SOFTGENETICS

Software *PowerTools* for Forensic Analysis

GeneMarker® HTS

Forensic NGS Analysis Software

Analyze mtDNA and STR/Y-STRs Simultaneously

Now Includes:

Mitochondrial DNA Analysis

- ✓ Whole Genome or HV1/HV2 and Control Region Analysis
- ✓ Unique Alignment Technology[1,2] • **Motif • Consensus**
- ✓ Forensic Nomenclature
- ✓ Easy uploads to EMPOP[3]

STR Analysis

- ✓ Autosomal & Y-STR
- ✓ Forensic Nomenclature
- ✓ Genotype & SNP Reporting

Validated

Easy-to-Use Windows® Interface

Compatible with major Chemistries & Platforms

Audit Trail & User Control

Comprehensive Reporting Options

SOFTGENETICS®

Software *PowerTools* for Forensic Analysis

Release Information

Document Version Number	*GMHTS-2.0-UG001*
Software Version	*2.0*
Document Status	*Final*

Copyright

Limit of Liability

SoftGenetics, LLC has used their best effort in preparing this guide. SoftGenetics makes no representations or warranties with respect to the accuracy or completeness of the contents of this guide and specifically disclaims any implied warranties of merchantability or fitness for a particular purpose. Information in this document is subject to change without notice and does not represent a commitment on the part of SoftGenetics or any of its affiliates.The accuracy and completeness of the information contained herein and the opinions stated herein are not guaranteed or warranted to produce any particular results, and the advice and strategies contained herein may not be suitable for every user.

The software described herein is furnished under a license agreement or a non-disclosure agreement. The software may be copied or used only in accordance with the terms of the agreement. It is against the law to copy the software on any medium except as specifically allowed in the license or the non-disclosure agreement.

Trademarks

The name "*SoftGenetics*," the SoftGenetics logo, and *GeneMarker HTS* are trademarks or registered trademarks of SoftGenetics, LLC. All other products and company names mentioned herein might be trademarks or registered trademarks of their respective owners.

Customer Support

Customer support is available to organizations that purchase *GeneMarker HTS* and that have an annual support agreement. Contact SoftGenetics at:

SoftGenetics, LLC
100 Oakwood Ave, Suite 350
State College, PA 16803
(814) 237-9340
(888) 791-1270 (US Only)
tech_support@softgenetics.com
www.softgenetics.com

Table of Contents

Preface **9**

Chapter 1: Getting Started with GeneMarker HTS **11**

GeneMarker HTS System Requirements 13

Installing GeneMarker HTS 14

To install GeneMarker HTS 14

Installing the SoftGenetics License Server 16

To install the SoftGenetics license server 16

To register GeneMarker HTS 18

To apply the license change 21

To configure the license server 22

Starting GeneMarker HTS 24

The GeneMarker HTS Main Window 25

Title bar 25

Application Settings icon 25

Project options 25

About button 26

Online Help button 26

Specifying GeneMarker HTS Application Settings 27

To specify GeneMarker HTS application settings 27

Chapter 2: GeneMarker HTS Project Setup and Management **35**

Overview of the New Project Dialog Box 37

Using a motif 41

Soft clipping based on quality scores 44

Specifying amplicon regions 45

Masking personal health information (PHI) for a subject 46

Working with Projects in GeneMarker HTS 49

To add a new project 49
To work with an existing project 53
GeneMarker HTS Project Output Files 57

Chapter 3: GeneMarker HTS Viewers 59

Mito Alignment Viewer Layout 61
Title bar 62
Main menu 62
Global Viewer pane 63
Pane displays 64
Variants table 67
Status Bar 68
Mito Alignment Viewer Navigation and Options 69
Mito Alignment Viewer navigation 69
Mito Alignment Viewer options 70
Mito Alignment Viewer Filter Settings 75
To specify the filter settings for an mtDNA project display 75
To specify the table settings for an mtDNA project display 79
To specify alternate alignment filters 81
Mito Alignment Viewer Reports 83
STR Analysis Viewer Layout 86
Sunburst Plot pane 87
Sample pane 88
STR Analysis Viewer Filter Settings 90
To specify the filter settings for an STR analysis project display 90
Comparison Viewer 93
Sample to Sample Comparison pane 93
Variant Comparison pane 95

Glossary **97**

Index **99**

Preface

Welcome to the *GeneMarker HTS User's Manual*. The purpose of the *GeneMarker HTS User's Manual* is to answer your questions and guide you through the procedures necessary to use the GeneMarker HTS application efficiently and effectively.

Using the manual

You will find the *GeneMarker HTS User's Manual* easy to use. You can simply look up the topic that you need in the table of contents or the index. Later, in this Preface, you will find a brief discussion of each chapter to further assist you in locating the information that you need.

Special information about the manual

The *GeneMarker HTS User's Manual* has a dual purpose design. It can be distributed electronically and then printed on an as-needed basis, or it can be viewed online in its fully interactive capacity. If you print the document, for best results, it is recommended that you print it on a duplex printer; however, single-sided printing will also work. If you view the document online, a standard set of bookmarks appears in a frame on the left side of the document window for navigation through the document. For better viewing, decrease the size of the bookmark frame and use the magnification box to increase the magnification of the document to your viewing preference.

Conventions used in the manual

The *GeneMarker HTS User's Manual* uses the following conventions:

- Information that can vary in a command—variable information—is indicated by alphanumeric characters enclosed in angle brackets; for example, <Project Name>. Do not type the angle brackets when you specify the variable information.
- A new term, or term that must be emphasized for clarity of procedures, is *italicized*.
- Page numbering is "online friendly." Pages are numbered from 1 to x, *starting with the cover* and ending on the last page of the index.

Although numbering begins on the cover page, this number is not visible on the cover page or front matter pages. Page numbers are visible beginning with the first page of the table of contents.

- This manual is intended for both print and online viewing.
    - If information appears in blue, it is a hyperlink. Table of Contents and Index entries are also hyperlinks. Click the hyperlink to advance to the referenced information.

