved, the software will connect to the database automatically.

Accessing the database: *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>7</td><td>Click Open to connect to the database.

<table><tr><th>If the login was...</th><th>Then...</th></tr><tr><td>successful</td><td>the text in the dialog box becomes grayed-out.</td></tr><tr><td>unsuccessful</td><td>an alert dialog box appears.

Check that:

♦ All the login information was entered correctly and in the correct case.

♦ Your interfaces file is correctly configured (page 3-7).

If the connection is still not open, consult Appendix D, “Troubleshooting the BioLIMS Database.” Also refer to “About Server Names” on page 3-4; if the server name is not correct, connection will fail.</td></tr></table></td></tr>
<tr><td>8</td><td>Click OK to close the Preferences dialog box.</td></tr>
</table>

Switching Between Sample File and BioLIMS Modes

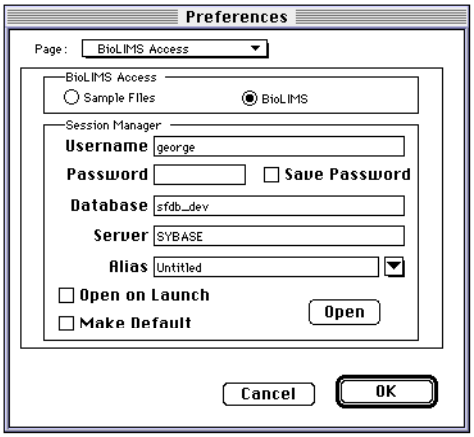
Introduction

The first time you open the Sequencing Analysis program, the software opens in Sample File mode.

To switch from Sample File mode to BioLIMS mode (or vice versa), follow the steps in the table below.

Switching Modes

To switch modes in Sequencing Analysis:

Step	Action
1	Choose Preferences from the Edit menu and BioLIMS Access from the submenu. The Preferences dialog box appears. 
2	To change to Sample File mode, click the Sample Files button. The Session Manager fields become unavailable. To change to BioLIMS mode, click the BioLIMS button. The fields in the Session Manager subpanel are activated. For more information about these fields, see “To Access the BioLIMS Database” on page 3-14.
3	Click OK to affect the changes to the Preferences dialog box and close it. Click Cancel to close the Preferences dialog box without changing the mode.

Using the Collection Browser Window

About the Collection Browser Window

The Collection Browser window is common to the following BioLIMS-aware applications.

- AutoAssembler™ DNA Sequence Assembly Software
- Factura™ Feature Identification Software
- Sample2DB Software
- Sequencing Analysis Software
- GeneScan® Analysis Software

Using the Collection Browser from within Sequencing Analysis, you can search the BioLIMS database for specific collections and sequences.

Sequences are database objects that contain the same information as samples files. Collections contain groups of sequences and other database objects and are analogous to folders in the Macintosh system.

Search by:

- Up to 5 collection-specific criteria (described on page 3-23)
- Up to 10 sequence-specific criteria (described on page 3-24)

In This Section

This section includes the following topics:

Displaying the Window

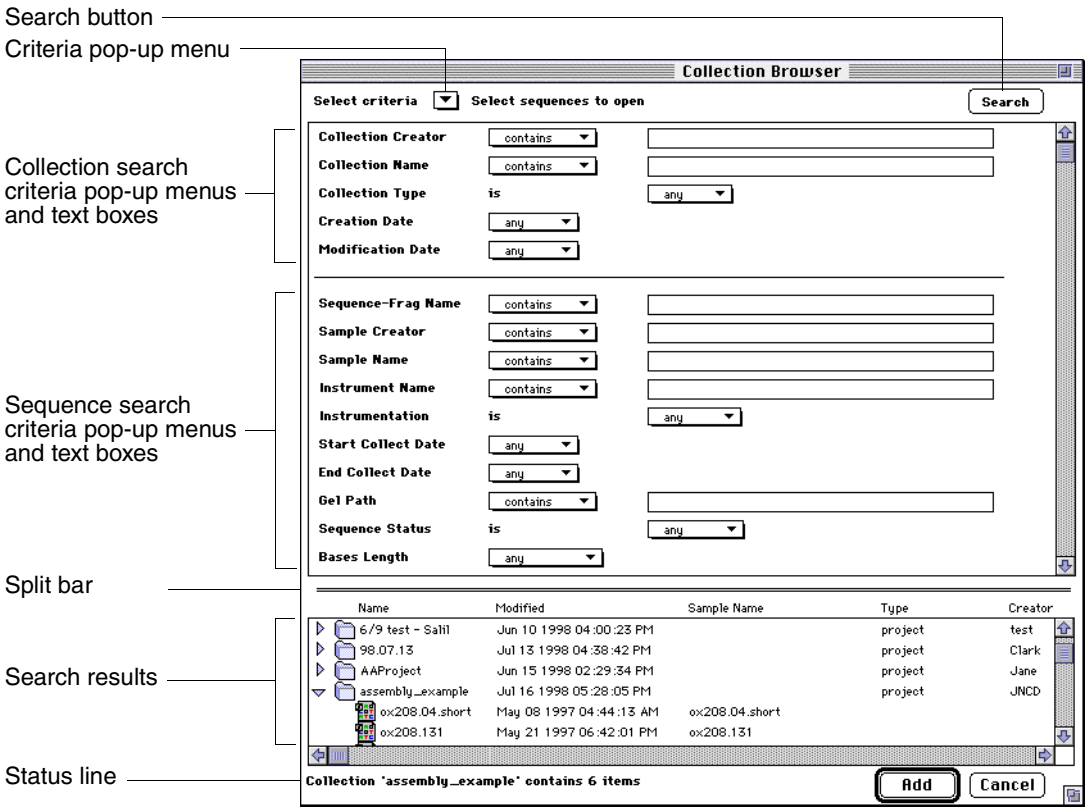
You can display the Collection Browser window as follows.

Note You must be in BioLIMS mode to view the Collection Browser window.

If you want to...	Then...	Result
open a sequence or sequences to view or edit	choose Open Sequence from the File menu.	The Collection Browser window appears (page 3-21).
add a sequence or sequences from the BioLIMS database to the Sequence Manager window	you can take the following action: ♦ Choose Add Sequences from the Manager menu. ♦ Click the Add Sequence button on the Sequence Manager window.	

The Window

When all search criteria are made visible, the Collection Browser looks like this.



Parts of the Window

The table below describes the parts of the Collection Browser window that were labeled in the figure above.

Item	Description
Criteria pop-up menu	Use this pop-up menu to specify the search criteria visible on the Collection Browser window. **Note** If you only intend to use a subset of criteria, setting only that subset visible helps to reduce clutter in the window. The search results are the same whether a criterion is invisible or blank and visible.

Item	Description
Search button	Click this button to query the BioLIMS database. **Note** You can also press the Return key to begin a search.
Collection search criteria pop-up menus and text boxes	Use these pop-up menus and text boxes to define the collection criteria of the search. **IMPORTANT** Only those sequences that match each and every criterion you specify are returned. That is, search criteria are combined using the logical AND operation. For more information, see “Collection Search Criteria” on page 3-23.
Sequence search criteria pop-up menu and text boxes	Use these pop-up menus and text boxes to define the sequence criteria of the search. **IMPORTANT** A collection is returned if one or more of the sequences contained in it fulfill all of the specified sequence criteria. For more information, see “Sequence Search Criteria” on page 3-24.
Split bar	Drag this bar to alter the relative amount of space allocated to the top and bottom portions of the Collection Browser window.
Search results	After a successful query, found collections are listed in this area as Name, Modification date, and Creator
Status line	Search results, error messages, and other important information are reported here. For example, the Status Line lists how many collections were returned in a search.

Collection Search Criteria

The table below shows the collection search criteria. The collections returned by the Collection Browser must match all of the collection criteria and contain at least one sequence that matches all of the sequence criteria.

Criterion	Pop-up Menu Choices	Allowed Text	Description
Collection Creator	♦ is ♦ starts with ♦ ends with ♦ contains	up to 255 characters	Name of the creator/owner of the collection
Collection Name	♦ is ♦ starts with ♦ ends with ♦ contains	up to 255 characters	Name of the collection
Collection Type	♦ any ♦ run ♦ project ♦ other	NA	Collection type Default is any menu item.
Creation Date	♦ any ♦ is ♦ before ♦ after ♦ between	date — set with arrow buttons (format: month/day/year)	Date the collection was created
Modification Date	♦ any ♦ is ♦ before ♦ after ♦ between	date — set with arrow buttons (format: month/day/year)	Date the collection was last modified

Sequence Search Criteria

The table below shows the sequence search criteria. The collections returned by the Collection Browser must contain at least one sequence that matches all of the specified sequence criteria.

Criterion	Pop-up Menu Choices	Allowed Text	Description
Sequence-Frag Name	♦ is ♦ starts with ♦ ends with ♦ contains	up to 255 characters	Name of the sequence This is the file name entered in the gel file Sample Sheet.
Sample Creator	♦ is ♦ starts with ♦ ends with ♦ contains	up to 255 characters	The User Name in the gel info window (page 4-13). (The Operator name set in the Data Collection software)
Sample Name	♦ is ♦ starts with ♦ ends with ♦ contains	up to 255 characters	Sample name from the Sample Sheet
Instrument Name	♦ is ♦ starts with ♦ ends with ♦ contains	up to 255 characters	Set in the General Settings Preferences of the Data Collection software
Instrumentation	♦ any ♦ gel ♦ capillary	NA	Whether the sample was run on a gel (377 or 373) or a capillary (310) instrument

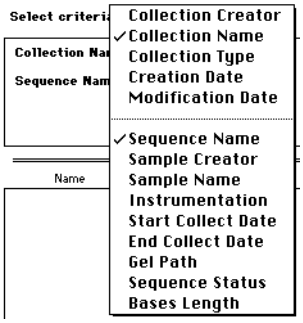
Criterion	Pop-up Menu Choices	Allowed Text	Description
Start Collect Date	♦ any ♦ is ♦ before ♦ after ♦ between	date — set with arrow buttons (format: month/day/year)	Date data collection began
End Collect Date	♦ any ♦ is ♦ before ♦ after ♦ between	date — set with arrow buttons (format: month/day/year)	Date data collection ended
Gel Path	♦ is ♦ starts with ♦ ends with ♦ contains	up to 255 characters	The full path name to the original gel file: e.g., Hard Disk:Data: GelRuns:L28t
Sequence Status	♦ any ♦ nascent ♦ prepare ♦ collect ♦ analysis ♦ cleanup ♦ assembly	NA	Status of the sequence. The pop-up menu lists the six stages of collection and analysis [a]
Bases Length	♦ any ♦ equal to ♦ less than ♦ greater than ♦ between	number	The length of the most recent version of the sequence in the database

a. Collect status is for unanalyzed data. Analysis status is for base-called data. Cleanup status is for data processed by Factura™ software. Nascent, Prepare, and Assembly status are not currently applied.

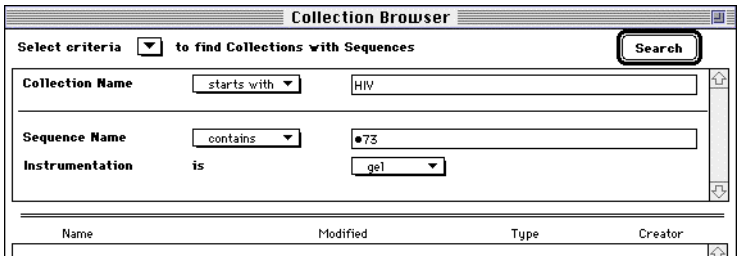
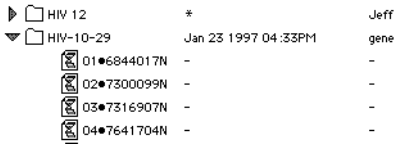
Searching the BioLIMS Database

Follow these steps to use the Collection Browser window to search the BioLIMS database for specific collections and sequences. For information on how to display the Collection Browser window, see "Displaying the Window" on page 3-20.

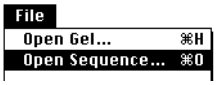
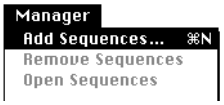
To search the BioLIMS database:

Step	Action
1	From the Criteria pop-up menu, select the criteria by which you want to search.
	
2	To use the pop-up menu: **Choose menu items...** / **To define the search by...** above the horizontal line / collection criteria below the horizontal line / sequence criteria **Note** As you choose items from the pop-up menu, a check mark appears next to the item on the menu and the item is added to either the collection search criteria or the sequence search criteria section of the window.
3	Use the pop-up menus and text fields to define your search query. Refer to the tables on 3-23 and 3-24 for details about the search criteria. When you are satisfied with the search setup, click Search. The results of the search appear in the lower portion of the window. Collections returned by the Collection Browser window must match all of the collection criteria and contain at least one sequence that matches all of the sequence criteria. **Note** To list all of the items in the BioLIMS database, perform the search with the text fields blank and the default pop-up menus. For large databases, this process may be slow.

To search the BioLIMS database: *(continued)*

Step	Action
	
4	To view the sequences contained in the collections, click the small triangle to the left of the collection name.  To view the electropherogram of a sequence, you need to add the sequence to the Sequence Manager and then open the sequence from the Sequence Manager window. **Note** If you double-click on the collection, all the sequences in the collection are added or opened in the Sequencing Analysis program.

To search the BioLIMS database: *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>5</td><td>You can take the following actions.
<table>
<tr><th>If you opened the Collection Browser window using...</th><th>Then...</th></tr>
<tr><td>Open Sequence from the File menu
 
</td><td>a. Select a sequence.
You can select:
– a range of sequences by selecting with the shift key down
– multiple sequences by selecting with the command key down
b. Click Open.</td></tr>
<tr><td>Add Sequences from the Manager menu or the Add button
 

Note You can add an entire collection by selecting the folder.</td><td>a. Select a sequence.
You can select:
– a range of sequences by selecting with the shift key down
– multiple sequences by selecting with the command key down
b. Click Add.
Sequences are added to the Sequence Manager.</td></tr>
<tr><td>either method, you can close the Collection Browser window</td><td>Click the
♦ Select button or
♦ Cancel button</td></tr>
</table></td></tr>
</table>

Note If you open a text-only sequence with the Collection Browser window, you are able to view the text sequence, but not analyze the data because the electropherogram information is not present.

Working with the Gel File 4

Overview

In This Chapter

This chapter contains information about how to view and edit the gel file, and how to generate sample files or BioLIMS database records after editing the gel file.

~~310~~
373
377

Note This entire chapter applies to the gel file, which is generated by the ABI 373™ DNA Sequencer Instrument and the ABI PRISM® 377 DNA Sequencer Instrument. If you are using an ABI PRISM® 310 Genetic Analyzer or working only with analyzed sequence data, ignore this chapter.

The Gel File Stores Raw Data

The gel file stores the raw data collected during the entire run of a ABI 373 or ABI PRISM 377 instrument. Initially, the file contains the raw data collected during the run, a gel image (a color picture similar to an autoradiogram), a copy of the data collection Sample Sheet, and a copy of the instrument file. After lane tracking and editing, the file also contains the lane tracking information and any changes made to the original information in the file.

Neural Net Tracker

The Neural Net Tracker is a stand-alone application that is called as needed by the Sequencing Analysis application. When called, the Tracker:

- Reads the sequencing gel file
- Locates the center of each lane
- Derives a tracker line down the center of each lane

The tracker line is used to extract the signal intensities from the gel. The tracker line can also be manipulated if manual correction is needed.

Also associated with the Tracker application are Tracker settings files that contain various tracker parameters optimized according to the number of channels and lanes in the gel file and the comb type set in the Gel Preferences. (See “Gel Preferences” on page 6-23.)

Note Gels must be multicomponented in order to be auto-tracked by the Neural Net Tracker program. For information on how to set preferences for multicomponenting, see “Multicomponent Gel Image” on page 6-24.

Problem Gel Files

If gel aberrations or weak sample signals exist, or if the comb was not properly centered in the gel, the Sequencing Analysis software may misinterpret the gel data. The program may completely miss a lane, declare a lane where none exists, or recognize a lane but be unable to follow it.

Each of these errors can cause lane data to be written to the wrong file or database record. For example, if the program mislabels lane 2 as lane 1, it will write the lane 2 data into the sample file for lane 1, the lane 3 data into sample file for lane 2, etc.

Data Extraction

During data extraction, the software generates a sample file for each tracked lane by averaging the data from the tracked channel and the number of channels (specified in Preferences) on either side of it. The default is three channels, the tracked channel and one channel on either side of it. Through the Gel Preferences dialog box, you can choose to use simple channel averaging or weighted channel averaging. (For more information, see “Gel Preferences” on page 6-23.)

The software also copies all the information required to identify and analyze (base call) the sample.

Gel Processing Parameters

The parameter values to use for gel file processing are set in the Gel Preferences page of the Preferences dialog box.

Note Specify whether or not you want automatic lane tracking and extracting in the instrument’s data collection program before starting the instrument run.

Review the Gel File

In most cases, if the run was successful, the gel file should be properly tracked, and only needs to be reviewed. If, however, the tracking is incorrect or the signal is weak and the Tracker missed a lane, you might need to make some changes and then re-extract the sample data.

After the gel image has been generated and the lanes tracked, perform the following review steps. Do this either before or after extracting the sample data from the gel file:

- Check the gel image. (See page 4-12.)
- If necessary, adjust the appearance and content of the gel image. (See page 4-19.)
- Check the lane assignment confidence value that is written to the Error Log. (See page 6-26.)
- If necessary, adjust the lane markers to correct lane numbering errors. (See page 4-25.)
- If necessary, adjust the placement of the tracker lines. (See page 4-39.)
- If you change any of the gel file or Sample Sheet information after extracting the sample data, re-extract the data from the edited lanes to regenerate the information in the sample files. (See page 4-43.)

IMPORTANT Review the gel file to see that all lanes were correctly labeled and tracked before data extraction. Do this before viewing or editing any analyzed data from this gel.

Displaying the Gel File in the Gel File Window

The Gel File Window

The Gel File window allows you to observe sample migration, lane tracking, and signal strength in the gel image. The gel image can be adjusted to improve visibility, individual lane markers can be realigned, and the positions of the tracker lines can be edited.

For details about the Gel File window, see “About the Gel File Window” on page 4-6.

Displaying the Gel File after Automatic Analysis

After automatic analysis, the Gel File window opens and displays the newly created gel file.

To set up automatic gel file analysis, select Autoanalyze with Sequencing Analysis (on ABI PRISM 377 instruments) or Analyze All Samples (on ABI 373 instruments) in the data collection program before starting the sequencing run.

Note If the Gel File window is not visible on the screen after automatic processing, it may be hidden behind the Sample Manager window. Choose the gel file from the Window menu to bring it to the front.

Opening a Gel File Manually

There are three ways to open a gel file manually:

- Double-click on the icon for the gel file
- Drag the gel file icon onto the Sequencing Analysis program icon
- Choose Open Gel from the Sequencing Analysis File menu, select the name of the file, and then choose Open

Note If the Sequencing Analysis v. 3.4 software has not already created a gel image for the file, it creates one when the file is opened. This process can take anywhere from a few seconds to a minute. Even if an image for the gel file was created in an earlier version of Sequencing Analysis, a new image is created when the gel file first opens in Sequencing Analysis v. 3.4. This occurs as Sequencing Analysis v. 3.4 displays the expanded image at 600 scans/window where previous versions of the software display at 350 scans/window.

What Next?

Once the gel file is opened, check the quality of the image. (See “Checking the Gel File” on page 4-12.)

To learn more about the gel window, see “About the Gel File Window” on page 4-6.

About the Gel File Window

Introduction

The gel image displayed in the Sequencing Analysis Gel File window is different from what is displayed by the data collection program during a run.

In the Data Collection Program

~~310~~
~~373~~
373XL
377

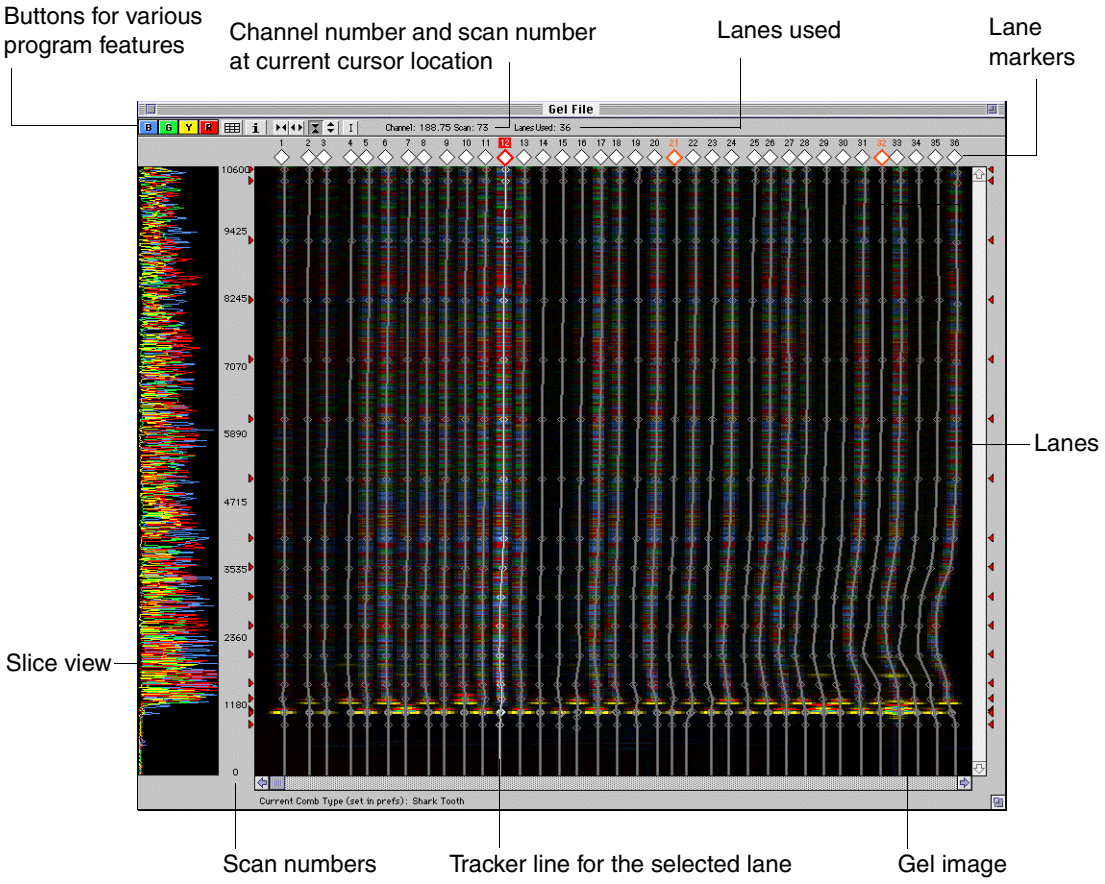
During the run, the Gel File window for the 377 Data Collection program shows real-time data as it is being collected. New data appears at the bottom of the screen as it is collected, so the top of the screen shows the start of the run. This data has not yet been saved to a file.

In the Sequencing Analysis Program

The Gel File window for the Sequencing Analysis software displays an image of the gel after data collection is finished. This image is inverted, so the bottom of the window displays the start of the run. The smallest fragments appear near the bottom of the window, just as they would on an autoradiogram.

Parts of the Gel File Window

The Gel File window includes the following parts:



Description of Parts

This table describes the parts of the Gel File Window.

Item	Description
Gel image	Represents a time history of all fluorescence detected during the run. Each base peak appears as a brightly colored band within the sample lane. Each position on the gel image is defined by a scan number and a channel number.

Item	Description
Slice view	Displays a graphical view of the data values in the tracked channel(s) of the selected lane. The display changes as the tracker line moves from one channel to another. Each peak in the Slice view corresponds to a band in the gel image and indicates a base in the DNA sequence. These bands and peaks do not represent analyzed data, but provide an overview of the relative signal intensity between the bands in that lane and, thus, allow a qualitative evaluation of the run. The Slice view is empty (black) when no lane is selected.
Channels	Theoretical divisions across the read region of a gel where the data collection software samples the data. The number of available channels depends on the instrument and run mode used. For more information about run modes, see your instrument manual.
Lane	The path followed by the sample as it migrates through the gel. A sample lane is several channels wide. The number of wells in the loading comb determines the approximate number of channels assigned per lane of the gel. For instance, on a 377 instrument with a 36-well comb, one lane includes approximately five channels.
Channel / Scan	The channel number (horizontal scale) and scan number (vertical scale) at the current cursor position. These numbers change as the cursor moves over the image.
Lanes Used	The number of lanes marked as used in the Sample Sheet. This is originally determined by which lanes had names typed in the corresponding cell in the Data Collection Sample Sheet.
Lane numbers	The numbers across the top of the gel image that show the lane number currently assigned to each lane on the gel.

<table>
<tr><th>Item</th><th>Description</th></tr>
<tr><td>Lane markers</td><td>The diamond-shaped markers (♦) that show the current status of each lane:

<table><tr><th>Item</th><th>Description</th></tr><tr><td>White</td><td>Lane is marked for extraction</td></tr><tr><td>Blue</td><td>Lane is not marked for extraction or the lane was previously extracted</td></tr><tr><td>Yellow</td><td>Lane was edited and extracted, but the gel file was not saved with the new information</td></tr><tr><td>Gray</td><td>Lane is not marked Used in the Sample Sheet. The Tracker software does not expect to find a lane here, if it does, it will confuse it and lane assignment confidence will be low.</td></tr><tr><td>Orange border</td><td>Lane was inferred by the Tracker software. If an inferred lane tracker line is moved or reshaped, it ceases to be inferred and the orange border is lost.</td></tr></table>
Note If the “Save Gel After Extraction” option is selected, the white and yellow lane markers revert to blue after new sample data is generated.</td></tr>
<tr><td>Vertical scale (Scan numbers)</td><td>The scale between the gel image and the Slice view, which shows the scan number at each location on the gel image and the Slice view.</td></tr>
<tr><td>Tracker lines</td><td>Lines that the Tracker application draws on the gel image to mark the position that represents the center of the lane. Placement of these lines can be done manually, if you are not satisfied with their location.</td></tr>
</table>

Description of Buttons

This table describes the buttons on the Gel File window.

Button	Name	Function
	Colors	Allow toggling on and off of the display of one or more colors in the gel image and Slice view.
	Sample Sheet	Allows display of the data collection Sample Sheet associated with the gel file. You can use the Sample Sheet information to check that lanes are correctly labeled on the gel image. (For more information about the Sample Sheet, see “Reviewing the Sample Sheet Information” on page 4-14.)
	Gel Info	Allows display of the Gel Info window which contains information about the run conditions when the gel file was created. (For more information about the Gel Info window, see “Review the Gel Info Window” on page 4-13.)
	Horizontal Shrink	Compresses the gel image horizontally so you can see all the gel lanes in a standard-size window. There are four levels of horizontal zoom: 1X, 2X, 4X, and 8X.
	Horizontal Expand	Expands the gel image horizontally so you can more easily adjust the tracker lines. There are four levels of horizontal zoom: 1X, 2X, 4X, and 8X.
	Vertical Shrink	Returns the vertical scale to normal after the Vertical Expand button was used (see below).
	Vertical Expand	Expands the gel image vertically so you can more easily adjust the tracker lines. There are two levels of vertical zoom: full scale and 600 scans. (Previous versions of the Sequencing Analysis software displayed at 350 scans in vertical expand mode.)

Button	Name	Function
	Interpolation Mode	Puts you into tracker-line interpolation mode. In this mode if two lanes are selected by clicking the tick marks at the bottom of the gel image, the tracker lines for the lanes between the selected lanes are interpolated from the tracker lines of the two selected lanes. (For more information see “Interpolating Tracker Lines” on page 4-36.)

Checking the Gel File

Introduction

Before looking at the lane markers and tracker lines, you should:

- Inspect the gel image (see below)
- Review the Gel Info window (see “Review the Gel Info Window” on page 4-13)
- Check the information contained in the data collection software Sample Sheet (see “Reviewing the Sample Sheet Information” on page 4-14)

Inspecting the Gel Image

An inspection of the gel image can give a general measure of the quality of the run and its extracted sample data. Review of the gel image should include the following:

- Inspect the general condition of the bands in the lanes in the gel image. Are the fluorescent signals displayed as discrete horizontal bands? If not, this may be indicative of a poor gel.
- Are any of the colors too bright or too dark? Is there a green or red haze? Is this something that can be fixed by adjusting the gel image contrast? (See “Adjusting the Gel Image” on page 4-19.)
- Inspect the lane markers. Look for data lanes without lane markers, and for lane markers between data lanes. (See “Adjusting Lane Markers” on page 4-25.)
- Verify that each lane marker corresponds to a sample as designated in the Sample Sheet. (See “Reviewing the Sample Sheet Information” on page 4-14)
- If necessary, adjust the locations of the lane markers. (See “Moving Misplaced Lane Markers” on page 4-25.)
- Inspect the tracker lines. Each tracker line should be in the center of the lane it tracks. If necessary, adjust the tracker line placement. (See “About Optimizing Tracker Line Locations” on page 4-31.)

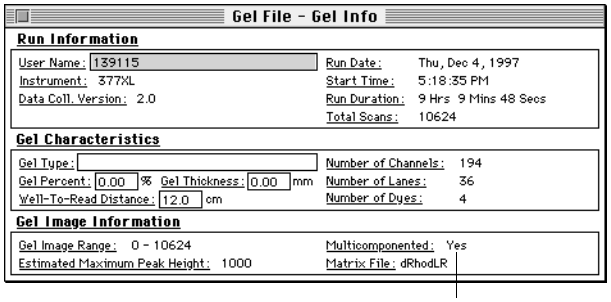
IMPORTANT If you change the lane markers or their Sample Sheet designation, or reposition any tracker lines **after** extracting the sample data from the gel file, you must re-extract the data to include the new information in the sample files.

Review the Gel Info Window

The Gel Info window displays information about the run conditions when the file was created. Reviewing the conditions under which the samples ran can give helpful information for reviewing the gel file.

You can view and edit the contents of this window whenever the gel file is open in the Gel File window. Changes you make in this window are stored in the gel file.

To view and edit the Gel Info Window:

Step	Action
1	Click the Gel Info button (i) near the top of the Gel File window, or select Gel Info from the Gel menu. The Gel Info Window appears.  This gel has been multicomponented.
2	Check that the gel has been multicomponented. The Tracker application cannot track gels that have not been multicomponented. (If the gel has not been multicomponented, refer to “Regenerating the Gel Image with Different Option Values” on page 4-22 and “Installing New Matrix Information” on page 4-23.) The Gel Info window displays information about the run, gel characteristics, and the gel image. This information is saved with each sample file generated from the gel file.
3	To edit information in the Gel Info window, click on any text field that is surrounded by a black rectangle. Then type in the new information. **Note** Not all the information in the gel file window can be edited. Some information, such as the date of the run, cannot be changed.
4	To close the window, click the close box at the top left corner of the window.

Reviewing the Sample Sheet Information

The Sample Sheet contains the sample information that was recorded in the data collection program before the run started. Information entered in the Sample Sheet is extracted into the sample file and used by the base-calling algorithms to create the analyzed sample data. In the Gel File window, only tracker lines for the sample lanes marked as used are displayed. Initially lanes are marked as used when a name is entered in the Data Collection Sample Sheet. A copy of the Sample Sheet for the run is embedded in the gel file during data collection.

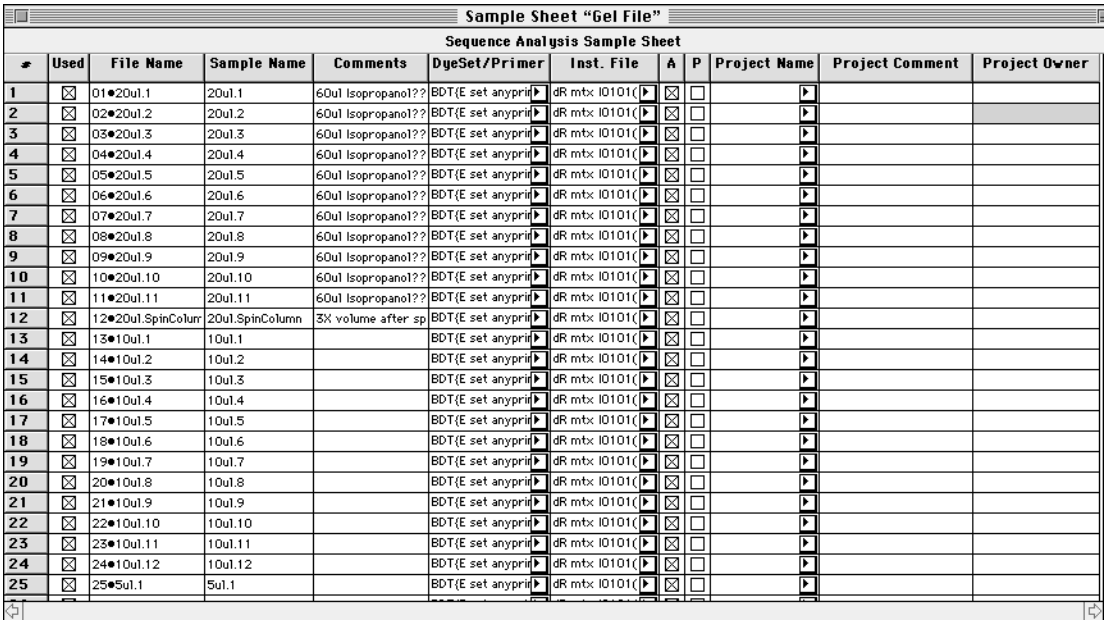
You can view, edit, and print this Sample Sheet copy whenever the gel file is open in the Gel File window. Changes made in this copy are stored in the gel file; they do not affect the original Sample Sheet file.

IMPORTANT If the Sample Sheet is edited after extracting the sample information from the gel file, you must re-extract the sample data to include the new information in the sample files. Be sure the Used check box is selected for each sample you wish to extract. Information is only extracted for samples that have the Used check box selected.

Note Changes to the Sample Sheet only affect extraction, not tracking. The Tracker does not refer to the Sample Sheet until after tracking. It then sets lane assignment confidence values according to how well the tracked lanes match the Sample Sheet.

About the Sample Sheet

The Sample Sheet looks like this:



Sample Sheet "Gel File"

Sequence Analysis Sample Sheet

#	Used	File Name	Sample Name	Comments	DyeSet/Primer	Inst. File	A	P	Project Name	Project Comment	Project Owner
1	☒	01•20ul.1	20ul.1	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
2	☒	02•20ul.2	20ul.2	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
3	☒	03•20ul.3	20ul.3	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
4	☒	04•20ul.4	20ul.4	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
5	☒	05•20ul.5	20ul.5	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
6	☒	06•20ul.6	20ul.6	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
7	☒	07•20ul.7	20ul.7	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
8	☒	08•20ul.8	20ul.8	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
9	☒	09•20ul.9	20ul.9	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
10	☒	10•20ul.10	20ul.10	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
11	☒	11•20ul.11	20ul.11	60ul Isopropanol??	BDT(E set anyprir	dR mtx I0101(	☒	☐			
12	☒	12•20ul.SpinColum	20ul.SpinColumn	3X volume after sp	BDT(E set anyprir	dR mtx I0101(	☒	☐			
13	☒	13•10ul.1	10ul.1		BDT(E set anyprir	dR mtx I0101(	☒	☐			
14	☒	14•10ul.2	10ul.2		BDT(E set anyprir	dR mtx I0101(	☒	☐			
15	☒	15•10ul.3	10ul.3		BDT(E set anyprir	dR mtx I0101(	☒	☐			
16	☒	16•10ul.4	10ul.4		BDT(E set anyprir	dR mtx I0101(	☒	☐			
17	☒	17•10ul.5	10ul.5		BDT(E set anyprir	dR mtx I0101(	☒	☐			
18	☒	18•10ul.6	10ul.6		BDT(E set anyprir	dR mtx I0101(	☒	☐			
19	☒	19•10ul.7	10ul.7		BDT(E set anyprir	dR mtx I0101(	☒	☐			
20	☒	20•10ul.8	10ul.8		BDT(E set anyprir	dR mtx I0101(	☒	☐			
21	☒	21•10ul.9	10ul.9		BDT(E set anyprir	dR mtx I0101(	☒	☐			
22	☒	22•10ul.10	10ul.10		BDT(E set anyprir	dR mtx I0101(	☒	☐			
23	☒	23•10ul.11	10ul.11		BDT(E set anyprir	dR mtx I0101(	☒	☐			
24	☒	24•10ul.12	10ul.12		BDT(E set anyprir	dR mtx I0101(	☒	☐			
25	☒	25•5ul.1	5ul.1		BDT(E set anyprir	dR mtx I0101(	☒	☐			

Item	Description
#	The lane number for the sample. The Sequencing Analysis software assigns lane numbers to the gel file lanes based on the numbers in this column. If a lane in the gel is empty, there must be a corresponding empty row in the Sample Sheet.
Used	When this check box is selected, the corresponding lane in the gel image is marked Used. When the Sequencing Analysis software extracts the sample data from the gel file, it creates sample data only from lanes marked Used. If you change the setting for this check box, the corresponding lane marker in the gel file is automatically changed.

Item	Description
File Name	The name of the sample file to be created for the data in this lane. The sample file is saved to the hard disk, into the Run folder next to the gel file. If the data is extracted into a BioLIMS® database, then this is the name assigned to the database record for that sample data.
Sample Name	The name of the sample in this lane. This can be edited, but initially it is the sample name entered in the sample sheet of the Data Collection software.
Comments	Comments about the sample. These can be edited, but initially they are the comments entered in the sample sheet of the Data Collection software.
DyeSet/Primer	The dye set and primer file to be used during analysis. The information in this file is used to designate the type of chemistry, dye/base relationships, and mobility correction to be applied to this sample.
Inst. File	The instrument file to be used during analysis. The Sequencing Analysis software uses the matrix information in the instrument file to adjust for spectral overlap in the dyes.
A	When this check box is selected, the Sequencing Analysis software automatically creates and analyzes the sample data after data collection is finished.
P	When this check box is selected, the Sequencing Analysis software automatically prints the sample information after it analyzes the sample data.
Project Name	If the samples are extracted into the BioLIMS® 2.0 Database, this is the name of the collection that will contain them. If a collection of that name does not exist in the database, one is created. (See “Editing or Adding Project Names” on page 4-18.)
Project Comment	Comment text associated with the BioLIMS 2.0 collection name, specified in the Project Name field described above.
Project Owner	Collection Creator text associated with the BioLIMS 2.0 collection name, specified in the Project Name field described above.

Reviewing the Sample Sheet

To review, edit, and print the Sample Sheet:

Step	Action
1	Click the Sample Sheet button near the top of the Gel File window ()or choose Gel Sample Sheet from the Gel menu. The Sample Sheet window appears.
2	Confirm that the information in the fields is correct. If necessary, edit the Sample Sheet information (instrument file, sample names, comments, etc.) that is automatically transferred to the sample files or BioLIMS database.
3	To change the width of a column in the Sample Sheet in order to see more of the information in that column: a. Put the cursor on the divider line to the right of the column title. b. When the cursor changes to two arrows, hold down the mouse button and drag the line to the desired location.
4	To edit the information in the Sample Sheet: ♦ Double-click in a text field and type in new text ♦ Use standard Edit menu commands ♦ Select/deselect check boxes Select filenames from pop-up menus
5	To print the Sample Sheet: choose Print from the File menu while the Sample Sheet window is active.
6	If you have made changes to the Sample Sheet, choose Save from the File menu.
7	When you are finished, click the close box to close the window.

Installing a New Sample Sheet

If you install a Sample Sheet with fewer rows than the current Sample Sheet, blank rows will be added to make up the difference. (For example, if you install a 36-lane Sample Sheet in a 48-lane gel, 12 blank lines will be added to the bottom of the Sample Sheet.) If you

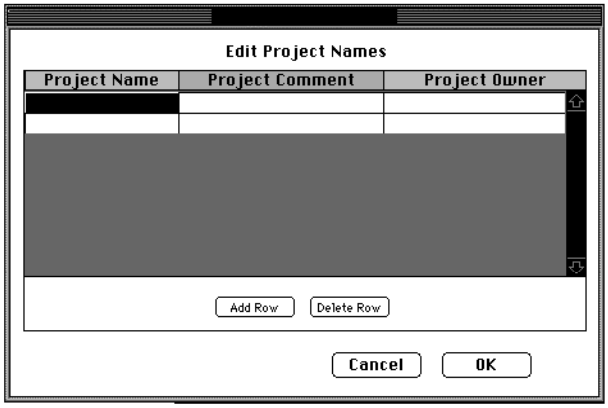
install a Sample Sheet with more rows than the current Sample Sheet, rows will be deleted from the bottom of the new Sample Sheet.

To replace the current Sample Sheet with the contents of a saved Sample Sheet file:

Step	Action
1	Close the current Sample Sheet window.
2	Choose Install New Sample Sheet from the Gel menu.
3	Select a Sample Sheet from the directory dialog box that appears. The Gel's Sample Sheet will be filled in with information from the Sample Sheet that you selected.

Editing or Adding Project Names

To edit project names, click Edit Project Name on the pop-up menu and the Project Name Editor dialog box appears.



You can take the following action.

If you want to...	Then...
add a name	click Add Row and enter a project name, project comment, and project owner.
edit a row	select a row and make the necessary changes.
delete a row	select a row and click Delete Row.

Adjusting the Gel Image

Introduction

Although the data shown in the gel image is not analyzed, the displayed information allows you to evaluate the quality of the run. You can adjust the content and appearance of the gel image in the following ways:

- Display or hide selected dye colors in the gel image.
- Adjust the color contrast in the gel image.
- Regenerate the gel image using a different data range, maximum peak height, and multicomponenting option.
- Install new matrix information in the gel file and use the new information to regenerate the gel image.

Note None of these options change the raw fluorescence data contained in the gel file nor the way data is extracted from the gel file.

Note The raw display color for a base is often different from its analyzed display color. For more information, see “Colors in Real-Time Data Display Windows” on page E-3.

Displaying and Hiding Selected Dye Colors

You can control the display of the colors in the gel image. For example, if you want to display only the blue bands, you can suppress the display of all green, yellow, red or orange (fifth dye only) bands.

Step	Action
1	Click the colored boxes near the top left corner of the Gel display to turn on or off display of each dye color.  The default is to have all colors displayed. Any changes you make in the button settings are saved in the gel file and used the next time that file is opened.

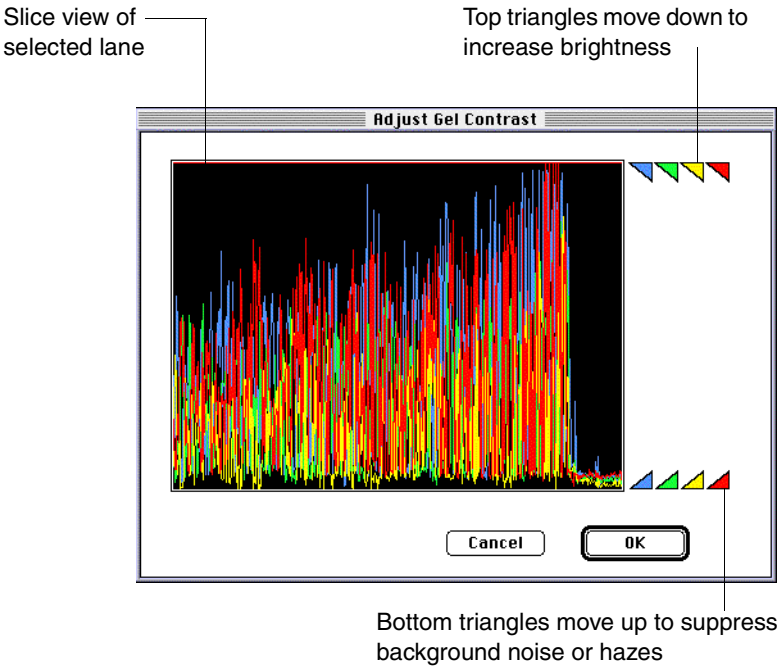
Adjusting the Contrast

You can adjust the color contrast in the gel image to increase or reduce the intensity of individual colors. These kinds of adjustments can make it easier to see the data in the gel and can improve the appearance of the gel image for publication.

The adjusted color values are saved in the gel file and used each time you open the file in the future. If you regenerate the gel image, the changes are discarded and the colors revert to their default values.

To adjust contrast for the gel image:

Step	Action
1	In the Gel File window, select any lane that contains the color(s) you want to adjust. The changes you make in this dialog box affect the entire gel, not just the selected lane.
2	Choose Adjust Gel Contrast from the Gel menu. The Adjust Gel Contrast dialog box appears.



Note The data shown is from the selected lane. If no lane is selected, the data from lane 1 is used.

To adjust contrast for the gel image: *(continued)*

Step	Action
3	Put the cursor on the triangle for the color you want to adjust, then hold down the mouse button and drag the triangle up or down to a new position. ♦ *To increase the intensity of a color,* pull the top triangle for that color down. For example, blue is sometimes hard to see on a gel. To correct this, pull the top blue triangle down until it is somewhat above the tallest blue peaks in the displayed slice view. ♦ *To suppress background noise of a particular color*, pull the bottom triangle for that color up. For example, there is sometimes a red background haze because of signal noise or because the signal baseline is not flat. To correct this, pull the bottom red triangle up until it is just above the baseline and noise in the displayed Slice view. It is best to adjust one color, apply the change, and view the effect in the gel image before you adjust another color.
4	Changes take place immediately. Choose OK to close the dialog box.
5	After the Gel File window is redrawn, note both the change in contrast in the gel image and the corresponding change in the peak heights in the Slice view. **Note** If you do not like the contrast adjustment, immediately choose Undo Adjust Contrast from the Edit menu to remove the change.

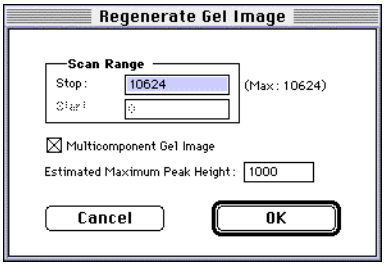
Regenerating the Gel Image with Different Option Values

Use Sequencing Analysis software to:

- Change the range of data included in the gel image (e.g,. if there is unusable data near the end of the run)
- Set a different maximum peak height
- Change the Multicomponent Gel Image setting

Note When it regenerates the image, the Sequencing Analysis software saves any tracker line changes you made in the original image, but does not extract the data.

To regenerate the gel image:

Step	Action
1	Choose Regenerate Gel Image from the Gel menu. The Regenerate Gel Image dialog box appears and displays the values that were used to create the current gel image. 

To regenerate the gel image: *(continued)*

Step	Action
2	Make any required changes in the dialog box values, as described below: **Item** / **Description** Stop — The last scan value to be included in the gel image and when extracting sample data. Start — The first scan value to be included in the gel image is always zero. Multi-component Gel Image — Causes the Sequencing Analysis software to apply the matrix information in the attached instrument file to the raw data, to adjust for spectral overlap of the dyes, before creating the gel image. It is usual to view the gel image multicomponented. Estimated Maximum Peak Height — The maximum signal level expected from samples in the run. This can be an approximate number, based on your typical run conditions and samples.
3	Choose OK to close the dialog box and start regenerating the gel image. You can use Command–[period] to cancel the regeneration process at any time.

Installing New Matrix Information

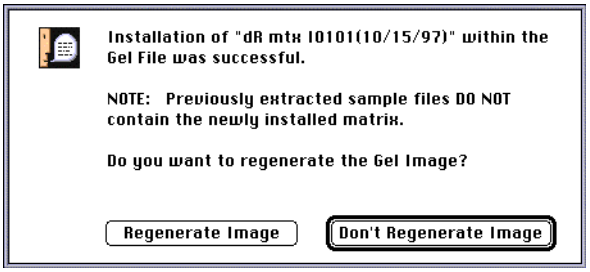
The data collection program copies the matrix information in the specified instrument file (matrix file) to the gel file during data collection. The Sequencing Analysis software uses this matrix information to generate the gel image. It is also copies this information to each sample file for use during data analysis.

Installing a new matrix changes the appearance of the gel image. If you install a new instrument file in the gel file, change the instrument file in the Sample Sheet also to ensure that the extracted samples are analyzed with the new instrument file.

If the instrument file in the gel file and the Sample Sheet are different, the Sample Sheet instrument file (if present) is applied when sample files are extracted. If the Sample Sheet instrument file is missing from the ABI Folder, the instrument file in the gel file is applied when sample files are extracted.

IMPORTANT There is no undo or cancel for the Install New Gel Matrix operation. When the new matrix is installed it is written immediately to the gel file. Even if you close the file without saving, the new matrix is incorporated into the gel file.

To install new instrument file information in the gel file:

Step	Action
1	Choose Install New Gel Matrix from the Gel menu. A directory dialog box appears. It shows only the names of folders and instrument (matrix) files.
2	Find and select the desired instrument file. Then choose Open. A dialog box like the following appears. 
3	Choose Regenerate Image or Don't Regenerate Image. The new Matrix is installed regardless of which option you choose. The new Matrix name is shown on the Gel Info Window.

Adjusting Lane Markers

Introduction

When the Sequencing Analysis software first opens a gel file, it adds lane numbers, lane markers, and tracker lines to the gel image. In most cases, the gel file should be properly tracked. If, however, the tracking is incorrect, you might need to make some changes. The Sequencing Analysis program allows you to:

- Move misplaced lane markers (page 4-25)
- Mark lanes used or unused (page 4-28)
- Mark and unmark lanes for extraction (page 4-29)
- Show and hide tracker lines (page 4-31)
- Position and reshape tracker lines so they more accurately track the samples (page 4-32)

Using the Keyboard to Move Between Lanes

The Sequencing Analysis software allows use of the Shift, Tab, and Arrow keys to move quickly between lanes and channels.

Keyboard shortcuts to switch quickly from one lane to the next:

- At the top of the display, click the Lane Marker of the lane where you want to move.
- Press the Tab key to move one lane to the right.
- Press Shift–Tab to move one lane to the left.

To move from one tracker line control point to the next:

- Press the Up Arrow key to move up one control point.
- Press the Down Arrow key to move down one control point.

Moving Misplaced Lane Markers

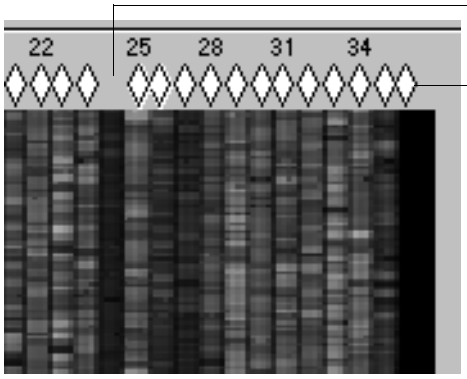
If gel aberrations or weak sample signals exist, or if your comb was not properly centered in the gel, the Sequencing Analysis software may misinterpret the gel data. The program may completely miss a lane, declare a lane where none exists, or recognize a lane but be unable to follow it.

Each of these errors can cause lane data to be written to the wrong place. For example, if the program mislabels lane 2 as lane 1, it will write the lane 2 data into the sample file for lane 1, the lane 3 data into sample file for lane 2, etc.

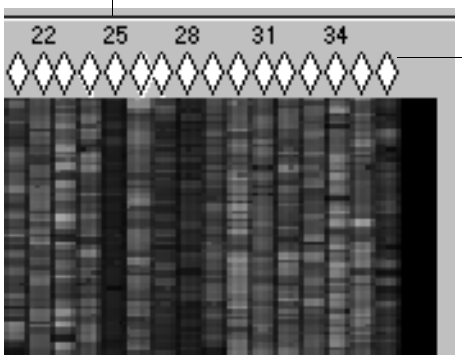
To correct these problems, you can compare the lane markers on the gel image to the Sample Sheet, then rearrange the lane markers so the lane numbers are properly aligned with the actual rows of information in the Sample Sheet. Then, when you later regenerate the sample data, the lane data will be written out correctly.

Note For information on how to view the data collection Sample Sheet to confirm that the gel lanes are properly labeled, see "Reviewing the Sample Sheet Information" on page 4-14.

To rearrange the lane markers:

Step	Action
1	Inspect the gel image for incorrectly labeled lanes. For example, in the following illustration, the Sequencing Analysis software missed the faint signals from lane #25. As a result, lanes #26–36 are mislabeled, and the lane #36 marker is over an unused area to the right of the lanes.
	
2	Click the incorrectly placed marker to select it. The selected marker becomes outlined in red.

To rearrange the lane markers: *(continued)*

Step	Action
3	Hold down the mouse button and drag the lane marker to the correct location. Lane markers always remain in numerical order from left to right and are attached to their respective tracker line within a few channels to either side. If you drag a lane marker across another marker, all the affected markers are renumbered accordingly. For example, if you drag the #36 marker (in the preceding illustration) to the real lane 25, lanes #25–36 all become correctly marked (as shown in the figure below). **Note** If the Sequencing Analysis software missed the first lane in the gel and put all the lane markers one position too far right, you can drag the right-most lane marker to the left-most lane (the missed lane) to label all the lanes correctly. **Note** Lane markers are aligned with their respective lanes at the top of the gel image. If the gel contains lanes which drift, the drifted portion of the lane may seem to be incorrectly aligned with its lane marker when viewed in zoomed-in mode.
	
4	To move multiple lane markers at once: a. Shift-click (to select) the lane markers you want to move. b. Drag the markers to the correct location.

To rearrange the lane markers: *(continued)*

Step	Action
5	Sometimes the Tracker detects noise in the gel and tracks that noise as one or more extra lanes. In particular, this may occur on the left side of the gel, leaving lanes on the right side of the gel without tracker lines. The fast way to correct this is: a. Select all the markers of the extra lanes in the gel (shift-click to select multiple markers). b. Select Force Selected Lanes to Right from the Gel menu.  The selected lanes are straightened and moved to the far right of the gel, and the lane markers renumbered. The markers forced right are stacked on top of each other. To view them individually, move them left one at a time.

Marking Lanes as Used

The Sequencing Analysis software only extracts sample data if the lane is marked Used in the data collection Sample Sheet.

IMPORTANT By marking lanes as Used or Unused, you specify which lanes should be used to generate sample data. This allows the Sequencing Analysis software to correctly number the used lanes and put the extracted sample data from each lane into the correct sample file. It also ensures that sample data is generated from only the intended lanes, and not from empty lanes.

When a lane is marked Used, its lane marker (◆) is colored blue, white, or yellow. Unused lanes have gray lane markers.

There are three ways to mark a lane as used (or unused):

- Click the lane marker, then choose Mark Lane Used/Unused from the Gel menu.
- Hold down the Command key while you click the lane marker.

- Click the  button to open the Sample Sheet that is attached to the gel file, then select or deselect the Used check box for the lane.

Note When you use mark a lane Used or Unused, the Sequencing Analysis software changes the setting in both the gel image and the Sample Sheet.

Marking / Unmarking Lanes for Extraction

During the extraction process, the Sequencing Analysis software only extracts data from gel lanes that are marked for extraction. This allows you to control which lanes to extract when using the Extract Lanes command.

Lane Marker Rules

The Sequencing Analysis software uses the following rules to mark lanes:

- Lanes identified with sample names in the Sample Sheet of the data collection software are automatically marked Used (blue, white, or yellow marker).
- All unidentified lanes are marked Unused (gray marker).
- When opening a gel file the first time, all Used lanes are marked for extraction (white marker).
- After the data is extracted from a lane, the lane is unmarked for extraction (blue marker).
- If you adjust the position of a lane marker or tracker line, the lane is automatically marked for extraction (white marker).
- If you modify a tracker line and then extract sample data from the lane before you save the gel image, the lane marker becomes yellow. This serves as a warning that the latest generated sample does not reflect the saved tracking information.
- If a lane marker is selected, it has a red border.
- If the Tracker has inferred a lane for some reason, the lane marker has an orange border.

To mark all lanes for extraction:

Step	Action
1	From the Gel menu, select Mark All Lanes For Extraction. The markers for *all* Used lanes turn white. If you choose the Extract Lanes command, the Sequencing Analysis software will use the current tracker line locations to extract the data in all Used lanes, and put the extracted data into sample files.

To mark a single lane for extraction:

Step	Action
1	Click the lane marker to select the lane that you want to mark for extraction.
2	Either choose Mark Lane For Extraction from the Gel menu, or Press the Option key and click the lane marker. The lane marker becomes white with a red outline. When you choose the Extract Lanes command, the Sequencing Analysis software will use the current tracker line locations to extract the data in this and other similarly marked lanes, and put the extracted data into sample files. **Note** This option is useful if you want to re-extract data from only selected lanes or if you want to extract data from a lane that was not automatically extracted.

To unmark a lane that is marked for extraction:

Step	Action
1	Click the lane marker for the lane that you want to unmark.
2	Select Unmark Lane For Extraction from the Gel menu. When you unmark the lane, the lane marker turns from white to blue. **Note** An alternate method for unmarking a lane is to press the Option key and click the lane marker.

Adjusting Tracker Lines

Introduction

This section tell you how to adjust tracker lines:

- Show and hide tracker lines (page 4-31)
- Position and reshape tracker lines so they more accurately track the samples (page 4-32)
- Interpolating tracker lines (page 4-36)

Showing and Hiding Tracker Lines

The first time that the Sequencing Analysis software opens a gel file, it adds tracker lines to the gel image. Any time you open a gel file, all tracker lines are displayed. To make it easier to view the gel image and edit individual tracker lines, you can turn this line display on and off.

To turn off the display of unselected tracker lines:

Step	Action
1	Choose Hide Tracker Lines from the Gel menu. All unselected tracker lines disappear. If a lane is selected, the white tracker line for only that lane remains visible.

To select one tracker line to display:

Step	Action
1	Click the lane marker (♦) for the tracker line you want to view. Any time you select a lane or tracker line, the program either hides or grays out all the other tracker lines.

About Optimizing Tracker Line Locations

Because the Sequencing Analysis software normally calculates the data values for each lane by averaging the data from multiple channels, it is important that tracker lines be positioned over areas of data that display the strong fluorescent signal.

Automatic tracking may misinterpret lane positions or fail to follow the path of a lane completely under certain conditions. These conditions include the following:

- Failure to complete the Sample Sheet correctly will almost always cause a problem. The Tracker uses the Used lane information to determine if what it found corresponds to what the Sample Sheet

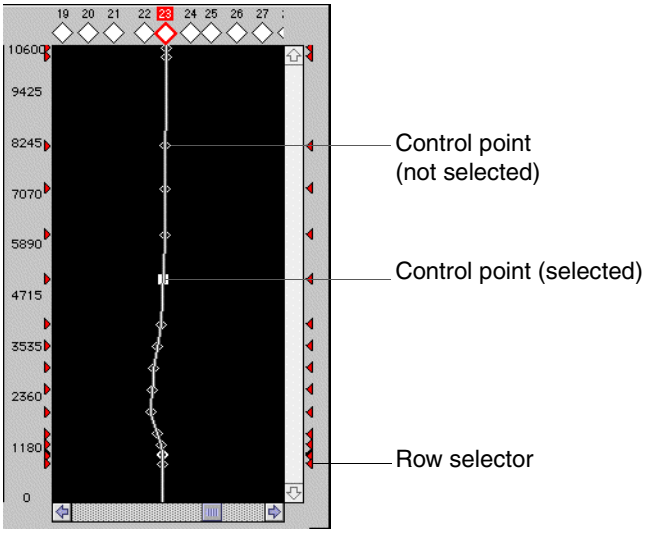
says was loaded. If too few or too many lanes are marked Used, the software must estimate which lanes to throw out and which to keep.

- Weak signals might cause the software to completely miss or be unable to follow a lane, especially if the gel ran aberrantly. Although the software creates a track for each used lane on the gel, the tracker lines might be incorrectly placed, indicating lane positions that do not exist or that are located elsewhere.
- You can verify optimal channel tracking by examining peak heights in the Slice view of the Gel File window. If the tracker line for a band is not optimally located, you may need to adjust it and re-extract the affected lane.

Reshaping Tracker Lines Using Control Points

A tracker line consists of a series of linked control points. You can optimize the position of the tracker line in the lane by moving the control points.

The control points are displayed on the line as hollow diamonds; selected control points are displayed as filled squares.



The operations that you can perform on control points are:

- Selecting and deselecting (page 4-33)
- Moving (page 4-33)
- Adding and deleting (page 4-33)

Selecting Control Points

Before you can move a control point, you must select it. Methods for selecting are:

- Click on the control point. This method only allows you to select one control point at a time (no shift-select allowed).
- Drag to select one or more control points.
- Select a complete row of control points by clicking on the row selector (red triangle) to the left or right of the gel image.
- If a control point is already selected, you can select the point immediately above or below it using the up or down arrow key.
- In vertical expand mode: if a control point is already selected, you can select the point immediately above or below it *and* scroll to that new point by holding the shift key down when pressing the up or down arrow key.

Moving Control Points

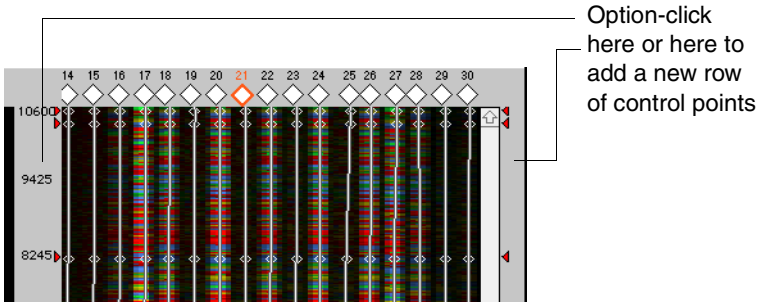
To optimize data extraction from the gel file, move the control points so as to center the tracker line over the most intense signal in the lane. Use the slice view as a guide. Methods for moving control points are:

- Drag the control point with the mouse. Only one point at a time can be moved with this method.
- Use the left and right arrow keys to move the control points in one channel increments. This method applies to all points selected.
- Hold the option key down and use the left and right arrow keys to move the control points in 0.1 channel increments. This method applies to all points selected.

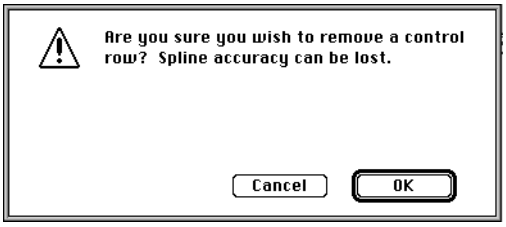
Adding and Deleting Control Points

Extra rows of control points can be added if necessary for finer control. Rows of control points can also be deleted:

- To add a row of control points, hold the option key down and click in the area between the existing row selectors.



♦ To delete a row of control points, hold the option key down and click on the selector ( ). Select OK in the dialog box that appears.



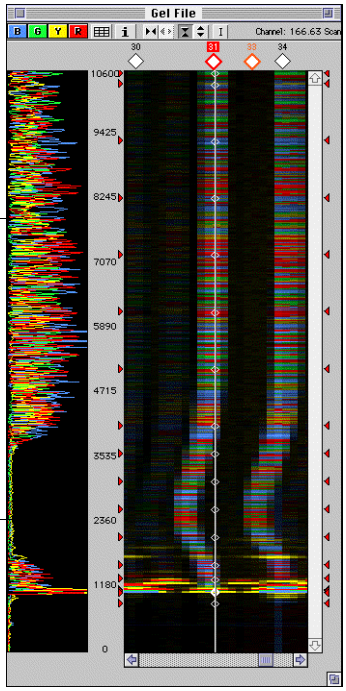
How to Reshape Tracker Lines

The following procedure describes how to reshape tracker lines to better follow the signal intensity in the lane.

To review and edit tracker line placement:

Step	Action
1	In the Gel File window, click the lane marker for the lane you want to review. You can also press Tab to move one lane to the right, or Shift–Tab to move one lane to the left, from the currently selected lane.
2	Inspect the placement of the tracker line on the lane. Does the tracker line seem to follow the brightest part of the lane (usually the center) for the full length of the lane? Or does the lane drift to the side while the tracker line does not? You may need to horizontally expand the gel image to see this clearly.

To review and edit tracker line placement: *(continued)*

Step	Action
3	Inspect the Slice view of the lane data. Are the peak intensities consistent along the full length of the lane? Or are some peaks very low because the tracker line is not located correctly on the lane? High peaks in upper part of Slice view because tracker line is in center of lane. No peaks in lower part of Slice view because tracker line is not on lane.  If you find that the location of the tracker line in the lane does not result in optimal peak heights, you can move the entire tracker line right or left (step 4). Or you can reshape the tracker line (step 5).
4	To move an entire tracker line: ♦ Click the lane marker for the tracker line you want to move. The lane marker becomes outlined in red, and the tracker lines for all other lanes become grayed out. ♦ Press the Left (or Right) Arrow key to move the entire line left (or right) one channel.

To review and edit tracker line placement: *(continued)*

Step	Action
5	Move the control points to follow the contour of the signal and to fall in the brightest portion of the lane. Use the methods described on pages 4-33 to 4-34 to adjust the control points. Add extra control points if necessary. Use the horizontal and vertical expand buttons to zoom in.
6	After you finish adjusting the tracker line positions for all lanes of interest, you can re-extract the data in those affected lanes. **Note** It is easier to adjust the tracker lines when you can see the entire length of the lane. Before extracting sample data, verify the exact tracker line positions in an expanded view.

Interpolating Tracker Lines

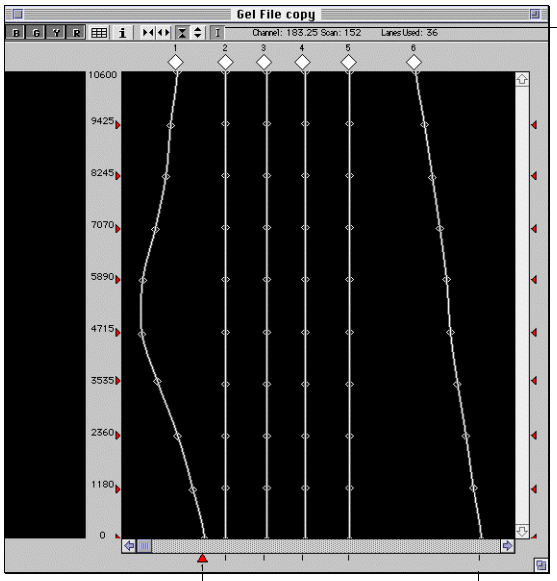
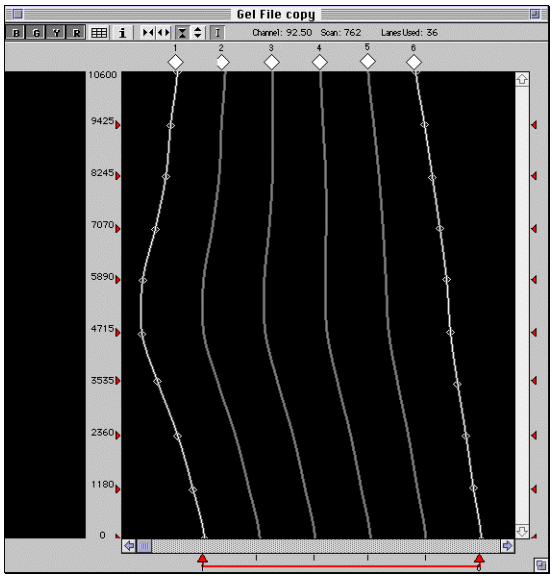
The shape of one lane is nearly always very similar to its neighbor lane. Curves or tilts in lanes tend to occur gradually across a gel, with each lane a little more curved or tilted than the next until the lane at the right is quite different in shape from the one on the left side of the gel.

Use the interpolation mode to quickly optimize the positions of a set of adjacent tracker lines.

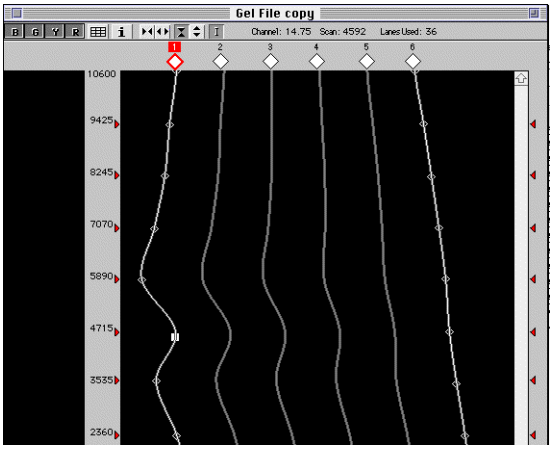
To interpolate tracker lines:

Step	Action
1	Click the interpolation mode button. 
2	Select the first of two lanes to be used as the interpolation guides. The positions of the tracker lines between the guide lines on the right and left will be interpolated from the guide lines. To select the lane, click in the tick mark at the bottom of the lane as shown below.

To interpolate tracker lines: *(continued)*

Step	Action
	 The Interpolation Mode button is shown depressed and tick marks appear at the bottom of the gel, indicating that the window is in interpolation mode. The interpolation guides are lanes 1 and 6. Selected guide lanes are indicated by the red triangle mark and the lane number. Click tick marks here and here to select the guide lanes.
3	Click the second of the two guide lines. The lanes between are interpolated. 

To interpolate tracker lines: *(continued)*

Step	Action
4	While in interpolation mode, if you move the guide control points, all the interpolated lines move proportionately. 
5	When you have finished, click the Interpolation Mode button to exit from the interpolation mode. Or click a different tick mark to choose a different guide lane, and further adjust tracker line positions. Individual tracker lines can be further adjusted as described in “Reshaping Tracker Lines Using Control Points” on page 4-32 above.

Note An alternative way to use interpolation mode is to adjust the tracker lines of the guide lanes *before* entering interpolation mode. When the Interpolation button is selected, the tracker lines are immediately interpolated between the already-adjusted guide lanes.

Tracking Lanes in the Gel File and Extracting the Data

Introduction

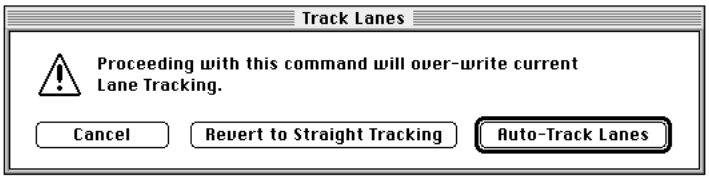
The Sequencing Analysis software provides three options for tracking and extracting the data in the gel file:

- To have the Sequencing Analysis software track the gel without extracting sample data, use the Track Lanes command (below).
- To have the Sequencing Analysis software both track the lanes and extract the data into sample files or BioLIMS database records, use the Track & Extract Lanes command (page 4-40).
- To have the Sequencing Analysis software extract sample data without changing the current lane tracking, use the Extract Lanes command (page 4-43).

Tracking the Gel File

When you choose the Track Lanes command, the Neural Net Tracker application calculates the tracking, but does not extract sample data. This command is useful to view the results of auto-tracking and, if necessary, to correct any tracking errors, before the Sequencing Analysis software extracts the lanes and generates sample data. It is also useful if you change your mind about the edited tracker information, and wish to redraw the lines based on original tracking information. If you have not saved the edited information, you can also simply close the gel file.

To track the gel file:

Step	Action
1	Choose Track Lanes from the Gel menu. The Track Lanes dialog box appears. 

To track the gel file: *(continued)*

Step	Action
2	Choose one of the three buttons. **Button** / **Function** Cancel — Cancels the tracking operation. Choose this if you do *not* want to lose the current tracking information. Revert to Straight Tracking — Adds straight, evenly spaced tracker lines to the gel. These lines can then be moved and reshaped as necessary to follow the center of each lane. Auto-Track Lanes — Places a tracker line on the center of each used lane in the gel. If necessary, you can reposition these lines after auto-tracking. **Note** To cancel the tracking process at any time, press Command–[period] and choose Cancel in the alert box that appears.

About Straight Line Tracking

For straight line tracking, the Sequencing Analysis software draws straight, evenly spaced tracker lines on the gel while ignoring any data that is present. The software applies straight line tracking the first time you open a gel file that has *not* been through automatic processing.

To change this tracking, you can either manually adjust the tracker lines or choose the Track command from the Gel menu to have automatic tracking applied.

Tracking the Gel and Extracting the Data

When you choose the Track and Extract Lanes command, the Tracker application calculates tracker lines, then extracts sample information from the tracked gel. This command is useful if you expect that automatic tracking will be satisfactory and there will be no need to correct tracking errors.

IMPORTANT When you Track and Extract a gel file, be sure the Sample Sheet associated with the gel file has the check box labeled “Used” selected for each sample you want to extract. The Sequencing Analysis software only extracts data from used lanes.

Disk Space Check Before Extraction

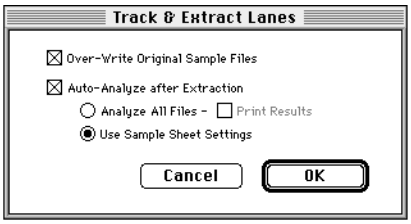
Before Sequencing Analysis begins extracting data from the gel file into sample files, it checks that there is sufficient space on the local hard disk to contain the sample files. If there is insufficient disk space, a message is written to the Error Log: "Error -34. Could not do this task because the disk is full."

The amount of disk space required depends upon the number of sample files you want to extract. Use the table below as a guide.

To extract this many lanes…	You need approximately this much disk space…
36	8 MB
64	16 MB
96	24 MB

Tracking and Extracting Data in Sample File Mode

To track and extract the gel file:

Step	Action
1	Choose Track and Extract Gel from the Gel menu. The Track & Extract Lanes dialog box appears. 

To track and extract the gel file: *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>2</td><td>Select the settings you want to use for this operation.

<table>
<tr><th>Setting</th><th>Description</th></tr>
<tr><td>Over-Write Original Sample Files</td><td>Overwrites any existing sample files that have the same name with the new sample data. Deselect this if sample files with these names already exist, and you want the Sequencing Analysis software to preserve the existing files and create new sample files for this data.

If you track and extract the gel file a second time without overwriting the original sample files, the software creates new sample files with a dot and a number appended to the original names (for example, if a file named MySample exists, the program will create files named MySample.1, MySample.2, etc.). (See also the Note on page 4-43.)</td></tr>
<tr><td>Auto-Analyze after Extraction</td><td>Analyzes the sample data after extraction is finished. De-select this to have the sample files extracted but not analyzed.

If Auto-Analyze after Extraction is selected, choose to Analyze All Files or Use Sample Sheet Settings. (See below.)</td></tr>
<tr><td>Analyze All Files</td><td>Analyzes all the sample files created from the gel file.</td></tr>
<tr><td>Print Results</td><td>After processing, prints the processing results for all the new sample files.</td></tr>
<tr><td>Use Sample Sheet Settings</td><td>Analyzes and prints only those files that are marked for analysis (and printing) in the Sample Sheet for the gel file. Select this option if you do not want to override sample sheet settings.</td></tr>
</table>
</td></tr>
<tr><td>3</td><td>Click OK to start tracking and data extraction. Sample files are written to the run folder in the same folder as the gel file. If the run folder does not exist, one is created.

Note To cancel the Track and Extract process at any time, press Command–[period] and choose Cancel in the alert box that appears.</td></tr>
</table>

Note If a series of numbered files exist, you discard one of the files, and you then have the Sequencing Analysis software re-extract that sample, the Sequencing Analysis software uses the first available number for the new file. For example, if SAMPLE01.1, SAMPLE01.2, and SAMPLE01.3 exist and you discard SAMPLE01.2, the Sequencing Analysis software names the next file SAMPLE01.2 (not SAMPLE01.4).

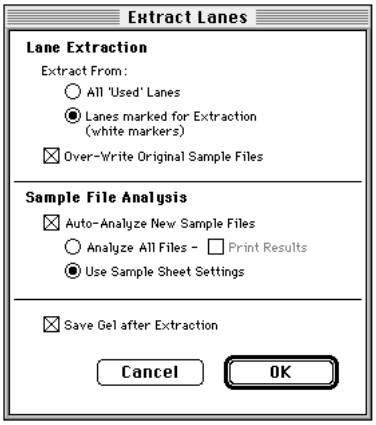
Extracting the Sample Data

Each time you change the tracker line positioning in a lane, or the information in the Sample Sheet, you must extract the sample data again to incorporate the new information into the sample files. Normally, you do this with the Extract Lanes command.

Disk Space Check Before Extraction

Before Sequencing Analysis begins extracting data from the gel file into sample files, it checks that there is sufficient space on the local hard disk to contain the sample files.

To extract the sample data:

Step	Action
1	Choose Extract Lanes (Command–L) from the Gel menu. The Extract Lanes dialog box appears. 

To extract the sample data: *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>2</td><td>Select the settings you want to use for this operation.

<table>
<tr><th>Setting</th><th>Description</th></tr>
<tr><td>All “Used” Lanes</td><td>Generates a new sample file for every lane marked Used in the Sample Sheet.</td></tr>
<tr><td>Lanes marked for Extraction</td><td>Generates a new sample data only for each lane with a white lane marker.

Note If a lane is not marked with a white lane marker and you want to mark it as modified, click the lane marker while holding down the Option key or select the marker and choose Mark Lane for Extraction from the Gel menu. The marker then turns white.</td></tr>
<tr><td>Over-Write Original Sample Files</td><td>If this check box is selected, the newly generated files will have the same names as the old files and the old files are lost. If this check box is deselected, a number is appended to the name for each newly generated file, and the original files are preserved.</td></tr>
<tr><td>Auto-Analyze New Sample Files</td><td>Automatically analyzes the sample data after extraction is finished. Deselect this to have the sample data extracted, but not analyzed.

If Auto-Analyze after Extraction is selected, choose to Analyze All Files or Use Sample Sheet Settings. (See below.)</td></tr>
<tr><td>Analyze All Files</td><td>Analyzes all the sample data created from the gel.</td></tr>
<tr><td>Print Results</td><td>After analysis, prints the analysis results for all the new sample data.</td></tr>
<tr><td>Use Sample Sheet Settings</td><td>Analyzes and prints the sample data as designated in the Sample Sheet.</td></tr>
<tr><td>Save Gel after Extraction</td><td>Saves tracker lines and other gel file modifications to the gel file after data extraction. If you do not select this, the settings used for the extraction are discarded when you close the gel file without saving manually.</td></tr>
</table>
</td></tr>
</table>

To extract the sample data: *(continued)*

Step	Action
3	When all of the information in the Extract Lanes dialog box is correct, click OK to begin extracting data. **Note** To cancel the Extract Lanes process at any time, press Command–[period] and choose Cancel in the alert box that appears.

Naming Sample Files

The following table shows the process the Sequencing Analysis software follows when naming generated sample files.

Note The table below refers to the Sample Sheet embedded in the gel file. When the Sample Sheet originally is set up in Data Collection software before the run, a default file name is set in the Collection Preferences. (For more information about Data Collection software, see your instrument user's manual.)

If the generated sample file...	Then the software names the file...
has an associated filename in the Sample Sheet \# | Used | File Name | Sample Name 1 | ☒ | 01•MySample | MySample	the filename in the Sample Sheet. For example: 01•MySample
does *not* have an associated file name \# | Used | File Name | Sample Name 1 | ☒ | | 2 | ☒ | | MySampleB	the lane number of the sample is placed after the name "Sample File". For example: Sample File 1, Sample File 2
has the same name as a previous sample file in the run sheet \# | Used | File Name | Sample Name 1 | ☒ | Test1 | 20u1.1 2 | ☒ | Test1 | 20u1.2 3 | ☒ | Test1 | 20u1.3	the lane number of the sample in parentheses added before the filename. For example: Test1, (02)Test1, (03)Test1
has the same name as a previously generated sample file in the sample file folder	the original filename with a dot and a number appended to it. For example, if lane 2 in the example above were re-extracted a second and third time without over-writing: (02)Test1.1, (02)Test1.2

Saving Gel Files

Saving the Gel File After Adjusting Tracking

IMPORTANT If you save changes to the gel file, the original tracking information is overwritten. You can retrieve the originally calculated tracking by choosing Track Gel from the Gel menu to retrack the gel.

If you select the check box labeled “Save Gel File Before Extraction” in the Generate New Samples dialog box, you need not manually save the gel file using the Save command.

Saving Selected Information and Archiving Gel Files

Because a gel file normally contains the raw data acquired by the data collection program, a gel image created by the Sequencing Analysis software, a copy of the data collection Sample Sheet, and a copy of an instrument file, the size of the file is normally 20–90 MB. You do not usually need to keep a gel file once the tracking is verified and the sample data are extracted from it. If desired, you can keep parts of the information, while discarding the image.

IMPORTANT Do not discard any gel file until you have verified the tracking and taken any required corrective action.

To Store a Gel File Temporarily

If you are running the ABI 373 with Data Collection version 1.2, the gel file is overwritten on the hard disk every time a new gel file is created, because the default name for all new gel files is *Gel file*.

If you want to temporarily save the current gel file on your hard disk, give the current file a new filename before you begin the next data collection run. It is not recommended to keep more than one gel file on the hard disk if you are using the original computer, as there is limited disk space.

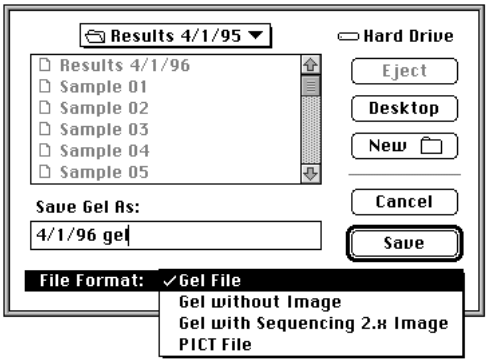
To Archive a Gel File

Gel files are too large to fit on floppy disks. For long-term storage, use magnetic tapes, removable cartridge drives, or optical drives to archive gel files.

To Save Selected Information From a Gel File

Use the Save As command from the File menu. The default filename is the original filename plus the word “copy”.

In the Save Gel As dialog box you specify the file format. The file formats are described in the table below.



File Format	Description
Gel File	Saves the entire gel file, including the raw run data, Sample Sheet data, instrument file information, and gel image with tracker lines. This option typically creates a 20–90 MB file.
Gel Without Image	Saves everything in the gel file except the gel image. A gel image can be recreated from this file later if it is needed. This option typically reduces the file size by about one third.
Gel With Sequencing 2.x Image	Saves the gel file in a special format, so it can be displayed by the Sequencing Analysis software version 2.1.2 (and earlier versions). This option typically reduces the file size by about one third. **Note** Only gels with 194 channels and 36 or fewer lanes can be saved in this way.
PICT File	Saves the gel image as a PICT file. The PICT file can be viewed on screen and printed, using Simple Text or other Macintosh programs. It cannot be opened by the Sequencing Analysis program or used to track or extract data later. This option typically creates a 300K file.

Printing the Gel Image

To Print a Gel Image

The Print command on the File menu is disabled when the Gel window is active.

To print the gel image, use this work around:

Step	Action
1	Take a screen capture by typing Command–Shift–3.
2	Print the resulting Picture file from another application such as SimpleText.

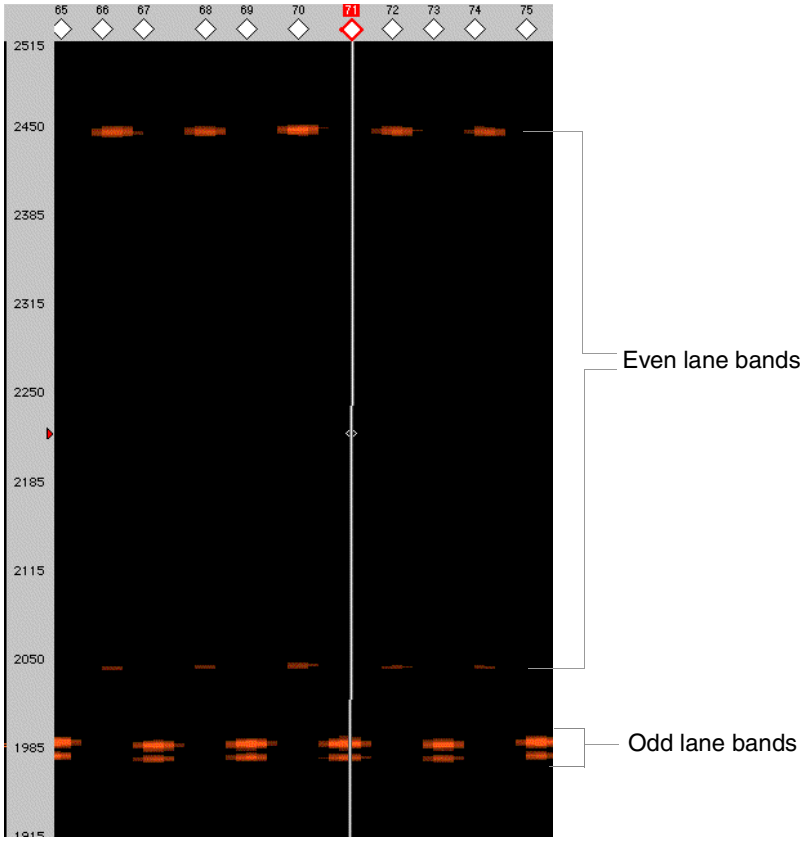
When a Fifth Dye Is Used

Lane Guide

The ABI PRISM® Lane Guide™ Lane Identification Kit may be used to improve the lane tracking process with a fifth dye. The kit contains two unique, fluorescently labeled DNA ladders, one of which is added to odd -numbered samples, the other to even-numbered samples. It is very important to follow the loading instructions in the protocol entitled *ABI PRISM Lane Guide Lane Identification Kit* (P/N 4313804).

Orange Bands

After electrophoresis, when you display the gel file in Sequencing Analysis software, the fifth dye shows up as a fifth color (orange). The display of orange can be toggled on and off just like blue, green, yellow, and red. The odd lanes have two bright orange bands close together (at 100 and 112 base pairs). Even lanes have two bright orange bands somewhat farther apart (112 and 150 base pairs). This is easiest to see when the entire gel image is displayed. For purposes of clarity, a close-up of the image is shown below.



The Tracker looks for the presence of Lane Guide bands after it tracks each lane. The Lane Guide bands help the Tracker to eliminate false positive or false negative lanes that were found during normal tracking.

Lane Confidence Value

When the Tracker has constructed a set of lanes, it compares tracking results with the information stored in the Sample Sheet. It then develops a confidence value based on the lane number assignment match with the Sample Sheet. If the lane confidence value is 70% or greater, the lanes found are considered equal to the lanes expected. (See below for more information about the lane confidence value.)

Reviewing a Gel File with a Fifth Dye

When you review the gel file of a fifth dye image:

- Check the lane assignment confidence value that is written to the Error Log. (See page 6-26.)

  If the lane confidence value is 70% or greater, we recommend a quick check to verify that the Tracker has identified the lanes correctly.

  If the lane confidence value is less than 70%, this is an indication that you should check the lane assignment. Visually check the gel image to ensure that odd and even codes appear in alternate lanes.
- If necessary, adjust the lane markers to correct lane numbering errors. (See page 4-25.)
- If necessary, adjust the placement of the tracker lines. (See page 4-39.)
- If you change any of the gel file or Sample Sheet information after extracting the sample data, re-extract the data from the edited lanes to regenerate the information in the sample files. (See page 4-43.)

Processing Sample Files 5

Overview

In This Chapter

This chapter explains how to set up batches of sequence data for processing and how to change the processing parameter values.

Note To learn about each of the processing parameters in detail and how to select the best parameter values for your situation, see Chapter 6, “The Processing Parameters.”

Note When Sequencing Analysis is in BioLIMS® mode, the Sequence Manager window replaces the Sample Manager window, and sample files are replaced by sequence records in the database.

About Base Calling

Sample file analysis (base calling) is the primary activity of the Sequencing Analysis software. During base calling, the software identifies each base in the sample and the order in which the bases are arranged. The software also marks locations where there is some question about the base identification, as when two bases seem to occur at the same position. This allows you to determine whether the ambiguity is caused by uneven base migration, a heterozygote condition, or some other irregularity.

Factura Processing and Printing

The Sequencing Analysis software can also manage further processing of files by the Factura™ Feature Identification Software and the printing of all processing results. All this is done from the Sequencing Analysis Sample Manager window.

Note In BioLIMS mode, Sequencing Analysis does not do Factura processing. To process sequence records with Factura, open the Factura application itself. Factura can be set to regularly scan the BioLIMS database for sequences that need Factura processing. For more information about Factura, see *ABI PRISM Factura Feature Identification Software User's Manual* (P/N 904946)

Overview of Sample File Processing

To process one or more sample files:

- Add the files to the Sample Manager window. (See "Adding Sample Files to the Sample Manager Window" on page 5-8.)
- If necessary, change the processing parameter values. (See "Changing the Processing Parameter Values" on page 5-14.)
- Start the processing operation. (See "Processing the Sample Files" on page 5-20.)
- Check for any problems that might have occurred during file processing. (See "Checking for Processing Problems" on page 5-22.)

Reasons to Reprocess Files After Automatic Processing

There are many reasons to reprocess a sample file or a group of files after automatic file processing is finished. Some reasons are:

- To correct initial setup errors (for example, the wrong instrument file was specified in the data collection program Sample Sheet).
- To change the point where the software stops calling bases either to exclude poor-quality data near the end of the run or to improve base calling of short PCR products.
- To use a different Basecaller or change the spacing estimate in order to improve the analysis results.
- To analyze the new sample data after adjusting the tracking on a gel that ran poorly.

About the Sample Manager Window

Introduction

The Sample Manager window allows you to list sample files you want processed by the Sequencing Analysis software, and to choose various analysis parameter values. This section describes the parts and functions of the Sample Manager window.

About Auto-Analyze

If Auto-Analyze was chosen in the data collection program, this list is automatically filled out after the instrument run, and the samples are processed using the values specified in the Data Collection Sample Sheet and the information entered on the Preferences page in the Sequencing Analysis software.