Assumptions for the manual

The *GeneMarker HTS User's Manual* assumes that:

- You are familiar with Windows-based applications and basic Windows functions and navigational elements.
- References to any third-party standards or third-party software functions were current as of the release of this version of GeneMarker HTS, and might have already changed.

Organization of the manual

In addition to this Preface, the *GeneMarker HTS User's Manual* contains the following chapters:

- Chapter 1, "Getting Started with GeneMarker HTS," on page 11 details the installation requirements for GeneMarker HTS, installing GeneMarker HTS, and activating your account. It also explains how to launch GeneMarker HTS and provides an overview of the major navigational elements for GeneMarker HTS. It also details how to customize the layout and display of your GeneMarker HTS instance.
- Chapter 2, "GeneMarker HTS Project Setup and Management," on page 35 details the New Project dialog box, which provides all the options for adding a new project in GeneMarker HTS. It also explains how to open and work with existing projects in GeneMarker HTS.
- Chapter 3, "GeneMarker HTS Viewers," on page 59 details the three viewers that are available in GeneMarker HTS application for viewing and working with the results of an STR/mtDNA projects – the Mito Alignment Viewer, the STR Results Analysis Viewer, and the Comparison Viewer. It also details the different reports that are available for your mtDNA alignment results.
- The Glossary on page 97 provides detailed definitions of specialized terms for GeneMarker HTS.

Chapter 1
Getting Started with GeneMarker HTS

The GeneMarker HTS software application provides a streamlined workflow for mtDNA data analysis from massively parallel sequencing (MPS) methods. The software also includes analysis of targeted amplicons from STRs. This chapter details the installation requirements for GeneMarker HTS, installing GeneMarker HTS, and activating your account. It also explains how to launch GeneMarker HTS and provides an overview of the major navigational elements for GeneMarker HTS. It also details how to customize the layout and display of your GeneMarker HTS instance.

This chapter covers the following topics:

- "GeneMarker HTS System Requirements" on page 13.
- "Installing GeneMarker HTS" on page 14.
- "Installing the SoftGenetics License Server" on page 16.
- "Starting GeneMarker HTS" on page 24.
- "The GeneMarker HTS Main Window" on page 25.
- "Specifying GeneMarker HTS Application Settings" on page 27.

Chapter 1
Getting Started with GeneMarker HTS

GeneMarker HTS System Requirements

GeneMarker HTS is currently available only for the Windows operating system. You must have Administrator rights for the computer on which you are installing GeneMarker HTS.

Component	Description
OS	64-bit Windows 7 or 10
Processor	Intel Core i5 3rd generation (or equivalent) or better
Memory	16 GB RAM (Minimum)
Hard Drive	250 GB of free space (SSD recommended)
Monitor	1366 x 768 or higher resolution monitor

Installing GeneMarker HTS

You install GeneMarker HTS as a standalone application on a Windows client.

To install GeneMarker HTS

You can install GeneMarker HTS from a disc that SoftGenetics supplies, or you can download the executable (GeneMarkerHTS_<Version>_Installer.exe) from SoftGenetics's website (http://softgenetics.com/downloads.php). The GeneMarker HTS installation requires the Microsoft Visual C++ 2012 Redistributable. If this redistributable is already installed on the computer on which you are installing GeneMarker HTS, then no additional action is required during installation; otherwise, the GeneMarker HTS Installation wizard guides you through the steps that are necessary to install this redistributable. GeneMarker HTS is installed in the following directory, and you cannot select a different location:

C:\Program Files (x86)\SoftGenetics\GeneMarker HTS

After GeneMarker HTS has been installed on your computer, a shortcut icon for the application is placed on your desktop. An option for GeneMarker HTS is also available from your Start menu.

See Figure 1-1 below through Figure 1-4 on page 15.

Figure 1-1: GeneMarker HTS Installation Wizard, Page 1

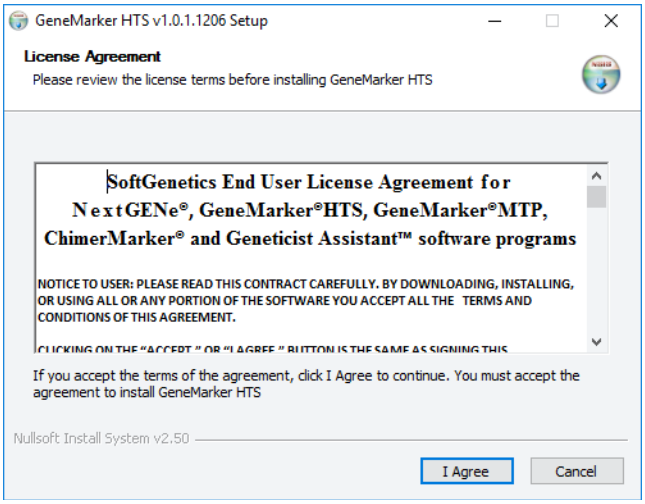


Figure 1-2: GeneMarker HTS Installation Wizard, Page 2 (Optional)

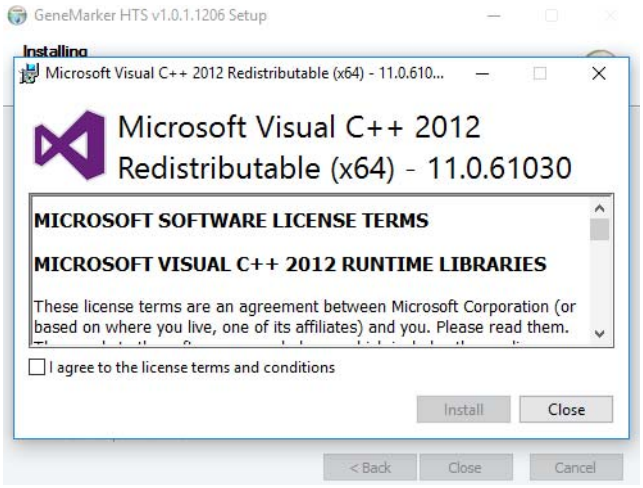


Figure 1-3: GeneMarker HTS Installation Wizard, Page 3 (Optional)

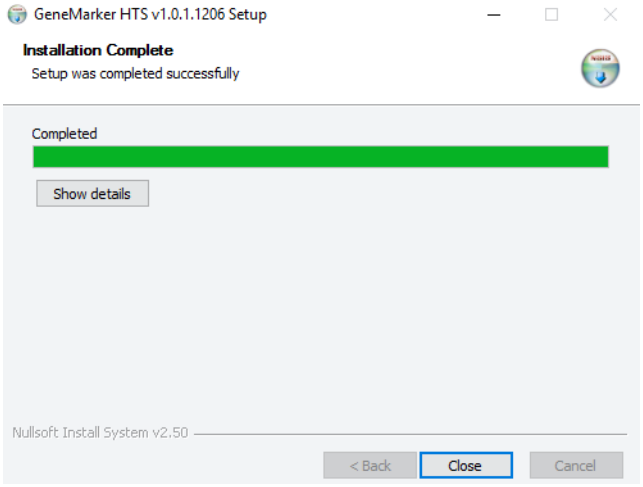
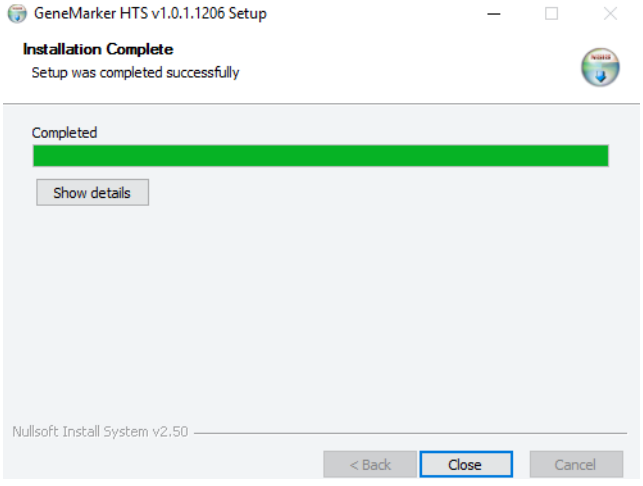


Figure 1-4: GeneMarker HTS Installation Wizard, Page 4



Installing the SoftGenetics License Server

If you have installed a *purchased* version of GeneMarker HTS, then you must also install the SoftGenetics License Server. The *license server* is a central location either locally or on a network where software application licenses are stored and can be accessed by GeneMarker HTS users. After you install the license server, you must register GeneMarker HTS for the server, and then you must configure GeneMarker HTS to connect to the SoftGenetics License Server so that you and your users can use GeneMarker HTS.

If you are running a trial version of GeneMarker HTS, then you do not need to install and configure the SoftGenetics License Server. By default, GeneMarker HTS is installed with a Trial license that is valid for 35 days. After 35 days, to continue using GeneMarker HTS, you must purchase a commercial license and install the license server.

Note the following about the SoftGenetics License Server:

- The executable to install the SoftGenetics License Server is named SrvSetup_<>_exe and it is available from SoftGenetics's ftp site. Contact SoftGenetics for assistance in obtaining this file.
- You must install the SoftGenetics License Server on a Windows client or server that is always on and running. You can install the SoftGenetics License Server on the same computer that is running GeneMarker HTS.
- You must be an administrator of the Windows client or server on which you are installing the license server. To satisfy the UAC settings on the client or server, you might also need to run the license server setup file as an administrator. (After you download the file, you can right-click it, and then on the context menu that opens, select Run as Administrator.)
- To allow for online registration of the SoftGenetics License Server, the server/client on which you are installing the server must be connected to the Internet; otherwise, offline registration is required.

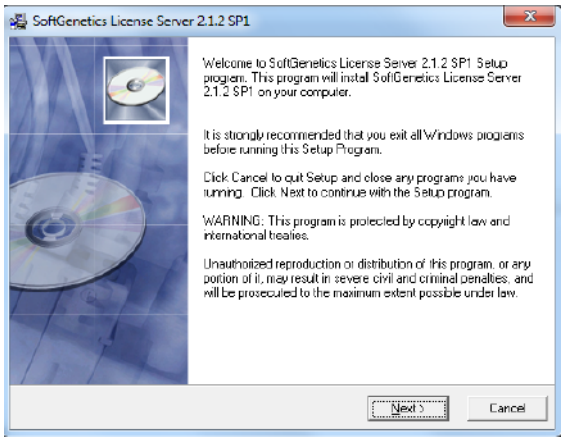
To install the SoftGenetics license server

1. Obtain the SoftGenetics License Server installation file from SoftGenetics.
2. Download the file to the Windows computer or server on which you are installing the server, and run the executable as an Administrator.

You can install the SoftGenetics License Server on the same computer that is running GeneMarker HTS.

The SoftGenetics License Server Installation wizard opens. See Figure 1-5 on page 17.

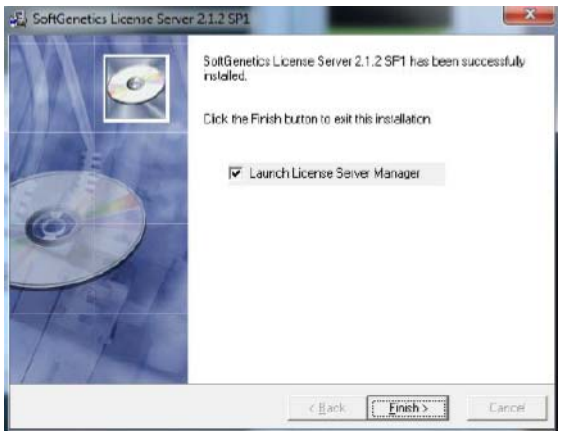
Figure 1-5: SoftGenetics License Server Installation wizard



3. Click Next.

   The SoftGenetics License Server is installed. After the license server is successfully installed, the installation wizard prompts you to launch the License Server Manager, which is the GUI for the SoftGenetics License Server.