Opening and Closing the Window

When the Sequencing Analysis program starts, an empty Sample Manager window appears on the screen.

To Open

To open the Sample Manager window, choose Show Sample Manager from the Window menu.

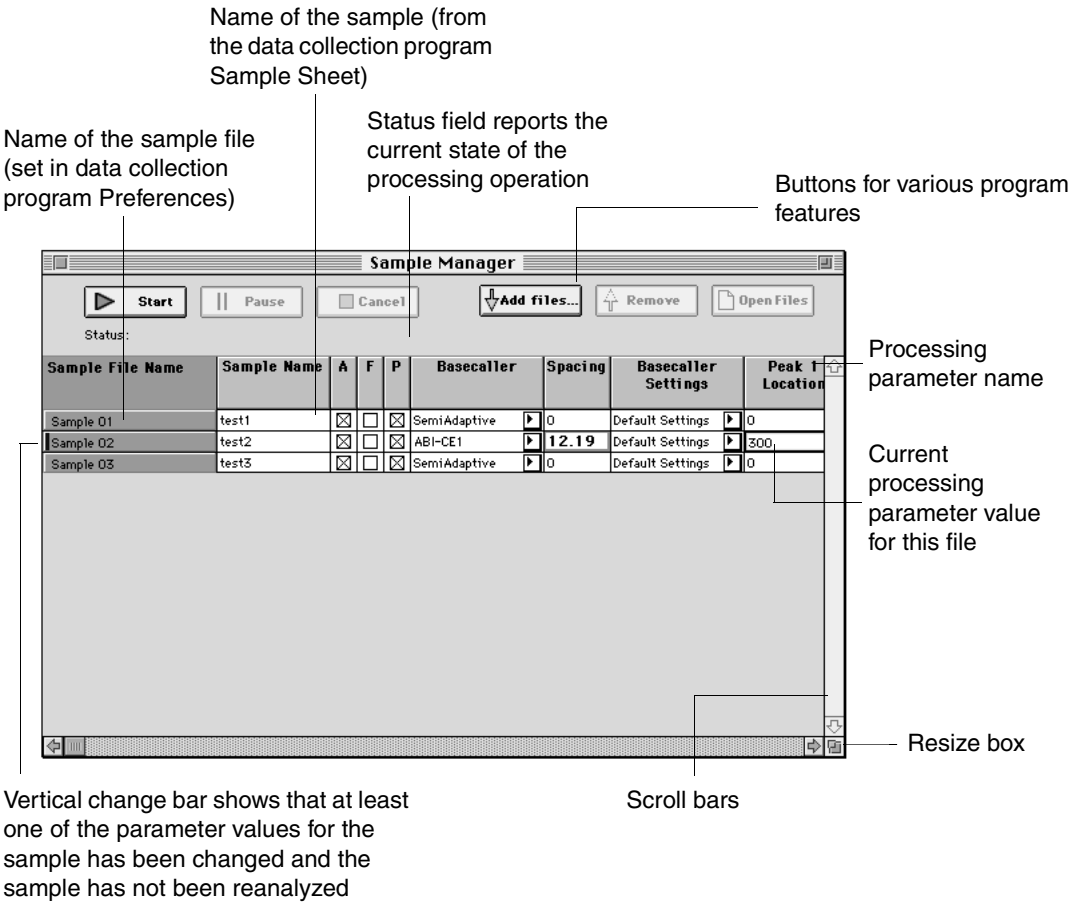
To Close

To close the Sample Manager window, click the Close box at the top left corner of the window, or choose Close from the File menu, or choose Hide Sample Manager from the Window menu.

Note If the Sample Manager window is closed and then reopened during a single Sequencing Analysis session, the contents of the window remain the same. The current list contents are always discarded when the Sequencing Analysis program quits.

Parts of the Sample Manager Window

When an empty Sample Manager window appears on the screen, only the upper left portion of the Sample Manager is visible in the window. After files are added, the window shows the current processing parameter values for each file.



Description of Parts

This table describes the parts of the Sample Manager Window. Most of the parts are labeled in the figure above.

Item	Description
Start button	Starts processing of the files in the list.
Pause button	Temporarily stops processing of the current file.
Resume button	Starts file processing, beginning at the point where processing was temporarily paused. The Resume button becomes visible only after the Pause button is selected.
Cancel button	Immediately stops processing of the current file, and cancels the entire processing operation.

Item	Description
Add Files button	Opens a directory dialog box in order that sample files can be added to the list.
Remove button	Removes the selected file(s) from the list.
Open Files button	Opens the selected file(s).
Status field	Displays messages about the current state of the processing operation.
Change bar	The thick vertical line that appears to the left of the Sample File Name if any processing parameter value is changed for that file before reanalyzing the file.
Scroll bars	Use the scroll bars to bring other parts of the window contents into view.

Processing Parameters

The processing parameters are defined briefly in the table below. For more details about processing parameters, see Chapter 6, “The Processing Parameters.”

Note If you change the value for a processing parameter in the Sample Manager window, or if the software encounters a problem with the selected value, the condition is reflected by a change in the appearance of the value in the window. This is explained in the table, “Sample Manager Field Status Indicators,” on page 5-7.

Parameter	Description
Sample File Name	The name of the sample file as it appears on the hard disk. This is originally taken from the File Name field in the Sample Sheet. The Sample File Name cannot be changed from within the Sample Manager window, but after the file is created, it can be changed in the Finder using normal Macintosh® operations. If you double-click on the Sample File Name for a file, the Sequencing Analysis software opens that sample file.
Sample Name	The name of the sample, as it is recorded in the sample file. The name is originally taken from the data collection Sample Sheet. The name can be edited in this window, but changing the name will un-link the sample file from the Sample Sheet.
A	If this box is checked, the file is analyzed when the Start button in the Sample Manager window is selected. The color of the check box indicates whether analysis was successful (green), failed (red), or has not yet been started (no color). The default setting for this check box is selected in the Sample Manager Defaults page of the Preferences dialog box; click the check box to change the setting.

Parameter	Description
F (This parameter is not visible in BioLIMS mode.)	If this box is checked, the file is passed to the Factura program for further processing after processing by the Sequencing Analysis software. The color of the check box indicates whether the processing was successful (green), failed (red), or has not yet been started (no color). The default setting for this check box is selected in the Sample Manager Defaults page of the Preferences dialog box; click the check box to change the setting. For more information about the Factura software, see the *Factura Feature Identification Software User's Manual*.
P	If this box is checked, the selected information for this file is printed after all processing is complete. The color of the check box indicates whether printing was successful (green), failed (red), or has not yet been started (no color). The default setting for this check box is selected in the Sample Manager Defaults page of the Preferences dialog box; click the check box to change the setting.
Basecaller	The Basecaller program used to identify bases during the most recent file analysis. Select from the pop-up menu to change the Basecaller. (For more information about Basecallers, see page 6-45.)
Spacing	The average number of data points between peaks. This is defined by the Basecaller program during analysis. Type in a new value to manually change the Spacing.
Basecaller Settings	The user-created set of rules and values that is used by the Basecaller program to decide the endpoint for file analysis. Select a different rule set from the pop-up menu or define a new rule set in the Basecaller Settings page of the Preferences dialog box.
Peak 1 Location	The scan number that marks the beginning of the first real base peak in the file. This is defined by the Basecaller during analysis. Type in a new value to change the Peak 1 Location.
Start Point	The point in the sample file at which the Basecaller program starts analyzing data. Type in a new value to change the Start Point. 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.
Stop Point	The point in the file at which the Basecaller program stops analyzing data. This is controlled by the values selected for the Basecaller Settings. Type in a new value to change the Stop Point. The number must be the scan number, not the base number, for the point where analysis is to stop.
DyeSet/Primer file	The DyeSet/Primer (mobility) file used for analysis. This is originally taken from the data collection Sample Sheet. Changes made to the Sample Sheet in the gel file should appear here. Select from the pop-up menu to change the DyeSet/Primer file.

Parameter	Description
Instrument file	The instrument file used during analysis. The instrument file contains the matrix information that is used to correct for spectral overlap of the fluorescent dyes. This is originally taken from the data collection Sample Sheet, but changes made to the Sample Sheet in the gel file should appear here. Select from the pop-up menu to change the instrument file.

Sample Manager Field Status Indicators

The font style used to display a processing parameter value provides additional information about that value, as shown in the following table.

Indicator Font	Meaning	Parameters Affected
Plain Text	Default, no action taken by user	All
Black Bold Text	User defined variable	Start, Peak 1 Location, Stop, Spacing
Red Bold Text	Basecaller could not calculate	Spacing
Outline Font	File not found	DyeSet/Primer, Instrument file (File not found in ABI folder in the System Folder)
Blue Bordered Text	User modified, but not yet analyzed	All

Adding Sample Files to the Sample Manager Window

Introduction

Sample files can be added to the Sample Manager window in three ways, as described below. In addition, the Sequencing Analysis software can open the window and add files as part of automatic analysis. Files are arranged in the list in the order in which they were added to the list.

To Add One or More Files from the Finder

To add one or more sample files from the Macintosh Finder to the Sample Manager, follow the steps in the table below. (This applies in Sample File mode only.)

Step	Action
1	Drag the icon(s) for the file(s) you want to add onto the Sequencing Analysis program icon. A sample window opens for each file. If the Sequencing Analysis application is not already open, dragging the sample files onto the icon will open the application.
2	For each sample window: a. Click on the sample file window to make it active. b. From the Sample menu, choose Add to Sample Manager. The sample is then listed in the Sample Manager window.

To Add a File from the Sample Window

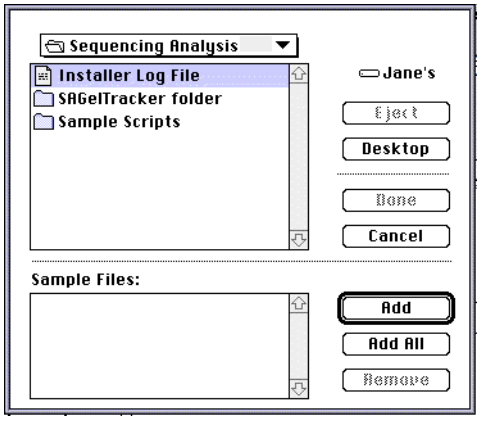
To add an open sample file to the Sample Manager:

Step	Action
1	Make sure the sample window is active.
2	From the Sample menu, choose Add to Sample Manager. The sample is then listed in the Sample Manager window.

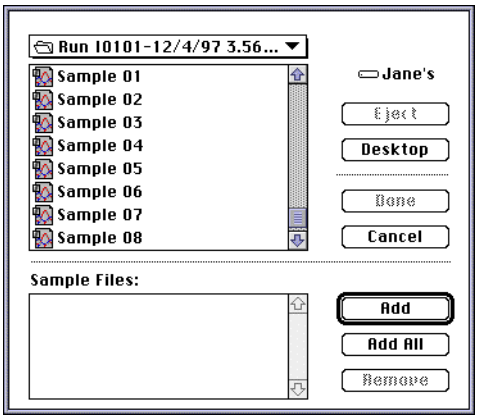
To Add Files from Within the Sample Manager Window

Follow this procedure to add sample files from within the Sample Manager Window.

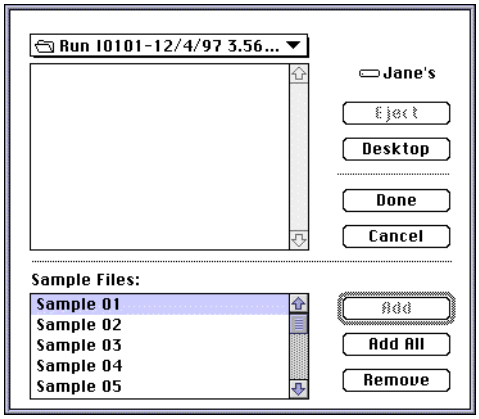
To add sample files from within the window:

Step	Action
1	Click the Add files button in the Sample Manager window, or choose Add Files from the Manager menu. A directory dialog box appears.  The upper part of the dialog box is similar to a standard directory dialog box, but it displays only the names of folders and sample files.

To add sample files from within the window: *(continued)*

Step	Action
2	In the upper list box, locate and open the folder that contains the files you want to add to the Sample Manager.  **Note** If a file is already included in the Sample Manager window, that file name is not visible in the upper list.
3	Add the files that you want in the Sample Manager to the Sample Files list at the bottom of the dialog box. **To add…** / **Do this:** a single file to the list — select the file, then choose Add, or double-click the name of the file. all the files to the list — choose Add All. some of the files to the list — either add them individually, or choose Add All, then use the Remove button to remove the files you do not want in the list **Note** To find a file in the list quickly, highlight the name of any file in the list. Then begin typing the name of the file you want. As you type, the highlight moves to the first file name that matches the character typed.

To add sample files from within the window: *(continued)*

Step	Action
4	When all the files you want are in the lower lists, click the Done button to close the dialog box and add the files to the Sample Manager. 

Note **In BioLIMS mode:** For information about adding sequence data from the BioLIMS to the Sequence Manager window, see “Using the Collection Browser Window” on page 3-19.

Moving and Removing Sample Files from the Sample Manager Window

Introduction

A sample file can be removed from the Sample Manager window at any time except when the program is currently processing that file. You can also rearrange the order that the sample files appear in the Sample Manager window.

Note It is not necessary to remove a file from the list in order to avoid processing it. The Sequencing Analysis software decides whether or not to process files based on the current information in the A, F, and P check boxes. If the check box for a processing option is empty, that process for that sample is skipped.

To Move a File Within the Window

Follow this procedure to move a file to a new location in the Sample Manager list.

Step	Action
1	Click the name of the file in the Sample File Name column. The entire row becomes highlighted.
2	Hold down the Option key while dragging the Sample File Name to the new location in the column.

To Remove a Sample File

Follow this procedure to remove a single sample file from the Sample Manager window.

Step	Action
1	Click the name of the file in the Sample File Name column. The entire row becomes highlighted.
2	Press the Delete key, or click the Remove button at the top of the window, or choose Remove Files from the Manager menu. The Sequencing Analysis software removes that file from the list.

To Remove Multiple Files

To remove multiple files from the Sample Manager window:

To remove…	Do this:
all the files	a. Choose Select All from the Edit menu. b. Choose the Remove button or the Delete key.
several adjacent files	a. Click the Sample File Name of the first file in the group. b. Hold down the Shift key and click the Sample File Name of the last file in the group. c. Choose the Remove button or the Delete key.
multiple files that are not next to each other	a. Hold down the Command button while clicking the File Name of each file to be removed. b. Choose the Remove button or the Delete key.

Changing the Processing Parameter Values

Introduction

Processing parameters are instructions and program settings that are used by the Sequencing Analysis software during file processing. The parameter values that are used for each file are the values currently displayed in the Sample Manager window.

For example, if the A check box is selected, that file is analyzed (base called) during file processing. If ABI-CE2 is selected for the Basecaller parameter, the base calling is done by the ABI-CE2 Basecaller.

You can change some processing parameter values in the Preferences dialog box, some in the Sample Manager window, and some at either location. For all parameters except Basecaller Settings, the value entered in the Sample Manager window always overrides the value in the Preferences dialog box.

Changing Parameter Values in the Sample Manager Window

In the Sample Manager window, processing parameter values for individual sample files or for groups of files can be changed. These changes affect only the files currently listed in the window.

This section explains how to change parameter values. Chapter 6, “The Processing Parameters,” explains how to decide which values are appropriate for your situation.

To Change a Parameter Value for One File

If…	Then…
the field has a check box	click once in the check box to select or de-select it.
the field has a pop-up menu	point to the pop-up menu icon, and press the mouse button to open the menu. Highlight the value you want to select, then release the mouse button.
the field has neither a check box nor a menu icon	double-click the field to activate the text-entry cursor. Then type in the new value.

To Change the Same Value for Several Files

Step	Action
1	Change the parameter value in the first field where you want to make the change.
2	While the new value is still highlighted, hold down the Shift key (group select) or the Command key (multiple select) and select the remaining fields that you want to change to the new value.
3	Choose Fill Down from the Edit menu. The Fill Down command copies the value in the first selected field to all the other selected fields in the column.

To Revert to the Original Values for the Parameters

Step	Action
1	Select the Sample File Name for the sample that you want to change back to the original values. The entire row becomes highlighted.
2	Choose Pre-Analysis Settings from the Manager menu. All the values for the selected sample are changed to the value specified in the sample file at the time of original analysis (no matter how many times the values have been modified since they were entered in the file).

To Clear the Contents of One or More Fields

Step	Action
1	Select all the fields that you want to clear.
2	Choose Clear from the Edit menu.

Changing Parameter Values in the Preferences Dialog Box

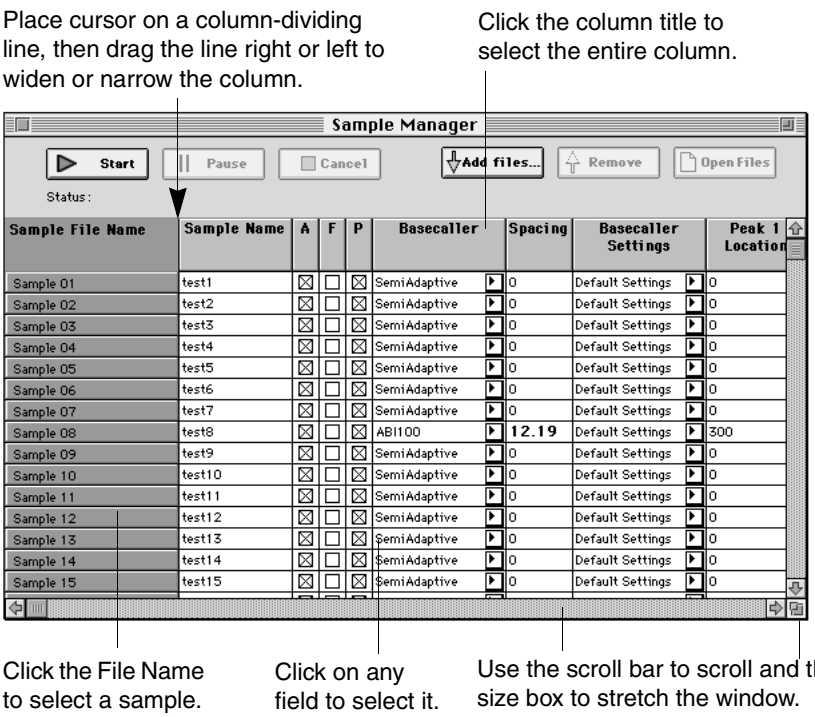
When you change a processing parameter value in the Preferences dialog box, the new value is used for all future processing until you change the value again or temporarily override that value for selected files in the Sample Manager window.

Note Changes that you make in the Preferences dialog box affect only files that you add to the Sample Manager window **after** you make the change. Files that are already listed in the window are not affected.

For how to change parameter values in the Preferences dialog box and how to decide which values are appropriate for your situation, see Chapter 6, “The Processing Parameters.”

Navigating the Sample Manager Window

Introduction The Sequencing Analysis software includes an assortment of keyboard shortcuts to enable you to easily move around, and make changes, in the Sample Manager window.



Moving Within Rows and Columns

There are keyboard shortcuts for moving from column to column and row to row in the Sample Manager window.

To Move from Column to Column Within One Row

Step	Action
1	Press the Tab key or the Right Arrow key (→) to move to the next right column.
2	Press the Shift+Tab keys or the Left Arrow key (←) to move to the next left column.

To Move from Row to Row Within One Column

Step	Action
1	Press the Return key or the Down Arrow key (↓) to move down one row.
2	Press the Up Arrow key (↑) to move up one row.

Selecting Fields and Samples

To Select a Field to Edit

Click once to highlight the field you want to edit.

To Select Samples

When you select a sample in the Sample Manager window, the entire row containing the sample is selected. Select samples as follows:

- To select one sample and to de-select all other samples, click the Sample File Name for the sample.
- To select a consecutive range of samples:
  - Click the Sample File Name of the first sample in the group.
  - Hold down the Shift key and click the Sample File Name of the last sample in the group.
- To select some samples and leave other (discontinuous) samples un-selected, hold down the Command key and click the Sample File Names for the samples you want to select.

To Select Entire Columns

Click the title of the column.

To Change Column Width

Place the cursor over the vertical line to the right of the column title. When the cursor symbol changes from  to  , hold down the mouse button and drag the line to the right to widen the column or to the left to narrow the column width.

You can change the width of all columns *except* the Sample File Name column.

To View Additional Information in the Window

Some of the rows or columns may not be visible on the Sample Manager window. To display additional information in the Sample Manager window, do either of these things:

- Use the size box  in the bottom right corner of the Sample Manager to stretch the window.
- Click on the vertical or horizontal scroll bar to see another part of the window contents.

Processing the Sample Files

Introduction

Once you start the processing operation, the Sequencing Analysis software processes each of the sample files in the list according to the parameter values selected for the file.

Starting File Processing

There are two ways to begin the processing of files in the Sample Manager list:

- Click the Start button in the Sample Manager window, or
- Choose Start from the Manager menu.

During Processing

During processing, the status of the processing operations appears in the Status field (above the file list). While a sample file is being processed, the Start button becomes inactive and the Pause, Stop, and Cancel buttons become active.

Note When the Sequencing Analysis software performs base calling, it stores the base calls as the “original” results. If you edit the sequence, the “original” results are kept and the edited sequence is kept as the “most recent” sequence. Each additional time that you edit the sequence the “most recent” sequence is updated.

Pausing Processing

There are two ways to temporarily pause processing:

- Click the Pause button, or
- Choose Pause from the Manager menu.

The program pauses processing on the current file. When processing is paused, the Pause button becomes the Resume button. When you are paused, you can edit the contents of the fields in the Sample Manager. For example, if you started the analysis and noticed that you had set the P field in error, you could pause, edit the print boxes, and resume processing.

Resuming Processing

There are two ways to resume processing:

- Click the Resume button, or
- Choose Resume from the Manager menu.

The program resumes processing, beginning at the point where it was paused.

Cancelling Processing

There are three ways to cancel processing:

- Click the Cancel button
- Press Command–[period]
- Choose Cancel from the Manager menu

The program immediately stops processing the current file and cancels the whole processing operation. The phrase “Sample___: Analysis Cancelled” appears in the Status field.

(To process the remaining files, you must click the Start button or choose Start from the Manager menu.)

Checking for Processing Problems

Introduction

After processing is completed, look at the A, F, and P columns in the Sample Manager window. These three narrow columns are the parameters that specify the processing actions to be taken on the listed files. (See summary table below or page 5-5 for details.)

Parameter	Processing Action
A	Analyze the sample file in the Sequencing Analysis program.
F	Process the sample file with the Factura program.
P	Print the specified views of the sample file.

Check Boxes After Processing

Following processing, each check box is colorless, green, or red, depending on the processing outcome.

Check Box Color	Processing Outcome
No color	The selected action was cancelled or never started
Green	The selected action was successfully completed
Red	The selected action failed

If the Analysis Box Is Red

If file analysis fails (the analysis box is red), you can:

- Change one or more parameters in the Sample Manager window and reanalyze the affected sample files, either individually or as a group.
- Check the Sequencing Analysis Error and Command Logs for information about problems that occurred during analysis (for details, see “Reviewing the Sequencing Analysis Error Log” on page C-13 and “Reviewing the Sequencing Analysis Command Log” on page C-15).

Check the run conditions to see if any problems occurred during data collection.

If the Factura Box Is Red	If Factura processing failed, you can check the Factura Log and Factura user's manual for possible causes.
If the Printing Box Is Red	If printing failed, you can check your printer manual and the Troubleshooting chapter of this manual for possible causes, or call Applied Biosystems Technical Support.

The Processing Parameters 6

Overview

In This Chapter

This chapter explains the sequence data processing parameters, how certain values are calculated by the Sequencing Analysis software, and how to select parameter values that are appropriate for your data.

About Processing Parameters

A processing parameter is a word, phrase, or check box that tells the Sequencing Analysis software what to do at a certain point during file processing. The Sequencing Analysis parameters described in this chapter generally determine how base calling, Factura™ analysis, and printing are carried out.

For example, if you select the A check box for a file in the Sample Manager window, that file will be analyzed (base called) during file processing. If you select ABI-CE2 for the Basecaller parameter, the base calling will be done by the ABI-CE2 Basecaller.

You can change some processing parameter values in the Preferences dialog box, some in the Sample Manager window, and some at either location. For all parameters except Basecaller Settings, the value entered in the Sample Manager window overrides the value in the Preferences dialog box.

Parameters in the Sample Manager Window

About Sample Manager Parameters

You can change many of the processing parameter values in the Sample Manager window. You can apply these changes to a single file, some of the files, or all of the files in the window.

The following sections (pages 6-3 to 6-20) describe each of the parameters in the Sample Manager window, and discuss the various factors to consider before you select a new value. For an explanation of how to select and change a parameter value for one or more files, see “Changing the Processing Parameter Values” on page 5-14.

The Sample Manager Parameters Listed

The table below lists the parameters that are visible in the Sample Manager Window.

The Sample File Name Parameter

About the Sample File Name Field

This is the name of the file that contains the sample information. This name appears with the icon for the sample file when viewed in the Macintosh® Finder on the hard disk.

Changing a Sample File Name

You cannot change the Sample File Name from within the Sample Manager window. If you double-click on that field, the Sequencing Analysis software opens that sample file.

If you need to change the Sample File name, do it through the Finder, changing the Sample File name as you would change the name of any other file.

Note If you are using Sequencing Analysis in BioLIMS® mode, you cannot change the name of the sequence in the database.

The Sample Name Parameter

About Sample Name Field

This is the name of the sample. Unless you have edited it, this is the sample name that was originally entered in the Sample Name column of the Sample Sheet for the run, in Data Collection.

Note This is distinct from the name of the sample file. However, the same name may be given to both the Sample File and the Sample.

Changing a Sample Name

You can edit the Sample Name in the Sample Manager window. The new name is recorded in the sample file when you move the cursor to a different field. Changing the Sample Name is immediate and permanent

Note The information in the sample file is normally connected to the Sample Sheet information through the Sample Name. You break this connection when you change the Sample Name. If necessary, you can use the run time and lane number to find the source of the sample file information. It is simpler to keep the original Sample Name until you no longer need the connection.

The A Parameter

About the Analysis Check Box

When this check box is selected, the Sequencing Analysis software analyzes (base calls) the file as a part of the processing operation.

When a file is added to the Sample Manager as part of automatic file processing, the software sets this check box (selected/de-selected) to match the A check box in the Sample Sheet. When you manually add a file to the Sample Manager window, the software sets this check box to match the A check box on the Sample Manager Defaults page of the Preferences dialog box (see page 6-33).

Check Box Status

The color of this check box indicates the analysis status.

If the check box is…	Then the analysis…
green	succeeded
red	failed
no color	has not been started since the sample was added to the Sample Manager window, or the sample was analyzed previously and is still in the Sample Manager.

If the check box is red, see “Changing the Processing Parameter Values” on page 5-14 and “Reviewing the Analysis Results” on page 7-21.

The F Parameter

About the Factura Check Box

When this check box is selected, the Factura program processes the file as part of the processing operation.

The Factura program, which is bundled with the Sequencing Analysis software, allows identification of heterozygous base positions and quick cleanup of sequences before alignment. For more details, see the *ABI PRISM Factura Feature Identification Software User's Manual*.

If both the A and F check boxes are selected, the file is processed by Factura *after* it is analyzed by the Sequencing Analysis software.

When a file is added to the Sample Manager, the software sets this check box (selected/de-selected) to match the F check box on the Sample Manager Defaults page of the Preferences dialog box (see page 6-33).

Check Box Status

The color of this check box indicates the status of Factura processing.

If the check box is…	Then Factura processing…
green	succeeded
red	failed
no color	has not been started since the sample was added to the Sample Manager window, or the sample was analyzed previously and is still in the Sample Manager.

If the check box is red, review the Factura Log. For more details, see the *ABI PRISM Factura Feature Identification Software User's Manual*.

Applying Factura to BioLIMS Data

When the Sequencing Analysis program is in BioLIMS mode, the Factura check box is not accessible from the Sequence Manager window. To use Factura to process sequences in a BioLIMS database, refer to the *ABI PRISM Factura Feature Identification Software User's Manual*.

The P Parameter

About the Printing Check Box

When this check box is selected, information from the file is printed as part of the processing operation. If you also select the A and/or F check boxes, printing is done *after* all other processing of that file is complete.

This check box controls *whether or not* printing occurs; you can specify *which* pages to print in the Printing Preferences page of the Preferences dialog box (see page 6-35).

When a file is added to the Sample Manager as part of automatic file processing, the software sets this check box (selected/de-selected) to match the P check box in the Sample Sheet. When you manually add a file to the Sample Manager window, the software sets this check box to match the P check box on the Sample Manager Defaults page of the Preferences dialog box.

Check Box Status

The color of this check box indicates the printing status.

If the check box is…	Then printing…
green	is completed
red	failed
no color	has not been started since the sample was added to the Sample Manager window, or the sample was analyzed previously and is still in the Sample Manager.

If the check box is red, check your printer connections, referring to your printer manual if necessary.

The Basecaller Parameter

About the Basecaller Program

The Basecaller is the program that identifies the bases in a sample file. You can choose the Basecaller for a file from the Basecaller pop-up menu in the Sample Manager.

For a detailed explanation of Basecallers and how to select the best Basecaller for a sample file, see “About Basecallers and Base Calling” on page 6-45.

The Spacing Parameter

About the Spacing Value

Spacing is defined by the number of scan points from the crest of one peak to the crest of the next peak. To calculate spacing, the Basecaller averages the peak-to-peak distance between scan 1000 and 2000 in the raw data relative to the Peak 1 Location.

Spacing can be changed only in the Sample Manager window. You may or may not be able to use the same spacing value for all samples in the run.

Note If you set this value to 0, the Basecaller recalculates the spacing.

Changing the Spacing for a Sample

Note Since the Basecaller program calculates spacing based on an average later in the run, if you are sequencing short PCR products, you are particularly likely to benefit from entering spacing calculated from early in the run as described in the procedure below.

To change the spacing for a sample:

Step	Action
1	Open the sample file.
2	Click the Raw Data view button 
3	Use the Zoom command to enlarge the view until peak spacing is easy to see.
4	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.
5	Enter the spacing value into the Spacing field for that sample. The Spacing field is outlined in blue to indicate that you have overridden a calculated value and the value entered is in bold, indicating that the value has been changed in this session.

The Basecaller Settings Parameter

About the Basecaller Settings

The Basecaller Settings are features of the Basecaller program which automatically truncate sample file analysis.

To change the Basecaller Settings for a file, choose the name of a parameter value set from the pop-up menu. The available parameter value sets are the ones you created in the Preferences dialog box. For more information, see “Basecaller Settings” on page 6-28.

During base calling, the Basecaller considers both the Basecaller Settings and any Stop Point value in the Sample Manager window, and stops as soon as it meets one of the endpoint criteria or the Stop Point value — whichever comes sooner.

To Revert to Default

To revert back to default settings for a Basecaller, you must also redetermine the Peak 1 Location, Start Point, and Stop Point. See “Making the Software Recalculate” on page 6-16.

The Peak 1 Location Parameter

About the Peak 1 Location

The Peak 1 Location is the data point that marks the beginning of the first base peak in the data. This is initially calculated by the Sequencing Analysis software. It is the reference point for the spacing and mobility corrections performed by the base calling software.

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

Follow the instructions in the table below to find the beginning of the first base peak for a sample. Then, if necessary, you can enter a new Peak 1 Location value in the Sample Manager.

Note Because the Peak 1 Location is linked to the mobility correction, changing the Peak 1 Location value affects the way the DyeSet/Primer file is applied to correct for mobility shifts. If you want to start analysis farther along than the actual location of the first base peak, change the value for the Start Point, not the Peak 1 Location value.

Note If you reset the Peak 1 Location value to “0”, the software recalculates the Peak 1 Location, Start Point, Stop Point, and Spacing.

Finding the Peak 1 Location Value

The ABI 373 and ABI PRISM® 377 Data Collection Software samples data 194 times[1] each time it scans across the gel. The ABI PRISM® 310 Data Collection Software samples data at one second intervals. Each sampling is stored as a data point. The Peak 1 Location value is defined as the first data point in the file that is from the sample — not including primer.

If you are using dye primer chemistry, follow the instructions in the table below.

If you are using dye terminator chemistry, follow the instructions in the table on page 6-14.

For dye primer chemistry, to find the Peak 1 Location value:

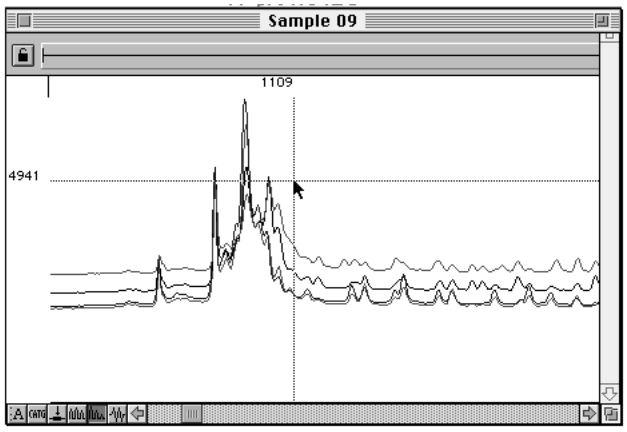
Step	Action
1	Open the sample file.

1. 388 times in 377 XL and 373 XL mode and 480 times in 377 96-Lane mode

For dye primer chemistry, to find the Peak 1 Location value: *(continued)*

Step	Action
2	Click the Raw Data view button.  Peaks are normally present in four colors on the display. They extend throughout the width of the window.
3	If the colored lines representing the bases do not appear, use the following steps to display them: a. Choose Display Options from the Window menu. b. Click to select the check boxes for all four bases. c. Choose OK.
4	Use the scroll bar at the bottom of the Raw Data view window to scroll along the sequence and find the large Primer peak near the beginning of the data. Use the Zoom In (Command – =) and Zoom Out (Command – –) commands in the Window menu for better views of the data. If you get lost in a zoomed-in view, choose Full View from the Window menu to see all the data.

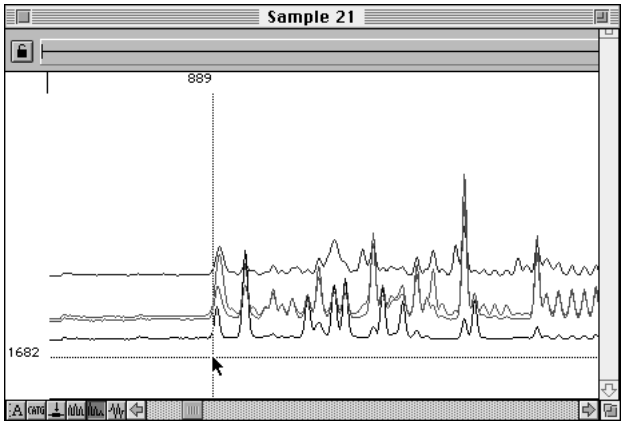
For dye primer chemistry, to find the Peak 1 Location value: *(continued)*

Step	Action
5	Find the beginning of the first base peak (the Peak 1 Location value). The general appearance of this peak depends on whether you used dye primer or dye terminator chemistry. (If you used dye terminator chemistry follow the table on page 6-14.) **Note** The scan number at which the first base peak occurs varies with the instrument, gel, electrophoresis conditions, and separation distance used to generate the data. If you used dye primer chemistry to prepare your samples, the initial peaks in the data are small, and a much taller (primer) peak appears at the beginning of the sequencing run. The beginning of the first base peak is on the downslope of this tall primer peak. The following figure shows the correct Peak 1 Location value (at scan 1109) for a sample prepared with dye primer chemistry and run on a 48 cm gel. 
6	Find the location on the downward slope of the primer peak where the first base peak begins. The blue peak (C) is usually the highest.
7	With the mouse cursor, point to the beginning of the peak and hold the mouse button down to display locator lines. Note the cursor position on the x axis — this is the scan point number at the top of the vertical locator line. This number is the Peak 1 Location value to use for analysis.

For dye terminator chemistry, to find the Peak 1 Location value:

Step	Action
1	Open the sample file.
2	Click the Raw Data view button.  Peaks are normally present in four colors on the display. They extend throughout the width of the window.
3	If the colored lines representing the bases do not appear, use the following steps to display them: a. Choose Display Options from the Window menu. b. Click to select the check boxes for all four bases. c. Choose OK.
4	Use the Zoom In (Command – =) and Zoom Out (Command – –) commands in the Window menu for better views of the data. If you get lost in a zoomed-in view, choose Full View from the Window menu to see all the data. Use the scroll bar at the bottom of the Raw Data view window to scroll along the sequence and find the first true base peak near the beginning of the data.

For dye terminator chemistry, to find the Peak 1 Location

Step	Action
5	Find the beginning of the first base peak (the Peak 1 Location value). The general appearance of this peak depends on whether you used dye primer or dye terminator chemistry. If you used dye terminator chemistry to prepare your samples, the raw data might show peaks before actual data that can be erroneously designated as the Peak 1 Location value by the software. The correct Peak 1 Location value is at the beginning of the sample peaks. The following figure shows the correct Peak 1 Location value (at scan 889) for a sample prepared with Taq terminator chemistry. 
6	With the mouse cursor, point to the beginning of the peak and hold the mouse button down to display locator lines. Note the cursor position on the x axis — this is the scan point number at the top of the vertical locator line. This number is the Peak 1 Location value to use for analysis.

The Start Point Parameter

About the Start Point

The Start Point is the raw data point where you want base calling to start in the sample file. The Start Point is normally the same as the beginning of the first base peak (the Peak 1 Location value).

Changing the Start Point

If 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.

IMPORTANT If you want to start analysis further along than the actual 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.

Making the Software Recalculate

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.

The Stop Point Parameter

About the Stop Point

The Stop Point specifies the last raw data point to be included in the base calling. If the default Stop Point is used, this endpoint is the last data point in the file.

Changing the Stop Point

It is possible to stop base calling earlier if there is clearly unusable raw data at the end of the file, or if you want to analyze only a portion of the raw data in the file.

Set an earlier Stop Point either by changing the values on the Basecaller Settings page of the Preferences dialog box (page 6-28) or by entering an earlier Stop Point in the Sample Manager window.

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. Because the basecaller software calculates spacing based on the whole data range, setting an accurate stop point results in better data analysis.

The First Endpoint Encountered Is Used

During base calling, the Basecaller considers both the endpoint set in the Basecaller Settings preferences and any Stop Point value in the Sample Manager window. The Basecaller stops analysis at the earliest designated endpoint.

The DyeSet/Primer File Parameter

About the DyeSet/Primer Pop-Up

This pop-up menu allows you to specify which DyeSet/Primer file to use for base calling. The default DyeSet/Primer is the one specified in the Sample Sheet for the run in Data Collection.

IMPORTANT If you change the DyeSet/Primer file and then reprocess the file(s), the Basecaller recalculates the Peak 1 Location, Start Point, Stop Point, and Spacing. Any user-entered values for these parameters are overwritten during this operation.

The pop-up menu displays all the DyeSet/Primer files in the ABI Folder on your hard disk. If the filename is displayed in outline font in the DyeSet/Primer file field, this means that the file is not present in the ABI Folder and that the sample file cannot be processed.

DyeSet/Primer File Required

The DyeSet/Primer file is required for analysis. Sequencing Analysis will not analyze the sample if the DyeSet/Primer file field is set to <none> or if the specified file is not in the ABI Folder.

Shortening the Pop-Up Menu

To make the DyeSet/Primer list shorter, you can discard any DyeSet/Primer files that do not apply to your laboratory's procedures:

Step	Action
1	Open the ABI Folder.
2	Drag the unwanted DyeSet/Primer files to another folder or to the Trash. Before you drag to the Trash make certain that you have a backup copy of the ABI Folder.

DyeSet/Primer Sets Mobility Shift Correction

The Basecaller algorithm needs the DyeSet/Primer information to be able to apply the proper mobility shift corrections. If you specified the wrong DyeSet/Primer (mobility) file in the data collection software, or used a different chemistry from the one for the selected DyeSet/Primer file, you can change this setting for each affected sample file by choosing from the pop-up menu and reanalyze the files.

Choosing the Correct DyeSet/Primer File

This table provides guidance for choosing the correct DyeSet/Primer file for the concentration of your gel or polymer. For an explanation of DyeSet/Primer file names, see “DyeSet/Primer File Naming Conventions” on page B-8.

% Gel or Polymer	Type of Gel or Polymer	Instrument	DyeSet/Primer File
	POP-6™	ABI PRISM 310	DT or DP POP6{XX}
	DNA Sequencing Polymer	ABI PRISM 310	DT DSP{dR Set-AnyPrimer}, or DT or DP5%CEHV{XX}
6%	Acrylamide	ABI 373 or ABI PRISM 377	DT or DP6%Ac{XX}
4.75%	Acrylamide	ABI 373 or ABI PRISM 377	DT or DP6%Ac{XX}
4.25%	Acrylamide	ABI 373 or ABI PRISM 377	DT or DP4%Ac{XX}
4%	Acrylamide	ABI 373 or ABI PRISM 377	DT or DP4%Ac{XX}
5%	Long Ranger™	ABI 373 or ABI PRISM 377	DT or DP4%Ac{XX}
5%	Long Ranger	ABI PRISM 377	DP5%LR{XX}

The Instrument File Parameter

About the Instrument File

The instrument file (sometimes referred to as the “matrix file”) is the file used to adjust for spectral overlap between the fluorescent dyes. Each analysis instrument normally has one instrument file associated with it per dye type (*i.e.* one instrument file for Rhodamine dyes and one for dRhodamine dyes), and that file contains all the matrices you may need for the instrument. For more information about how to create the matrices for an instrument file, see Appendix E, “Creating Instrument Files.”

Information in the instrument file is copied to the gel file when it is made, and to each sample file before base calling is done. If you selected the wrong instrument file during data collection setup, the instrument file information in the gel file and the sample files will be wrong and base calling will be inaccurate.

Changing the Instrument File

To change the instrument file information for a sample file, choose the correct instrument file from the pop-up menu. The menu shows all instrument files in the ABI Folder in the System Folder on your computer.

If the filename is displayed in outline font in the Instrument file field in the Sample Manager, this means that the file is not present in the ABI Folder.

The Instrument File in the Gel File

The filename in the Instrument file column in the Sample Manager window is taken from the Sample Sheet. If the Instrument file specified in the Sample Sheet is not found in the ABI Folder when the sample file is analyzed, Sequencing Analysis will try to use the instrument file saved in the gel file. (The Sample Manager window will be updated to show the name of the instrument file used to analyze the sample files.)

Parameters in the Preferences Dialog Box

Introduction

You can select preferred values (preferences) for most of the processing parameters used by the Sequencing Analysis software. Some of these parameters will guide the software's actions during auto-analysis. Some will be the parameter values automatically listed for each sample file that you manually add to the Sample Manager window.

When desired, you can override these preference values for individual files or groups of files.

The following sections (pages 6-23 to 6-48) describe each of the parameters that can be set through the Preferences dialog box. For a general explanation of how to select and change a Preferences parameter value, see "Changing Parameter Values in the Preferences Dialog Box" on page 6-22.

About the Default Values

The default values in the Preferences dialog box are the values most commonly used by Applied Biosystems customers. You can open the Preferences dialog box and change these values at any time.

Note Changes you make in the Preferences dialog box take effect as soon as you close the dialog box. However, the changes do not affect the values already defined for sample files currently listed in the Sample Manager window.

To return to the default values for the preferences, delete the file Seq Analysis v3.4 Prefs from the Preferences folder in the System Folder. All existing preference settings will be lost.

The Preferences Parameters Listed

The preferences are grouped into the following categories, which are available on separate pages of the Preferences dialog box.

Changing Parameter Values in the Preferences Dialog Box

About Changing Parameter Values

When you change a processing parameter value in the Preferences dialog box, the new value is used for all future processing until you change the value again or temporarily override that value for selected files in the Sample Manager window.

Note Changes you make in this dialog box take effect as soon as you close the dialog box. However, the changes do not affect the values already defined for sample files currently listed in the Sample Manager window.

This section provides a generic explanation of how to change Preference values. The following sections explain how to decide which values are appropriate for your situation.

Changing a Preference Parameter

To change a preference value:

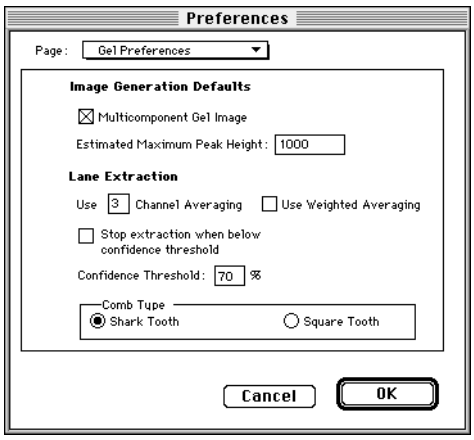
Step	Action
1	Highlight the Preferences… command on the Edit menu to open the Preferences submenu, then select one of the commands from the Preferences submenu. The Preferences dialog box appears. It displays the current value(s) for the Preferences item that you selected. This is one page of Preferences that you can view and change in this dialog box.
2	Use the check boxes, text fields, and pop-up menus in the dialog box to change the preference value(s). For an explanation of each Preferences page, see the following sections of this chapter.
3	After you make any required changes on the page, either select a different page from the Page pop-up menu at the top of the dialog box, or choose OK to close the dialog box. The changes take effect as soon as you close the dialog box.