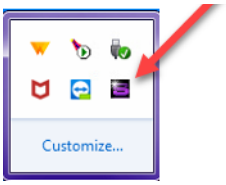
Figure 1-6: SoftGenetics License Server Installation wizard



4. Click Finish.

   The SoftGenetics License Server Installation wizard closes and an icon for the License Server Manager (LSM) is placed in the system tray.

Figure 1-7: LSM icon in system tray



5. Continue to "To register GeneMarker HTS" on page 18.

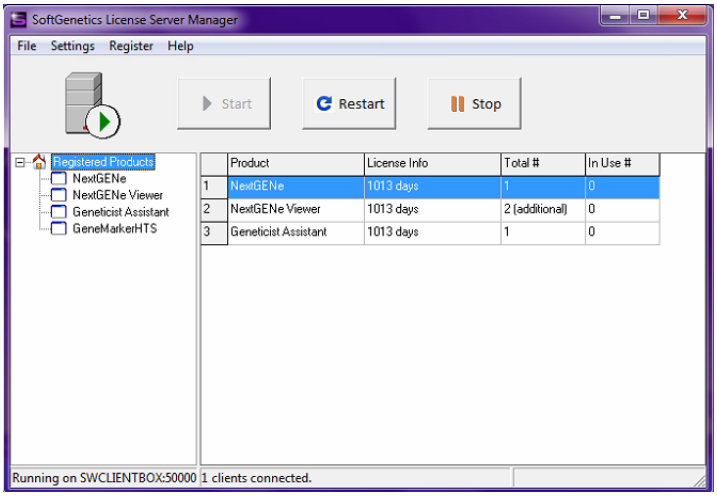
To register GeneMarker HTS

The type of license that you register determines whether your GeneMarker HTS instance can be used for adding new projects and opening and viewing projects, or simply opening and viewing projects.

1. In the system tray, click the LSM icon.

   The License Server Manager opens. If you have installed any other SoftGenetics products that require the SoftGenetics License Server, then these products and their licensing information are displayed in the LSM; otherwise, the display is blank.

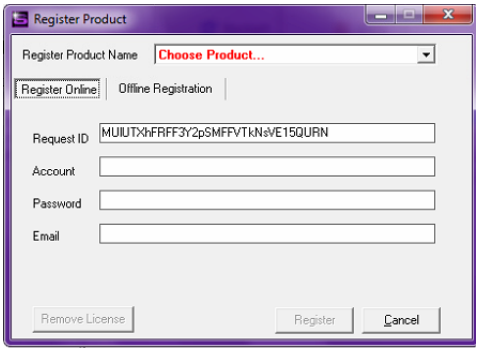
Figure 1-8: License Server Manager



2. On the LSM main menu, click Register.

   The Register Product dialog box opens. The Request ID field is pre-populated.

Figure 1-9: Register Product dialog box



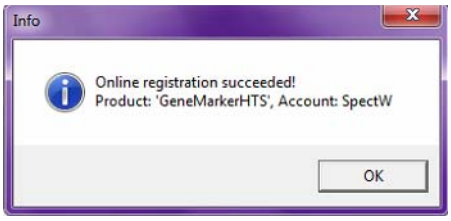
3. For the Register Product Name, select GeneMarker HTS.
4. Enter the account credentials (Account and Password) that SoftGenetics supplied to you.
5. Enter the appropriate email address for the account.

6. Click Register.

   After successful registration, an Info dialog box opens indicating this.

If the online registration is not successful, then see "To register GeneMarker HTS offline" on page 20.

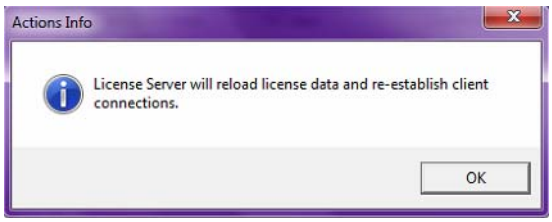
Figure 1-10: Info dialog box with successful registration message



7. Click OK.

   The Info dialog box closes, and an Actions Info dialog box opens, indicating that the license server will reload the license data and re-establish client connections.

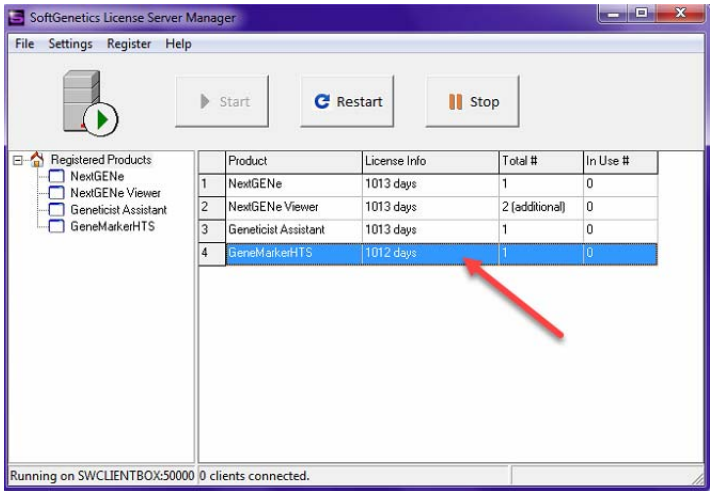
Figure 1-11: Actions Info dialog box



8. Click OK.

   The Actions Info dialog box closes and the LSM remains open. The newly registered GeneMarker HTS product is displayed in the LSM.

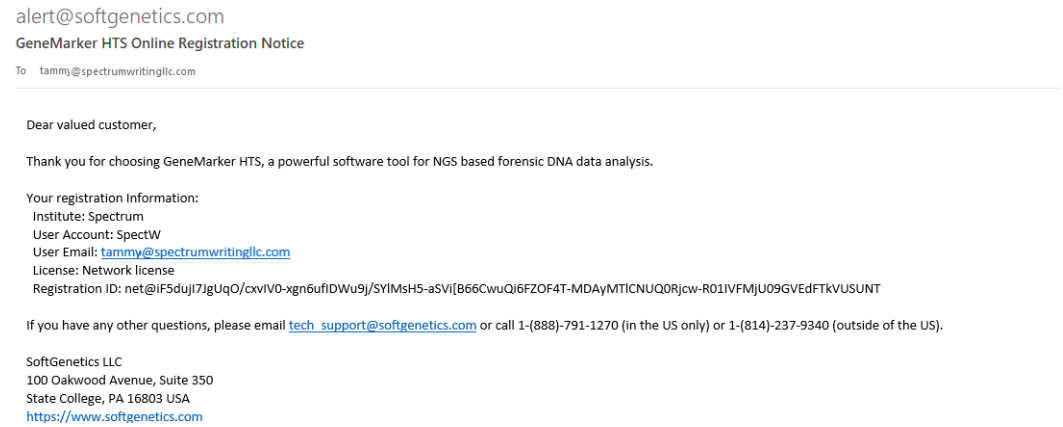
Figure 1-12: License Server Manager



After the Actions Info dialog box closes, a Message dialog box might open, prompting you to restart the LSM for the changes to take effect. Click Restart Now to complete the registration process.

A registration email is sent to the email address that you specified. The email acknowledges your successful registration and also details your account information.

Figure 1-13: Registration email



alert@softgenetics.com

GeneMarker HTS Online Registration Notice

To tammy@spectrumwritingllc.com

Dear valued customer,

Thank you for choosing GeneMarker HTS, a powerful software tool for NGS based forensic DNA data analysis.

Your registration Information:
Institute: Spectrum
User Account: SpectW
User Email: tammy@spectrumwritingllc.com
License: Network license
Registration ID: net@iF5dujI7JgUqO/cxvlV0-xgn6uflDWu9j/SYlMsH5-aSVi[B66CwuQi6FZOF4T-MDAyMTlCNUQ0Rjcw-R01IVFMjU09GVEdFTkVUSUNT

If you have any other questions, please email tech_support@softgenetics.com or call 1-(888)-791-1270 (in the US only) or 1-(814)-237-9340 (outside of the US).

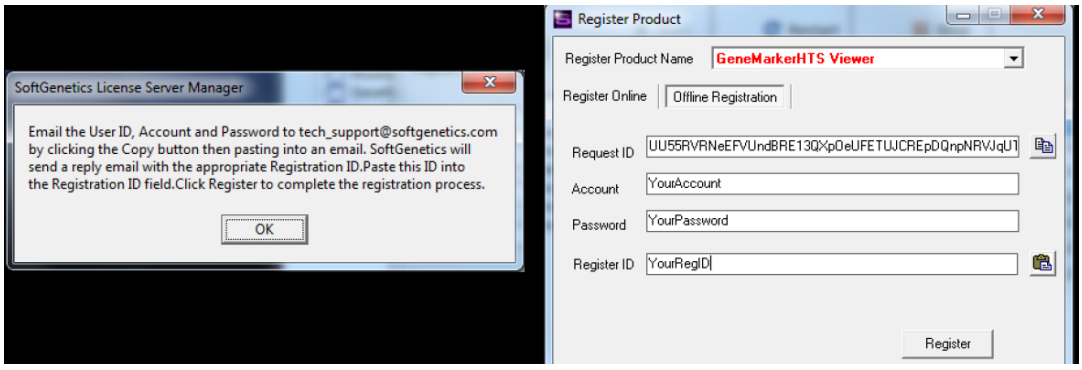
SoftGenetics LLC
100 Oakwood Avenue, Suite 350
State College, PA 16803 USA
https://www.softgenetics.com

9. Close the LSM and then continue to "To apply the license change" on page 21.

To register GeneMarker HTS offline

If online registration for GeneMarker HTS fails, then the Product Registration dialog box is toggled to show the Offline mode, and you must register the application offline.

Figure 1-14: Register Product dialog box, Offline mode



1. Copy the entire Request ID string that is displayed in the dialog box.
2. Paste the copied Request ID string into an email that includes your account credentials (Account, Password, and Email) and send the email to tech_support@softgenetics.com.

   SoftGenetics will email you a new Registration ID.
3. Copy the new Registration ID, and then paste it in the Request ID field.

4. Enter your account credentials, and then click Register.

   After GeneMarker HTS is registered successfully, you are prompted to restart the LSM for the license change to be effective. Continue to "To apply the license change" below.

To apply the license change

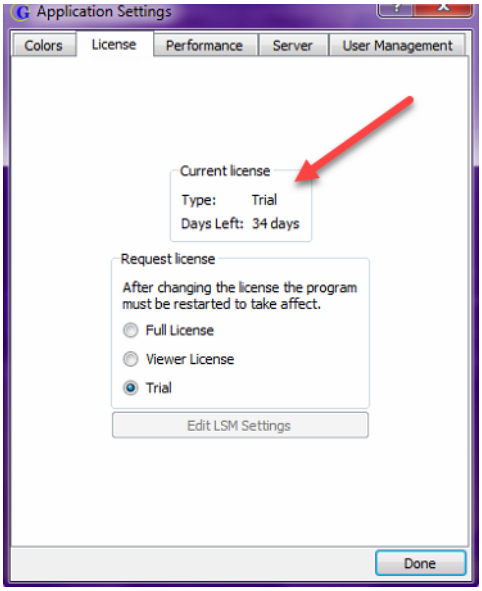
The following procedure assumes that GeneMarker HTS has already been installed. If you have not yet installed GeneMarker HTS, see "Installing GeneMarker HTS" on page 14.

1. To start GeneMarker HTS, double-click the desktop icon, or select the option from your Start menu (Start > GeneMarkerHTS).
2. On the GeneMarker HTS main window, click the Application Settings icon .

   The Application Settings dialog box opens. The Performance tab is the open tab.
3. Open the License tab.

   The tab indicates that a Trial license is in use and that 35 days remain on the trial.

Figure 1-15: Application Settings dialog box, License tab showing Trial license



4. Select Full License or Viewer License as appropriate.
5. Click Done.
6. Close the Application Settings dialog box.
7. To complete the licensing change, close, and then re-open GeneMarker HTS.
8. Continue to "To configure the license server" on page 22.