Gel Preferences

About the Gel Preferences Page

The Sequencing Analysis software uses the values on this page when it tracks and extracts information from ABI 373 and ABI PRISM 377 gel files.

~~310~~
373
377



There are seven Gel Preferences parameters:

- Multicomponent Gel Image check box
- Estimated Maximum Peak Height text box
- Use ___ Channel Averaging text box
- Use Weighted Averaging check box
- Stop extraction threshold check box
- Confidence Threshold text box
- Comb Type radio button

These are described in detail below.

Multicomponent Gel Image

Select the Multicomponent Gel Image if you want the Sequencing Analysis software to adjust for spectral overlap of the dyes when generating the gel image.

This process affects only the gel image, not the raw data.

Note Gels must be multicomponented in order to be auto-tracked by the Neural Net Tracker program.

Estimated Maximum Peak Height

In this text box, enter the maximum signal level you expect from samples in the run. This can be an approximate number, based on your typical run conditions and samples.

The Estimated Maximum Peak Height value affects the brightness of the colored bands that represent the base peaks. All bands with a data value at or above the Estimated Maximum Peak Height value are assigned the brightest dye color. The dye colors for bands with values below that level are dimmed proportionally.

In general, the lower you set this value, the brighter the bands appear in the Gel File window. A value of 1000 is satisfactory for most gel files. If the gel image is very dim, try 500; if it is too bright, try 2000.

This value also determines the scale of the peaks in the Slice view of the Gel File window. For the highest quality gel image, the highest sample peaks (*not* the primer peak) should just reach the top of the scale in the Slice view. If you find that many of the peaks are cut off, you may want to readjust the Estimated Peak Height value to a higher number.

Note This option affects only the appearance of the gel image (not the raw data) when the image is generated the first time the gel file is opened. To change the image appearance for any other gel file, use the Regenerate Gel Image command (“To regenerate the gel image:” on page 4-22).

Use ___ Channel Averaging

The number of channels to be averaged for each lane when extracting data from the gel file is normally set to 3. Averaging reduces the amount of noise in the sample file. Further data smoothing may be achieved by using the Weighted Average (page 6-25 below).

Each tracker line in the Gel File window marks the channel where the Sequencing Analysis software located the strongest fluorescent signal for that lane. If you use the default three-channel average, the raw data

in each sample file is an average of the data in the channel marked by the tracker line and one channel on either side of it.

Altering the Channel Averaging

If you choose two-channel averaging, data is taken from the tracked channel and the channel to the right of it. You can include data from up to nine channels. Three-channel averaging is recommended.

You might choose to use one channel (no averaging) if the gel bands are severely tilted. For example, if the left channel of the lane ran faster than the right, a better result would be obtained by taking the center channel alone, rather than averaging three channels.

Note When you use multiple-channel averaging, be sure each tracker line marks the center of its lane. If a tracker line is near the right or left edge of its lane, empty channel(s) between lanes may be included in the average and cause an erroneously low value. Also signal from neighboring lanes may be included in the average.

Use Weighted Averaging

Weighted channel averaging feature was introduced in Sequencing Analysis v. 3.2. Weighted averaging is now possible because the new Tracker interface allows tracker line placement to within a tenth of a channel. The Use __ Channel Averaging field applies to both weighted and non-weighted averaging.

No Weighted Averaging

If the Use Weighted Averaging box is not checked, data averaging is done per channel. For example, if the tracker line falls within channel 10, and 2-channel averaging is set:

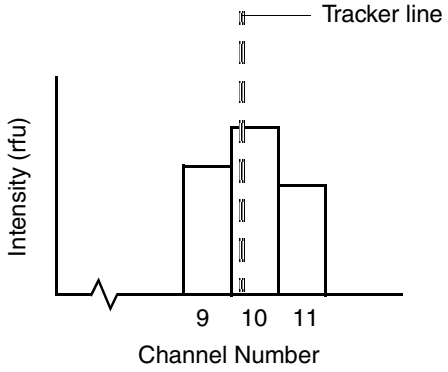
$$\text{channel average} = \frac{I_{\text{ch. 10}} + I_{\text{ch. 11}}}{2}$$

Where I is the intensity for a given channel and scan number.

Weighted Averaging

If the Use Weighted Averaging box is selected, data averaging is done to the tenth of a channel. For example, if the tracker line falls 20% into channel 10 (see diagram below), and 2-channel averaging is set:

$$\text{channel average} = \frac{(I_{\text{ch. 9}} \times 0.8) + I_{\text{ch. 10}} + (I_{\text{ch. 11}} \times 0.2)}{2}$$



If you use weighted averaging, the annotation view of the sample file indicates this by specifying “(Weighted)” in the Channels Ave field.

Stop Extraction When Below Confidence Threshold

If the Stop Extraction When Below Confidence Threshold box is checked, when the lane assignment confidence level is below that specified in the Confidence Threshold, the lane extraction will not be carried out and a warning dialog box appears.

The dialog box gives you the option to cancel or continue the gel file extraction and analysis.

Confidence Threshold

Lane extraction is not to be carried out after tracking, if

- The Stop Extraction When Below Confidence Threshold box is checked

  and

- The lane assignment confidence value is less than this confidence threshold

The default value for the Confidence Threshold is 70. You can enter any number between 0 and 100 for the Confidence Threshold.

Lane Assignment Confidence Value

After a gel is auto-tracked, a lane assignment confidence value is written to the Error Log. This value indicates the Tracker’s confidence in how well the assigned lanes match the Sample Sheet. (This value is no indication of how well the tracker lines follow the fluorescence intensity within the lanes.)

Lane assignment confidence values tend to be extreme numbers: very low or very high. Although, a value of 70 or more generally indicates that the lane assignment for the gel is correct, it is recommended that you check the tracker lane assignment anytime the reported lane assignment confidence value is less than 100%.

Comb Type

~~310~~
373
377

The Neural Net Tracker uses special tracker settings files that are optimized according to the number of channels and lanes in the gel file and the comb type (shark-tooth).

It is important that you set the correct comb type to Shark Tooth in the Gel Preferences page so that the Tracker applies the correct tracker settings file.

Basecaller Settings

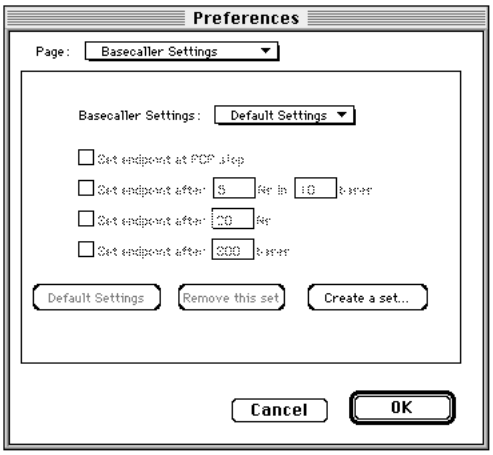
About the Basecaller Settings Page

The Basecaller Settings tell the Basecaller program what rules to use to decide the analysis endpoint for each sample.

IMPORTANT The Basecaller stops when it reaches the Stop Point set in the Sample Manager window *or* an endpoint specified in the Basecaller Settings page — whichever it meets first.

This page allows you to create multiple sets of Basecaller Settings, then to select one as the preference.

310
373
377



The Default Setting, Recommended

The default setting is to have no endpoint. This means that the basecaller tries to basecall to the end of the sample file. This is the recommended setting except for sequencing short PCR fragments when an earlier end point should be set using the procedure described on page 6-29.

Selecting a Set of Basecaller Settings

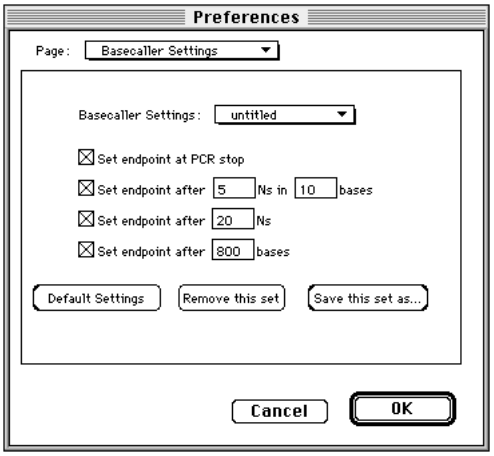
To select a predefined set of Basecaller settings, use the Basecaller Settings pop-up menu.

See step 2 on page 6-30 for a description of the settings.

Note Predefined settings are stored in the Seq Analysis v3.4 Prefs file in the Preferences folder, which is inside the System Folder. If this Prefs file is deleted, the predefined settings, along with all other preferences, are lost.

Creating a Set of Basecaller Settings

To create and save a new set of Basecaller settings:

Step	Action
1	Choose the Create a set… button. The button name changes to Save this set as… The check boxes can now be edited. 

To create and save a new set of Basecaller settings: *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>2</td><td>Use the check boxes and text fields to create a parameter value set as explained in the table below. You can set more than one endpoint condition. The Basecaller will stop when any one of the conditions are met.

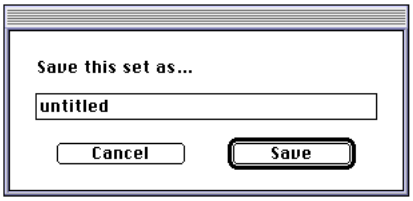
<table>
<tr><th>Item</th><th>Description</th></tr>
<tr><td>Basecaller Settings</td><td>The name for this value set.</td></tr>
<tr><td>Set endpoint at PCR stop</td><td>Use this only if you are sequencing short PCR products using dye primer chemistry.

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. 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, this is considered as signal and the stop point is incorrectly calculated to be after the noise.</td></tr>
<tr><td>Set endpoint after __ Ns in __ bases</td><td>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).</td></tr>
<tr><td>Set endpoint after __ Ns</td><td>Sets the analysis endpoint after a certain number of Ns occur (for example, after 20 Ns are detected).</td></tr>
<tr><td>Set endpoint after ___ bases</td><td>Sets the analysis endpoint after a certain number of bases (for example, after 800 bases are detected).

Note This should be used when sequencing short PCR products using dye terminator chemistry to ensure an earlier stop than the end of the run.</td></tr>
<tr><td>Default Settings button</td><td>Changes the values for the current value set to the default values (all check boxes de-selected).</td></tr>
</table>
</td></tr>
</table>

To create and save a new set of Basecaller settings: *(continued)*

Step	Action
3	Choose the Save this set as… button. A Save dialog box appears. 
4	Type a descriptive name for this parameter value set in the text field.
5	Choose Save to save the new Basecaller Settings, close the dialog box, and add this name to the Basecaller Settings pop-up menu. **Note** The set that appears in the pop-up menu will be the set specified for any files added to the Sample Manager window after the Preferences dialog box is closed.

Editing a Parameter Value Set

To edit an existing set of Basecaller Settings:

Step	Action
1	Choose the set that you want to edit from the Basecaller Settings pop-up menu.
2	Edit the check boxes and text fields as needed.
3	Choose the Save this set button.
4	Choose OK to close the dialog box.

Removing a Parameter Value Set

To delete an existing parameter value set from the Basecaller Settings pop-up menu:

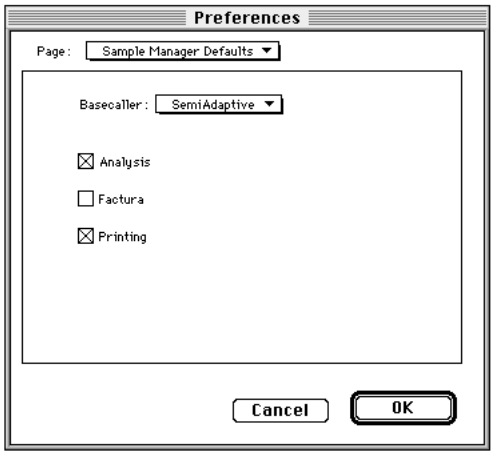
Step	Action
1	Choose the set that you want to remove from the Basecaller Settings pop-up menu.

Step	Action
2	Choose the Remove this set button. The set is removed from the list and deleted from the program. **Note** The set that appears in the pop-up menu will be the set specified for any files added to the Sample Manager window after the Preferences dialog box is closed.

Sample Manager Defaults

About the Sample Manager Defaults Page

This page allows you to select which Basecaller to use when processing samples, and to specify whether or not all the files will be base called, processed in Factura, and/or printed.



There are four parameters on the Sample Manager Defaults page:

- Basecaller pop-up menu
- Analysis check box
- Factura check box (not present in BioLIMS mode)
- Printing check box

These are described in detail below.

Basecaller

The Basecaller is the program that determines the individual base identities in a sequence. To choose a Basecaller, select the name from the Basecaller pop-up menu.

The Sequencing Analysis software package includes the Basecaller stand-alone program. During analysis, the Sequencing Analysis program automatically calls the Basecaller program.

For an explanation of Basecallers and how to choose the best one for your data, see “About Basecallers and Base Calling” on page 6-45.

Note The Basecaller program must be stored in the same folder as the Sequencing Analysis program. At installation, it is placed in the same folder as the Sequencing Analysis program.

Analysis If the Analysis check box is selected, the Sequencing Analysis software selects the A check box in the Sample Manager window for each file that *you* manually add to the Sample Manager.

When files are added to that window as part of automatic analysis, the software matches the A check box setting in the window to the A check box setting in the Sample Sheet, *not* to the setting in this dialog box.

When the A check box is selected in the Sample Manager window, the software analyzes (base calls) the file as part of file processing.

Factura If the Factura check box is selected, the software selects the F check box for *all* files added to the Sample Manager window.

When the F check box is selected in the Sample Manager window, the file is submitted to the Factura program for further processing after base calling and before printing.

Note If the Sequencing Analysis program is in BioLIMS mode, the Factura check box does not appear on the Sample Manager Defaults preferences page.

Printing If this check box is selected, the Sequencing Analysis software selects the P check box in the Sample Manager window for each file that *you* manually add to the Sample Manager.

When files are added to the window as part of automatic analysis, the software matches the P check box setting in the window to the P check box setting in the Sample Sheet, *not* to the setting in this dialog box.

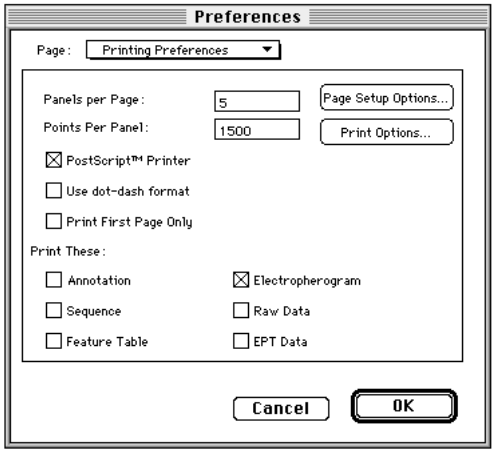
When the P check box is selected in the Sample Manager window, the file is printed after all other requested processing is finished.

Printing Preferences

About the Printing Preferences Page

This page allows you to choose

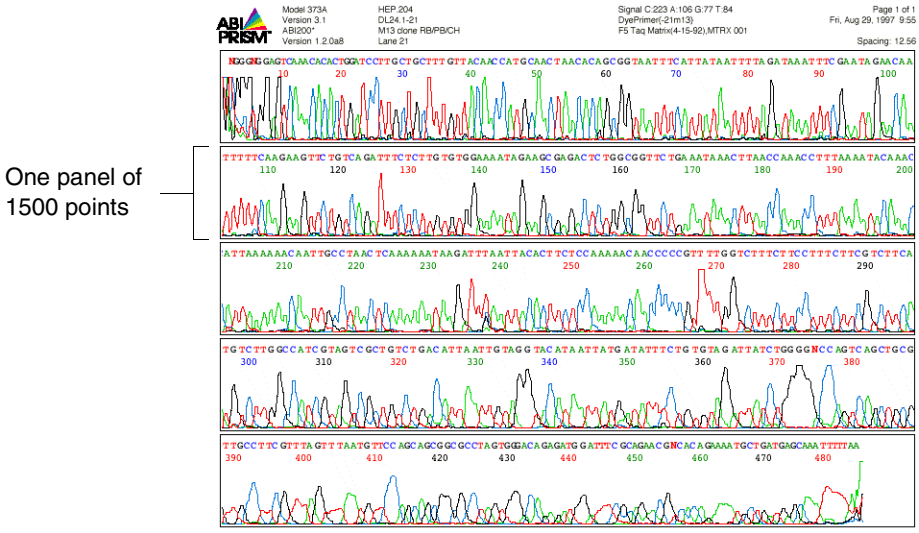
- The arrangement of the information on the page
- The appearance of the data on the page
- Page size and other standard page setup options
- How many copies are printed and other standard print options
- What will be printed when printing from the Sample Manager window



The Printing Preferences parameters are discussed in detail below.

Panels Per Page Text Box

The number of panels to print on each page of graphical (Electropherogram, Raw Data, EPT) views. The default is five panels of 1500 points. A sample print out with five panels is shown below.

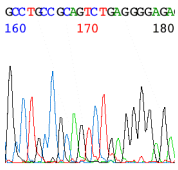


Points Per Panel Text Box

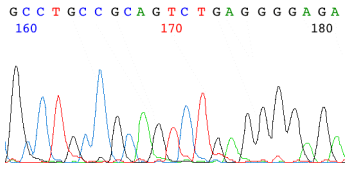
The number of data points in each panel. The default is 1500 data points per panel. If you decrease the number of data points per panel, the peaks are broader, with fewer bases per panel.

Use this text box to effect a zoom-in (fewer points per panel) or a zoom-out (more points per panel) in the printed graphical data.

1500 points per panel



700 points per panel



PostScript™ Printer Check Box

Select this if using a PostScript-compatible printer. De-select this if your printer is *not* a PostScript printer. See your printer manual to determine if you have a PostScript printer.

Use Dot-Dash Format Check Box

Select this to print the lines in Electropherogram, Raw Data, and EPT views as dotted or dashed lines, using a different pattern for each line. This option is available only for PostScript printers.

Print First Page Only Check Box

Select this to print only the first page of the specified sample file data. De-select this to print all the pages.

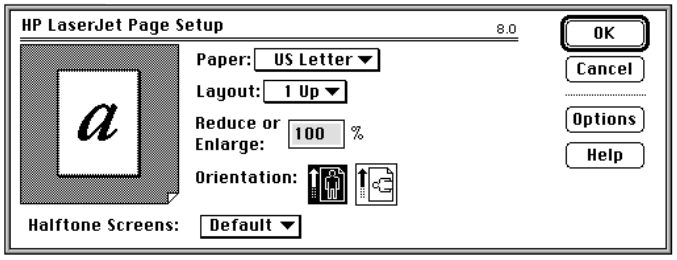
Print These Check Boxes

Select the check box next to each Sample window view (Annotation, Sequence, Feature Table, Electropherogram, Raw Data, EPT Data) to print when printing is started from the Sample Manager window or through auto-analysis.

IMPORTANT Do **not** select all six views. The exact number of views you can print at one time without overloading your printer will depend on your printer and the views selected for printing.

Page Setup Options Button

Opens the standard Page Setup dialog box for your printer so you can select paper size and orientation, screening options, etc.



The exact contents of the Page Setup dialog box depend on your printer. Normally, you should select your standard paper size and landscape orientation.

Choose OK to save the selected page setup to the Seq Analysis v3.4 Prefs file and close the Page Setup dialog box.

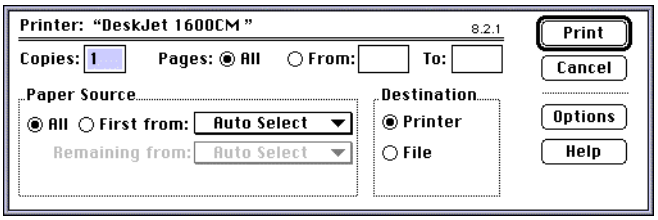
Note If your Sequencing Analysis prefs file is lost or discarded, the print options dialog box appears when you start the Sequencing Analysis software.

IMPORTANT Each time the printer selection in the Apple Chooser window is changed, you must open the Page Setup dialog box to reestablish the default selection.

Print Options Button

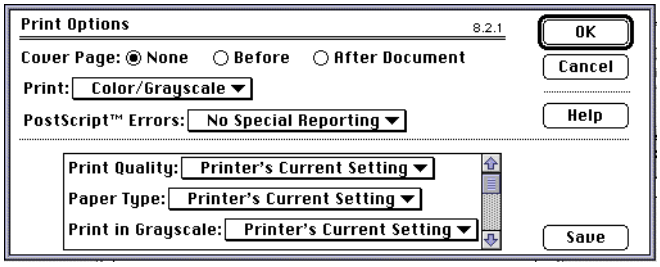
Opens the standard Printer dialog box for your printer, so you can select a default number of copies to print, paper source, etc.

When printing from the Sample Manager window, the Sequencing Analysis software does not ask for printer information. The software uses the values you select in this Printer dialog box. If your Sequencing Analysis Prefs file is lost or discarded, this dialog box will appear the next time you start the Sequencing Analysis software.



The exact contents of this dialog box depend on your printer.

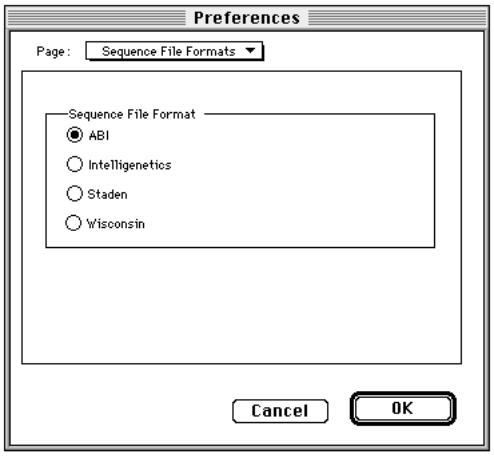
To ensure base letters on the electropherogram print in color:

Step	Action
1	Click the Print Options… button to open the Print dialog box.
2	Click the Options button of the Print dialog box to open the Print Options dialog box. 
3	Select Color/Grayscale for the Print option, then click OK to close the Print Options dialog box.
4	Click Print to close the Print dialog box. **IMPORTANT** If you do not choose Print to close the box, the settings are not saved. Nothing is actually printed at this time.

Sequence File Formats

About the Sequence File Formats Page

The Sequence File Formats page allows you to specify which file format to use when saving sequence (.Seq) files. The .Seq file is a text file that includes only the base sequence for the corresponding sample file.



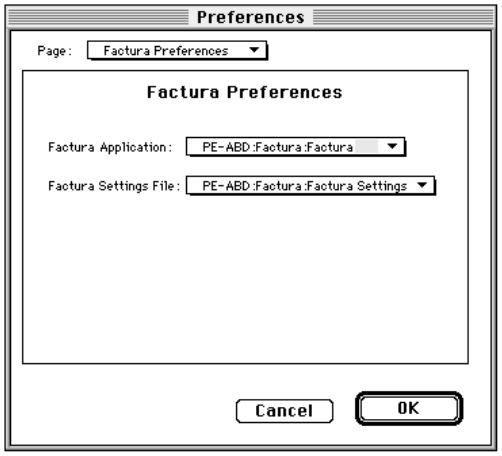
To use other ABI software that reads files in ABI text format, choose ABI.

If you want to import the text sequence files into other programs, select the appropriate sequence file format for that software.

Factura Preferences

About the Factura Preferences Page

This page allows you to specify the Factura program and settings file used when automatic Factura Feature Identification processing is selected for a sample file.



Note If Sequencing Analysis is in BioLIMS mode, the Factura Preferences have no effect. For information about Factura processing of sequences in the BioLIMS database, see the *ABI PRISM Factura Feature Identification Software User's Manual*.

Factura Application

Specifies the version and location of the Factura program to be used for further sample processing. If you do not want to use the Factura program, select <none> from the pop-up menu.

The Factura program is used to further process sequence files produced by Applied Biosystems instruments.

Raw DNA sequences often contain vector sequence and ambiguously called bases at both ends that should be removed prior to assembly or final analysis. This program allows you to clean up sequence files by identifying designated vectors and ambiguous regions, and flagging these features in the sequence file. Assembly and analysis applications can then disregard these ambiguous regions, using only the target DNA data.

You can also use Factura software to identify and assign IUPAC/IUB codes to mixed base-positions in a sequence.

Factura Settings File

Specifies the Factura settings file to use for processing when the F check box is selected. If you do not want automatic Factura processing to occur, select None. For more information about this file, see *ABI PRISM Factura Feature Identification Software User's Manual*.

Adding a File to the Pop-Up Menu

To add a program (or settings file) to the pop-up menu:

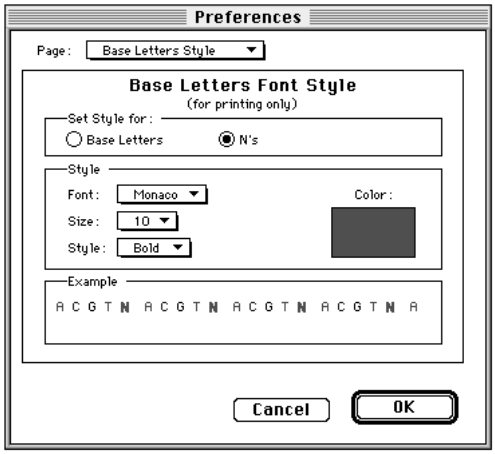
Step	Action
1	Select Other from the pop-up menu. A directory dialog box appears. Only the names of folders and files of the specified type are visible in the directory lists.
2	Locate and select the name of the program (or settings file) you want to add to the list. Then choose Open. When you choose Open, the complete file name and path name are added to the corresponding pop-up menu.

Base Letters Style

About the Base Letters Style Page

On the Base Letters Style page, specify the font, size, and style for the base letters and Ns that appear *on printed Electropherogram views*. For the Ns, select the color of the letters. Scaling of the printed base peaks is adjusted according to the font size selected.

IMPORTANT If you pick an extremely large font, the base-call letters may not line up correctly with their corresponding electropherogram peaks.



Note The settings on this page have no effect on the appearance of base letters on screen.

Descriptions of Base Letter Parameters

The Base Letter Style parameters are described in the following table.

Item	Description
Base Letters	Select this if you want to change the appearance of the four base letters (C, A, G, and T) on printed electropherograms.
Ns	Select this if you want to change the appearance of the letter “N” on printed electropherograms.
Font	The font used for the base letters and Ns on printed electropherograms. The options are Monaco and Courier. The default is Monaco.

Item	Description
Size	The size of the base letters and Ns on printed electropherograms. The options are 9, 10, 12, 14, 18, and 24 points. For Base Letters the default is 9 point, and for Ns the default is 10 point.
Style	The style of the base letters and Ns on printed electropherograms. The options are plain and bold. For Base Letters the default is plain, and for Ns the default is bold.
Color	The color of the Ns on printed electropherograms. The default is red.

Selecting a Color

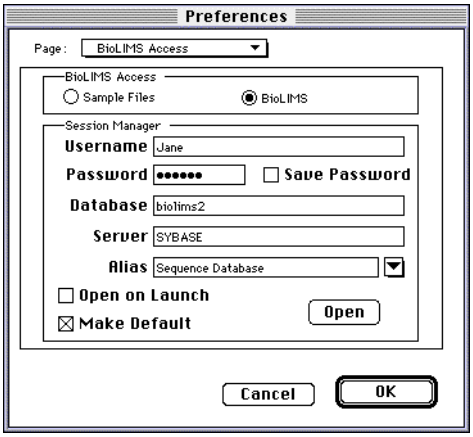
To set the color for the Ns in electropherogram prints, follow these steps:

Step	Action
1	Click the Ns radio button.
2	Click on the colored rectangle to open the ColorPicker dialog box.
3	In the dialog box, either enter values in the numeric fields or click on the color wheel to select the color you want.
4	Click OK to save your choice and close the ColorPicker dialog box.

BioLIMS Access

About the BioLIMS Access Page

If you have a BioLIMS database, the final page of the Preferences dialog box is the BioLIMS Access page. Information on how to set these preferences is given in “Accessing BioLIMS” on page 3-13.



About Basecallers and Base Calling

Introduction

Base calling is the primary function of the Sequencing Analysis software. For accurate base calling, it is important to understand the process and to select the best Basecaller for your data.

How Base Calling Works

The Sequencing Analysis program analyzes the data in the sample files for signal strength to evaluate whether the data should be analyzed and printed, and performs base calling. The following describes the processes involved:

- Preprocessing includes noise filtering, signal strength analysis, and finding the first base peak in the sample.
- A first pass of the software determines the spacing between peaks. This includes the following processes:
  - *Multicomponent analysis* adjusts for the spectral overlap of the dyes. This function utilizes the instrument file.
  - *Mobility shift* adjusts for differences in mobility between the dyes. This function utilizes the DyeSet/Primer file.
- The raw data is re-spaced, based on information computed in this initial processing.
- A second processing is based on the re-spaced raw data. The software begins again with location of the primer peak, multicomponent analysis, and mobility shifts, then performs the following processes:
  - *Peak height normalization* normalizes the signal strengths between the colors. Each dye exhibits different levels of fluorescence; this process scales each color so the total signals of each are equal.
  - *Signal enhancement* enhances the peak shape by applying a bandpass filter to the data.
  - *Initial base calling* locates the best candidate peak in each interval of 12 data points. If none is found or the data is conflicting, the software calls an N.
- A final pass of the software adds or deletes bases, based on the distance between each peak and its nearest neighbors.
- The analyzed data and other information are stored in the sample file.

The actual base calling is performed by the Basecaller program. There are seven types of basecaller. These are described under separate headings later in this section.

Choosing a Basecaller

Choosing the most effective Basecaller for any given sample file depends on the quality of the data, the type of run, and the run and gel conditions. The following information can help you decide which Basecaller is most suitable. In addition, you can try each Basecaller with some typical data to see which works best under your laboratory conditions.

If you have a...	Then use...
run on the ABI PRISM 310 that used POP-6 with d-Rhodamine terminators or BigDye™ primers or BigDye™ terminators or DNA Sequencing Polymer with d-Rhodamine, rhodamine, or fluorescein/rhodamine dye chemistries	ABI-CE1
run on the ABI PRISM 310 that used POP-6 with rhodamine dye terminators	ABI-CE2
24- or 34-cm well-to-read, Full or XL Scan run on the ABI 373	ABI50
BaseSprinter or 377-18 run on the ABI 373	ABI100
average 100-bph, 2X, 1200 scans/hr, 36-cm run on the ABI PRISM 377	ABI100
average 200-bph, 4x, 2400 scans/hr, 36-cm run on the ABI PRISM 377	ABI200
48-cm run on the ABI 373	ABI50
48-cm run on the ABI PRISM 377	ABI100
run with many insertions or deletions near the end of the run (for example, if the sample is a PCR product)	Semi-Adaptive
spacing that is a negative number	Semi-Adaptive
spacing that is still a negative number with SemiAdaptive	Adaptive
problems with run conditions	Adaptive

The ABI Basecallers

The labels 50, 100, and 200 on the ABI 373 and ABI PRISM 377 Basecallers refer roughly to the bases per hour (bph) separated on the slab gel electrophoresis instruments.

Select one of the ABI Basecallers to perform base calling on a standard run. The names indicate the instrument and type of run for which the Basecaller is optimized.

- ABI-CE1 is optimized for runs on the ABI PRISM 310 instrument (CE refers to capillary electrophoresis) that use DNA Sequencing Polymer (any chemistry) or d-Rhodamine terminator chemistry and BigDye™ DNA Chemistry with POP-6™ polymer. You should not use it with ABI 373 or ABI PRISM 377 data.
- ABI-CE2 is optimized for runs on the ABI PRISM 310 instrument (CE refers to capillary electrophoresis) that use the rhodamine dye terminator chemistry and POP-6 polymer. You should not use it with ABI 373 or ABI PRISM 377 data.
- ABI100 is optimized for data collected at approximately 100 bph. The ABI PRISM 377 instrument runs at approximately 100 bph using the 2X (1200 scans/hr) run module (including 48 cm wtr runs) as defined for the data collection software in the *ABI PRISM DNA Sequencer User's Manual*. The ABI 373 runs at approximately 100 bph when in BaseSprinter or 373–18 mode. You should not use this Basecaller with ABI PRISM 310 data.
- ABI200 is optimized for 200 bph runs on the ABI PRISM 377 instrument using the 4X (2400 scans/hr) run module as defined for the data collection software in the *ABI PRISM DNA Sequencer User's Manual*. You should not use this Basecaller with ABI 373 or ABI PRISM 310 data.
- ABI50 is optimized for data collected using a 24 or 34 cm separation distance and Full Scan or XL Scan mode on the ABI 373 instrument. It is the base calling method (called Standard) in previous versions of the Analysis software. You should not use this Basecaller with ABI PRISM 310 data.

The ABI Basecallers differ from each other primarily in the shape of the internal spacing curves. All of them use DyeSet/Primer files stored in the ABI Folder.

The SemiAdaptive Basecaller

Use the SemiAdaptive Basecaller for data from any of the genetic analysis instruments. This Basecaller dynamically measures spacing from the data. Unlike the ABI Basecallers, it does not use standard spacing curves. It does, however, use the DyeSet/Primer files stored in the ABI Folder, as do the ABI Basecallers.

Use the SemiAdaptive Basecaller when:

- Accuracy with the ABI Basecallers is low, especially when many insertions or deletions exist near the end of the run.
- The gel ran too slow or too fast.
- The spacing value is a negative number.

The spacing value appears in the Annotation view of the Sample window, on the electropherogram printout, and in the Spacing column of the Sample Manager.

Note If length of read (total number of bases) is critical to your run, try using the SemiAdaptive Basecaller. Under some run conditions, it is able to accurately read more bases than the other Basecallers (with the potential loss of losing some basecalls at the beginning of the run).

The Adaptive Basecaller

Use the Adaptive Basecaller for data from any of the genetic analysis instruments. This Basecaller dynamically measures both mobility shifts and spacing from the data for each sample. It uses this information to calculate the data preprocessing before calling the bases.

Use the Adaptive Basecaller if:

- You performed the run with nonstandard conditions, such as a different gel type or speed.
- The results from using the ABI Basecaller or the SemiAdaptive Basecaller are not satisfactory.
- You experienced problems during the run.

Often the Adaptive Basecaller can correct problems that occurred during a run.

IMPORTANT Although each ABI Basecaller is tuned for a specific type of run, depending on your run conditions you might get stronger data using a different Basecaller. Analyze your data with different Basecallers to determine which one works best for your run conditions.

If you reanalyze a sample file, the previous analysis results are overwritten by the new results. To avoid erasing the previous analysis results, save a copy of the sample file under a different name on the hard disk before you perform the second analysis.

Viewing and Editing Sample Files 7

Overview

In This Chapter This chapter explains how to view, edit, and print the analyzed sequence data in the six views of the Sample window.

What Information Do Sample Files Contain?

Sample files contain the following information about the DNA sequence:

- Raw data, as captured by the instrument before any post-collection processing
- The first sequence called by the Basecaller program
- Any edited base calls which have been saved to the file
- Annotation information describing the instrument run and analysis conditions
- Analysis settings
- Processed (analyzed) electropherogram information which visually describes the intensity of each fluorescent signal
- Summary of electrophoresis conditions (voltage, temperature, power) during the run
- Features added by the Factura™ program

All of this information can be viewed in graphical and text formats. Thus, sample files contain the target DNA sequence plus all of the historical information about the ABI PRISM analysis necessary to interpret the data and processing parameters.

BioLIMS Records Vs. Sample Files

BioLIMS® database sequence records contain the same information as sample files (see list above). Analyzed sequence data can be viewed and further processed by other ABI PRISM DNA sequencing software and by third-party software. Some programs read both sample files and BioLIMS sequence records, others read only sample files (refer to the table below).

ABI PRISM Software	Reads Sample Files?	Reads from BioLIMS Database?
Factura	yes	yes
AutoAssembler	yes	yes
Sequence Navigator	yes	no
EditView	yes	no

If you want to use software that does not read from the BioLIMS database, then

- Extract gel file data in sample file mode (not BioLIMS mode), or
- Use the Sample2DB utility software to download sequence database records into sample files.

Opening a Sample File in a Sample Window

Introduction

There are several ways to open a sample file in a Sample window. The number of sample files that you can have open at one time depends on the amount of free memory on your computer. Typically, the maximum number of sample files that you can have open at one time is in the range of 25–30.

Opening a Sample File from the Finder

There are two ways to open a sample file from the Finder:

- Double-click the name or icon of the file you want to open.
- Drag the icon for the file you want to open onto the Sequencing Analysis program icon.

Opening a Sample File Using Menu Commands

To open a sample from within the Sequencing Analysis program using the pulldown menus:

Step	Action
1	Choose Open Sample from the Sequencing Analysis program File menu.
2	Locate and select the desired sample file in the directory dialog box that appears. Then choose Open.

Note In BioLIMS mode, the Collection Browser window appears instead of the directory dialog box when Open Sequence is chosen from the file menu or the Add Seq. button on the Sequence Manager is clicked. Use the Collection Browser to find and select BioLIMS sequence records. For more information about the Collection Browser, see “Using the Collection Browser Window” on page 3-19.

Opening Sample Files from the Sample Manager

To Open One Sample File

There are three ways to open a single sample file from the Sample Manager window:

- Double-click the Sample File Name for the file.
- Click once on the file name, then click the Open Files button.

- Click once on the file name, then choose Open Files from the Manager menu.

To Open Multiple Sample Files

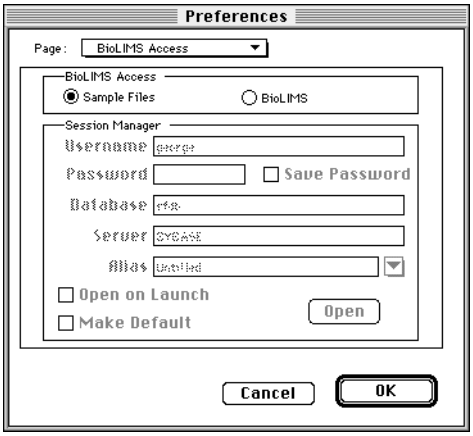
Step	Action
1	Hold down the Command key while you click the Sample File Names of the files you want to open.
2	Click the Open Files button at the top of the window. Or, choose Open Files from the Manager menu.

The Sequencing Analysis software displays the Sample window for as many of the selected files as memory allows. You can choose any of the six views for each file.

BioLIMS Users

If you have an installation of the BioLIMS database and want to open a sample file (not a sequence from the database), the Sequencing Analysis program must be set in Sample File mode.

Use the BioLIMS Access page of the Preferences dialog box (page 3-13) to check that Sequencing Analysis is in sample file mode (as shown below).



The Six Sample Window Views

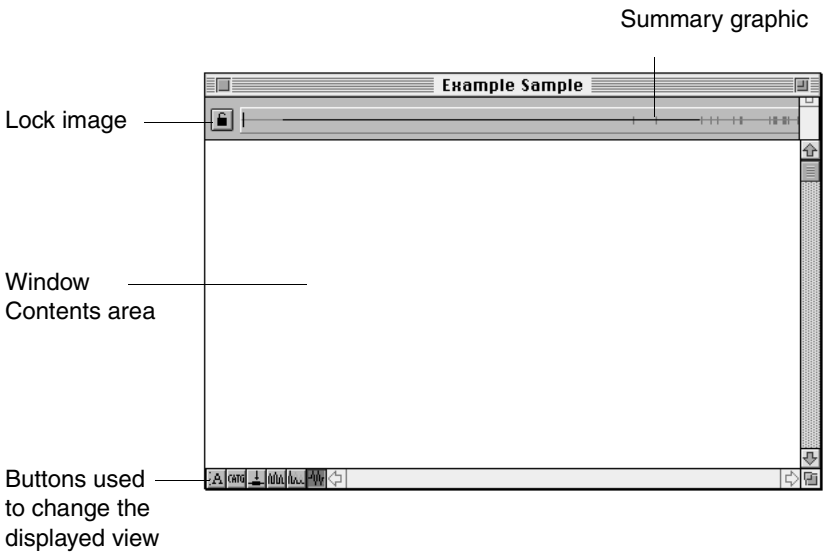
Introduction

The Sample window is used to view or edit the sequence data. There are six different views available in the Sample window.

Common Features in Views

Certain features of the Sample window are available in all six window views (as shown below). The contents of the window and the available menu choices depend on the view selected.

Note To see the on-line help for these windows, choose Show Balloons from the Balloon menu near the right end of the main menu bar.



Lock Image

When the lock image appears closed (locked), the sample file is protected from edits. You cannot cut from or paste to the sample file (using the Edit menu). To open or close the lock, click on the lock image.

Summary Graphic

Immediately below the window name and to the right of the Lock Image is a horizontal line inside a frame. This line represents the length of the sequence. The larger tick mark shows the cursor position as you move it to different places in the sequence. The smaller tick marks, if present, show the location of color-marked features in the sequence. The arrowhead at either the right or left end of the line indicates the orientation of the sequence.

If you select an area of the sequence in Sequence view, that area is shown as a rectangle on the Summary graphic in all views.

Window Contents Area

The main portion of the Sample window contains the information pertaining to the sequence. You can display six different data views in the Sample window. To change views, click the button for the view you want to see.

Button	View	Short-cut	Description
	Annotation	Command-E	Summary sample information entered in the data collection program, and additional information entered by the data collection and Sequencing Analysis programs. (See page 7-9.)
	Sequence	Command-R	The nucleotide (base) sequence text called for the data. The Sequencing Analysis software displays an empty window if base calling has not yet occurred. Edit the sequence after analysis in either this view or in the Electropherogram view. (See page 7-10.)
	Feature	Command-T	The features that were found in the sequence. Features are added to the analyzed sequence data by the Factura software, which is used for further processing of the data. If features are visible, the sequence has been processed by Factura. If no features are available, the Sequencing Analysis software displays an empty window. (See page 7-12.)
	Electro-pherogram	Command-Y	A four-color picture of analyzed data, with peaks representing the bases. It is the default view that appears when an analyzed sequence opens. *This view is available only after base calling is done.* (See page 7-13.)

Button	View	Short-cut	Description
	Raw Data	Command-U	The raw data collected by the instrument. This is the default view if base calling has not been done. (See page 7-16.)
	EPT	Command-I	A plot of run voltage, temperature, current, and power values. (See page 7-19.)

Annotation View

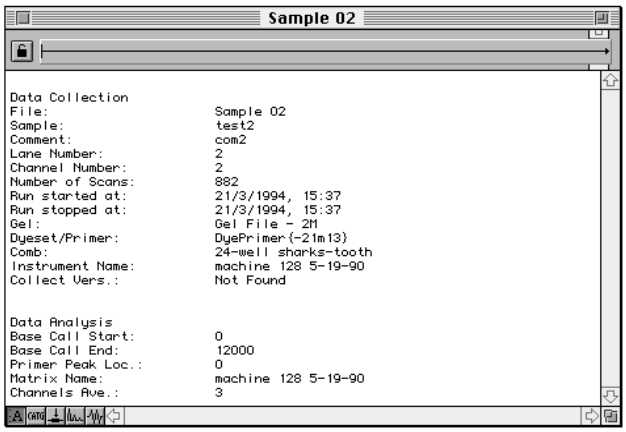
Displaying

To display Annotation view:

- Type Command-E, or
- Click the button shown below

About Annotation View

Annotation view shows the sample information you entered in the data collection program, additional information entered by the data collection and analysis programs (for example, the start time and stop times), and changes that you made to the original information.



Note The information displayed in Annotation view depends in part on the instrument used to generate the data.

View and Print

In Annotation view, you can:

- View the information in the window, but not edit it.
- Print the contents of the window (for details, see “Printing the Sample Window Views” on page 7-33).

Sequence View

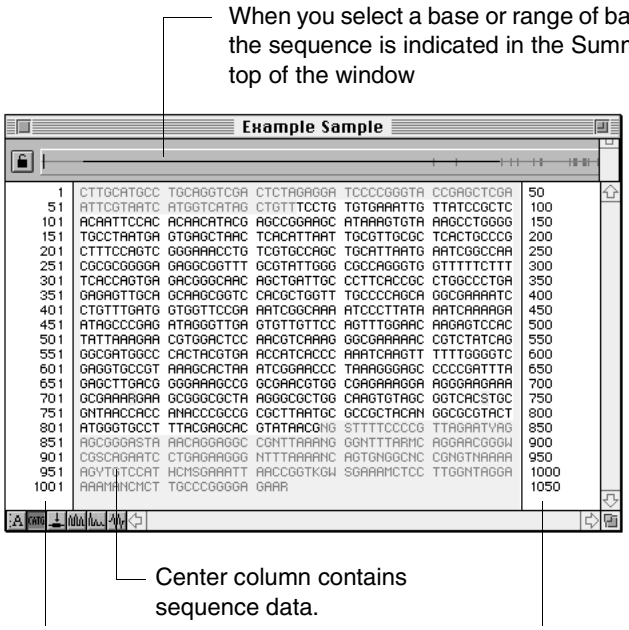
Displaying

To display Sequence view:

- Type Command-R, or
- Click the button shown below

About Sequence View

Sequence view shows the nucleotide sequence called for 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.