To configure the license server

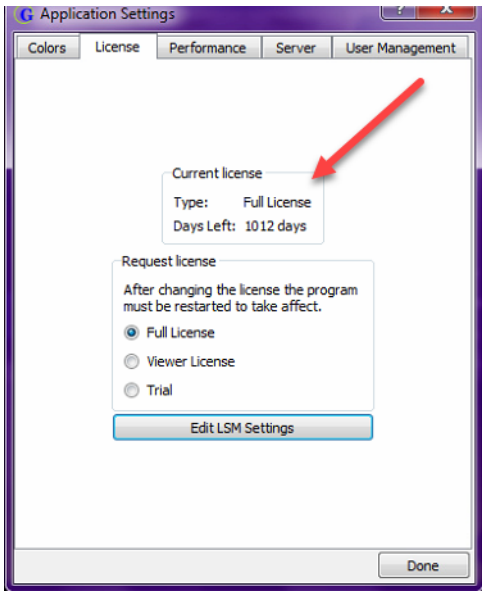
1. On the GeneMarker HTS main window, click the Application Settings icon .

   The Application Settings dialog box opens. The Performance tab is the open tab.

2. Open the License tab.

   The tab now indicates the type of license that was configured for GeneMarker HTS (Full or Viewer) and the number of days for which the license is active.

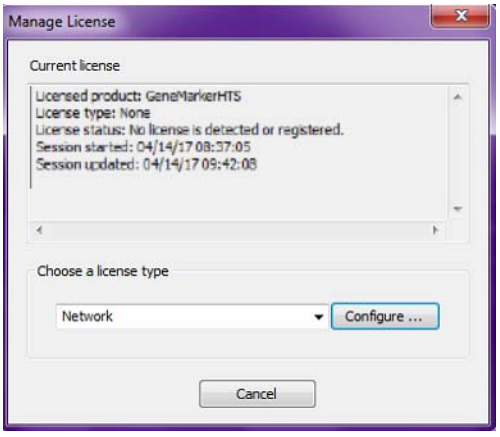
Figure 1-16: Application Settings dialog box, License tab showing Full license



3. Click Edit LSM Settings.

   The Manage License dialog box opens.

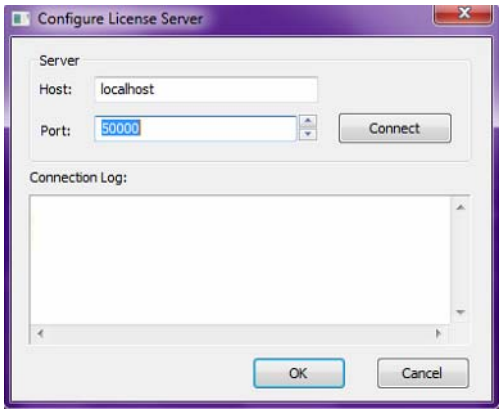
Figure 1-17: Manage License dialog box



4. On the Choose a license type dropdown list, select Network.
5. Click Configure.

   The Configure License Server dialog box opens.

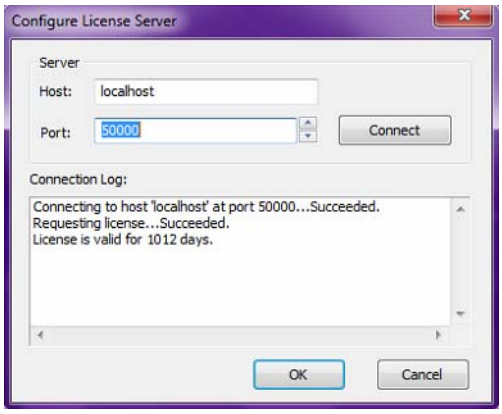
Figure 1-18: Configure License Server dialog box



6. In the Host field, enter the Fully Qualified Domain Name or IP address for the license server. (If the license server and GeneMarker HTS are installed on the same computer, then you can leave Host set to localhost.)
7. Leave the Port set to 50000.
8. Click Connect.

   Messages are displayed in the Connection Log pane indicating that connection request was successful and that GeneMarker HTS has accessed the commercial license.

Figure 1-19: Configure License Server dialog box



9. Click OK.

   The Configure License Server dialog box closes. GeneMarker HTS or the GeneMarker HTS Viewer is now ready for use.

Starting GeneMarker HTS

You can double-click the desktop icon to launch GeneMarker HTS, or you can select the option from your Start menu (Start > GeneMarkerHTS).

Figure 1-20: GeneMarker HTS desktop icon

See "The GeneMarker HTS Main Window" on page 25.

The GeneMarker HTS Main Window

The GeneMarker HTS main window is your starting point for the GeneMarker HTS application. The GeneMarker HTS main window has five major components—the title bar, the Application Settings icon, the Project Launcher options, the About button, and the Online Help button.

Title bar

The name "GeneMarker HTS" is displayed in the title bar at the top of the GeneMarker HTS main window. The version of GeneMarker HTS that you are running is displayed on the left side of the title bar.

Figure 1-21: Title bar

In the top right corner of the main window underneath the title bar, click About to open the About GeneMarker HTS dialog box, which displays information about your GeneMarker HTS installation, including the version.

Application Settings icon

The Application Settings icon is the icon that is displayed in the left corner of the main window underneath the title bar  . Click the icon to open the Application Settings dialog box, which provides all the options for specifying the settings to customize your GeneMarker HTS instance, including Colors, License, and Performance. See "Specifying GeneMarker HTS Application Settings" on page 27.

Project options

The Project options—New, Open, and Manage—are displayed in the center of the main window underneath the title bar.

- Click New to open the New Project dialog box, and set up a new analysis project in GeneMarker HTS.
- Click Open to open the Open Project dialog box and browse to and select an existing GeneMarker HTS project with which to work.
- Click Manage to open the Project Management dialog box, and then click an option.