Find, Edit, and Print

In Sequence view, you can:

- Use the Find command to search for a base character, a range of bases, or a specified base pattern (for details, see “Finding Patterns in Sequence View” on page 7-25).
- Use any of the standard Macintosh® commands to edit the sequence (for details, see “Editing Bases in Sequence View” on page 7-29).
- Print the contents of the window (for details, see “Printing the Sample Window Views” on page 7-33).

Feature View

Displaying

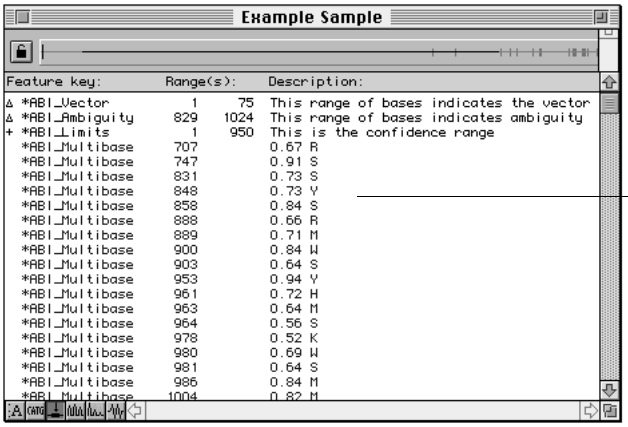
To display Feature view:

- Type Command-T, or
- Click the button shown below

About Feature View

The Feature view shows features that were added to the analyzed sequence data by the Factura software.

If the sequence data has not been processed in Factura, this window is empty. After processing in Factura, the window displays the features list for the file.



View and Print

In Feature view, you can:

- View the window contents (for an explanation of the information displayed in this view, see the *ABI Prism Factura Feature Identification Software User's Manual*).
- Print the window contents (for details, see "Printing the Sample Window Views" on page 7-33).

Electropherogram View

Displaying

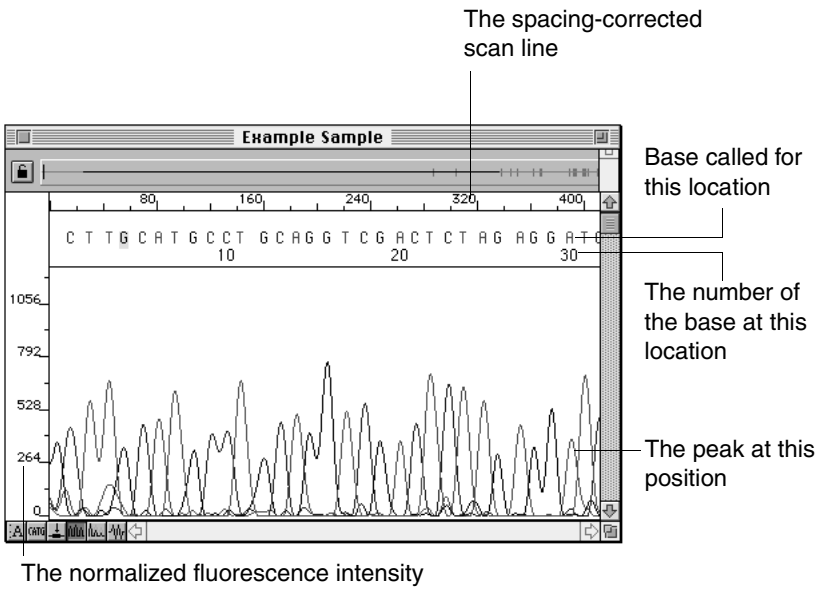
To display Electropherogram view:

- Type Command-Y, or
- Click the button shown below

Note If the raw data has not been analyzed, Electropherogram view is not available.

About Electropherogram View

Electropherogram view shows a four-color picture of the analyzed sample data, with peaks representing the bases called for the sample.



Trace and Base Colors

The trace lines and the letters above the peaks are colored to represent the four bases.

Base	Color
C	Blue
A	Green
G	Yellow or Black[a]
T	Red

a. G is shown as black when printed or shown against white on the screen.

Ambiguous Bases

An N above a peak means that the software could not confirm that base, or that there is more than one base at that position (for example, a heterozygote).

The default style is to have Ns displayed in bold type, making them easier to find. (To alter this style, see “Base Letters Style” on page 6-42.)

If the sequence has been processed by the Factura program, you may also see IUB codes. If you are using the default Mark Style in Factura, IUB codes are displayed underlined and in red type. (IUB codes are listed in the Glossary.)

Edited Bases

Bases that have been edited since the last basecall are underlined.

Correlation with Other Views

Sequence View and Electopherogram View

If you click on the sequence in Sequence view, then switch to Electropherogram view, the electropherogram shows the area of the sequence around the point where you clicked.

If you highlight a range of bases in Sequence view, that range of bases is also highlighted in Electropherogram view.

Raw Data View and Electopherogram View

The scan numbers of the Electropherogram view do not map directly to scan numbers of the Raw Data view due to the application of the

basecaller algorithm, which alters the scan number-to-data-point relation.

The raw and analyzed data views show the same dye colors. These are the converted colors based on the DyeSet/Primer that was used to analyze the sample file.

Raw data files that have not been analyzed show the colors of the actual dyes detected (just like the gel file). The colors are only converted when base calling is performed.

View, Edit, and Print

In Electropherogram view, you can:

- Zoom in or out to see the data at different magnifications (see “Using the Zoom Commands” on page 7-41).
- Use the right arrow keys to move to the next base or the left arrow key to move the previous base.
- Use the Tab key to find the next occurrence of an N, or Shift-Tab to find the previous occurrence.
- Edit the bases one at a time (see “Editing Bases in Electropherogram View” on page 7-30).
- Display the original, unedited base calls while you edit the bases (see “Showing Original Data in Electropherogram View” on page 7-32).
- Change the horizontal and vertical ruler display (see “Changing the Displayed Lines and Scales” on page 7-43).
- Change the colors of the trace lines that represent the bases, or hide one or more trace lines (see “Changing the Displayed Lines and Scales” on page 7-43).
- Hold down the mouse button, while the cursor is in the data area of the window, to display cross hairs and the coordinates for the current cursor location.
- Print the window contents (for details, see “Printing the Sample Window Views” on page 7-33).

Raw Data View

Displaying

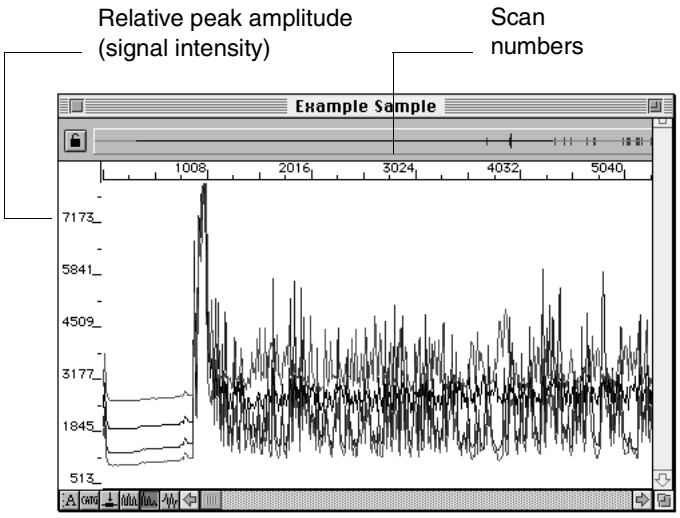
To display Raw Data view:

- Type Command-U, or
- Click the button shown below

About Raw Data View

The Raw Data view shows the raw data for the sample, before any processing is performed. This is the first view you see if you open a sample file before the bases are called. (After bases are called, you see Electropherogram view first when you open the file.)

If the sample was run on a ABI PRISM® 310 Genetic Analyzer, this is the information collected by the data collection software. If the sample was run on a ABI 373 or ABI PRISM 377 DNA Sequencer, this is the data that the Sequencing Analysis software extracted from the gel file created during the instrument run.



When you first select Raw Data view, the Sequencing Analysis software displays the data in full view, with all the data compressed into one normal-sized window.

The four colored trace lines represent the fluorescence data from the four fluorescent dyes. The base represented by each color depends on the chemistry and filter set used. For more details, see “Summary of the Instruments and Chemistries” on page E-2.

The Importance of the Raw Data View

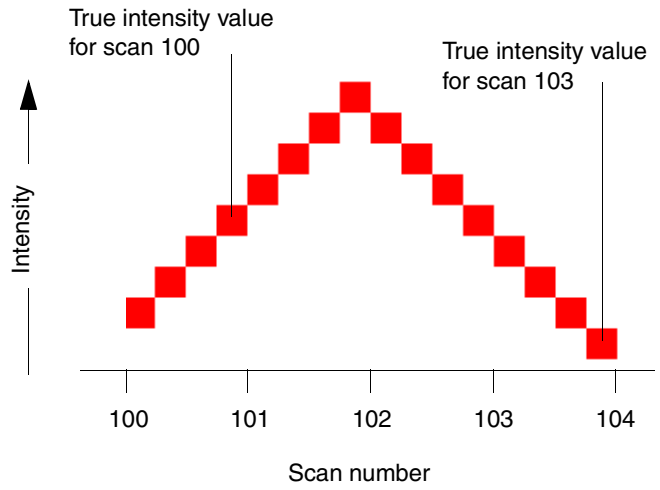
The Raw Data enables you to perform a number of important checks and troubleshooting tasks. You can use the Raw Data view to:

- Verify the point used by the Sequencing Analysis software to start and stop base calling.
- Measure 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 base calling or could indicate instrument problems.
- For data from ABI 373 or ABI PRISM® 377 instruments, find areas with lower signal that could indicate bad tracking of the gel file.
- Determine the scan number that corresponds to a given location in the data (for details, see “Determining the Value for a Data Point” on page 7-23).
- Estimate base spacing by measuring the scan points which define two adjacent peaks.

Intensity Values in Raw Data 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.

For example, a peak in raw data view at maximum magnification might look like this. The true intensity points for two scan are indicated in the figure below.



View, Edit, and Print

In Raw Data view, you can:

- Zoom in or out to see the data at different magnifications (see “Using the Zoom Commands” on page 7-41).
- Change the colors of the trace lines that represent the fluorescent dyes, or hide one or more trace lines (see “Changing the Displayed Lines and Scales” on page 7-43).
- Hold down the mouse button, while the cursor is in the data area of the window, to display cross hairs and the coordinates for the current cursor location.
- Print the window contents (for details, see “Printing the Sample Window Views” on page 7-33).

EPT View

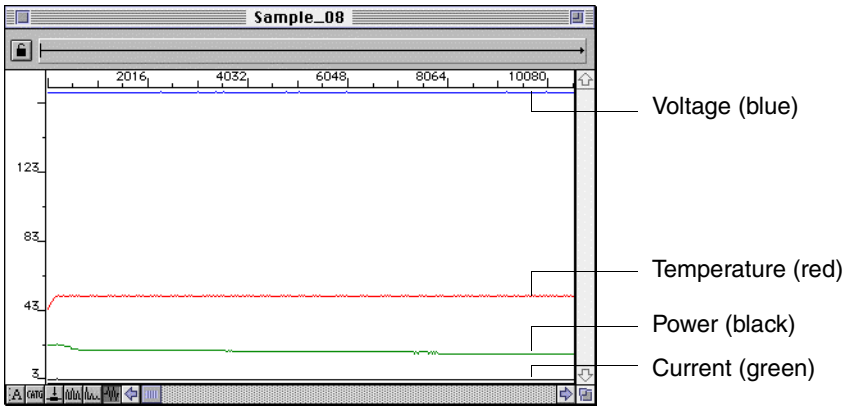
Displaying

To display EPT view:

- Type Command-I, or
- Click the button shown below

About EPT View

EPT view shows values for run voltage, temperature, power, and current.



The colors indicated in the table below are the default colors. You can edit the colors using the Display Options dialog box (page 7-43)

Measurement Plotted	Default Color	Units 377/373 Data	Units for 310 Data
Voltage	Blue	V/10	V/100
Current	Green	mamp	μamp
Power[a]	Black	W	mW x 10
Temperature	Red	°C	°C

a. For 310 data, this is laser power. For 377 and 373 data, this is electrophoresis power.

View and Print

In EPT view, you can:

- Hold down the mouse button, while the cursor is in the data area of the window, to display cross hairs and the data values at the current cursor location.
- Choose Display Options… from the Window menu to open a dialog box and determine the type of information represented by a particular color of line.
- Print the window contents (for details, see “Printing the Sample Window Views” on page 7-33).

Reviewing the Analysis Results

Introduction

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

Reviewing the Error Log

If a problem occurred during processing, the Error Log will be displayed in front of the other windows. If the log is visible, determine the source of the problem and take appropriate action.

Make the Error Log visible by choosing Show Error Log from the Window menu.

Reviewing the A/F/P Check Boxes

Review the A, F, and P check boxes in the Sample Manager window (for details, see “Checking for Processing Problems” on page 5-22). If the Analysis check box for any file is red, determine the source of the problem by reviewing the sample file and Error Log. If necessary, reanalyze the file. (A check box without color means that the processing step did not occur; it does not indicate a problem.)

Reviewing the Analyzed Data

The following review steps are recommended for each sample file:

Review the Spacing Values

Review the spacing values in the Sample Manager window. If a value is displayed in bold red text, the Basecaller encountered a problem while calculating the value and was unable to resolve the problem.

Review the Files Used in Processing

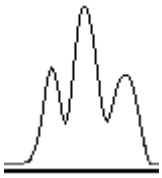
Review the files specified for use during processing. If the name of a file appears as outlined text, the software could not find that file in the expected location.

For analysis to proceed, you must specify a Basecaller that is present in the same folder as the Sequencing Analysis software, and a DyeSet/Primer file that is present in the ABI folder in the System Folder on your hard disk.

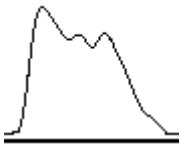
Scroll through the Electropherogram

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

If there are problems, see Appendix C, “Troubleshooting Sequencing Analysis.”

Check Base Calls in the Electropherogram

Look at the base calls in the Electropherogram view. Where two peaks are close together, or the peak is low, or the background noise level is high, compare each peak to the bases called for that peak. If necessary, edit incorrect base calls manually.

Selecting Bases

Bases can be selected either with the mouse or using the arrow keys. The arrow keys are the easier way to select bases. If necessary, the mouse can be used to insert a base between two bases called by the software.

Search for Ns in the Electropherogram

Use the Tab key to move forward and Shift-Tab to move backwards and search for Ns. If you can visually determine the correct base call at an N location, manually change the N to the correct character.

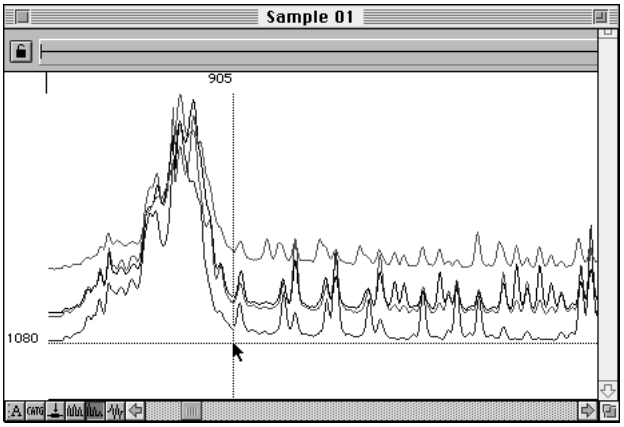
Determining the Value for a Data Point

Introduction

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 Data, or EPT view of the Sample window.

Determining Values of a Given Point

To determine the values for a given data point:

Step	Action
1	Put the cursor near the point of interest. Then hold down the mouse button. The scan number ruler disappears and the crosshair locator lines appear. 
2	Drag the cursor across the window until the locator lines intersect the point of interest.

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>3</td><td>Note the values at the top of the vertical and the left of the horizontal locator lines.

<table>
<tr><th>For the view…</th><th>Value at the top of vertical line represents the…</th><th>Value at the left of horizontal line represents the…</th></tr>
<tr><td>Electropherogram</td><td>re-spaced scan number</td><td>normalized fluorescence intensity</td></tr>
<tr><td>Raw Data</td><td>raw scan number</td><td>normalized fluorescence intensity</td></tr>
<tr><td>EPT</td><td>raw scan number</td><td>parameter value</td></tr>
</table>
</td></tr>
</table>

Finding Patterns in Sequence View

Introduction

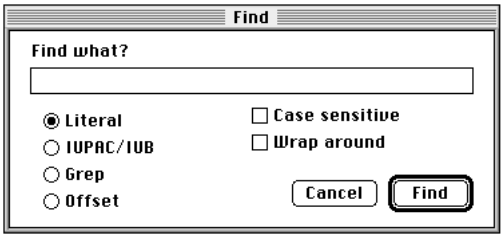
You can use the Find and Find Again commands in the Edit menu to search for a particular base, or pattern of bases, in a sequence. The search operation must be done in the Sequence view of the Sample window.

Note You cannot use the Find command in Electropherogram view. Instead, use the Find command in Sequence view, and when the pattern is highlighted, switch to Electropherogram view.

Note To find the next occurrence of an N, use the Tab key to search forward, or Shift-Tab to search backward.

Searching for a Pattern in a Sequence

To find a pattern in a sequence:

Step	Action
1	Display the Sequence view of the Sample window.
2	Click at the position in the sequence where you want to start the search. **Note** The search begins at the cursor position. If the pattern is before the cursor, it is only found if “the Wrap around” check box is selected (see step 3 below). If you only want to find a pattern in the valid range, place the insertion point just before this range in the sequence.
3	Choose Find from the Edit menu. A special Find dialog box appears: 

To find a pattern in a sequence: *(continued)*

Step	Action
4	In the Find What field, enter the search instruction. In addition to normal base character (G, A, T, C) patterns, the search string can include IUPAC/IUB characters, Grep search expressions, or offset instructions. (For details, see “About Search Expressions” on page 7-27.)
5	Select the radio button that matches the type of instruction entered in the Find What field. (For details, see “About Search Expressions” on page 7-27.)
6	Select or de-select the two check boxes as needed. ♦ Check Case sensitive to have the upper- and lower-case variants of a letter be recognized as different symbols. If this check box is not selected, the Sequencing Analysis software considers upper and lower case versions of a character to be the same (for example, “A” and “a”). ♦ Check Wrap around to have the search start again at the beginning of the sequence after it has reached the end. If the Wrap Around check box is *not* selected, the search stops at the end of the sequence.
7	Choose 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 window.
8	To find other occurrences of the same pattern: ♦ Choose Find Again from the Edit menu. This allows you to bypass the Find dialog box and search for the next occurrence of the specified pattern.

About Search Expressions

In the Find dialog box, choose one of four different types of search:

- Literal
- IUPAC/IUB
- Grep
- Offset

Literal

Choose Literal to search for patterns that match exactly what you typed in the Find What field.

IUPAC/IUB

Choose IUPAC/IUB if you included an IUB character as part of the pattern. The Find command locates all possible matches. For instance, if the pattern you enter is TAR, the Find command locates either TAG or TAA. IUB codes are listed in the Glossary.

Grep

Choose grep if you include an expression in the search string. The following table describes some of the expressions you can use and how they function.

Expression	Match Performed	Example
[ ] (brackets)	Any character inside the brackets	`AA[AC][GT]` matches `AAAG`, `AAAT`, `AACG`, or `AACT`. `[AGC]` matches `A`, `G` or `C`.
[^] (brackets with ^ as first character inside)	Any character EXCEPT the character(s) inside the brackets	`A[^AG]C` matches `ACC` or `ATC`.
* after character	Zero or more such characters	`AT[CG]*T` matches `ATT` or `ATCT` or `ATGGT`, and so on.
. (period)	Any character	`AA.A` matches `AAAA`, `AACA`, `AAGA`, `AATA`, `AANA`, and so on.
– (dash) enclosed by brackets	A range of characters	`AA[A-z]` matches `AAA`, `AAC`, `AAG`, `AAz`, and so on.

Offset

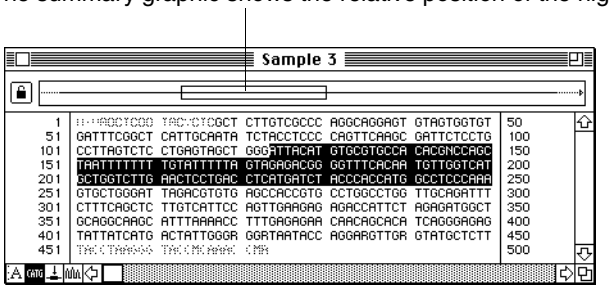
Choose Offset to move the cursor to the position or range of positions you specify. If you enter a number in the Find What field, the insertion point is moved to that base position. If you enter a range of numbers, the whole range is highlighted.

For example,

- Enter 123 in the Find What field to move the insertion point to the base character at position 123.
- Enter the range 123…250 to highlight all the base characters between number 123 and number 250.

IMPORTANT Type Option-[semicolon] to create the ellipsis dots (…) between the numbers.

The summary graphic shows the relative position of the highlighted range



Editing Analyzed Sequence Data

Introduction

You can use the Sequencing Analysis software to change a base that was called by the software during analysis, or to enter bases where the software called Ns.

To help you keep track of changes, you can display the original, unedited base calls in addition to the editable base characters.

You can edit the sequence directly in the Sample window, either in Electropherogram view or in Sequence view.

The original sequence data is the sequence calculated by the Basecaller. If the data is base called again, the old base-called data is overwritten by the new bases so that the most recently called bases are maintained as the “original”. Editing is done on a copy of the original data. Normally, only the editable copy is visible in the Sample window.

Editing Bases in Sequence View

In Sequence view, you can use the standard editing commands from the Edit menu to cut, copy, paste, and clear bases or ranges of the sequence in the active window. You can also use the Select All command to select the entire sequence.

Note If you add bases in Sequence view then change to Electropherogram view, the new bases are spaced as evenly as possible between the previously existing bases.

To add a base or range of bases to the sequence:

Step	Action
1	Place the insertion point at the position in the sequence where you want to add one or more bases. The program allows you to add any base-identification character that is recognized by the program, including IUPAC/IUB codes.
2	Type the character(s) you want to insert.

To delete a base or range of bases from the sequence:

Step	Action
1	Select the base or range of bases.
2	Use the Delete key or choose Clear from the Edit menu.

To change a base in the sequence:

Step	Action
1	Select the base you want to change.
2	Type the new character for that position.

Note If you edit data in Sequence view, the Electropherogram view is immediately updated to match the changed Sequence view data.

Editing Bases in Electropherogram View

In Electropherogram view, the Edit menu commands are *not* available. You can edit only one base character at a time.

In this view, the spacing of the characters is much more precise and approximately ten base positions are available between the displayed bases. If you place the insertion point between two characters and click, the software selects one of the available positions.

- To move from one displayed base to the next, use the left and right Arrow keys.
- To move from a base position to the next position (often pixel-by-pixel) hold down the Option key while you use the left and right Arrow keys.

To select a base in Electropherogram view:

Step	Action
1	Place the insertion point (cursor) to the left or the right of the character you want to select, then click the mouse button.
2	Use the right Arrow key or the left Arrow key to position the highlight directly on the appropriate base character. When you use the Arrow key, the cursor always moves to the next base character in the sequence. This procedure ensures that you have selected the base, not a position only one pixel away from it. Once you have selected the base, you can delete it using the Delete key or replace it by typing a new character.

To add a base in Electropherogram view:

Step	Action
1	Place the insertion point between the displayed base characters, where you want to insert the base character, then click the mouse button.

To add a base in Electropherogram view: *(continued)*

Step	Action
2	To move the insertion point closer to one of the flanking bases, hold down the Option key while you press the Left- or Right-Arrow key. Each time you press the Arrow key *while you hold down the Option key*, the cursor moves one scan point closer to the next base position.
3	When the insertion point is at the appropriate location, release the Option key, then type the new base character. The program allows you to add any base-identification character that is recognized by the program (including IUPAC/IUB codes).

Note If you edit data in Electropherogram view, the Sequence view is immediately updated to match the changed Electropherogram view data.

Showing Original Data in Electropherogram View

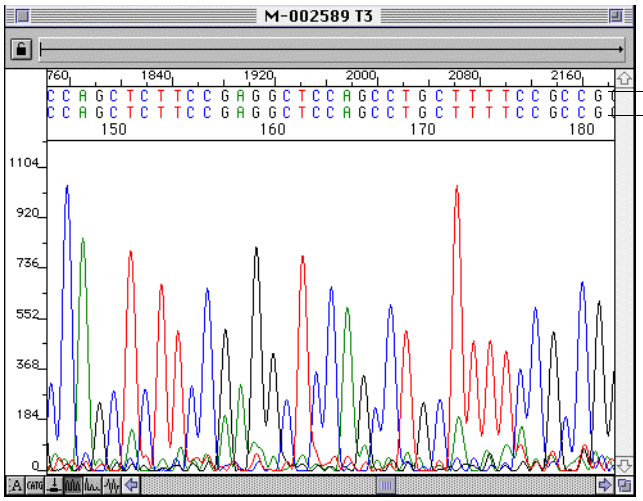
Introduction

In Electropherogram view, you can display the original sequence data in addition to the editable copy, so you can compare them. This is particularly helpful if you are editing bases in this view.

Showing the Original Data

To show the original data in Electropherogram view:

Step	Action
1	Make sure you are in Electropherogram view in the Sample window.
2	Choose Show Original from the Sample menu. A second line of base characters appears at the top of the window. The upper line is the original, uneditable data and the lower line is the editable copy.



Hiding Original Data

To hide the original data in Electropherogram view, choose Show Original from the Sample menu. This command acts as a toggle, either showing or hiding the original data. When the original data is visible, a check mark appears next to this command in the menu.

Printing the Sample Window Views

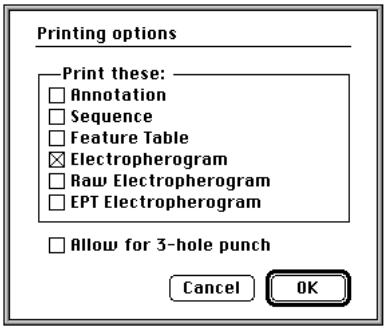
Introduction

Automatic printing is set up at the Printing Preferences page of the Preferences dialog box (for details, see “Printing Preferences” on page 6-35). You can use the steps below to temporarily change those settings and to print directly from the sample in the currently active Sample window.

Printing the Sample File

Follow the steps in the table below to print from an active Sample window.

To print the contents of a displayed sample file:

Step	Action
1	If you want to temporarily change the page orientation, paper type, panels/page, etc., choose Page Setup from the File menu to open a special Page Setup dialog box.
2	If you opened the Page Setup dialog box, adjust the settings as needed. Then choose OK to close the dialog box. The bottom part of the dialog box contains the four special setting options described on the table on page 7-34.
3	Choose Print from the File menu. A Printing Options dialog box appears. 
4	Check the view(s) you want to print.
5	If you want to leave an extra wide left margin to allow for three-hole punched paper, select “Allow for 3-hole punch”.
6	Choose OK to close the Printing Options dialog box and open the standard Printer dialog box for your printer.

To print the contents of a displayed sample file: *(continued)*

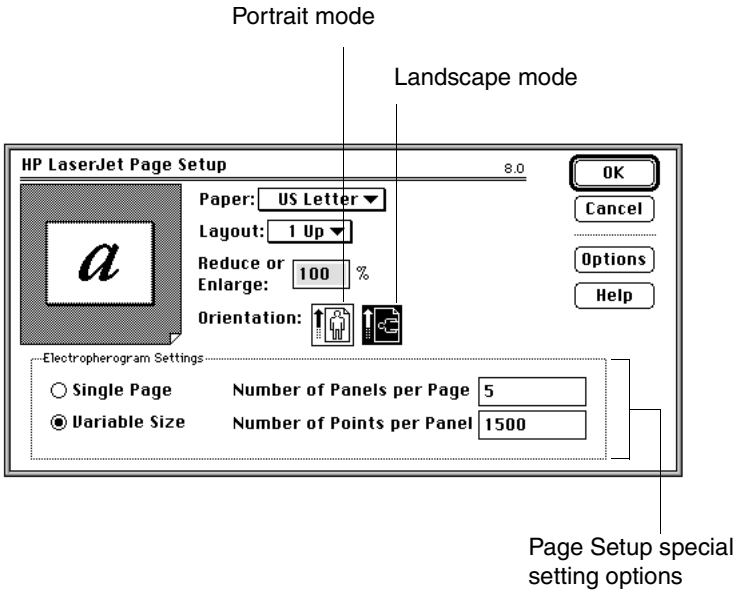
Step	Action
7	Make any required changes in the Printer dialog box, then choose Print to start printing.

About the Page Setup Dialog Box

If you open the Page Setup dialog box when a Sample window is active, there is a special section under the heading “Electropherogram Settings”. These are the Page Setup special settings options, which apply to all graphical data (sequence, raw, and EPT) that are printed from the Sample window.

These four options are described in the table below.

Note Your Page Setup dialog box may not look exactly like the one shown here, depending on your printer and installed printer driver.



The Page Setup Special Settings Options

Setting	Description
Single Page radio button	Prints all the information (for the Electropherogram view, Raw Data view, or EPT view) on a single page. If this button is selected, the Number of Panels and the Number of Points text boxes are grayed-out.

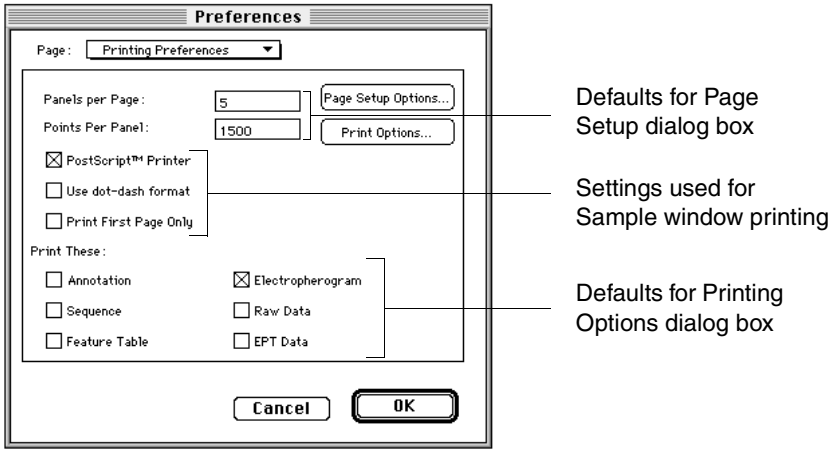
Setting	Description
Variable Size radio button	Specifies the number of panels and data points to print on any given page. If you select Variable Size, the two entry field options become available.
Number of Panels Per Page text box	Specifies the number of panels of data to print on a page when printing Electropherogram view, Raw Data view, or EPT view. These views are printed in tiled panels on the page. (For more information, see “Panels Per Page Text Box” on page 6-36.)
Number of Points per Panel text box	Specifies the number of data points to be included in each panel on the page. Because all panels are the same width, the peaks appear wider and flatter when you include fewer points. (For more information, see “Points Per Panel Text Box” on page 6-36.)

Note The default values for the text boxes are taken from the Printing Preferences Page, see “Page Setup and Printing Defaults” below.

Page Setup and Printing Defaults

The default values for the Page Setup special settings options and the Printing Options dialog box are determined by what is set in the Printing Preference dialog box (described on page 6-35).

The other Printing Preferences (PostScript Printer, Use dot-dash format, and Print First Page Only) are applied directly when printing from the Sample window, just as they are applied when automatic printing from the Sample or Sequence Manager is used.



Viewing Printed Electropherograms

Introduction

The printed electropherogram shows a four-color picture of analyzed data, with peaks representing the bases. The number of panels displayed on each page depends on the value selected for "Number of Panels Per Page" in the Page Setup dialog box. (Described in "Panels Per Page Text Box" on page 6-36.)

Why Print the Electropherogram?

Printed electropherograms have two advantages over electropherograms viewed on screen:

- The printed electropherogram includes information from the Sample Sheet that is only visible on screen in Annotation view.
- The printed electropherogram can display several panels of data on one page; on the screen you can only see one section of the data at a time.

Trace and Base Colors

For analyzed data (Electropherogram view), the four colors represent the individual bases in the sequence.

Base	Color
C	Blue
A	Green
G	Black[a]
T	Red

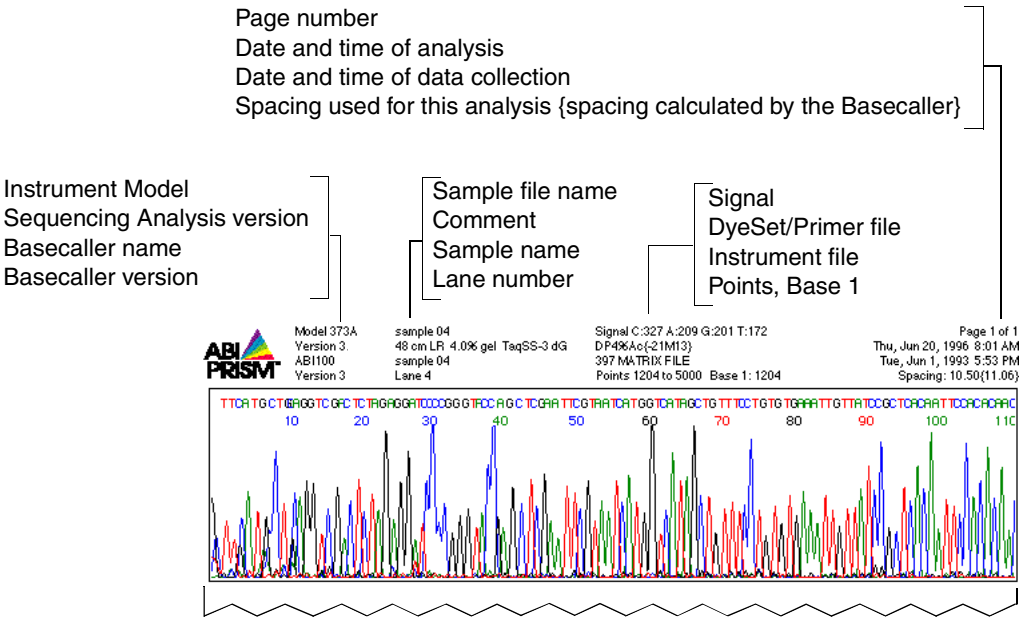
a. G is shown as yellow in AutoAssembler™ software.

Note For raw data, the meaning of each color depends on the chemistry (dyes) and filter set (physical or virtual). For details, see "Colors in Real-Time Data Display Windows" on page E-3.

The letters above the peaks are colored to represent the appropriate bases. An N above a peak means that the software could not confirm that base, or that there is more than one base at that position (for example, a heterozygote).

The Printer Header

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



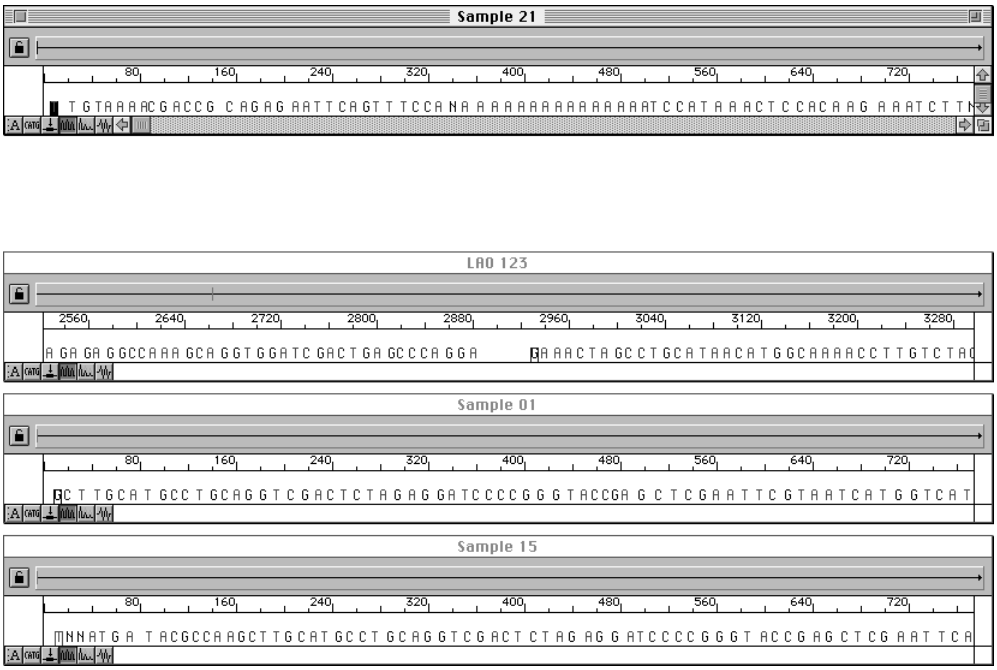
Column	Field	Description
First	Instrument Model	The instrument model used to collect the data.
	Sequencing Analysis version	The version of Sequencing Analysis software used to analyze the data.
	Basecaller name	The name of the Basecaller used to analyze the data.
	Basecaller version	The version of the Basecaller used to analyze the data.
Second	Sample file name	The name used for the sample file. (This is the name shown in the File Name column in the Sample Sheet.)
	Comment	The comments entered in the comment field of the Sample Sheet.

Column	Field	Description
	Sample name	The name entered in the Sample Name column of the Sample Sheet.
	Lane (not included for 310 runs)	The lane in which the sample ran on a 373/377 gel.
Third	Signal	The signal strengths or intensity of the fluorescence for each nucleotide in the sample.
	DyeSet/Primer file	The DyeSet/Primer file used during analysis to adjust for mobility shifts
	Instrument file	The instrument file used to analyze the data and adjust for spectral overlaps.
	Points, Base 1:	The range of the data points collected that were used to analyze the data. Base 1 is the data point where the analyzed data starts.
Fourth	Page x of x	The page number for this page and the total number of pages.
	Date and time of analysis	The date and time the analysis took place
	Date and time of collection	The date and time the data collection took place.
	Spacing {Basecaller Calculated Spacing}	Base spacing used for this analysis. {Spacing as calculated by the Basecaller.} If the two spacing values are different, the sample was analyzed with a user-defined value.

Tiling or Stacking Windows

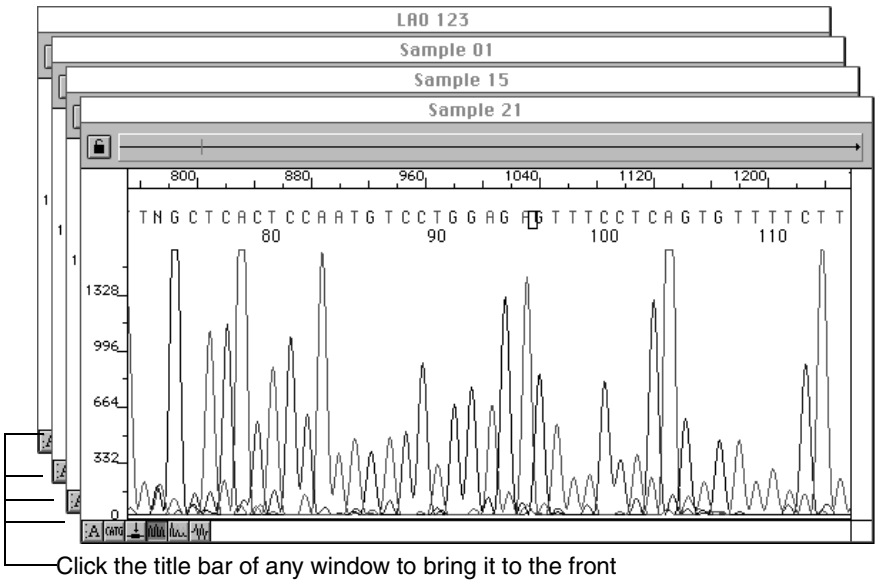
Introduction The number of Sample windows (sample files) that you can have open at one time is limited only by the amount of available computer memory (RAM). You can quickly organize multiple open windows by either tiling or stacking them.

About Tiling To arrange the open sample files so they do not overlap and a good sized portion of each is visible, choose Tile Windows from the Window menu. This method is useful when you have only a few samples open and you want to compare bases, as shown in the following figure.



About Stacking

To arrange a large number of open sample files so they are reduced in size and stacked from back to front, choose Stack from the Window menu. When you choose Stack, only a small amount of each window is visible.



Click the title bar of any window to bring it to the front

To bring any window to the front of the stack: Click the title bar of the window that you want to bring to the front.

If You Try to Open Too Many Windows

If you try to open too many windows for the available memory, the Error Log opens with an “A -108” error listed at the top and a warning dialog box appears.

Many of the menu options and analysis options on the Sample Manager are disabled when memory is low. To free up memory, close some of the open windows or increase the amount of memory allocated to the program (for details, see “Out of Memory dialog box …” on page C-9).

Using the Zoom Commands

Introduction

The Window menu contains four zooming commands that change the amount of data visible in any of the graphic views.

The effects of the four zoom commands are illustrated in the figure on page 7-42.

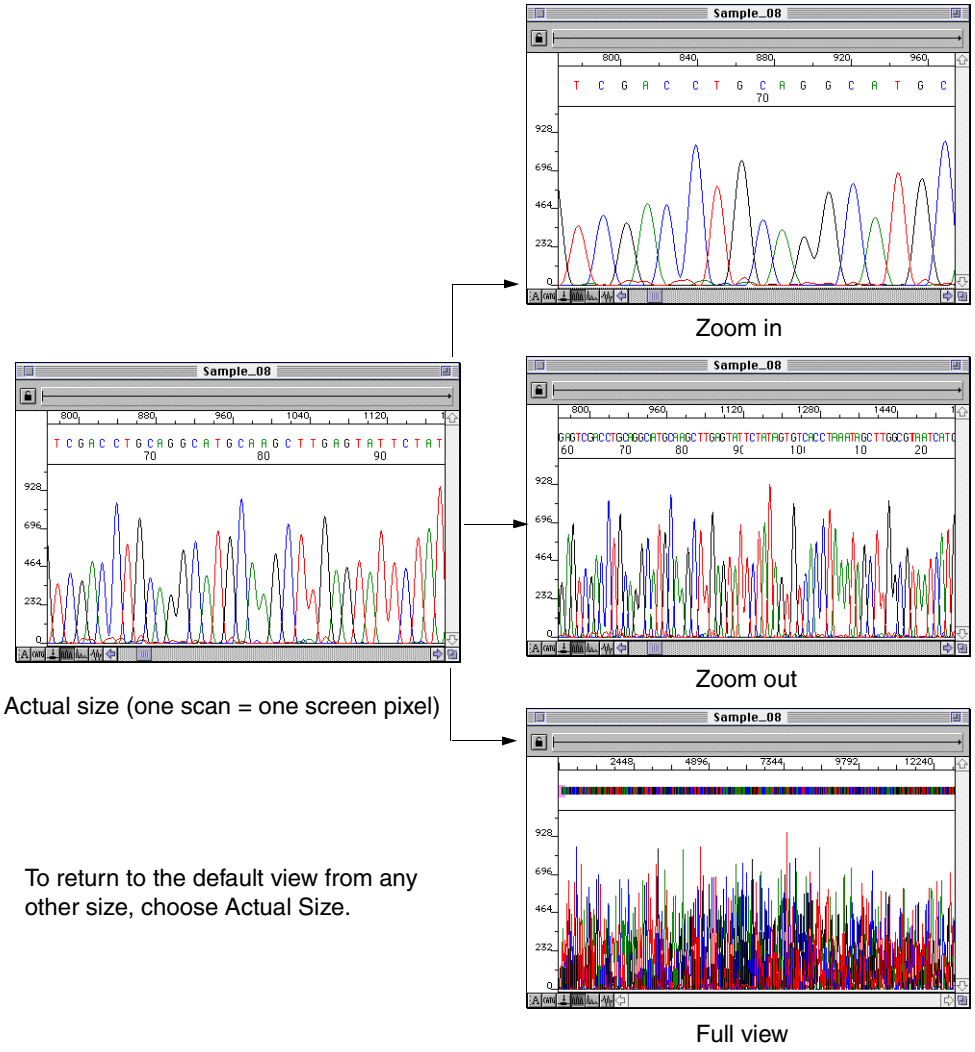
Zooming the View

To zoom a view:

Step	Action
1	Click in the data region that you want to view.
2	To see successively larger scale views of a part of the data, choose Zoom In from the Window menu.
3	To see successively smaller scale views of the data, choose Zoom Out from the Window menu.
4	To scale the data so that the entire length fits within the standard size view window, choose Full View from the Window menu.
5	To return the view to its original size (one scan = one screen pixel) after using Zoom In, Zoom Out, or Full View command, choose Actual Size from the Window menu.

Zoom Commands Illustrated

This example of zoom commands uses the Electropherogram view. The commands also work in Raw Data view and EPT view.



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

Changing the Displayed Lines and Scales

Introduction

For the Electropherogram view, Raw Data view, and EPT view of the Sample window, you can use the Display Options dialog box to:

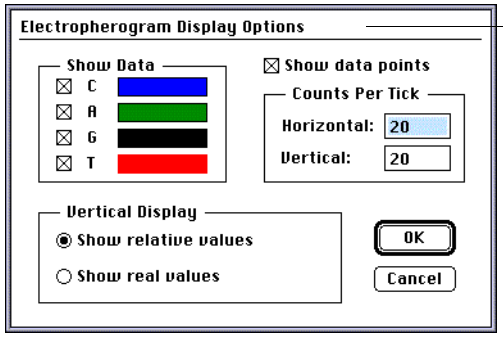
- Determine which color is used to represent each kind of data
- Change the colors of the trace lines to make them easier to see on screen
- Selectively turn off one or more trace line
- Change the type of scaling used for the display
- Change the labeling of the tick marks on the scale in the display

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. There is no return-to-default option for this dialog. The only way to automatically return to the default settings is by deleting the Seq Analysis Prefs file which resets *all* preferences to the default values.

Changing Trace Lines or Scale

Follow the steps below to change the trace lines or the scale on the electropherogram, Raw Data or EPT display.

To change trace line or scale display:

Step	Action
1	Click any Sample window which displays the view (Electropherogram, Raw Data, or EPT) where you want to make the changes.
2	Choose Display Options… from the Windows menu. A Display Options dialog box appears; the exact name and contents depend on which Sample window view is active. 

To change trace line or scale display: *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>3</td><td>To alter which of the four base lines are displayed, use the Show Data check boxes.

You can turn off any combination of lines. This can make it easier to identify heterozygotes, or to hide baseline or noisy data.

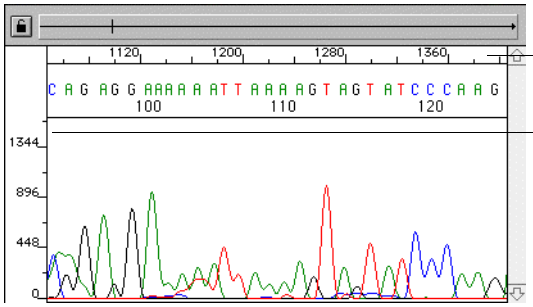
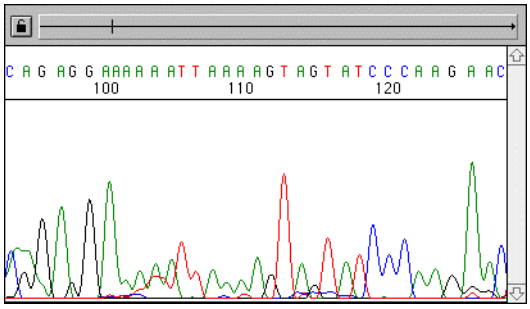
<table>
<tr><th>For the view…</th><th>The four colors represent the…</th></tr>
<tr><td>Electropherogram</td><td>four (analyzed) bases</td></tr>
<tr><td>Raw Data</td><td>detected raw fluorescent signals from the four dyes</td></tr>
<tr><td>EPT</td><td>voltage, temperature, power, and current during the run</td></tr>
</table>

If you turn off the screen display of a trace line, that trace line is also omitted when the view is printed.

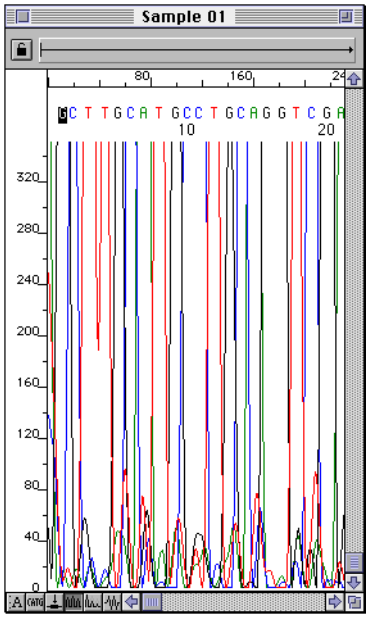
Note The base calls shown on the line at the top of the Electropherogram view window cannot be selectively turned off.</td></tr>
<tr><td>4</td><td>To change the color of a trace line, use the Show Data color bars:
a. Click the color bar to open the Color Picker dialog box.
b. Click the color you want to use, or enter numeric values in the text fields.
c. Choose OK to close the dialog box and change the color.

The color change applies to all sample files displayed until you again change the line color in this dialog box. This change does not affect the colors used on printed Electropherogram, Raw Data, or EPT views.</td></tr>
</table>

To change trace line or scale display: *(continued)*

Step	Action
5	To toggle the display of the vertical and horizontal rulers, use the Show data points check box.   Normally the show data points option is selected. If you select “Show Real Values” for the vertical scale (see “Show real values” below), the maximum vertical value is normally about 1200 full scale and the tick marks are too close together to be useful. Under those conditions, deselect “Show data points.”
6	To change the horizontal and vertical indexing of the rulers, edit the Counts Per Tick text boxes: ♦ Horizontal: The unit of measure (the count) on the horizontal axis is the number of scans performed on the instrument to obtain the data. ♦ Vertical: The vertical axis indicates signal intensity. The Counts Per Tick value is not initially applied when the vertical display is set to “Show relative values.” In this mode, the default Counts Per Tick value (20) would present too many tick marks. If you desire, you can change the Counts Per Tick value to recalibrate the vertical scale.

To change trace line or scale display: *(continued)*

Step	Action
7	Use the Vertical Display radio buttons to toggle between relative and real values for the intensity axis scale. The default setting is Show relative values. ♦ Show relative values compresses the vertical scale of the electropherogram display so that the electropherogram fits within a standard size Sample window. Always select this option unless you have a specific reason to select “Show real values.” ♦ Show real values displays the real scale of the fluorescence data, as shown in the following illustration. Only select this option if you need to see the real data values, for example, to resolve a problem.  **Note** The window above uses the original value (20) for vertical scale and “Show Real Values.” The window is stretched vertically to show a larger portion of the data; it would have to be stretched quite a bit farther vertically to show the tops of the highest peaks, which are at approximately 1200 on the scale.

Command Reference A

Overview

Introduction This appendix briefly describes all the commands on the Sequencing Analysis main menu, their corresponding keyboard shortcuts (if any), and where the main uses of each command are explained.

In This Appendix This appendix includes the following topics:

The File Menu

File Menu Commands

The table below lists and describes the commands accessible from the Sequencing Analysis program File menu.

Command	Keyboard Shortcut	Description	See Page
Open Gel…	Command-H	Opens a gel file. (Command is not present if the Sequencing Analysis software is installed for use with only ABI PRISM® 310 instruments.)	4-5
Open Sample…	Command-O	Opens a sample file. In BioLIMS® mode, this is the Open Sequence… command; it opens a sequence record from the BioLIMS database.	7-4
Close	Command-W	Closes the active window.	
Save	Command-S	Saves the contents of the active window.	4-46
Save As…		Saves the file with the name and format you specify.	4-46
Page Setup…		Allows you to choose page size, orientation, and other settings for printing.	
Print…	Command-P	Allows you to choose which pages to print, the number of copies, etc., and to start printing.	
Quit	Command-Q	Closes the Sequencing Analysis program.	

The Edit Menu

Edit Menu Commands

The table below lists and describes the commands accessible from the Sequencing Analysis program Edit menu.

Note The availability of the various Edit menu commands depends on which type of window is currently active (Gel File, Sequence view, Command Log, etc.).

Command	Keyboard Shortcut	Description	See Page
Undo	Command-Z	Undoes the effects of the most-recent command. (Some commands cannot be undone.)	
Cut	Command-X	Cuts the selected item from the window and puts it on the clipboard.	
Copy	Command-C	Copies the selected item in the window to the clipboard.	
Paste	Command-V	Copies the current clipboard contents to the current cursor location.	
Clear		Cuts the selected item from the window and discards it (without disturbing the current clipboard contents).	5-15
Select All	Command-A	Selects the entire contents of an active Sample window.	
Find…	Command-F	In Sequence view, searches for a specific base or a string of bases.	7-25
Find Again	Command-G	In Sequence view, searches for the next occurrence of the string specified in the Find dialog box.	7-26
Fill Down	Command-D	Copies the value in the topmost selected field to all other selected fields in the same column.	5-15
Preferences		Opens a submenu with options: one for each page in the Preferences dialog box. Allows you to change the values used by the Sequencing Analysis software when it processes gel files and sample files, when it passes a file to the Factura™ program for further processing, and when it prints.	6-21

The Gel Menu

Gel Menu Commands

The table below lists and describes the commands accessible from the Sequencing Analysis program Gel menu.

~~310~~
373
377

Note This menu contains commands that affect gel files. The menu is present only if the Sequencing Analysis software was installed for use with ABI 373 and ABI PRISM® 377 instruments.

Command	Keyboard Shortcut	Description	See Page
Extract Lanes…	Command-L	Copies the raw fluorescence data (and other run information) from the gel file to individual sample files.	4-43
Track Lanes…		Discards the current tracking information and retracks the entire gel image.	4-39
Track & Extract Lanes…		Tracks lanes in the gel image *and* extracts the sample file data.	4-40
Hide/Show Tracker Lines	Command-K	Turns the display of tracker lines off and on.	4-31
Gel Info		Displays information about the run, the gel used, and the gel image.	4-13
Gel Sample Sheet		Displays the information that was included in the data collection Sample Sheet.	4-14
Adjust Gel Contrast…	Command-J	Adjusts the contrast of each of the dyes in the gel image. This does not change the raw data.	4-19
Mark Lane for Extraction	Option-click lane marker	Sets the selected lane marker so that information in that lane will be extracted during the extraction process. Blue markers are used lanes that have already been extracted. White markers are for extraction.	4-30
Mark Lane Used/Unused	Command-click lane marker	Marks the lane as Used. A lane can only be marked for extraction if it is first marked Used. Gray markers are unused.	4-28
Mark All Lanes for Extraction		Sets the lane markers for all Used lanes in the gel so that the information in those lanes will be extracted during the extraction process.	4-30
Unmark All Lanes for Extraction		Sets the lane markers for all Used lanes in the gel so that information in those lanes will **not** be extracted during the extraction process.	4-30

Command	Keyboard Shortcut	Description	See Page
Mark All Lanes Used		Marks all lanes as Used. A lane can only be marked for extraction if it is first marked Used.	
Mark All Lanes Unused		Marks all lanes as Unused. (If you want to extract just one or two lanes, mark all lanes unused, then select and mark for extraction the one or two lanes that you want.)	
Straighten Selected Lanes		Sets the tracker lines straight for the selected lanes. The lane marker position does not alter. The tracker line is drawn as a straight line from the marker with no regard to lane position.	
Force Selected Lanes to Right		Moves selected tracker lines to the far right of the gel image. Lanes are renumbered accordingly. All lanes moved right are stacked on top of each other.	4-28
Regenerate Gel Images…		Regenerates the gel image from the raw data.	4-22
Install New Sample Sheet…		Replaces the current Sample Sheet contained in the gel file with the contents of the Sample Sheet file you select. If you select the wrong Sample Sheet at the time of data collection, this is the best way to repair the error.	4-17
Install New Gel Matrix…		Attaches new instrument file information to the gel file, but does not change the instrument file name in the Sample Sheet.	4-23

The Sample Menu

Sample Menu Commands

The table below lists and describes the commands accessible from the Sequencing Analysis program Sample menu.

Command	Keyboard Shortcut	Description	See Page
Add To Sample Manager	Command-B	Adds the file in the active Sample window to the current Sample Manager.	5-8
Show Original		Displays the original base calls on a separate line above the editable base calls in the Electropherogram view.	7-32

Note In BioLIMS mode, this menu is the Sequence menu and the first item is “Add to Sequence Manager”.

The Manager Menu

Manager Menu Commands

The table below lists and describes the commands accessible from the Sequencing Analysis program Manager menu.

Command	Keyboard Shortcut	Description	See Page
Add Files…	Command-N	Adds files to the Sample Manager.	5-9
Remove Files	Delete	Removes files from the Sample Manager.	5-12
Open Files		Opens all files currently selected in the Sample Manager window.	7-4
Pre-Analysis Settings		For the currently selected file(s) in the Sample Manager, replaces any values that you changed with the original values from the Sample Sheet. (Be careful: there is no undo for this command.)	5-15
Start	Command-M	Starts processing of files in the Sample Manager.	5-20
Pause		Temporarily stops the processing of files in the Sample Manager.	5-20
Resume		Continues a paused processing of files in the Sample Manager.	5-20
Cancel	Command-Period (.)	Cancels the processing of files in the Sample Manager.	5-21

The Window Menu

Window Menu Commands

The table below lists and describes the commands accessible from the Sequencing Analysis program Window menu.

Command	Keyboard Shortcut	Description	See Page
Zoom In	Command- =	Enlarges an area of the active window so that more detail is visible. (Available for the three graphical views of the Sample window.)	7-41
Zoom Out	Command- –	Reduces the scale of the active window, so that you can see a larger area. (Available for the three graphical views of the Sample window.)	7-41
Full View	Command- [	Displays all the data in a standard-size window. (Available for the three graphical views of the Sample window.)	7-41
Actual Size	Command-]	Displays the contents of the window at 1:1 scale, no matter what scale is displayed at the time you select this command. (Available for the three graphical views of the Sample window.)	7-41
Display Options…		Changes display options (e.g. ruler indexing, relative or real values, visible base traces). (Available for the three graphical views of the Sample window.)	7-43
Show/Hide Sample Manager	Command-1	Opens (or closes) the Sample Manager window.	5-3
Show/Hide Command Log	Command-2	Opens (or closes) the window that displays a list of commands performed by the Sequencing Analysis software.	C-15
Show/Hide Error Log	Command-3	Opens (or closes) the window that displays a list of all errors that occurred during analysis.	C-13
Tile Windows		Arranges the open Sample windows so they do not overlap and a good sized portion of each is visible.	7-39
Stack Windows		Arranges windows so they are the same size and stacked from back to front, with only the title of each visible.	7-39
(window names)		Lists all currently open Sequencing Analysis program windows.	

Keyboard Shortcuts for the Gel File Window

Lane Marker Shortcuts

Try these shortcuts to make manual gel tracking faster and easier.

To…	Do this…
select a lane marker	click on the marker.
select multiple lane markers	Shift-click on the marker.
select a lane marker when another lane marker is selected	use the Tab key to select a marker to the right or use Shift-Tab to select a marker to the left.
move a lane marker	click the marker then use the left or right arrow keys to move the lane marker and tracker line.
toggle a lane marker between used and unused	Command-click on the lane marker.
mark or unmark a lane for extraction	Option-click on the lane marker.

Control Point Shortcuts

Try these shortcuts for moving, adding, deleting, and viewing control points.

To…	Do this…
select a control point	click on the control point.
select one or more control points	drag diagonally across the point(s) to select.
select the control point above or below the current point	use the up or down arrow key.
in vertical expand mode, to select the control point above or below the current point and scroll to that point	hold down the Shift key and use the up or down arrow key.
move one or more control points left or right in one-channel increments	select the control point(s), then use the left or right arrow key.
move one or more control points left or right in 0.1-channel increments	select the control point(s), then hold down the Option key and use the left or right arrow key.
add a row of control points	Option-click in the area between the row selector.
delete a row of control points	Option-click on a row selector and select OK in the dialog box.
to move between guide lines in interpolation mode	use the Tab key.

Keyboard Shortcuts for Sample Window Views

Sample Window Views

The table below shows the buttons and keyboard commands that set the view of the Sample window.

Button	Command	Keyboard Shortcut	Description	See Page
	Change to Annotation View	Command-E	Switches from any other Sample window view to Annotation view.	7-7
	Change to Sequence View	Command-R	Switches from any other Sample window view to Sequence view.	7-7
	Change to Feature View	Command-T	Switches from any other Sample window view to Feature view.	7-7
	Change to Electrophero-gram View	Command-Y	Switches from any other Sample window view to Electropherogram view.	7-7
	Change to Raw Data View	Command-U	Switches from any other Sample window view to Raw Data view.	7-7
	Change to EPT Data View	Command-I	Switches from any other Sample window view to EPT Data view.	7-7

Input and Output Files

B

Overview

Introduction

This appendix describes the files that contribute information for the operation of the Sequencing Analysis software (input files) and the files created by the software (output files).

Some of these files *must* be located in the System Folder on your computer. Others can be kept in various locations, depending on the type of instrument used and your personal preference.

In This Appendix

This appendix includes the following topics:

Input and Output Files in the System Folder

Introduction

The System Folder on the hard disk of your Macintosh® computer contains assorted files that are used by the Sequencing Analysis software, as well as the preferences file and log files which are created by the software.

With the exception of the preferences file, the Sequencing Analysis system files are contained in the ABI Folder within the System Folder.

About the ABI Folder

When ABI PRISM software is installed, a special folder, named the ABI Folder, is created in the System Folder. To this folder are added various important system files required for running ABI PRISM software.

Note When Sequencing Analysis software is installed, if there is already an ABI Folder (used by other ABI PRISM software) then a new folder is not created — the Sequencing Analysis system files are added directly to the existing ABI Folder.

ABI Files in the System Folder

The following table lists the Sequencing Analysis files that must be present in the System Folder.

File Type	Folder Location in System Folder	Description
Sequence DyeSet/Primer files (input)	ABI Folder	Contain dye and primer mobility information. Applied Biosystems supplies these files, which are used by both the Data Collection software and the Sequencing Analysis Basecaller program.
Instrument file (also called the matrix file) (input)	ABI Folder	Contains three mathematical matrices that correct for spectral overlap. The matrix to be applied to the data is specified by the user prior to analysis, based on the dyes and the chemistry used to prepare the samples.
Command Log (output)	ABI Folder	Lists all commands performed by the Sequencing Analysis software, either as requested directly, or in the course of analysis. **Note** If this file is deleted or removed from the System Folder, a new log file is generated by the Sequencing Analysis program.
Error Log (output)	ABI Folder	Lists all errors that occurred during analysis. **Note** If this file is deleted or removed from the System Folder, a new log file is generated by the Sequencing Analysis program.
Preferences file (input)	Preferences folder	Record Preferences selected in the Sequencing Analysis program. **Note** If this file is deleted or removed from the System Folder, a new default preferences file is generated by the Sequencing Analysis program.

Input Files Not Located in the System Folder

Introduction

The following table describes the input files that are *not* located in the System Folder. The locations shown are the system defaults. You can change the locations of ABI PRISM® 310 and ABI PRISM® 377 files, and specify the new locations of these in the Preferences Folder Locations dialog box (in the Data Collection software).

Input Files External to the System Folder

The following table lists the input files for Sequencing Analysis software that are not stored in the System Folder.

File Type	Folder Location	Description
Program files (Sequencing Analysis, Basecaller, DataUtility)	Sequencing Analysis 3.4 folder	Provide the primary input that analyzes data. The Sequencing Analysis program analyzes the data sent from the ABI PRISM instrument after the run is complete. The Basecaller and Tracker are opened automatically by the Sequencing Analysis program as needed; you specify which to use. The DataUtility allows you to make and copy matrices. **IMPORTANT** Do not move or rename these files.
Gel file from ABI 373 instrument	Sequencing Analysis 3.4 folder	A large file created by the Data Collection program. The gel file contains all of the original raw data from all channels of the gel. For a typical run, a gel file can be very large (20–90 MB).
The gel file from ABI PRISM 377 (or ABI 373 XL) instrument	Individual Run folder inside the Runs folder inside the ABI PRISM 377 or (ABI 373 XL) folder[a]	
Sample AppleScript® scripts	Sample Scripts folder inside the Sequencing Analysis 3.4 folder	Can be used to develop AppleScript routines that are tailored to your site.
Tracker program file	SAGelTracker[b] folder inside the Sequencing Analysis 3.4 folder	Neural Net Tracker program launched by the Sequencing Analysis v. 3.4 software to track gel files.

File Type	Folder Location	Description
Tracker Settings Files	SAGelTracker[b] folder inside the Sequencing Analysis 3.4 folder	Tracker settings files: ♦ SA194Tracker34SHK.mat (for 194 channels and 32–36 lane shark-tooth gels) ♦ SA388Tracker48SHK.mat (for 388 channels and 48 lane shark-tooth gels) ♦ SA388Tracker64SHK.mat (for 388 channels and 64 lane shark-tooth gels) ♦ SA480Tracker96SHK.mat (for 480 channels and 96 lane shark-tooth gels) **IMPORTANT** Do not move or rename these files.
Tracker Extensions	SAGelTracker[b] folder inside the Sequencing Analysis 3.4 folder	These three extensions are required for the Tracker program to run: ♦ libmatlb ♦ libmcc ♦ libtbx **IMPORTANT** Do not rename or move these files.

a. This location is the default location for a run folder, but the folder can be placed anywhere on the hard disk.

b. Do not move or change the name of this folder.

Output Files Not Located in the System Folder

Introduction

The Sequencing Analysis program creates two types of data files for analyzed data, and two log files.

Additionally, when Sequencing Analysis software is installed, the installer program creates a file named “Installer Log File” and places it in the Sequencing Analysis 3.4 folder.

The two log files (Command Log and Error Log) are located in the System Folder, as described on page B-3.

Note When Sequencing Analysis is in BioLIMS mode, no data files are created by the program.

Output Files External to the System Folder

The following table describes the three output files created by Sequencing Analysis software that are not stored in the System Folder.

File Type	Location for ABI PRISM 310	Location for ABI PRISM 377 and ABI 373XL	Description
Sample Files	Individual Run folder in the Runs folder inside the ABI PRISM 310 folder[a]	Next to the gel file in a folder in a Run folder marked with the date and time the gel was run[a]	The sample file contains six parts: annotation, features table, sequence, chromatogram (electropherogram), raw data, and EPT data (electrophoresis conditions). It combines information from the Sample Sheet, raw data, and analysis conditions and results. The annotation has Sample Sheet information and analysis results like basespacing and signal strength. The features table contains the results from Factura™ processing.
.Seq Files	Individual Run folder in the Runs folder inside the ABI PRISM 310 folder[a]	Next to the gel file in a folder in a Run folder marked with the date and time the gel was run[a]	Text files that contain the base letter sequence only. You can create these files in several formats, and can open and print them from word processing programs. You can also import .Seq files into other programs that accept text files in these formats.

File Type	Location for ABI PRISM 310	Location for ABI PRISM 377 and ABI 373XL	Description
MatLab text files		SAGelTracker folder inside the Sequencing Analysis 3.4 folder	MatLab text files (with extension .mat) are output to the SAGelTracker whenever a gel is tracked. You can ignore these files; if you throw them away, new ones are created when you next track a gel.

a. This location is the default choice of run folder location, but the individual run folders can be stored anywhere on the hard disk.

DyeSet/Primer File Naming Conventions

Introduction

The files installed by the Sequencing Analysis program have specific filenames that provide information about the files. Many filenames are self-explanatory (log files, for instance). The naming conventions for the DyeSet/Primer files are less straightforward.

File Naming Conventions

The DyeSet/Primer File names use a combination of characters to indicate the chemistry (*e.g.*, dye primer, dye terminator), gel concentration, and gel type. The abbreviations are as follows:

Abbreviation	Meaning
DP	Dye Primer chemistry was used.
DT	Dye Terminator chemistry was used.
X%	The approximate percent of the gelling agent that was used.
Ac	(Acrylamide) — For ABI 373 and ABI PRISM 377 runs, the type of gel used. Currently Applied Biosystems offers files which are compatible with acrylamide type gels.
LR	(Long Ranger™) — For ABI 373 and ABI PRISM 377 runs, the type of gel used.
POP6	For ABI PRISM 310 runs that use Performance Optimized Polymer, POP-6™ polymer.
DSP	For ABI PRISM 310 runs that use DNA Sequencing Polymer.
CEHV	(Capillary Electrophoresis High Viscosity) — For ABI PRISM 310 runs that use DNA Sequencing Polymer.
{XX}	Additional information such as the filter set (or virtual filter set), primer, and chemistry.

Example 1 The filename DP4%Ac{–21M13} indicates:

Dye Primer chemistry (DP)
4% acrylamide gel (4%Ac)
The –21 M13 primer ({–21M13})

Example 2 The filename DT POP6{BD Set-Any Primer} indicates:

Dye Terminator chemistry (DT)
Performance Optimized Polymer (POP6)
BigDye terminator using any custom primer ({BD Set-Any Primer})

Troubleshooting Sequencing Analysis

C

Overview

Introduction

This appendix describes various problems that can occur when using the Sequencing Analysis software, and what to do about each problem.

If You Do Not Find Help Here

For additional information about troubleshooting sequence data, see the *ABI PRISM Automated DNA Sequencing Chemistry Guide*, (P/N 4305080) or call Technical Support (see page 1-20).

In This Appendix

This appendix includes the following topics:

General Troubleshooting Hints

Two Suggestions

The following two general suggestions can be helpful in a variety of situations:

- Check the Error Log for recent error messages.
- Check the size of the sample and gel files as compared to the usual size for similar files at your site.

Checking the Error Log

In the Error Log window, note the message number and description for any error messages you find. Look for the error message in this troubleshooting chapter. If it is not here, call Applied Biosystems Technical Support and tell them both the number and description of the error message. (See “Technical Support” on page 1-20.)

For information about viewing and printing the Error Log, see “Reviewing the Sequencing Analysis Error Log” on page C-13.

Checking the Size of Data Files

Compare the sizes of the gel file and sample files with the size of data files for similar files at your site.

Gel File

If a gel file is much smaller than normal, the run data is probably missing.

Sample File

Sample files are normally 70–80 KB when they contain only raw data, and up to 250 KB after analysis. If the file is either too small or too big, there is probably something wrong with the data.

Troubleshooting Error Log Messages

Introduction

Errors that occur during analysis appear in the analysis software Error Log. When you encounter an error message in the log, consult the table below for the meaning of the error message and suggested corrective action.

For more details about the Error Log, see “Reviewing the Sequencing Analysis Error Log” on page C-13.

Error Message Listing

The following table lists some of the more common error messages and what action to take for each.

For help with error messages that do *not* appear in this table, contact Applied Biosystems Technical Support. (See “Technical Support” on page 1-20.)

Table of Common Error Messages

Error Message	Observed Symptoms	Recommended Action
Error #0. Could not do this task because of a program error. The tracker application could not be loaded. A shared library may be missing.		Make sure that the three extensions: libmatlb, libmcc, and libtbx, are within the SAGelTracker folder with the SAGelTracker program.
Error #0. Changed gel image resolution. Regenerate gel image.	Gel was previously opened with another version of Sequencing Analysis.	Ignore this message. The Sequencing Analysis program regenerated the gel image.
Error #22. Could not do this task because of a program error. Multicomponent matrix error. Bad data…		Reanalyze the sample file with a better instrument file.
Error -43. Could not do this task because file not found. The tracker application is missing.	Sequencing Analysis failed to launch the Tracker program.	Check that the Tracker program, SAGelTracker, is in the SAGelTracker folder in the Sequencing Analysis v. 3.4 folder and that it is correctly named.
Error -61. Could not do this task because the edition is not a publisher.	Occurs when making changes to the Sample Sheet.	Ignore this message. Changes to the Sample Sheet are saved despite the error message.

Table of Common Error Messages *(continued)*

Error Message	Observed Symptoms	Recommended Action
Error –108 error	Some Sample Manager window menu commands become unavailable. Program produces erratic results or crashes.	Close unneeded windows and other programs. Use the Get Info window to allocate more memory (see “To allocate more memory:” on page C-9 for details). Do not open more than 25 Sample (or Sequence) windows at one time.
Error -40, SetMark Error Power Failure.	No gel image; no scroll bar, blank gel image. In some cases, the gel file is present with analyzed data. Some cases will have truncated gel file and sample files (less scan points than a normal run).	Choose the Regenerate Gel Image command from the Sequencing Analysis Gel menu.
Error -2700 Could not do this task because the tracking process failed. Runtime error: Couldn’t open file…	Tracker settings file missing from the SAGelTracker folder.	Find the setting file, check the name and that it is in the SAGelTracker folder. (See list of settings files on page B-5.) If the settings file is correct, it could be that the gel is too difficult for the Tracker program and you should track the gel manually. (See “Adjusting Lane Markers” on page 4-25.)
Error –10023, TDGetBuff: Tag not found.	Gel image present; no tracker lines or lane markers. Gel file is present and correct size: no analyzed data.	Choose the Regenerate Gel Image command from the Sequencing Analysis Gel menu. If this does not recover the file, the instrument file may be corrupted; re-install it from backup. Or, the instrument file may be incomplete and need to be remade. (See “Running Standards and Viewing Raw Sample Files” on page E-10.)

Table of Common Error Messages *(continued)*

Error Message	Observed Symptoms	Recommended Action
Error –10024, TDGetBuf: DATA error, Tag not found; TDPutBuf: Bad Data Size; Cannot process file, data length less than 50.	Possibly none; gel file of the correct size. Sample files created (2K–6K in size); no raw or analyzed data.	Choose the Regenerate Gel Image command from the Sequencing Analysis Gel menu.
TDOpen: Mobility file error, File not found.	Analysis fails. No effect on gel file or sample file.	Select the file for analysis again. Update the DyeSet/Primer file in the Sample Manager window. Reanalyze the file. The DyeSet/Primer file must be located in the ABI Folder in the System Folder of the computer where analysis is performed.
Any Limit Check PostScript error.	Printing fails.	Reduce number of pages printed or number of data points per page. Perhaps add memory to printer.
Sample file input error, File not found.	Some Sample Manager window menu commands become unavailable. Program produces erratic results or crashes.	Either find the missing instrument file and place in the ABI folder or pick a different instrument file in the Sample Manager from the pop-up menu.

Troubleshooting Other Types of Sequencing Analysis Software Problems

Introduction This section describes various problems that can occur, and how to resolve each type of problem.

Table

Troubleshooting Table

Observation	Possible Cause	Recommended Action
Tracking fails	The gel is not multicomponented	Check that the gel is multicomponented with the correct instrument (matrix) file. (See step 1 on page 4-13.) If not, then: a. Install a new instrument file if necessary (See "Installing New Matrix Information" on page 4-23.) b. Regenerate the image with the Multicomponent Gel Image box checked. (See "Regenerating the Gel Image with Different Option Values" on page 4-22.)
Tracking fails	Comb Type on Gel Preferences page not set to Shark Tooth	Select Preferences… Gel Preferences from the Edit menu. Set Comb Type to Shark Tooth.
Tracking fails	Tracker program or settings files are misplaced or misnamed	Check that the Tracker program, SAGelTracker, and the required settings file are in the SAGelTracker folder in the Sequencing Analysis v. 3.4 folder. Check that the files are correctly named — refer to the list of settings filenames on page B-5.
Tracking fails	Not enough memory to run the Tracker program.	Choose About This Computer… under the Apple menu. Check to see that the largest block of unused memory is at least 20 MB. If it is less than 20 MB, you may need to defragment your hard disk using a defragmenting utility program (e.g. Norton Utilities).
Tracking fails	Sample lane does not contain red data.	Track lane by hand.

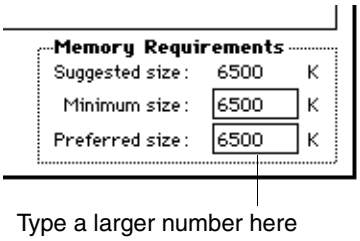
Troubleshooting Table *(continued)*

Observation	Possible Cause	Recommended Action
No base calling occurred when you chose Start to begin analysis.	No Basecaller is running	Carry out the following steps in the order listed: a. Check that the correct type of Basecaller is selected in the Sample Manager window. b. Check that the Basecaller program is installed in the same folder as the Sequencing Analysis program. c. Rebuild the desktop (hold down the Command and Option keys while you restart the computer) to update the desktop database so the Sequencing Analysis software will be able to recognize and start the correct Basecaller program. d. If there is an extensions conflict, hold down the Shift key while you restart the computer to turn off all extensions. Then start the Sequencing Analysis program, add the files to the Sample Manager, and start analysis.
Analysis with Adaptive Basecaller: Status Bar of Sample Manager says that Analysis is successful and A check box is green. Error Log gives message: “Error#23. Could not do this task because of a program error. "Sample Name" Error computing raw mobility shifts, using file. Adaptive Processing Failed.”	The adaptive basecaller algorithm could not compute a good mobility shift with the raw data of the sample file. Sequencing Analysis then uses the mobility (dye/primer set) file that is attached to the sample file.	Understand that the Adaptive Basecaller is using the mobility file that was provided with the sample. If the Status message is “successful analysis”, then check your data as usual. If the Status message is “failed analysis”, then reanalyze your data with different parameters.

Troubleshooting Table *(continued)*

Observation	Possible Cause	Recommended Action
Only the first portion of the sequence is called, or there is no analysis even though the raw data shows signal.	The endpoint defined in the Basecaller Preferences page is "Set endpoint after __ Ns in __ bases", and there is noisy data at the start of the run.	Carry out the following steps: ♦ Uncheck the Set endpoint box in the Basecaller Preferences page (page 6-28). Or, change the Start Point to skip the noisy "false start" data (page 5-6). ♦ Set Peak 1 Location or Start Point to zero. ♦ Reanalyze the sample file. **Note** This "false start" behavior is seen more often with terminator chemistry than with primer chemistry.
Base spacing value is a negative number or appears as a large, bold, red number.	The base spacing value is a measure of speed of DNA migration through the gel. If the reported value is a negative number, the base spacing value was outside the allowed range (8.5–16) so the program used a default spacing of 9 to call the bases. A reported value that is a negative number indicates a possible gel or running buffer problem, or weak or noisy data.	The optimal value for a sample depends on your instrument and run configuration (see the *ABI PRISM Automated DNA Sequencing Chemistry Guide*, (P/N 4305080) for more details). To see the base spacing value for a sample, look at Annotation view of the Sample window, the top of the printed electropherogram, or the Spacing column of the Sample Manager. If the reported negative value occurred while the ABI PRISM instrument was running under normal conditions, look for a gel or running buffer problem, or weak or noisy data. Also, you can try using a different Basecaller. (See "Choosing a Basecaller" on page 6-46.) For more information, see "Negative Spacing" on page C-17 and "The Spacing Parameter" on page 6-9.
Menu and analysis options are disabled in Sample Manager window.	Many of the menu options and analysis options in the Sample Manager window are automatically disabled when memory is low.	See the following problem, "Out of Memory dialog box appears..."

Troubleshooting Table *(continued)*

Observation	Possible Cause	Recommended Action
Out of Memory dialog box appears during tracking or extracting of a gel.	Too many windows are open and there is not enough memory for tracking and extracting. During the Track and Extract phase of gel processing, almost all of the program memory is required for the gel processing.	(a) Do *not* try to extract data while you have two gel files, or one gel file and several sample files, open. (b) If you must have those windows open at the same time, allocate more memory to the program, as described below. To allocate more memory: a. Quit the Sequencing Analysis program. b. In the Finder, click the Sequencing Analysis program icon once to select it. Then choose Get Info (Command–I) from the File menu.  **IMPORTANT** Do not double-click the icon. The program must remain closed. c. Highlight the number in the Preferred Size entry field in the lower right corner of the Get Info window, then type a larger number in its place.  d. Click the close box at the top left corner to close the Get Info window. When you start the Sequencing Analysis software, the Finder will allocate the newly specified amount of memory to the program.

Troubleshooting Table *(continued)*

Observation	Possible Cause	Recommended Action
Printing is slow.	Graphical view pictures (gel image, and Electropherogram, Raw Data, and EPT views) contain many bytes of data.	To increase printing speed, make the following changes. ♦ Turn off background printing in the Chooser dialog box. ♦ Select the fast print or draft printing option (if available on your printer). ♦ On the Printing Preferences page of the Preferences dialog box, and in the special Page Setup dialog box that is available whenever a Sample window is open, do the following: a. Select the Print First Page Only check box. b. Select to print only the Electropherogram view, rather than multiple views. c. Decrease the Number of Panels per Page value and/or the Points per Panel value. At most, print no more than five panels of 1500 points each for the Electropherogram view.
Printer crashes while printing sample file views.	Printing analyzed data and raw data at the same time can overload some types of printers and cause printing to fail.	If the problem is lack of printer memory, add more memory to your printer. See also the options under “Printing is slow,” above.
Trouble printing sample files on A4 paper		Instead of printing by the Print command from the File menu, add the sample files to the Sample Manager and print by selecting the Print box (only) and clicking the Start button.

Troubleshooting Table *(continued)*

Observation	Possible Cause	Recommended Action
Printed sample file format is incorrect.	Printing page of the Preferences dialog box is incorrectly set.	Check that all the printer settings are correct. ♦ Select the printer and printer driver through the Chooser. ♦ On the Printing Preferences page of the Preferences dialog box: a. For most printers, make sure the PostScript Printer check box is selected. If you have a non-PostScript printer, you may need to de-select this option. b. Check that the Panels per Page and Points per Panel values are correct. c. Click the Page Setup Options button to open the Page Setup dialog box. Check that all options are set correctly in that dialog box, then close the dialog box. d. Click the Print Options button to open the Printer dialog box. Check that all options are set correctly in that dialog box. e. In the Printer dialog box, click the Options button to open the Print Options dialog box. Check that the Color/Grayscale button is selected in that dialog box. Then close the Print Options dialog box and the Printer dialog box.
Sequencing Analysis software unexpectedly quits.	Running the Sequencing Analysis software and another RAM-intensive program at the same time.	Do not run the Sequencing Analysis software at the same time that you run another RAM-intensive program.
Sequencing Analysis quits with a low memory error.	Too many Sample (or Sequence) windows are open at once.	Do not open more than 25 Sample (or Sequence) windows at one time.
Sequencing Analysis (in BioLIMS mode) quits with a low memory error.	Too many sequences are in the Sequence Manager window.	Do not put more than 300 sequences into the Sequence Manager window.

Troubleshooting Table *(continued)*

Observation	Possible Cause	Recommended Action
Signal strength is below 40. It is important that the signal from your sequence samples be higher than the background fluorescence of the plate and gel; an average signal strength of 40 or above is generally adequate. If the signal strength number for any of the four bases in the sample is below forty, there *might* be a problem with the data. (Signal strength numbers are shown in Annotation view of the Sample window and in the header of the printed electropherogram.)	Two possible causes are: a. The reported signal strength number for each base is an average value, calculated over the range of analyzed data points. If the Start point and Stop point are not defined correctly, data values on either end of the run can skew the averaged value. b. The sample or reaction did not work well or the data is too weak.	Examine both your raw data and your analyzed data closely for possible problems. ♦ If appropriate, change the analysis Start and Stop points, then reanalyze the data. ♦ If the sample or reaction was not satisfactory, rerun the sample. **Note** Signal strength is very dependent on the chemistry. For example, the dRhodamine terminator chemistry typically has weaker fluorescence, but signal strengths below 40 are rarely a problem with this chemistry. **Note** For more information, see the *ABI PRISM Automated DNA Sequencing Chemistry Guide*, (P/N 4305080).
Sequencing Analysis (with BioLIMS) crashes on launch or gives an error message, 'Could not open " " '.	Insufficient memory to load the Oracle® or Sybase® libraries.	Be sure that there is at least 2 MB of free memory in addition to the preferred memory requirements set on the Sequencing Analysis Info box. (To access the Info box, select the Sequencing Analysis and choose Get Info from the File menu.)
Sequencing Analysis (with BioLIMS) crashes when attempting to connect to a Sybase-based BioLIMS database.	SybaseConfig control panel or libtcp extension file is missing or disabled.	Use the Extension Manager to check that these files are present and turned on. If either file is missing, reinstall the BioLIMS Client software.

Reviewing the Sequencing Analysis Error Log

Introduction

The Error Log lists all errors that occurred during analysis and can be useful for troubleshooting.

Where Is the Error Log File?

The Error Log is maintained in a file called “Seq Analysis Error File”, which resides in the ABI Folder in the Macintosh® System Folder. If this file is removed from the ABI Folder, a new error file is created automatically when the Sequencing Analysis program next opens.

Can the Error Log Become Full?

No, the Error Log never becomes full. When it reaches its maximum length of 200 lines, the oldest messages are automatically deleted from the log as new messages are added to the top.

Lane Assignment Confidence Values

In addition to errors, the lane assignment confidence value calculated during gel tracking is also written to the Error Log. You should make a point of checking the lane assignment confidence value for each gel tracked. For more information about lane confidence values, see “Stop Extraction When Below Confidence Threshold” on page 6-26.

Reviewing the Error Log

If the Error Log is not already visible, choose Show Error Log (Command–3) from the Window menu.

The menu command changes to Hide Error Log (so you can choose the command again to hide the log) and the Error Log window appears.

In the Error Log window, the newest entry is at the top of the list.

Note You cannot select multiple lines or edit the Error Log.

Printing the Error Log

To print a copy of the Error Log:

Step	Action
1	Choose Print from the File menu while the Error Log window is active.
2	In the Printer dialog box, either select All to print all the log pages, or type in the page numbers for the range of pages you want to print. **Note** Because the most recent entry is at the top of the log file, most often, it is enough to print only the first one or two pages of the file.
3	Choose Print.

Reviewing the Sequencing Analysis Command Log

Introduction

The Command Log lists all commands performed by the Sequencing Analysis software, either as requested directly, or in the course of analysis. The newest command appears at the top of the list.

This log can be very useful during troubleshooting. It also can help you remember where you stopped if you are interrupted while using the software.

Where Is the Command Log File?

The Command Log is maintained in a file called “Seq Analysis Command File”, which resides in the ABI Folder in the Macintosh System Folder. If this file is removed from the ABI Folder, a new command file is created automatically when the Sequencing Analysis program next opens.

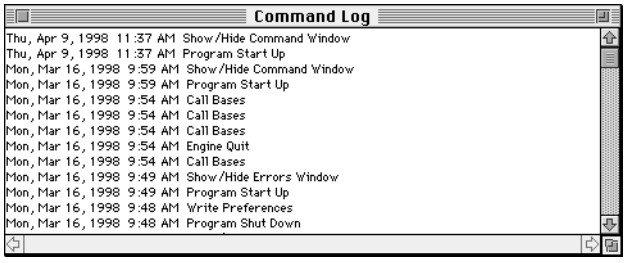
Can the Command Log Become Full?

No, the Command Log never becomes full. When it reaches its maximum length, the oldest messages are automatically deleted from the log as new messages are added to the top.

Reviewing the Command Log

Choose Show Command Log (Command–2) from the Window menu.

The menu command changes to Hide Command Log (so you can choose the command again to hide the log) and the Command Log window appears.



In the Command Log window, the newest entry is at the top of the list.

Note You cannot select multiple lines or edit the Command Log.

Printing the Command Log

To print a copy of the Command Log:

Step	Action
1	Choose Print from the File menu while the Command Log window is active.
2	In the Printer dialog box, either select All to print all the log pages, or type in the page numbers for the range of pages you want to print. **Note** Because the most recent entry is at the top of the log file, most often, it may be enough to print only the first few pages of the file.
3	Choose Print.

Troubleshooting with the Printed Electropherogram

Introduction

Two items of information in the header of the printed electropherogram can be especially useful for troubleshooting:

- Signal Strength
- Spacing

This information is also found in the annotation view of the Sample window.

Signal Strength Values

The signal strength numbers at the top of the third column, when below forty (40) for any one base *might* indicate a problem with the data. (Signal strength is very chemistry dependent.) If any of the signal strength numbers falls below 40, you should examine both your raw data and your analyzed data closely for possible problems.

Spacing Values

Negative Spacing

When the Basecaller could not properly analyze your data, it displays the spacing value for the sample as a negative number (Annotation view) or as a bold, red value in the Sample Manager window. When you see out-of-range spacing displayed this way, try analyzing this data with the SemiAdaptive or Adaptive Basecaller to see if it produces a better result. Or try changing the spacing in the Sample Manager. (See “Changing the Spacing for a Sample” on page 6-9.)

Negative spacing is a signal from the Basecaller program that it could not properly analyze your data. Try analyzing this data with the SemiAdaptive or Adaptive Basecaller to see if it produces a better result. Or try changing the spacing in the Sample Manager. (See “Changing the Spacing for a Sample” on page 6-9.)

Negative spacing values can be caused by:

- The DNA running through the gel too quickly or too slowly.
- A high level of noise in the sample due to low template concentrations, fluorescent contaminants, secondary priming, two primers, or poor priming, which gives rise to weak signal.
- Short PCR fragments where less than 3000 data points of actual sequence are present.

Basecaller Default Spacing

When the Basecaller cannot accurately measure the peak spacing in a sample, it assumes a default spacing value and examines the data based on that value. For example, when the ABI100 Basecaller cannot accurately measure the spacing, it assumes the spacing is 9.00, corrects the data based on a spacing of 9.00, enters a bold, red 9.0 in the Sample Manager window, and enters a -9.0 in both the Annotation view of the Sample window and in the header information of the printed analyzed chromatogram.