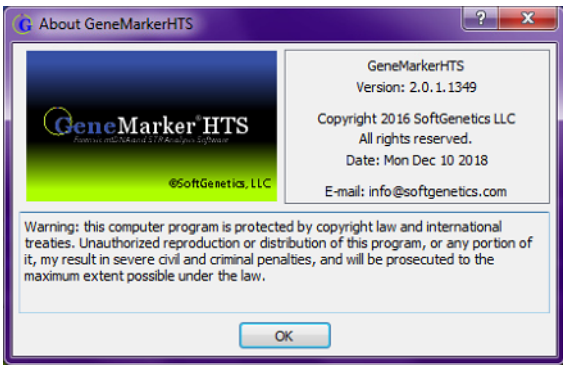
  See Chapter 2, "GeneMarker HTS Project Setup and Management," on page 35.

The full functionality of project management in GeneMarker HTS is currently under development.

About button

The About button is displayed in the top right corner of the main window underneath the title bar. Click About to open the About GeneMarker HTS dialog box, which displays information about your GeneMarker HTS installation, including the version.

Figure 1-22: About GeneMarker HTS dialog box



Online Help button

The Help button is displayed in the top right corner of the main window underneath the title bar. Click the Help button to open the online help for GeneMarker HTS.

Specifying GeneMarker HTS Application Settings

The Settings dialog box provides all the options for specifying the settings to customize your GeneMarker HTS instance. When you make any change to any GeneMarker HTS setting, the change is persistent within the GeneMarker HTS instance and across GeneMarker HTS instances. For some settings, an option is available to reset all values to their default values in a single step.

Typically, you specify your License settings *before* you run a project in GeneMarker HTS and you do not adjust the default values for the Color settings and/or Performance settings until after you open one or more projects in a viewer. Moreover, after you configure GeneMarker HTS for a commercial license, you typically do not need to edit the license type and/or LSM settings again.

Server and User Management are currently under development and therefore, are not discussed.

To specify GeneMarker HTS application settings

1. On the GeneMarker HTS main window, click the Application Settings icon  .

   The Application Settings dialog box opens. The Performance tab is the open tab.

2. Continue to one of the following:

   - "To specify the colors settings" below.
   - "To specify the performance settings" on page 34.

To specify the colors settings

You can do the following on the Colors tab of the Application Settings dialog box:

- You can specify the colors to use for background/highlighting and the nucleotides in the Pane displays for the Mito Alignment Viewer.
- You can specify the colors that indicate coverage (forward, reverse, and/or low) and variant type (major or minor) in the Global Viewer pane of the Mito Alignment Viewer.
- You can specify the colors to use for highlighting the different statuses of edited variants (added, removed, and so on) in the Variants Table of the Mito Alignment Viewer.
- You can specify the highlighting colors to use in the two panes (Sample to Sample Comparison and Variant Comparison) of the Comparison Viewer.

For detailed information about the viewers and their layouts that are available in GeneMarker HTS, see Chapter 3, "GeneMarker HTS Viewers," on page 59.

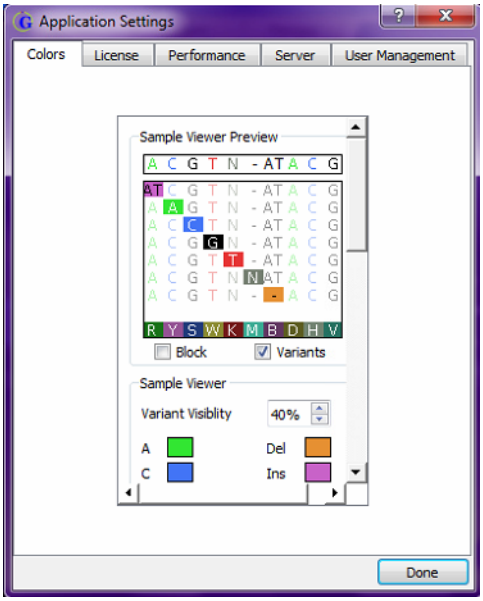
1. On the Application Settings dialog box, make sure that the Colors tab is open.

   The tab is divided in to five panes, which, from top to bottom, are the following:

   - The Sample Viewer Preview pane, which shows what the display currently looks like for the panes in the Mito Alignment Viewer.
   - The Sample Viewer pane, which shows the currently selected colors for the Pane displays in the Mito Alignment Viewer.
   - The Global Viewer pane, which shows the currently selected colors for the Global Viewer pane display in the Mito Alignment Viewer.
   - The Report Table Category pane, which shows the currently selected colors for the Variants Table pane in the Mito Alignment Viewer.
   - The Comparison Viewer pane, which shows the currently selected colors for the two panes (Sample to Sample Comparison and Variant Comparison) in the Comparison Viewer.

Use the scroll bar on the right side of the tab display to scroll through and view all the available panes.

Figure 1-23: Application Settings dialog box, Colors tab



2. Continue to any of the following:

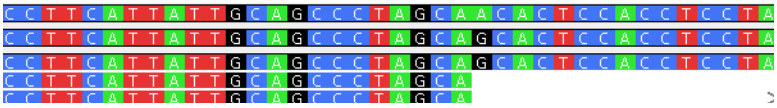
   - “To edit the Panes display colors” on page 29.
   - “To edit the Global Viewer colors” on page 31.
   - “To edit the Report Table Category colors” on page 32.
   - “To edit the Comparison Viewer highlighting colors” on page 32.

To edit the Panes display colors

If you edit the Panes display colors, then the Preview pane is dynamically updated with your settings. Note the following when editing the Pane display colors:

- If Blocks is selected, then all nucleotides are displayed in white on a colored background.

Figure 1-24: Blocks selected for Pane displays

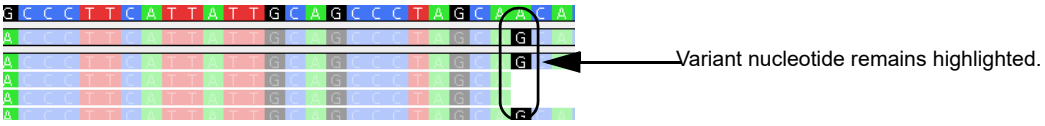


- If Variants is selected (the default value), then all nucleotides *other* than variant nucleotides are displayed in color on a white background. Conversely, variant nucleotides are displayed in white on a colored background (referred to as a highlighted variant).