The default spacing values for ABI Basecallers are shown below.

Basecaller	Default Spacing
ABI50, ABI100, ABI200, SemiAdaptive, Adaptive	9.00
ABI-CE1, ABI-CE2	12.00

Troubleshooting the BioLIMS Database

D

Introduction

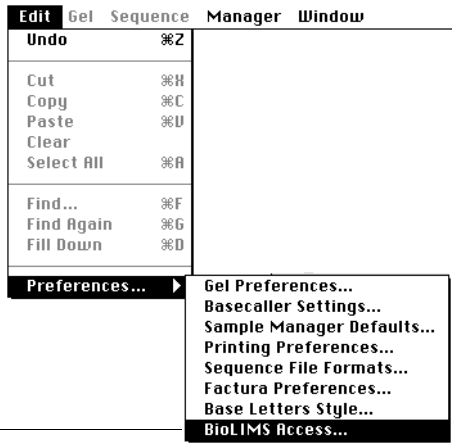
In This Appendix

This appendix describes how to troubleshoot connection problems between the Macintosh client and the BioLIMS database server.

This appendix includes the following topics.

If the BioLIMS Preference Page Does Not Appear

Problem If the BioLIMS® Access page is not present after “Base Letters Style” in the Preferences dialog box, the Sequencing Analysis program is unable to find all the database support files and system extensions required to access the BioLIMS database.



This BioLIMS Access menu item will not appear if certain database support files are missing.

Solution Be sure that all the Oracle® or Sybase® database support files are installed correctly. If files are missing, reinstall the BioLIMS Client or Instrument package from the original CD-ROM.

About Troubleshooting the Client to Sybase Connection

Introduction

A common source of difficulty using the BioLIMS System is establishing connection between the BioLIMS programs running on Macintosh® client computers and the BioLIMS database on the Sybase server.

Note For information about troubleshooting an Oracle server connection, see "About Troubleshooting the Client to Oracle Connection" on page D-11.

SybPing and Telnet

To help with troubleshooting, two programs have been provided:

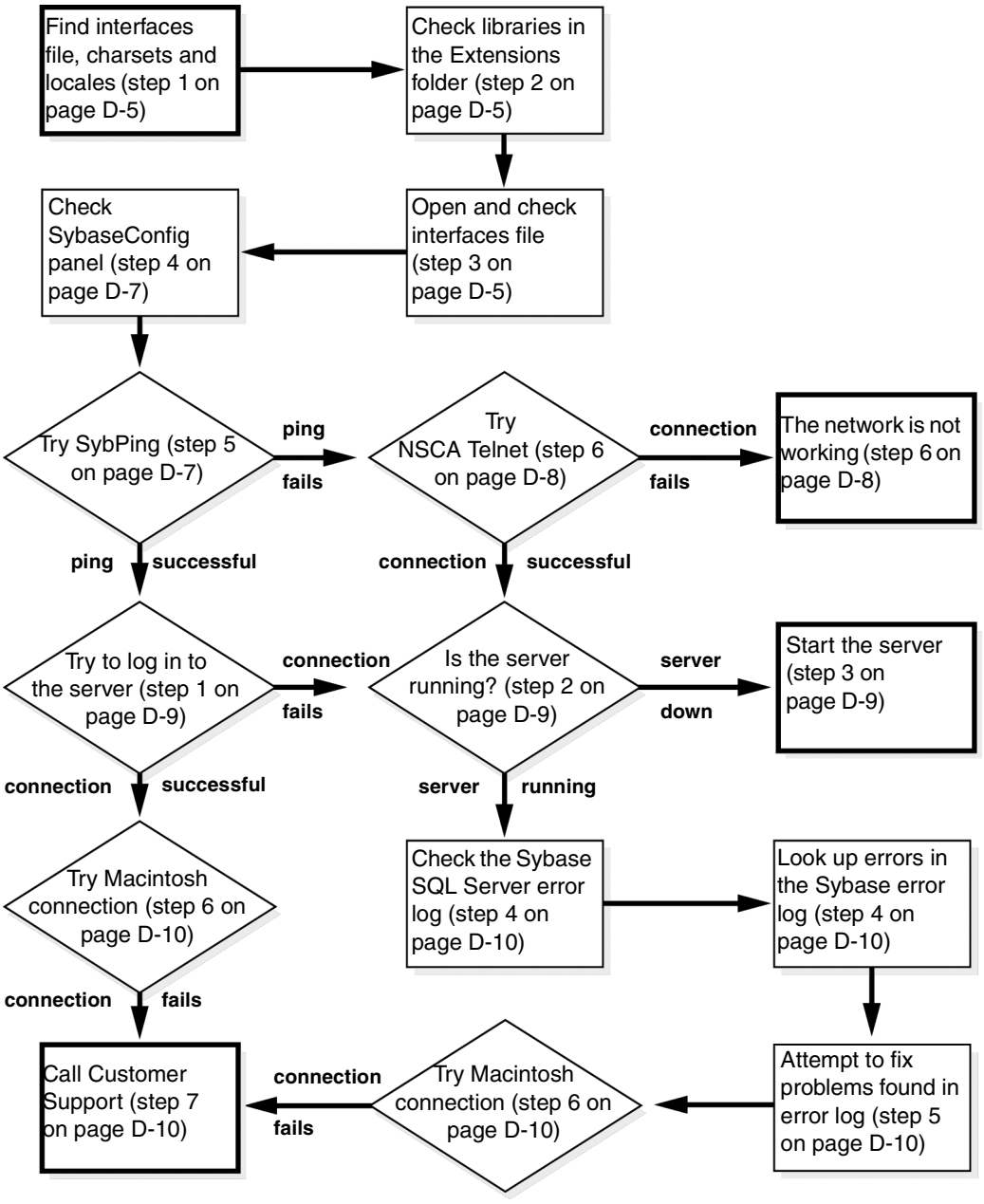
- **SybPing**—a Sybase tool for testing network connections. (Look for this application in the BioLIMS:BioLIMS Extras:Sybase:bin folder.)
- **NCSA Telnet 2.6**—a program used for interactive access from a Macintosh client to a telnet host on TCP/IP networks. NCSA Telnet was developed by the National Center for Supercomputing Applications at the University of Illinois in Urbana/Champaign. (Look for this application in the BioLIMS:BioLIMS Extras folder.) Applied Biosystems does not support NCSA Telnet.

  You can reach them at the address shown below:
  http://www.ncsa.uiuc.edu/SDG/Software/Brochure/Overview/MacTelnet.overview.html

The Troubleshooting Process for the Sybase Database Connection

Troubleshooting Flow Chart

On the following page is a flow chart illustrating the process for troubleshooting database connection between the Macintosh Client and the Sybase server. The step numbers given in the flow chart refer to the tables on pages D-5 to D-10 where the troubleshooting procedures are described in detail.


Find interfaces file, charsets and locales (step 1 on page D-5)
Check libraries in the Extensions folder (step 2 on page D-5)
Open and check interfaces file (step 3 on page D-5)
Check SybaseConfig panel (step 4 on page D-7)
Try SybPing (step 5 on page D-7)
ping
fails
Try NSCA Telnet (step 6 on page D-8)
connection
fails
The network is not working (step 6 on page D-8)
ping
successful
connection
successful
Try to log in to the server (step 1 on page D-9)
connection
fails
Is the server running? (step 2 on page D-9)
server
down
Start the server (step 3 on page D-9)
connection
successful
server
running
Try Macintosh connection (step 6 on page D-10)
Check the Sybase SQL Server error log (step 4 on page D-10)
Look up errors in the Sybase error log (step 4 on page D-10)
connection
fails
Call Customer Support (step 7 on page D-10)
connection
fails
Try Macintosh connection (step 6 on page D-10)
Attempt to fix problems found in error log (step 5 on page D-10)

Procedures for Troubleshooting the Client to Sybase Connection

Introduction

The procedure for troubleshooting the Client-Server connection is divided into two parts:

- Troubleshooting from the Macintosh client (below)
- Troubleshooting from the UNIX server (page D-9)

Troubleshooting from the Client

The step numbers in the following procedure correspond to the steps marked in the flow chart above.

IMPORTANT The documentation for each BioLIMS Macintosh program includes a section about setting up the database connection. Make sure you have followed that procedure carefully.

Troubleshooting the connection from the client (Macintosh):

Step	Action
1	Locate the interfaces file. Confirm that the folders charsets and locales are located in the same folder as the interfaces file. The default installation places interfaces, charsets, and locales in the BioLIMS:BioLIMS Extras:Sybase folder.
2	Confirm that the Sybase library files are installed into the Extensions folder in the System Folder. See “Required Sybase Extension Files” on page D-8. In particular, if the libtcp extension is missing or not turned on, the Sequencing Analysis application crashes when connection to the database is attempted.
3	Open the interfaces file and confirm the server information is correct. An example of the Sybase-based BioLIMS server entry is shown below: *SYBASE* `query MacTCP mac_ether` *neuron.apldbio.com 2500* **Note** Refer to the table immediately below for an explanation of the BioLIMS server entry.

Troubleshooting the connection from the client (Macintosh): *(continued)*

Step	Action

The BioLIMS server entry:

Where…	Represents…
SYBASE	the name you chose to call the BioLIMS Server. The server can have any name. (“About Server Names” on page 3-4.)
query MacTCP mac_ether	the part of the entry that is always the same. **Note** The tab preceding this phrase is required.
neuron.apldbio.com	the host and domain name of the BioLIMS Sybase SQL Server machine. In this example, neuron is the host name and apldbio.com is the domain name. You can also use an IP address. This information is available from your system administrator.

2500	the port number that the BioLIMS Sybase SQL Server is using to connect with the clients. This number is assigned to the server when it is installed. You can find the port number in the interfaces file, which is located in the home directory of the BioLIMS Sybase SQL Server. This is an example of the server entry in the UNIX server interfaces file:
	`## SYBASE on neuron`
	`##  Services:`
	`##  query  tcp  (2500)`
	…
	The third line of this entry shows the port number (2500) of the BioLIMS server.

Troubleshooting the connection from the client (Macintosh): *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>4</td><td>Confirm that the SybaseConfig control panel is set up correctly:
a. Open the SybaseConfig control panel. This control panel must be located in the Control Panels folder in the System folder. (If it is missing or disabled, Sequencing Analysis crashes when connection to the database is attempted.)
b. Click the Interfaces File button to open a browser box.
c. Use the browser box to locate and select the interfaces file that you located in step 1 above.
d. Confirm that the correct default server is chosen. Use the Default Server pop-up menu to select the default server. (This is the same name you chose to call the BioLIMS Server in the step above.)
e. Confirm that the Default Language pop-up menu is set to us_english.</td></tr>
<tr><td>5</td><td>Use the program SybPing to confirm communication with the BioLIMS Sybase SQL Server™.

The default installation places SybPing into the BioLIMS: BioLIMS Extras: Sybase: bin folder.
a. Start the SybPing program.
b. Select the server from the Servers pop-up menu. (The servers shown in the pop-up menu are the servers listed in the interfaces file.)
c. Click Ping.

The program responds with a message of whether or not the Ping was successful.

<table><tr><th>If the Ping is…</th><th>Then…</th></tr><tr><td>successful</td><td>skip to “Troubleshooting the client connection from the Sybase SQL Server:” on page D-9.</td></tr><tr><td>unsuccessful</td><td>continue to step 6.</td></tr></table></td></tr>
</table>

Troubleshooting the connection from the client (Macintosh): *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>6</td><td>If the Ping is unsuccessful:

Open the program NCSA Telnet. The default installation places NCSA Telnet into the BioLIMS: BioLIMS Extras folder.

a. Select Open Connection... from the File menu.

b. Enter the host name (from the interfaces file) into the Host/Session Name text field.

c. Click Connect.

<table>
<tr><th>If the connection is…</th><th>Then…</th></tr>
<tr><td>successful</td><td>a window is displayed with a UNIX login prompt. Skip to step 2 of “Troubleshooting the client connection from the Sybase SQL Server:” on page D-9.</td></tr>
<tr><td>unsuccessful</td><td>an error message is displayed. The network is not working. Call your network administrator.

Note If the host.domain name does not work, try the IP address of the host. This number can be found in the /etc/hosts file on the UNIX server.</td></tr>
</table>
</td></tr>
</table>

Required Sybase Extension Files

The following Sybase library files are placed in the Extensions folder in the System Folder by the BioLIMS Client and Instrument installers. These files are required for connection to the BioLIMS database.

- libblk
- libcomn
- libcs
- libct
- libctb
- libintl
- libsybdb
- libtcl
- libtcp

Troubleshooting from the Sybase Server

The step numbers in the following procedure correspond to the steps marked in the flow chart on page D-4. (Step 1 below corresponds to the diamond “Try to log in to the server…” on page D-4.)

You need to have the account name and password for the Sybase user on the UNIX system that runs the Sybase SQL Server. If you do not have access to the Sybase user account, you should ask your database administrator to carry out the following procedure.

Troubleshooting the client connection from the Sybase SQL Server:

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>1</td><td>Log in to the Sybase user account on the UNIX server.

Try to connect to the Sybase SQL Server with isql. Use the same client user name as the one entered in the BioLIMS access dialog box. For example:

% isql -U george -P george1 -S SYBASE
1> use sfdb
2> go
1> quit

<table>
<tr><th>Where…</th><th>Represents the…</th></tr>
<tr><td>george</td><td>client user name</td></tr>
<tr><td>george1</td><td>client user password</td></tr>
<tr><td>SYBASE</td><td>server name</td></tr>
<tr><td>sfdb</td><td>name of the BioLIMS database</td></tr>
</table>

<table>
<tr><th>If the Login is…</th><th>Then…</th></tr>
<tr><td>successful</td><td>skip to step 6.</td></tr>
<tr><td>unsuccessful</td><td>continue to step 2.</td></tr>
</table>
</td></tr>
<tr><td>2</td><td>Use the showserver script to find out if the Sybase SQL Server is running.

<table>
<tr><th>If the server is…</th><th>Then…</th></tr>
<tr><td>running</td><td>skip to step 4.</td></tr>
<tr><td>not running</td><td>continue to step 3.</td></tr>
</table>
</td></tr>
<tr><td>3</td><td>Ask your database administrator to restart the Sybase SQL Server.

After restarting the server, try to connect to the database using isql as in step 1 on page D-9. If the connection still fails, go to step 7.</td></tr>
</table>

Troubleshooting the client connection from the Sybase SQL

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>4</td><td>Check the Sybase error log and make a note of any error messages. (The error log can be found in the install directory of the sybase home directory and error messages are preceded by the string Msg.)</td></tr>
<tr><td>5</td><td>Refer to the Sybooks™ documentation for an explanation of the SQL Server Error Messages. Attempt to fix the problem following the Sybook instructions.

Try to connect to the database using isql as in step 1 on page D-9.

<table><tr><th>If the connection is…</th><th>Then…</th></tr><tr><td>successful</td><td>continue to step 6.</td></tr><tr><td>unsuccessful</td><td>skip to step 7.</td></tr></table></td></tr>
<tr><td>6</td><td>Try once more to connect to the BioLIMS database from the BioLIMS Macintosh program. Make sure that the user name, password, database and server names are all typed correctly and are in the correct case.

If the connection still fails, go to step 7.</td></tr>
<tr><td>7</td><td>Call Customer Support. See “Technical Support” on page 1-20.</td></tr>
</table>

About Troubleshooting the Client to Oracle Connection

Introduction

A common source of difficulty using the BioLIMS System is establishing connection between the BioLIMS programs running on Macintosh client computers and the BioLIMS database on an Oracle server.

Note For information about troubleshooting a Sybase SQL Server connection, see “About Troubleshooting the Client to Sybase Connection” on page D-3.

Telnet

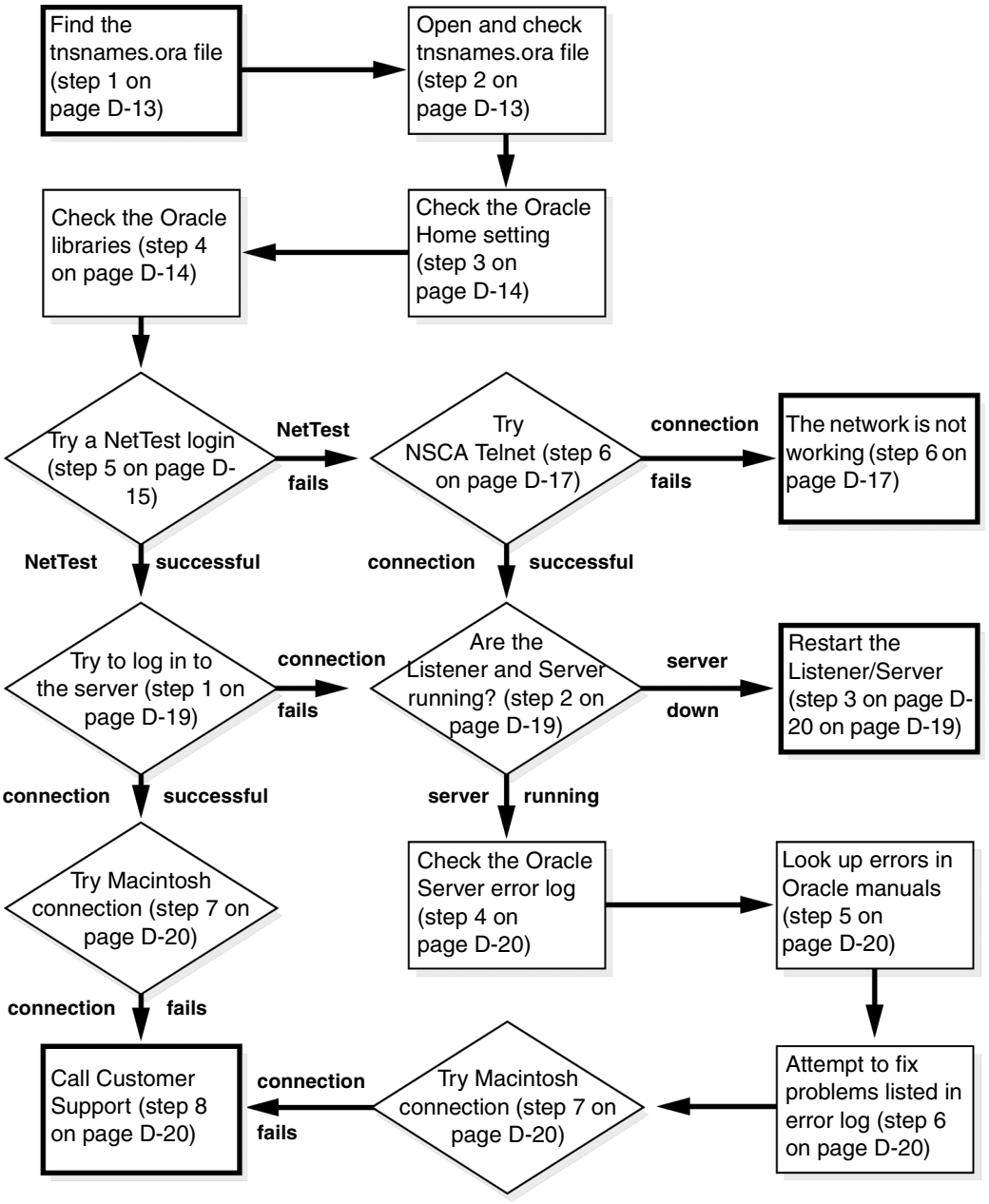
To help with troubleshooting, NCSA Telnet 2.6 has been provided in the BioLIMS Extras folder.

NCSA Telnet 2.6 is an program used for interactive access from a Macintosh client to a telnet host on TCP/IP networks. NCSA Telnet was developed by the National Center for Supercomputing Applications at the University of Illinois in Urbana/Champaign.

The Troubleshooting Process for the Oracle Database Connection

Troubleshooting Flow Chart

On the following page is a flow chart illustrating the process for troubleshooting the database connection between the Macintosh Client and the Oracle server. The step numbers given in the flow chart refer to the tables on pages D-13 to D-20 where the troubleshooting procedures are described in detail.


Find the tnsnames.ora file (step 1 on page D-13)
Open and check tnsnames.ora file (step 2 on page D-13)
Check the Oracle Home setting (step 3 on page D-14)
Check the Oracle libraries (step 4 on page D-14)
Try a NetTest login (step 5 on page D-15)
NetTest fails
Try NSCA Telnet (step 6 on page D-17)
connection fails
The network is not working (step 6 on page D-17)
NetTest successful
connection successful
Try to log in to the server (step 1 on page D-19)
connection fails
Are the Listener and Server running? (step 2 on page D-19)
server down
Restart the Listener/Server (step 3 on page D-20 on page D-19)
connection successful
server running
Try Macintosh connection (step 7 on page D-20)
Check the Oracle Server error log (step 4 on page D-20)
Look up errors in Oracle manuals (step 5 on page D-20)
connection fails
Call Customer Support (step 8 on page D-20)
connection fails
Try Macintosh connection (step 7 on page D-20)
Attempt to fix problems listed in error log (step 6 on page D-20)

Procedures for Troubleshooting the Client to Oracle Connection

Introduction

The procedure for troubleshooting the Client-Server connection is divided into two parts:

- Troubleshooting from the Macintosh client (below)
- Troubleshooting from the UNIX server (page D-19)

Troubleshooting from the Client

The step numbers in the following procedure correspond to the steps marked in the flow chart above.

IMPORTANT The documentation for each BioLIMS Macintosh program includes a section about setting up the database connection. Make sure you have followed that procedure carefully.

Troubleshooting the connection from the client (Macintosh):

Step	Action
1	Locate the tnsnames.ora file. The default installation places the tnsnames.ora file in the BioLIMS 2.0:BioLIMS Extras:Oracle:Network:Admin folder.
2	Open the tnsnames.ora file and confirm that the server information is correct. An example of the Oracle-based BioLIMS server entry is shown below.

```
Oramozart =   (DESCRIPTION=
                  (ADDRESS=
                  (PROTOCOL=TCP)(host=mozart)(port=1521))
                   (CONNECT_DATA=(SID=WG733)
                  )
          )
```

Note Refer to the table immediately below for an explanation of the BioLIMS server entry.

Troubleshooting the connection from the client (Macintosh): *(continued)*

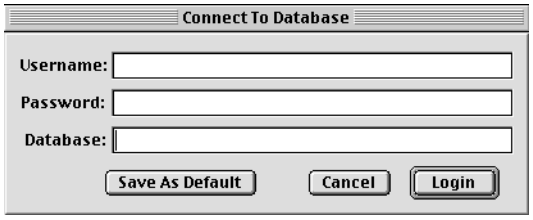
<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td colspan="2">The BioLIMS server entry:

<table>
<tr><th>Where…</th><th>Represents…</th></tr>
<tr><td>Oramozart</td><td>the name you chose to call the Oracle-based BioLIMS database connection. This can have any name. See “About Server Names” on page 3-4.</td></tr>
<tr><td>mozart</td><td>the host of the Oracle server for the BioLIMS database. In this example, mozart is the host name. You can also use a host and domain name such as mozart.apldbio.com or an IP address such as 167.116.100.61.

This information is available from your system administrator. The IP address can also be found in the /etc/hosts file on the UNIX server.</td></tr>
<tr><td>1521</td><td>the port number that the Oracle server for the BioLIMS database is using to connect with the clients. This number is assigned to the server when it is installed. You can find the port number in the tnsnames.ora file that is located in the $ORACLE_HOME/network/admin directory on the Oracle server.</td></tr>
<tr><td>WG733</td><td>the Oracle database system identifier (sid) defined at server installation.

You can find the port number in the tnsnames.ora file that is located in the $ORACLE_HOME/network/admin directory on the Oracle server.</td></tr>
</table></td></tr>
<tr><td>3</td><td>Confirm that the Oracle Home is set correctly:
a. Open the Set Oracle Home program. At installation this program is placed into the BioLIMS 2.0:BioLIMS Extras:Oracle:Applications folder.
b. Use the file browser to locate and select the Oracle folder.</td></tr>
<tr><td>4</td><td>Confirm that the Oracle library files are installed. See “Required Oracle Extension Files” on page D-18.</td></tr>
</table>

Troubleshooting the connection from the client (Macintosh): *(continued)*

Step	Action
5	Use the program NetTest to confirm communication with the Oracle server. At installation, NetTest is placed into the BioLIMS 2.0:BioLIMS Extras:Oracle:Applications:Networking folder. a. Open the NetTest program. **Note** No window appears for the NetTest program, but the menu bar changes. b. From the Database menu, choose Logon. The Connect To Database dialog box appears.  c. Enter the Username and Password for the database. **Note** These are the same Username and Password you use to log into BioLIMS (*e.g.* george, george1). d. In the Database field, enter the alias name of the Oracle Server from the tnsnames.ora file (*e.g.,* Oramozart).

Troubleshooting the connection from the client (Macintosh): *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td></td><td>e. Click Login. The message “Attempting Connection...” appears, then the Result Explanation dialog box appears.
<table>
<tr><th>If the login is…</th><th>Then…</th></tr>
<tr><td>successful</td><td>the Result Explanation dialog box displays the following:
♦ Result field: “ORA-00000: normal, successful completion”
♦ Connected To field: Oracle server version information
♦ Comments field: “Your SQL*Net setup appears to be in working order”
Skip to “Troubleshooting from the Oracle Server” on page D-18.</td></tr>
<tr><td>unsuccessful</td><td>the Result Explanation dialog box displays the following:
♦ Result field: the Oracle error encountered
♦ Comments field: A detailed explanation of the login failure
Continue to step 6.</td></tr>
</table></td></tr>
</table>

Troubleshooting the connection from the client (Macintosh): *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>6</td><td>If the NetTest login is unsuccessful:

Open the program NCSA Telnet. The default installation places NCSA Telnet into the BioLIMS 2.0:BioLIMS Extras folder.

a. Select Open Connection... from the File menu.

b. Enter the host name (from the tnsnames.ora file) into the Host/Session Name text field.

Note You can enter the host, host and domain name, or IP address here. See step 2, “mozart,” on page D-14 for more information.

c. Click Connect.

<table>
<tr><th>If the connection is…</th><th>Then…</th></tr>
<tr><td>successful</td><td>a window is displayed with a UNIX login prompt. Skip to step 2 of “Troubleshooting from the Oracle Server” on page D-18.</td></tr>
<tr><td>unsuccessful</td><td>an error message is displayed. The network is not working. Call your network administrator.

Note If you entered the host name in step 6b above and the connection fails, try entering the host.domain name or the IP address before calling your network administrator. See step 2, “mozart,” on page D-14 for more information.</td></tr>
</table>
</td></tr>
</table>

Required Oracle Extension Files

The BioLIMS Client and Instrument Installers place the following Oracle library files into the BioLIMS 2.0:BioLIMS Extras:Oracle:Libraries folder. Running Set Oracle Home places an alias to this folder into the Macintosh System Folder:Extensions folder.

- OracleCore23Lib
- OracleKernel71Lib
- OracleNetNLLib
- OracleNetTCPLib
- OracleNetTNSLib
- OracleNetTNSTCPLib
- OracleNLS23Lib
- OracleOci71Lib
- OracleOra71Lib
- OraclePlsql21Lib
- OraclePstd21Lib
- OracleRuntime13Lib
- OracleSql16Lib
- OracleTNSATKLib
- OracleVsoci71Lib

Troubleshooting from the Oracle Server

The step numbers in the following procedure correspond to the steps marked in the flow chart on page D-12.

You need to have the account name and password for the Oracle user on the UNIX system that runs the Oracle server. If you do not have

access to the Oracle user account, you should ask your database administrator to carry out the following procedure.

Troubleshooting the client connection from the Oracle server:

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>1</td><td>Log in to the Oracle user account on the UNIX server.

Try to connect to the Oracle Server with SQL*Plus. Use the same client user name as the one entered in the BioLIMS access dialog box. For example:

% sqlplus george/george1
SQL> exit

<table>
<tr><th>Where…</th><th>Represents the…</th></tr>
<tr><td>george</td><td>client user name.</td></tr>
<tr><td>george1</td><td>client user password.</td></tr>
</table>

<table>
<tr><th>If the Login is…</th><th>Then…</th></tr>
<tr><td>successful</td><td>skip to step 7.</td></tr>
<tr><td>unsuccessful</td><td>continue to step 2.</td></tr>
</table>
</td></tr>
<tr><td>2</td><td>Find out if the Listener Process is running.

As the Oracle user, type:

% lsnrctl status

<table>
<tr><th>If the Listener Process is…</th><th>Then…</th></tr>
<tr><td>not running</td><td>continue to step 3.</td></tr>
<tr><td>running</td><td>Check that the server is running. Type:

% ps -ef | grep ora_

<table>
<tr><th>If server processes are…</th><th>Then…</th></tr>
<tr><td>displayed</td><td>server is running. Skip to step 4.</td></tr>
<tr><td>not displayed</td><td>continue to step 3.</td></tr>
</table>
</td></tr>
</table>
</td></tr>
</table>

Troubleshooting the client connection from the Oracle server: *(continued)*

Step	Action
3	See your system administrator for help restarting the Listener Process or Server. After restarting, try to connect to the database using SQL*Plus as in step 1 on page D-19. If the connection still fails, go to step 8.
4	Check the Oracle error log and make a note of any error messages. (The error log can be found in the install directory of the oracle home directory and error messages are preceded by the string Msg.)
5	Refer to the Oracle documentation for an explanation of the error messages.
6	Attempt to fix the problem following the documentation instructions. Try to connect to the database using SQL*Plus as in step 1 on page D-19. If the connection still fails, go to step 8.
7	Try once more to connect to the BioLIMS database from the BioLIMS Macintosh program. Make sure that the user name, password, database, and server names are all typed correctly and are in the correct case. If the connection still fails, go to step 8.
8	Call Customer Support. See “Technical Support” on page 1-20.

Creating Instrument Files

E

Overview

Introduction This appendix describes how to create and change instrument files. Instrument files contain matrix information specific to each filter set and chemistry.

In This Appendix This appendix includes the following topics:

Summary of the Instruments and Chemistries

The Sequencing Chemistries

Five cycle sequencing chemistries are currently available to prepare DNA samples for ABI PRISM® genetic analysis instruments.

Chemistry	Applicable to…
Fluorescein/Rhodamine Dye Primers	373 with filter set A, and 310 and 377 with virtual filter set A
Rhodamine Dye Terminators	373 with filter set A, and 310 and 377 with virtual filter set A
dRhodamine Terminators	373 with BigDye filter wheel, and 310 and 377 with virtual filter set E
BigDye™ Terminators	
BigDye™ Primers	

Dye Labels Specific to each Chemistry

Each chemistry has a specified set of dye labels that emit fluorescence when excited by a laser. Each dye label in the set emits fluorescence at a different wavelength, and these emissions are detected during data collection.

On the ABI PRISM® 377 and the ABI PRISM® 310, the wavelengths are separated by a spectrograph into a predictably spaced pattern across a CCD camera.

On the ABI 373, the wavelengths are separated by physical band-pass filters, and the signal is amplified by a photomultiplier tube (PMT).

Colors in Real-Time Data Display Windows

ABI 373 Filter Wheel

~~310~~
373
~~377~~

On the ABI 373, a physical four-filter wheel or five-filter wheel is used to separate the wavelengths emitted by the fluorescent dyes.

There are three types of filter wheel.

Filter Wheel	Filter Set
Original four-filter wheel	Default (no choice)
Five-filter wheel	Filter set A and Filter set B
BigDye filter wheel	Filter set A

The CCD Camera

310
~~373~~
377

The ABI PRISM 377 and ABI PRISM 310 data collection programs collect the fluorescent signal from specific locations on a CCD camera. These locations correspond to different wavelengths of light. The result is the same as using a physical filter to separate the light wavelengths. This is referred to as a virtual filter, since no physical filtering hardware is used. (See the instrument *User's Manual* for more details.)

Real-Time Display Colors Vary

310
373
377

On the real-time displays (the Scan window and the ABI 373/ABI PRISM 377 Gel File window), the data collection program displays the light intensities, color-coded according to wavelength. Blue, green, yellow, and red (in that order) represent the wavelengths of the dye emissions within each dye set. Blue represents the shortest wavelength, and red represents the longest. The colors on the real-time displays therefore represent the wavelengths of the dyes being detected, rather than the bases being detected.

Colors Represent Relative Wavelengths

Different filter sets (both virtual and physical) use the same four colors to represent different wavelengths, so the colors do not represent actual wavelengths. They represent the *relative* wavelengths of the four dyes in each dye set. For example, Filter Set A uses the four colors to represent wavelengths within Dye Set 1 and Dye Set 2.

Each of the chemistries used for preparing DNA is associated with a dye set. Each dye set labels the four bases differently, so the relative wavelength, and therefore the color, associated with each base varies with the chemistry used to label it. Due to this, the four colors on the

real-time displays represent different bases, depending on the chemistry used for labeling.

The tables below describe the colors that represent each of the four bases on the real-time displays for the ABI PRISM instruments.

Color Guide for ABI PRISM 377 and 310

The following tables lists the raw data display colors and dyes for the ABI PRISM 377 (gel image and raw data) and ABI PRISM 310 (raw data). There are two virtual filter sets that are used with sequencing chemistry. Be sure to choose the correct run modules and DyeSet/Primer (mobility) files for the chemistry used.

310
~~373~~
377

Raw Data Colors for Virtual Filter Set A

	Fluorescein/Rhodamine Dye Primers		Rhodamine Dye Terminators	
Color	**Base**	**Dye**	**Base**	**Dye**
Blue	C	5-FAM	G	R110
Green	A	JOE	A	R6G
Yellow	G	TAMRA	T	TAMRA
Red	T	ROX	C	ROX

Raw Data Colors for Virtual Filter Set E

	dRhodamine Terminators		BigDye Primers		BigDye Terminators	
Color	**Base**	**Dye**	**Base**	**Dye**	**Base**	**Dye**
Blue	G	dR110	C	FAM-dR110	G	FAM-dR110
Green	A	dR6G	A	FAM-dR6G	A	FAM-dR6G
Yellow	C	dTAMRA	G	FAM-dTAMRA	T	FAM-dTAMRA
Red	T	dROX	T	FAM-dROX	C	FAM-dROX

Color Guide for ABI 373

The following tables list the raw data display colors and dyes for the ABI 373 gel image and raw data.

Raw Data Colors for Virtual Filter Set A

	Fluorescein/Rhodamine Dye Primers		Rhodamine Dye Terminators	
Color	**Base**	**Dye**	**Base**	**Dye**
Blue	C	5-FAM	G	R110
Green	A	JOE	A	R6G
Yellow	G	TAMRA	T	TAMRA
Red	T	ROX	C	ROX

~~310~~
373
~~377~~

Raw Data Colors for Filter Set A (BigDye Filter Wheel)

	dRhodamine Terminators		BigDye Primers		BigDye Terminators	
Color	**Base**	**Dye**	**Base**	**Dye**	**Base**	**Dye**
Blue	G	dR110	C	FAM-dR110	G	FAM-dR110
Green	A	dR6G	A	FAM-dR6G	A	FAM-dR6G
Yellow	C	dTAMRA	G	FAM-dTAMRA	T	FAM-dTAMRA
Red	T	dROX	T	FAM-dROX	C	FAM-dROX

After Analysis Color Guide

~~310~~
373
377

The Sequencing Analysis program converts the information collected by the data collection program, so that after analysis the colors representing each base are consistent regardless of the chemistry used. The colors on all displays of analyzed data, including printed electropherograms, are as follows:

Color Guide for All Analyzed Data

Base	Color
C	Blue
A	Green
G	Black[a]
T	Red

a. G is shown as yellow in AutoAssembler™ software.

ABI 373 Instrument Configurations

Three Filter Wheels

There are three filter wheels for the ABI 373 instrument.

Instruments With a Four-Filter Wheel

The four filters contained in the filter wheel are unique to your instrument. The wavelength centers of detection of the individual filters are 531, 560, 580, and 610 nm. Use this filter set only with fluorescein/rhodamine dye primers and rhodamine dye terminators. The table below summarizes the relationship between the filters and dyes for the individual sequencing reaction chemistries.

Four-Filter Wheel

Filter Center Band (nm)	Fluorescein/Rhodamine Dye Primers	Rhodamine Dye Terminators
531	C–Rxn	ddG
560	A–Rxn	ddA
580	G–Rxn	ddT
610	T–Rxn	ddC

For more details, see the *ABI PRISM Automated DNA Sequencing Chemistry Guide* (P/N 4305080) or the instrument *User's Manual*.

Instruments with a Five-Filter Wheel

The five filters contained in the filter wheel are unique to your instrument. The wavelength centers of detection of the individual filters are 531, 545, 560, 580, and 610 nm. Only four of the filters are used for each sequencing run (one for each dye).

Two filter sets (A and B) are available with the five-filter wheel. The filters used in set A are 531, 560, 580, and 610 nm. Use set A with the fluorescein/rhodamine dye primers and rhodamine dye terminators. The filters used in set B are 531, 545, 560, and 580 nm.

Use Filter Set B for GeneScan® applications that use different dyes. (Sequenase (T7) terminator chemistry, originally collected on Filter Set B is now obsolete.)

The table below summarizes the relationship between the filters and dyes for the individual sequencing reaction chemistries.

Five-Filter Wheel

Filter Center Band (nm)	Fluorescein/ Rhodamine Dye Primers	Rhodamine Dye Terminators
531	C–Rxn	ddG
545	—	—
560	A–Rxn	ddA
580	G–Rxn	ddT
610	T–Rxn	ddC

IMPORTANT You cannot use both filter sets on a single 373 run.

Instruments with the BigDye Filter Wheel

If you want to use dichlororhodamine (dRhodamine)-based sequencing chemistries exclusively, you can have the BigDye filter wheel installed on your ABI 373 instrument.

BigDye Filter Wheel

Filter Center Band (nm)	dRhodamine Terminators	BigDye Primers	BigDye Terminators
540	ddG	C–Rxn	ddG
570	ddA	A–Rxn	ddA
595	ddC	G–Rxn	ddT
625	ddT	T–Rxn	ddC

IMPORTANT Once you have installed the BigDye filter wheel, these three chemistries (dRhodamine Terminators, BigDye Primers, BigDye Terminators) are the only ones that you can use.

The Instrument File

Correction for Spectral Overlap

Although the dyes fluoresce at different wavelengths, there is some overlap in the spectra. To correct for this overlap when analyzing data, a mathematical matrix is created for each dye set and stored in a file called the instrument file. The instrument file must contain a matrix for each chemistry that you run on the instrument. During data analysis the appropriate matrix is applied to remove any spectral overlap.

What Does the Instrument File Contain?

The instrument file normally contains

- Three matrices
- A comment field
- An instrument name field

These can be seen in the Copy Matrix window in the DataUtility program. A copy of this instrument file is attached to every gel file and sample file when these files are first created. For this reason, each computer on which you use the Sequencing Analysis program must have an instrument file in the ABI Folder (which is located in the System Folder).

IMPORTANT Due to slight variations in the filters of the ABI 373 instruments, and the CCD cameras of the ABI PRISM 377 and ABI PRISM 310 instruments, the instrument file created for your ABI PRISM genetic analysis instrument is sub-optimal for other ABI PRISM genetic analysis instruments.
If you analyze sample files on a different computer from the one that was used to collect data, be sure to copy the correct instrument file(s) to the analysis computer.

When to Make a New Instrument File?

When each ABI PRISM genetic analysis instrument is installed, an instrument file is created specifically for that instrument. You should create a new instrument file if any of the optics in the instrument change either due to service or age. Some specific situations that require a new instrument file are:

- The filter wheel is replaced on an ABI 373.
- The CCD camera is replaced on an ABI PRISM 377 or ABI PRISM 310.
- A run shows consistent and proportional pull-up peaks, indicating poor or incorrect spectral separation. (Pull-up peaks appear as smaller peaks of one color directly under larger peaks of another.)
- Fluorescence and spectral overlap are affected by the media/gel used for the run, you *may* need to make a new matrix/instrument file if you change the type of acrylamide, or other gel reagents.

Note If a valid instrument file exists in the ABI Folder (inside the System Folder) on your Macintosh® computer, you need not create one. If you need to replace a lost or damaged instrument file, and you do not have a backup copy, see “Viewing and Copying Matrices” on page E-31 before you re-create the entire instrument file.

Running Standards and Viewing Raw Sample Files

First Obtain Raw Data

An instrument file can contain a Dye Primer matrix, a Taq Terminator matrix, and/or a T7 Sequenase Terminator matrix.

For information about how to create raw matrix data for an instrument file, see the instrument *User's Manual.* (Remember to deselect auto-analyze in the Data Collection software.)

After you have created the raw matrix data, use the following procedure (Verifying the Raw Data) to confirm that the data is satisfactory.

IMPORTANT We do not recommend making a matrix using analyzed data. To confirm that the sample file is not analyzed, open the file in the sample window and check that there is no electropherogram view available (page 7-10).

Once you have satisfactory raw data, you can use the DataUtility program to either make a new instrument file or add the matrix data to an existing instrument file, as described later in this chapter:

- "Making a New Instrument File" on page E-12
- "Adding or Replacing a Matrix in an Existing Instrument File" on page E-26 (Requires four standards for each matrix.)

Sample File Data Only (No Data from BioLIMS)

The raw matrix data must be contained in sample files. If you extracted the matrix data in BioLIMS® mode, export the data to sample files using the Sample2DB program. (For more information on exporting BioLIMS database records, see the *ABI Prism BioLIMS Sample2DB Software User's Manual,* (P/N 4304072).)

Verifying the Raw Data

After you run the matrix standards, the next step is to verify that the run was successful and you have raw data for the matrix.

To verify lane tracking and peaks in the raw data:

~~310~~
373
377

Step	Action
1	Start the Sequencing Analysis program.
2	For 373 and 377 runs, open the gel file and track and extract the standard lanes into sample files. Before beginning extraction, check that auto-analysis is deselected. **IMPORTANT** Because the Tracker program only recognizes red data, you have to adjust the tracker lines by hand for the green, blue, and yellow standards. For how to view gel files, and track and extract sample information, see Chapter 4, “Working with the Gel File.”
3	Open the sample file for the standard in a Sample window. For how to view files in the Sample window, see Chapter 7, “Viewing and Editing Sample Files.”
4	Choose Quit from the File menu to quit the Sequencing Analysis program.
5	Make backup copies of the standard sample files before you make the instrument file.

Making a New Instrument File

Introduction

Follow these instructions to make a new instrument file. (For information on how to add a matrix to an existing instrument file, see “Adding or Replacing a Matrix in an Existing Instrument File” on page E-26.)

Note If you need to replace a lost or damaged instrument file, and you do not have a backup copy, see “Viewing and Copying Matrices” on page E-31 before you use these instructions to re-create the entire instrument file.

Outline of New Matrix Procedure

The steps that you need to perform are outlined briefly below and described in detail later in this section. These steps include:

- Run the appropriate matrix standards for your instrument, verify that lane tracking is correct (373 and 377 runs only), and verify that peaks exist in the raw data — as described earlier in this chapter. (“First Obtain Raw Data” on page E-10)
- Use the DataUtility program to make the instrument file. (“To make the instrument file:” on page E-13)
- Backup the raw sample files for the standards.
- After making the instrument file, analyze each matrix standard (raw data) file with the new instrument file to confirm the accuracy of the instrument file. (“Verifying the Instrument File” on page E-20)
- Properly store the new instrument file. (“Storing and Backing Up the Instrument File” on page E-25)

The Worksheet

If you are new to making an instrument file or if you are not using the default Start Point and Number of Points, you may want to use the Worksheet on page E-17 to help you keep track of the standards files.

Placement of Standards Files in DataUtility

The table below shows the placement of standards in the DataUtility program for dyes used for fluorescein/rhodamine dye primer and rhodamine dye terminator chemistries.

Matrix Standard Tube Labels Corresponding to DataUtility Boxes

Box	Dye Primer Matrix	Taq Terminator Matrix	T7 Terminator Matrix
C…	FAM	Taq C-term.	not used
A…	JOE	Taq A-term.	not used
G…	TAMRA	Taq G-term.	not used
T…	ROX	Taq T-term.	not used

The table below shows the placement of standards in the Data Utility program for dyes used for dRhodamine chemistries.

Matrix Standard Tube Labels Corresponding to DataUtility Boxes

Box	Dye Primer Matrix	Taq Terminator Matrix	T7 Terminator Matrix
C…	dR110	dROX	dR6G
A…	dR6G	dR6G	dTAMRA
G…	dTAMRA	dR110	dROX
T…	dROX	dTAMRA	dR110

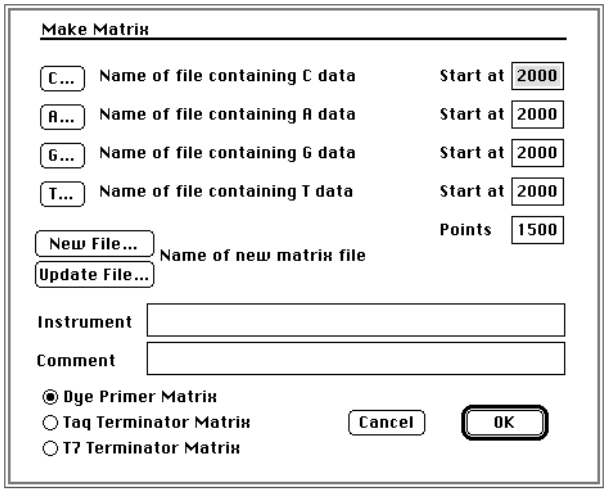
Making an Instrument File with the DataUtility Program

The DataUtility program has two main functions for users: to make instrument files and to copy matrices from one instrument file to another. The Measure Noise function of the program is used by Applied Biosystems Service personnel and is not discussed here.

To make the instrument file:

Step	Action
1	Open the DataUtility program. This program is located in the Utilities folder inside the Sequencing Analysis folder. The program icon looks like this: 

To make the instrument file: *(continued)*

Step	Action
2	Choose Make Matrix from the Utilities menu. The Make Matrix dialog box appears.  **Note** The files you select for the four nucleotides are the sample files you named on the Sample Sheet when you electrophoresed the matrix standards. **Note** You need to make backup copies of each set of four files: one copy for the Taq Primer, one for the Taq Terminator, and one for the T7 Terminator Matrix.
3	Specify the sample file to be used for each standard. (Refer to tables on page E-13.) a. Click the C… button. b. In the directory dialog box that appears, select the file that contains the data from the C standard, then choose Open. c. Repeat this selection process with for the A, G, and T standards.

To make the instrument file: *(continued)*

Step	Action
4	In each “Start at” text box, either accept the default value or type the correct start point value for that standard. **Note** For ABI 373 instrument data and ABI PRISM® 377 instrument data, the defaults of 2000 for start point and 1500 for data points is almost always appropriate. However, if your first attempt at matrix-making fails or if you want to reduce the chance of initial failure, you should follow the procedure on page E-17 that describes how to set Start and Stop points. For ABI PRISM 310 data, you should look at the raw data to determine the start point and number of data points to include; a start point of 2000 is usually satisfactory, but 1000 is sometimes better.
5	In the Points text box, either accept the default value or type the number of data points to be used for the matrix. (See Note above.)
6	Check that the correct radio button for the correct chemistry type is selected: Dye Primer, Taq Terminator, or T7 Sequenase Terminator.
7	Type (or edit) comment information in the Instrument and Comment text boxes.
8	Choose the New File… button.
9	Choose Save to close the dialog box and save the instrument file in the ABI Folder in the System Folder. Choose a descriptive name for the file. Since instrument files are specific to instruments and chemistries use these to name the file. E.g. “474-BigDye-InstFile”. **IMPORTANT** Only alpha-numeric characters, the period (.), the dash (–), and the comma (,) are permissible characters for instrument file names.
10	Choose OK to start the matrix calculation. The calculation takes about one minute. When the matrix is complete, the message “Make matrix successfully completed” appears. If an error message appears and the matrix is not made, see “Correcting Errors in Matrix Creation” on page E-29.
11	Choose OK to close the Make Matrix dialog box, or wait about 20 seconds for the dialog box to disappear.

To make the instrument file: *(continued)*

Step	Action
12	Repeat steps 2–11 for the Taq Terminator Matrix, except that at steps 8 and 9, choose Update File… instead of New File… to open the instrument file that you saved in step 9 on page E-15.
13	Repeat steps 2–11 for the T7 Terminator Matrix, except that at steps 8 and 9, choose Update File… instead of New File… to open the instrument file that you saved in step 9 on page E-15. You should now see matrices in all three boxes. **Note** The instrument file must contain a primer matrix, whether or not this chemistry will be used, in order to operate properly.
14	Quit the DataUtility program.

A Worksheet for Instrument File Matrices

About The Worksheet

You may want to photocopy (and enlarge) the worksheet on page E-18 and use it to help you keep track of which standards/sample file should be assigned to which base letter in which matrix in the DataUtility program.

Make All These Matrices

If you are making an instrument file for the dRhodamine or BigDye sequencing, you need only run four matrix standards. But you should make all three matrices. Leaving one blank, even if you think that you do not need that chemistry, can cause software and analysis problems.

If you are making an instrument file for fluorescein/rhodamine dye primer and rhodamine dye terminator sequencing (i.e. pre-dRhodamine and -BigDye chemistries), you may need to run eight standards, four to make the primer matrix and another four for the terminator matrices. You may leave the T7 Terminator matrix blank.

Setting Start Points and Number of Points

For each matrix do the following.

Choose start and end points for the data:

Step	Action
1	Open all the sample files for the standards in a Sample window in the Sequencing Analysis program.
2	For each sample file, identify a start point where there are no peaks and where the baseline is flat — beyond the primer peak or first large peaks that appear. You may want to note down this start point in the worksheet below. Each of the four standards can have different start and stop points but the number of points used for each sample must be the same.
3	Select a number of data points to analyze such that no peaks in the range are off-scale, *i.e.*, above 4000 relative fluorescence units (rfu), and that the baseline at the end of the range is flat. A typical number of data points is 1500. You may want to note down the number of data points in the worksheet below.
4	Calculate the stop point for each standard and verify that the baseline at the stop point is flat.

The Worksheet

You may wish to photocopy and fill in this worksheet before starting the procedure “Making an Instrument File with the DataUtility Program” on page E-13.

Data for Dye Primer Matrix

Box	Sample File Name	Start Point	Number of Points[a]	Stop Point
C…				
A…				
G…				
T…				

a. Number Points = (Stop Point) – (Start Point). Every number in this column must be the same.

Data for Taq Terminator Matrix

Box	Sample File Name	Start Point	Number of Points[a]	Stop Point
C…				
A…				
G…				
T…				

a. Number Points = (Stop Point) – (Start Point). Every number in this column must be the same.

Data for T7 Terminator Matrix

Box	Sample File Name	Start Point	Number of Points[a]	Stop Point
C…				
A…				
G…				
T…				

a. Number Points = (Stop Point) – (Start Point). Every number in this column must be the same.

Example Below is an example of how a worksheet might be filled in if you were making an instrument file using dRhodamine standards (for dRhodamine and BigDye chemistries). (The text that you would fill in is shown in italic.) If you are using other dye matrix standards, refer to your instrument *User's Manual*.

For the dRhodamine standards instrument file, the same four standard sample files are used for each of the three matrices, but the order that they are used is different for each matrix.

Data for Dye Primer Matrix

Box	Sample File Name	Start Point	Number of Points[a]	Stop Point
C…	*dR110*	*1500*	*1500*	*3000*
A…	*dR6G*	*2000*	*1500*	*3500*
G…	*dTAMRA*	*1450*	*1500*	*2950*
T…	*dROX*	*2000*	*1500*	*3500*

a. Number Points = (Stop Point) – (Start Point). Every number in this column must be the same.

Data for Taq Terminator Matrix

Box	Sample File Name	Start Point	Number of Points	Stop Point
C…	*dROX*	*2000*	*1500*	*3500*
A…	*dR6G*	*2000*	*1500*	*3500*
G…	*dR110*	*1500*	*1500*	*3000*
T…	*dTAMRA*	*1450*	*1500*	*2950*

Data for T7 Terminator Matrix

Box	Sample File Name	Start Point	Number of Points	Stop Point
C…	*dR6G*	*2000*	*1500*	*3500*
A…	*dTAMRA*	*1450*	*1500*	*2950*
G…	*dROX*	*2000*	*1500*	*3500*
T…	*dR110*	*1500*	*1500*	*3000*

Verifying the Instrument File

Introduction

There are two procedures to check the instrument file:

- Inspect the instrument file using the DataUtility program (page E-20)
- View and verify the matrix standard files in the Sequencing Analysis program (page E-22)

Inspect the Matrices Using the DataUtility Program

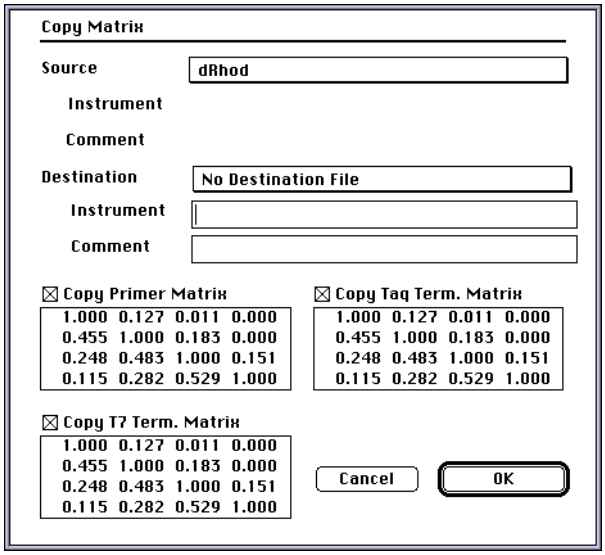
This operation allows you to:

- Check the quality of the matrices in the instrument file
- Verify that you have the matrix needed for the selected chemistry
- Determine if the matrix you used is responsible for poor data.

To view the instrument file:

Step	Action
1	Open the DataUtility program.
2	From the Utilities menu, choose Copy Matrix…
3	Under Source, select Instrument file and choose the new instrument file name. The three matrices within the instrument file appear as shown below. The numbers shown here are not representative values for all chemistries. **Note** For dRhodamine and BigDye chemistries, all three matrix boxes must be completed and the numbers for all three matrices are the same. **Note** For fluorescein/rhodamine dye primers and Rhodamine chemistries, only the Primer and Taq Terminator matrix boxes are completed and the numbers in the two matrices are different. **Note** If you find that the numbers in the matrix appear misaligned, change your System Font from Charcoal to Chicago. (From the Finder, choose Options in the Appearance control panel to do this.)

To view the instrument file: *(continued)*

Step	Action
	
4	Make sure that all three matrix boxes have numbers that range from 0–1. The numbers on the diagonals from top left to bottom right should be 1. If not, then repeat the matrix-making procedure starting with “Making a New Instrument File” on page E-12. **Note** If any numbers outside the diagonal are greater than 1, then the matrix may not work correctly.
5	Click Cancel.
6	Quit from the DataUtility program.

Inspect the Matrix Standard Files with Sequencing Analysis

Verify the matrix standard files in the Sequencing Analysis program:

To verify matrix standards:

Step	Action
1	Open the Sequencing Analysis program.
2	Open the standard sample files used to create the instrument file.
3	Use the electropherogram (analyzed) data view to confirm that the analyzed data looks OK: In each file, you should see one color trace with obvious peaks and all other color traces should be flat throughout the run. A pattern of pronounced peaks or dips in any of the other three colors indicate that something is wrong.
4	If all the data looks OK, go to “To properly store the instrument file:” on page E-25 below. If the data does *not* look OK, pick a different range of raw data points and remake the matrix. (Be sure to use the raw data files that you backed up in step 5 on page E-11. An analyzed file cannot be used to make an instrument file (See Important note on page E-10.)

Making an Instrument File from a Sample File

Introduction

An instrument file can be made from matrix standards as explained above, or it can be made from a sample file. This procedure requires fewer steps than running matrix standards; however, the matrix made from a sample file may not be as good as one made from matrix standards. The quality of an instrument file made from a sample file depends on the quality of the sample file used.

The best samples to choose for making a matrix have approximately 25% each of A, C, G, and T. A good example of this is the pGEM DNA with the –21M13 primer that is included as a control in every Ready Reaction Sequencing Kit.

Making the Instrument File from a Sample File

To create an instrument file from a sample file:

Step	Action
1	Before making the matrix, verify that lane tracking is accurate. Adjust if necessary.
2	Duplicate the unanalyzed sample file four times. Use the Duplicate command from the File menu in the Finder. You will have a total of four copies of the same sample file with the following names: ♦ Sample name ♦ Sample name Copy 1 ♦ Sample name Copy 2 ♦ Sample name Copy 3 ♦ Sample name Copy 4
3	These four sample file copies can now be used in the same way as the four matrix standard samples. The same instructions can be used with these four samples as with the four matrix standard samples.

To create an instrument file from a sample file: *(continued)*

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>4</td><td>For Filter Set A instrument files:</td></tr>
<tr><td></td><td>Follow the directions in your instrument user's manual.

Whenever the protocol indicates a specific matrix standard to be used, follow the table below:

<table>
<tr><th>Matrix Standard</th><th>Standard File</th></tr>
<tr><td>C…</td><td>Sample name Copy 1</td></tr>
<tr><td>A…</td><td>Sample name Copy 2</td></tr>
<tr><td>G…</td><td>Sample name Copy 3</td></tr>
<tr><td>T…</td><td>Sample name Copy 4</td></tr>
</table>
</td></tr>
<tr><td></td><td>For Filter Set E instrument files:</td></tr>
<tr><td></td><td>Follow the directions in Appendix A of the protocol for your sequencing chemistry.

Whenever the protocol indicates a specific matrix standard to be used, follow the table below:

<table>
<tr><th>Matrix Standard</th><th>Standard File</th></tr>
<tr><td>dR110</td><td>Sample name Copy 1</td></tr>
<tr><td>dR6G</td><td>Sample name Copy 2</td></tr>
<tr><td>dTAMRA</td><td>Sample name Copy 3</td></tr>
<tr><td>dROX</td><td>Sample name Copy 4</td></tr>
</table>
</td></tr>
<tr><td>5</td><td>Apply the newly made matrix back to the original sample file. There should be defined peaks and a flat baseline. If the baseline is not flat, if there are dips or pull-up peaks, then the instrument file is wrong and should not be used.</td></tr>
</table>

Storing and Backing Up the Instrument File

Introduction

The instrument file must be placed in the ABI Folder in the System Folder and obsolete instrument files should be deleted or archived. The new instrument file should be backed up.

Follow the steps in the table below after creating and verifying a new instrument file.

Storing the Instrument File

To properly store the instrument file:

Step	Action
1	Use the Finder to make sure the new instrument file is stored in the ABI Folder inside the System Folder. If you saved the file to a different location, drag it to the ABI Folder now. To be used by the Sequencing Analysis program, the instrument file must be in the ABI Folder.
2	Clean up the ABI Folder by deleting any invalid instrument files.
3	Put a backup copy of the instrument file on a server or a disk (and put the disk in a safe location). It is a good idea to put the raw sample files for the standards in the same place.

Adding or Replacing a Matrix in an Existing Instrument File

Introduction

Use the procedure described below to:

- Add a matrix to an incomplete instrument file.
- Replace an existing matrix.
- Make an additional instrument file for testing purposes.

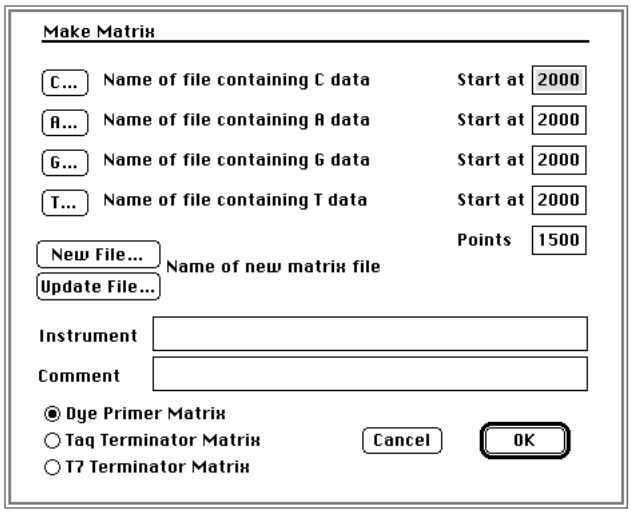
Note Be sure to make a backup copy of the original instrument file before you modify it.

Adding or Replacing a Matrix

To add or replace a matrix in an existing instrument file:

Step	Action
1	Open the DataUtility program.The program icon looks like this:  The program is located in the Utilities folder inside the Sequencing Analysis folder.

To add or replace a matrix in an existing instrument file: *(continued)*

Step	Action
2	Choose Make Matrix from the Utilities menu. The Make Matrix dialog box appears.  **Note** The files you select for the four nucleotides are the sample files you named on the Sample Sheet when you electrophoresed the matrix standards.
3	Specify the sample file to be used for each standard. a. Click the C… button. In the directory dialog box that appears, select the file that contains the data from the C standard, then choose Open to close the dialog box. b. Repeat this selection process with for the A, G, and T standards.

To add or replace a matrix in an existing instrument file: *(continued)*

Step	Action
4	In each "Start at" text box, either accept the default value or type the correct start point for analyzing that standard. If you do not use the defaults, use the numbers you wrote down during step 4-b on page D-19. **Note** For ABI 373 instrument data and ABI PRISM® 377 instrument data, the defaults of 2000 for start point and 1500 for data points is almost always appropriate. However, if your first attempt at matrix-making fails, you should follow the procedure on page E-17 that describes how to set Start and Stop points. For ABI PRISM 310 data, you should look at the raw data to determine the start point and number of data points to include; a start point of 2000 is usually satisfactory, but 1000 is sometimes better.
5	In the Points text box, either accept the default value or type in the number of data points to be used for the matrix. If you do not use the defaults, use the range of values you identified in the Raw Data view in step 4-b on page D-19.
6	Choose the Update File button.
7	In the directory dialog box that appears, select the name of the instrument file where you want to add the new matrix. Then choose Open.
8	Click the button for the appropriate matrix chemistry (Dye Primer, Taq Terminator, or T7 Sequenase Terminator) at the bottom of the Make Matrix dialog box. **IMPORTANT** When you add a new matrix to an instrument file, it overwrites any existing matrix of the same type. It has no effect on the other matrices in the file.
9	Choose OK to start the matrix calculation. The calculation takes about one minute.
10	When the message "Make matrix successfully completed" appears, choose OK or wait about 20 seconds for the dialog box to disappear. If an error message appears and the matrix is not made, see "Correcting Errors in Matrix Creation" on page E-29.
11	Analyze the new matrix standard, verify the accuracy of the instrument file, and properly store the new file, as described in "Verifying the Instrument File" on page E-20 and in "Storing and Backing Up the Instrument File" on page E-25.

Correcting Errors in Matrix Creation

Introduction

This section describes the two most common problems that can occur during matrix creation and how to resolve each problem.

- If the signal is too weak, see below.
- If an error message reports that the matrix was not made successfully, see page E-30.

Signal Too Weak

If the signal size for any of the data is too small, an error message appears and the matrix is not made.

To correct for weak signal:

Step	Action
1	Open the Sequencing Analysis program.
2	Open sample file for the standard in the Sample window.
3	Choose Raw Data from the Window menu.
4	Find a data range with about 1500 points with reasonable signal strength. Write down the start point and end point for the range. If the file does not contain enough good data, run a new set of matrix standards.
5	Repeat the Make Matrix process (or the Add/Replace Matrix process), using the new start point and data range numbers.

Incorrect Files or Chemistry

If any of the files selected are obviously incorrect, or you selected the wrong chemistry button, an error message appears and the matrix is not made.

To correct file or chemistry selection:

Step	Action
1	Repeat the Make Matrix process, selecting the correct chemistry button for the correct set of matrix sample files.
2	Use the gel file to verify that the matrix sample files contain the dye that the file indicates. The section “Colors in Real-Time Data Display Windows” on page E-3 explains the correlation between the colors in the gel file and the base that each color represents. **IMPORTANT** The gel file in the data collection program shows unconverted raw data, so the colors displayed represent different bases, depending on the chemistry. See “Colors in Real-Time Data Display Windows” on page E-3.

~~310~~
373
377

Viewing and Copying Matrices

Introduction

In addition to making matrices, you can use the DataUtility program to:

- View existing matrices (“Inspect the Matrices Using the DataUtility Program” on page E-20).
- Copy a matrix from a source file (sample, instrument, or gel) into a destination file (sample, instrument, or gel) (page E-31).

These operations are not necessary for normal operation of an ABI PRISM genetic analysis instrument. They are useful if you want to verify an existing matrix or re-create a lost matrix or instrument file.

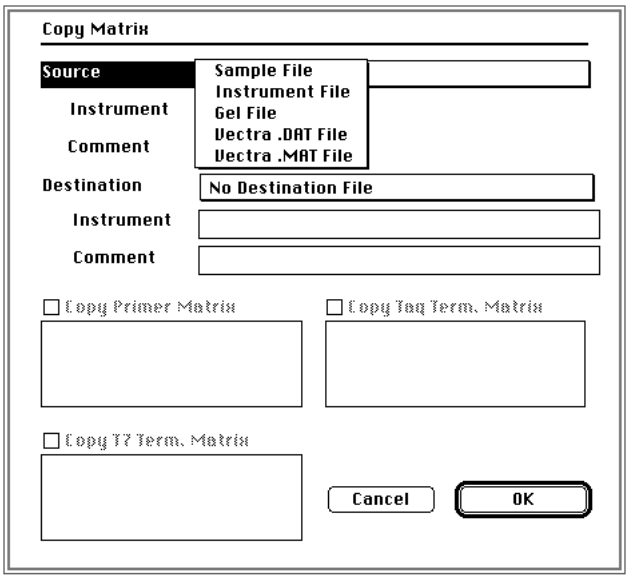
Why Copy Matrices from Source to Destination Files

The table below lists some common reasons you might copy files using the Copy Matrix dialog box.

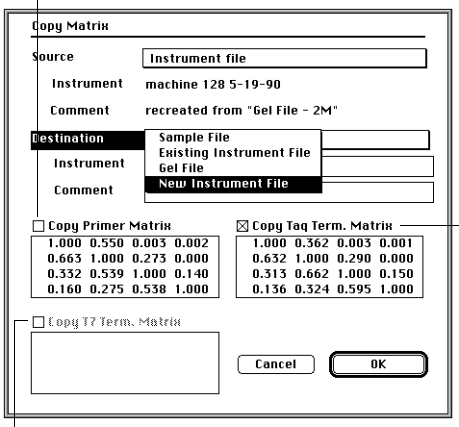
Reason for Use	Source	Destination
To recover the information for a lost instrument file from a sample file. The sample file need not contain good data for this to work.	Sample File	Instrument File (New)
To recover the information for a lost instrument file from a gel file. The gel file need not contain good data for this to work.	Gel File	Instrument File (New)
To copy the instrument file contents into a gel file. All sample files that are subsequently generated from the gel file will contain this instrument file information.	Instrument File (Existing)	Gel File
To copy the instrument file contents into a sample file for subsequent reanalysis.	Instrument File (Existing)	Sample File
To copy a matrix from one instrument file to another.	Instrument File (Existing)	Instrument File (Existing or New)

Copying Matrices Between Files

To copy a matrix from one file to another file:

Step	Action
1	Start the DataUtility program if it is not already running. The DataUtility program resides in the Utilities folder inside the Sequencing Analysis folder. The program icon looks like this: 
2	Choose Copy Matrix from the Utility menu.
3	From the Source pop-up menu, choose the type of file that contains the matrix or matrices you want to copy. 
4	In the directory dialog box that appears, locate and select the file that contains the matrix information you want to copy. Then choose Open. After the dialog box closes, the Copy…Matrix fields at the bottom of the Copy Matrix dialog box display the matrix information in the source file.

To copy a matrix from one file to another file: *(continued)*

Step	Action
5	Select the Copy…Matrix check box for each matrix you want to copy to the destination file. Be sure to de-select the check box for any matrix that you do *not* want to copy. **IMPORTANT** When you copy a new matrix to a file, the new matrix overwrites any existing matrix of the same type in that file.
6	From the Destination pop-up menu (see following figure), select the *type* of file into which you want to copy the selected matrices. De-select the check box for any matrix you do not want to copy  Select the check box to copy the matrix No T7 Sequenase matrix present in file

To copy a matrix from one file to another file: *(continued)*

Step	Action
7	In the directory dialog box that appears, do one of the following. For file types *Sample File, Existing Instrument File,* and *Gel File*, locate and select the name of the file into which you want to copy the selected matrices. Then choose Open. For file type *New Instrument File*, name the new instrument file with the instrument's serial number or another descriptive name, so this file is not confused with any other instrument file. Then, choose Save. **IMPORTANT** The only permissible characters for instrument file names are alphanumeric characters, the period (.), the dash (–), and the comma (,).
8	Choose OK to start the copy matrix procedure.
9	Follow the steps under "Inspect the Matrices Using the DataUtility Program" on page E-20 to view the destination file and verify that the matrices were successfully copied.

AppleScripting

F

Overview

Introduction

This appendix contains:

- A list of the AppleScript® commands that are supported by the Sequencing Analysis program
- Some sample scripts

About AppleScript

AppleScript is a simple programming language (a scripting language) that is part of the Macintosh® Operating System. Using AppleScript, you can automate many routine tasks. For example, you might write a five-to-ten line script that would take a folder full of sample files, add the files to the Sample Manager, turn on the check boxes for analysis and printing, and start processing.

For more information about how to create and use AppleScript scripts, see the *Applescript Language Guide* by Apple Computer Inc. or other similar books.

In This Appendix

This appendix includes the following topics:

AppleScript and Sequencing Analysis

Introduction

Sequencing Analysis v. 3.4 is Apple scriptable and recordable. It supports the Standard and Required Suite. In addition, an ABD defined ABI PRISM® Suite is supported.

It is through the ABD defined ABI PRISM Suite that Sequencing Analysis performs program-specific functions.

Limited Recordability

Scriptability is complete, but recordability is limited to basic functions, such as opening a file and other menu commands.

BioLIMS Access Through AppleScript

One limitation of scripting BioLIMS® database manipulations of Sequencing Analysis is that all additions and modifications to dbSequences (the AEOM object for sequence data) must occur within an open and close connection call to the database. This can be seen in the following valid example AppleScript:

```
open connection with username "user" with database
"db" …
add dbSequence 12345 of Collection "Klingon Genome
Project"
close connection
```

Further use of this dbSequence after the close connection will result in an error condition and is not supported.

Commands, Objects, and Events

AppleScript Commands and Objects

The Sequencing Analysis software supports the following AppleScript commands and objects.

General Commands

open	add	print
start	get	quit
run	set	close
pause	remove	process
resume	track	—
cancel	extract	—

General Objects

application	*sample file*	*window*
processor	*gel*	—

BioLIMS Specific Commands (Session Manager Suite)

open connection	open default connection
make new connection	select connection
close connection	delete connection
delete all connections	—

BioLIMS Specific Objects

collection	*dbSequence*
session manager	—

Supported Apple Events

Using the commands and objects above, you can create the following apple events:

- Open a sample file
- Add a sample file to the Sample Manager
- Open a gel file

- Remove a file from the Sample Manager
- Track gels files
- Start analysis
- Extract a gel into sample files
- Process (track and extract) a gel
- Cancel analysis
- Pause analysis
- Close a window
- Resume analysis
- Print
- Quit the application
- Set processing parameters of sample files

Examples of BioLIMS-specific apple events that you can create:

- Open a database connection
- Change a database connection
- Close a database connection
- Open a database sequence
- Add a database sequence to the Sequence Manager
- Remove a database sequence from the Sequence Manager
- Extract a gel into the database

Sample Script for Tracking Gel Files

About the Script

A script called “Track Folder full of Gels” is provided with Sequencing Analysis software. This script comes in two forms:

- The “Track Folder full of Gels” script that can be run using the Script Editor
- The “Track Folder full of Gels App” that can be run as a stand alone application without the Script Editor

If you want to…	Then…
run the default script without modification	use the “Track Folder full of Gels App”.
see how the script works and/or modify it	open the “Track Folder full of Gels” script within the Script Editor application.

Track Gel Files Overnight

Gel tracking is slow on some older Macintosh computers. This script allows you to batch sequence gel files for automatic tracking. For example, you might use this script for tracking a number of gel files overnight; in the morning, you could quickly extract and analyze from the tracked gels.

Before Running the Script

Before you run the script, be sure that the “Stop extraction when below confidence threshold” box in the Gel Preferences Dialog box is not checked.

If this box is checked and the Tracker encounters a lane assignment confidence value less than the threshold, an alert box appears and the script will halt and wait until you acknowledge the alert box.

Track Folder Full of Gels Script

The script:

- Prompts you to select a folder that contains sequencing gel files.
- For each gel file in the folder the script:
  - Opens the gel file in the Sequencing Analysis program
  - Tracks the gel file
  - Saves the gel file with the new tracking information

Note Place the gel files immediately inside the folder. Gel files nested within folders in the selected folder may cause the script to hang.

After the Script Runs

After the script is complete and before you extract data from the gel files, you should open the Sequencing Analysis Error Log and check the Lane Assignment Confidence Value for each gel file. If the value is less than 100, you should view the gel file and the Gel Sample Sheet to check the Tracker's lane assignment before extracting data.

Sample Scripts for Sequencing Analysis in Sample File Mode

The Scripts

Two sample scripts are provided for tracking/extracting and analysis with Sequencing Analysis software in Sample File mode:

- Process Gel Script
- Sample Manager Script

Process Gel Script

At installation, the Process Gel Script is placed in the Sample Scripts folder of the Sequencing Analysis folder.

This script:

- Prompts you to select a gel file for processing
- Opens the gel file
- Tracks the gel
- Extracts the sequence information into sample files
- Closes the gel file

During Sequencing Analysis software installation, this script is installed on your hard disk in the Sample Scripts folder inside the Sequencing Analysis folder.

Note To extract the gel data into the BioLIMS database instead of into sample files, use the script, “DB Process Gel Script” on page F-8.

Sample Manager Script

During Sequencing Analysis software installation, the Sample Manager Script is installed on your hard disk in the Sample Scripts folder inside the Sequencing Analysis folder.

This script:

- Prompts the user for a folder of sample files to be analyzed
- First analyzes all the files with a certain Basecaller (here, the ABI100 Basecaller)
- Then checks to see if the spacing value for each file is less than a specific value
- Uses a different Basecaller (SemiAdaptive) and reanalyzes the files, if the spacing value is too low

Note To analyze sequence data from the BioLIMS database instead of from sample files, use the script, “DB Sequence Manager Script” on page F-8.

Sample Scripts for Sequencing Analysis with BioLIMS

The Scripts

Two sample scripts are provided for using Sequencing Analysis software with BioLIMS:

- DB Process Gel Script
- DB Sequence Manager Script

DB Process Gel Script

During Sequencing Analysis software installation, the DB Process Gel Script is installed on your hard disk in the Sample Scripts folder inside the Sequencing Analysis folder.

This script:

- Prompts you to select a gel file for processing
- Sets the Sequencing Analysis program to BioLIMS mode
- Opens a database connection
- Opens a gel file
- Tracks the gel
- Extracts the sequence information into the BioLIMS database
- Closes the gel file

The tracking and extraction process is done separately in this script but it can be combined with the “process” verb if needed.

Note To extract the gel data into individual sample files instead of into the database, use the script, “Process Gel Script” on page F-7.

DB Sequence Manager Script

During Sequencing Analysis software installation, the DB Sequence Manager Script is installed on your hard disk in the Sample Scripts folder inside the Sequencing Analysis folder.

This script:

- Prompts the user for a collection of database sequence records to be analyzed
- Sets the Sequencing Analysis program to BioLIMS mode
- Opens a database connection

- First analyzes all the sequence records in the specified collection with a certain Basecaller (in this example script, the ABI100 algorithm is used)
- Saves the results to the database
- Then checks to see if the spacing value for each is less than a specific value
- Uses a different Basecaller (SemiAdaptive) and reanalyzes the sequence record, if the spacing value is too low
- Saves the results back to the database

Note To analyze sample files rather than data from the BioLIMS database, use the script, “Sample Manager Script” on page F-7.

Software License and Warranty

G

Applied Biosystems Software License and Limited Product Warranty

Warranty Statement

PURCHASER, CAREFULLY READ THE FOLLOWING TERMS AND CONDITIONS (THE "AGREEMENT"), WHICH APPLY TO THE SOFTWARE ENCLOSED (THE "SOFTWARE"). YOUR OPENING OF THIS PACKAGE INDICATES YOUR ACCEPTANCE OF THESE TERMS AND CONDITIONS. IF YOU DO NOT ACCEPT THEM, PROMPTLY RETURN THE COMPLETE PACKAGE AND YOUR MONEY WILL BE RETURNED. THE LAW PROVIDES FOR CIVIL AND CRIMINAL PENALTIES FOR ANYONE WHO VIOLATES THE LAWS OF COPYRIGHT.

Copyright

The SOFTWARE, including its structure, organization, code, user interface, and associated documentation, is a proprietary product of Applied Biosystems and is protected by international laws of copyright. Title to the SOFTWARE, and to any and all portion(s) of the SOFTWARE shall at all times remain with Applied Biosystems, which grants to you the limited rights contained in this Agreement.

License

1. You may install and use each original copy of the SOFTWARE on a separate single computer (or on a single network, if your software is designated as a network version). You may transfer the SOFTWARE to another single computer (or network, if a network version), so long as you first delete the SOFTWARE from the previous computer or network. You may never have operational SOFTWARE on more than one computer (or more than one network, if a network version) per original copy of the SOFTWARE at any time.

2. You may make one copy of the SOFTWARE for backup purposes. You may replace the SOFTWARE on any single computer (or on any single network, if applicable) with the backup copy if the need arises,

but you may not have operational SOFTWARE on more than one computer or network per original copy of the SOFTWARE at any time.

3. You may transfer the SOFTWARE to another party, but only if the other party agrees in writing with Applied Biosystems to accept the terms and conditions of this Agreement. If you transfer the SOFTWARE to another party, you must immediately transfer all copies to that party, or destroy those not transferred. Any such transfer terminates your license.

Restrictions

1. You may not copy, transfer, rent, modify, use, or merge the SOFTWARE, or the associated documentation, in whole or in part, except as expressly permitted in this Agreement.

2. You may not reverse assemble, decompile, or otherwise reverse engineer the SOFTWARE.

Limited Warranty

For a period of 90 days after purchase of the SOFTWARE, Applied Biosystems warrants that the SOFTWARE will function substantially as described in the documentation supplied by Applied Biosystems with the SOFTWARE. If you discover an error which causes substantial deviation from that documentation, send a written notification to Applied Biosystems. Upon receiving such notification, if Applied Biosystems is able to reliably reproduce that error at its facility, then Applied Biosystems will do one of the following at its sole option: (i) correct the error in a subsequent release of the SOFTWARE, which shall be supplied to you free of charge, or (ii) accept a return of the SOFTWARE from you, and refund the purchase price received for the SOFTWARE. Applied Biosystems does not warrant that the SOFTWARE will meet your requirements, will be error-free, or will conform exactly to the documentation. Any sample or model used in connection with this Agreement is for illustrative purposes only, is not part of the basis of the bargain, and is not to be construed as a warranty that the SOFTWARE will conform to the sample or model.

Limitation Of Liability

EXCEPT AS SPECIFICALLY STATED IN THIS AGREEMENT, THE SOFTWARE IS PROVIDED AND LICENSED "AS IS". THE ABOVE WARRANTY IS GIVEN IN LIEU OF ALL OTHER WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING THOSE OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. NOTWITHSTANDING ANY FAILURE OF THE CENTRAL PURPOSE OF ANY LIMITED REMEDY, APPLIED BIOSYSTEMS LIABILITY FOR

BREACH OF WARRANTY SHALL BE LIMITED TO A REFUND OF THE PURCHASE PRICE FOR SUCH PRODUCT. IN NO EVENT WILL APPLIED BIOSYSTEMS BE LIABLE FOR ANY OTHER DAMAGES, INCLUDING INCIDENTAL OR CONSEQUENTIAL DAMAGES, EVEN IF APPLIED BIOSYSTEMS HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Term

You may terminate this Agreement by destroying all copies of the SOFTWARE and documentation. Applied Biosystems may terminate this Agreement if you fail to comply with any or all of its terms, in which case you agree to return to Applied Biosystems all copies of the SOFTWARE and associated documentation.

Miscellaneous

1. Failure to enforce any of the terms and conditions of this Agreement by either party shall not be deemed a waiver of any rights and privileges under this Agreement.

2. In case any one or more of the provisions of this Agreement for any reason shall be held to be invalid, illegal, or unenforceable in any respect, such invalidity, illegality, or unenforceability shall not affect any other provisions of this Agreement, and this Agreement shall be construed as if such invalid, illegal, or unenforceable provisions had never been contained herein.

3. This Agreement shall be construed and governed by the laws of the State of California.

4. This Agreement and the Applied Biosystems Sales Quotation constitute the entire agreement between Applied Biosystems and you concerning the SOFTWARE.

Glossary

This glossary includes some of the special terms used in the Sequencing Analysis Software User's Manual. If a special term is not defined here, check the index to see if it is explained elsewhere in the manual.

Base Spacing Base Spacing is the number of data points from one peak to the next. The value is used to improve the accuracy of the base-calling algorithm used in analysis. Spacing values between 8.5 and 14 are generally acceptable. Spacing values outside this range may indicate errors in base calling. Spacing of a negative number indicates a problem with your samples, the gel, and/or the analysis parameters.

Basecaller The Basecaller is a program that determines the bases of a sequence during analysis. Seven types of Basecaller are included with this version of the Sequencing Analysis software: ABI-CE1, ABI-CE2, ABI50, ABI100, ABI200, SemiAdaptive, and Adaptive. They are described under separate headings in "About Basecallers and Base Calling" on page 6-45.

BioLIMS The BioLIMS® system is a genetic information management system that provides a relational database for storage and retrieval of DNA sequence and fragment data. In addition to the database itself, BioLIMS contains a set of software applications and tools for maintaining and interacting with the database. The BioLIMS database resides on a UNIX workstation and uses a Sybase® or an Oracle® database server. The client applications run on Power Macintosh computers and/or on UNIX workstations.

Channels Channels are the theoretical divisions across the read region of a gel where the 373 or 377 data collection software samples the data. The number of wells in the loading comb determines the approximate number of channels assigned per lane of the gel (for instance, with a 36-well comb, one lane includes approximately five channels). When the Sequencing Analysis Tracker software tracks a gel, it places the tracker line for each lane down the middle of the lane data. During extraction, it averages the data from that channel with the data from adjoining channels (as specified by the user) to determine the raw data for the sample file.

Chromatogram A chromatogram is a multi-color picture of a sequence, in which the bases are represented by peaks. The term is used interchangeably with electropherogram in this manual.

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.

cycle See *module*.

data point A sampling of fluorescence. Each data point has scan number and a channel number associated with it.

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. The Sequencing Analysis installer installs DyeSet/Primer files in the ABI Folder located in your System Folder. These files are sometimes referred to as mobility files.

editable data Sequencing Analysis base-called 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*, *original data*.

electropherogram A multi-color picture of a sequence, showing peaks that represent the bases. The term is used interchangeably with chromatogram in this manual.

feature A defined region in a sequence. Features are created and used by software programs that perform further analysis of sample files (for instance, the Factura™ and AutoAssembler™ programs). The Sample window includes a Feature view that displays feature information if any is present in the file.

genetic analysis instrument Used to refer generally to the instruments that provide data for analysis by the Sequencing Analysis software: the ABI PRISM® 310 Genetic Analyzer, the ABI 373 DNA Sequencer, and the ABI PRISM® 377 DNA Sequencer.

heterozygote "a diploid or polyploid individual that has inherited different alleles at one or more loci and therefore does not breed true" (King, R.C.; Stansfield, W.D. *A Dictionary of Genetics*; Oxford University: New York, 1990; p 147). In this manual, the term describes a position at which the electropherogram displays more than one nucleotide peak, indicating the possible presence of different alleles.

instrument file A file stored in the ABI Folder inside the System Folder. The file contains one to three matrices, a Comment field (where you can enter comments about the file), and an Instrument Name field (where you can enter the name of the instrument the file is to be used with). This file defines the matrices used to correct for the spectral overlap between the fluorescent dyes used on the ABI PRISM genetic analysis instruments. A mathematical matrix of the spectral overlaps is created and the inverse matrix is used to correct the data during analysis. This file is also sometimes called the matrix file because it contains one or more matrices.

IUB code An alphabetic character representing the occurrence of mixed bases at a given position in a sequence. These codes were 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.

IUB Codes:

Base	IUB Code	Complement
Adenosine	**A**	**T**
Cytidine	**C**	**G**
Guanosine	**G**	**C**
Thymidine	**T**	**A**

IUB Codes: *(continued)*

Base	IUB Code	Complement
Adenosine or Guanosine (pu**R**ine)	**R**	**Y**
Cytidine or Thymidine (p**Y**rimidine)	**Y**	**R**
Guanosine or Thymidine (**K**eto)	**K**	**M**
Adenosine or Cytidine (a**M**ino)	**M**	**K**
Guanosine or Cytidine (**S**trong — 3 H bonds)	**S**	**W**
Adenosine or Thymidine (**W**eak — 2 H bonds)	**W**	**S**
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 (a**N**y base)	**N**	**N**

IUPAC International Union of Pure and Applied Chemistry. This acronym is also used to refer to IUB codes, 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.

matrix file See *instrument file*.

mobility file See *DyeSet/Primer file*.

module A file that provides instructions about conditions of operation to the ABI PRISM instrument. You might use three different modules during a typical run to specify conditions for plate check, pre-run, and the run itself. For more details, see the ABI PRISM instrument user's manual.

Neural Net Tracker A neural network is a computational structure inspired by the study of biological neural processing. Neural networks learn by example. The Tracker application uses a neural network to learn which features in a gel correspond to the center of a lane and which features to ignore. The Tracker application has been shown hundreds of hand-tracked gels as examples of how to accurately track. It has been shown very noisy gels and told to ignore various types of "noise" including red rain, primer peaks, blobs, etc.

original data The sequence data created the last time the Basecaller was run. This base-called 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 base-called 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*, *sample file*s.

preferences Values that are selected by the user, stored in the program's Preferences file (in the Preferences folder inside the System Folder), and used by the program during normal operation. For example, you can specify that the Sequencing Analysis software should always use the "SemiAdaptive" Basecaller and save .Seq files in the "Wisconsin" format. You can change

preference values whenever your needs change; you can also temporarily override some types of preferences without changing their values in the preferences file.

sample files A sample file contains raw DNA sequence data (as read by the electrophoresis instrument), and the base calls, peak locations, and electropherogram created by the Sequencing Analysis software. After processing by the Factura software, or other similar programs, the file also contains additional analysis information (for example, features). For the 373 or 377 instruments, sample files are created by the Applied Biosystems software in the gel extraction process. For the 310 instrument, raw sample files are created by the 310 Data Collection software. Raw sample files are analyzed by Sequencing Analysis and saved as analyzed sample files.

scan number During an ABI 373 or ABI PRISM 377 run, the data collection software typically scans the gel some 15,000 times. It samples the data 194 times (in full scan mode) or 388 times (in XL mode) or 480 times (in 96-Lane mode) during each scan. Each sampling is stored as a data point that is described in terms of its scan number. A data point is represented by one line number on the Gel display. The scan number describes the location of the data point. On an ABI PRISM 310 instrument, one sampling is taken during each scan and the information is stored as a data point.

selected sequence A sequence that you have specified by clicking its name in the Sample Manager.

separation distance The length from the wells of the gel to the read region of the gel. Also called the well-to-read or WTR distance.

.Seq files Text files created by the Sequencing Analysis program. The .Seq files contain only the characters of the sequence and can be created in several formats (ABI, Intelligenetics, Staden, and Wisconsin) for use with other programs.

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. The chemistries and kits supplied by Applied Biosystems for performing such reactions are described in the *ABI PRISM DNA Sequencing Chemistry Guide*.

settings Values that you can select and which are then used by the program during program operations. Settings can be relatively permanent (see *preferences* earlier in this glossary) or temporary (as when you decide to print six copies of a sample file for a meeting).

shark-tooth comb A piece of flexible plastic material inserted into a gel that is used for a sequencing run. During gel polymerization, the comb is inserted with the flat edge down to form a single well (or a separate casting comb can be used). Later, the toothed edge is inserted to form wells into which samples are loaded. It is called a “shark-tooth” comb because it has pointed teeth along one side.

signal strength 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 Annotation view of the Sample File widow and in the header of the printed electropherogram.

spacing See *base spacing*.

square-tooth comb A piece of flexible plastic material inserted into a gel that is used for a GeneScan or sequencing run. During gel polymerization, the comb is inserted into the gel. Later, it is removed and

sample is loaded into the square holes formed by the comb. It is called a “square-tooth” comb because it has square teeth along one side.

summary graphic A horizontal line displayed near the top of the Sample window. It is used to show the location of features and the currently selected section of the sequence.

tracker line A line drawn on a gel display to track the migration of the DNA sample through the gel matrix during electrophoresis. You can edit the tracker line to correct for migration problems. You can also specify the number of channels on either side of the tracker line to be used when creating an averaged data value.

views Various displays provided in the Sample window. For information about the Sample window views, see “The Six Sample Window Views” on page 7-6.

WTR (well-to-read) See *separation distance*.

Index

F

S

T

U

V

W

Headquarters
850 Lincoln Centre Drive
Foster City, CA 94404 USA
Phone: +1 650.638.5800
Toll Free: +1 800.345.5224
Fax: +1 650.638.5884

Worldwide Sales Offices
Applied Biosystems vast distribution and service network, composed of highly trained support and applications personnel, reaches into 150 countries on six continents. For international office locations, 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 the USA, 02/2001
Part Number 4306158C

an Applera business