Figure 1-25: Variants selected for Pane displays

You can select both Variants and Blocks for the display. In this case, all nucleotides are still displayed in white on a colored background; however, all nucleotides other than variant nucleotides are dimmed and all variant nucleotides are highlighted to visually differentiate them from non-variant nucleotides.

Figure 1-26: Variants and Blocks selected for pane display in the Mito Alignment Viewer



- *Variant Visibility* determines how dimmed non-variant nucleotides are when compared to the variant nucleotides when *both* Variants and Blocks are selected for the pane display in the Viewer. The default value is 40%. The pane also displays the various colors that are to be used in the Pane displays where the *default* values are the following:
  - For non-variant nucleotides and SNPs: If Variants is selected, then Green, Blue, Black, and Red indicate the colors for the nucleotides A, C, G, and T, respectively. If Blocks is selected, then these indicate the colors for the nucleotide backgrounds.
  - For Deletions: No nucleotide is displayed. Instead, a dash highlighted in burnt orange is displayed.
  - For Insertions: If Variants is selected, then the nucleotide is displayed in purple. If Blocks is selected, then the nucleotide background is purple.

- If coverage areas fall below the Total Coverage value that is specified for the Filter Settings: If Variants is selected, then a gray "N" is displayed for the consensus sequence in these regions. If Blocks is selected, then the nucleotide background is displayed in gray and "N" is displayed in white.

For detailed information about Filter Settings, see "Mito Alignment Viewer Filter Settings" on page 75.

Figure 1-27: Default colors for the Pane displays in the Mito Alignment Viewer

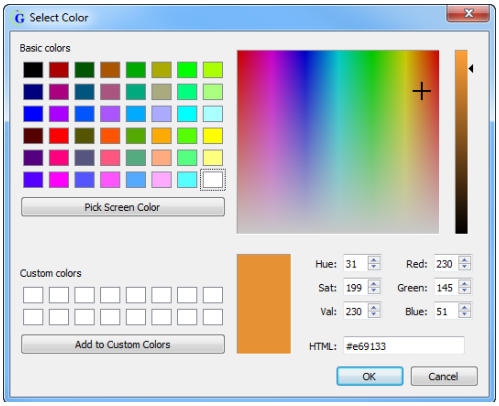
Figure 1-28: "N" indicating that coverage areas fall below the Total Coverage Value that is specified in the Filter Settings

1. Do any or all of the following as needed to change colors and/or the variant visibility:

If you change colors and/or the Variant Visibility value, you might need to refresh the display to view your changes. You can use your mouse scroll wheel, or you can zoom in/zoom out to refresh the display.

- Change one or more colors for the nucleotides and/or backgrounds. To do so, double-click the color swatch that is displayed for a nucleotide or variant type to open the Select Color dialog box, select a different color, and then click OK.

Figure 1-29: Select Color dialog box



The Select Color dialog box closes and you return to the Colors tab of the Application Settings dialog box. The newly selected color is displayed for the selected nucleotide or variant type.

- Adjust the value for Variant Visibility. To make the non-variant nucleotides more visible in the pane display when both Blocks and Variants are selected, increase the value. Conversely, to make the non-variant nucleotides less visible in the pane display when both Blocks and Variants are selected, decrease the value. You can manually enter a new value, or you can use the Up/Down arrows.

To reset all Color settings and/or the Variant Visibility to their default values in a single step, click Default.

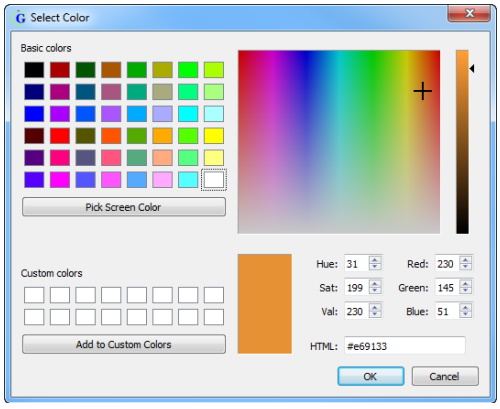
2. If you are done specifying your settings, then click OK to close the Application Settings dialog box and return to the GeneMarker HTS main window; otherwise, continue to add and/or edit the information on this tab or any of the other tabs as needed.

To edit the Global Viewer colors

You can specify the colors that indicate coverage (forward, reverse, and/or low) and variant type (major or minor) in the Global Viewer pane of the Mito Alignment Viewer.

1. Double-click the color swatch that is displayed for a coverage type or variant type to open the Select Color dialog box, select a different color, and then click OK.

Figure 1-30: Select Color dialog box



The Select Color dialog box closes and you return to the Colors tab of the Application Settings dialog box. The newly selected color is displayed for the selected coverage or variant type.

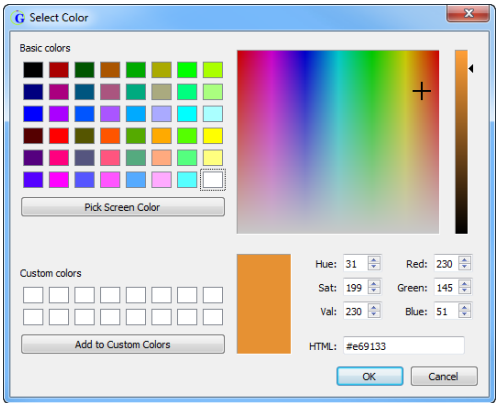
2. If you are done specifying your settings, then click OK to close the Application Settings dialog box and return to the GeneMarker HTS main window; otherwise, continue to add and/or edit the information on this tab or any of the other tabs as needed.

To edit the Report Table Category colors

You can specify the colors to use for highlighting the different statuses of edited variants (added, removed, and so on) in the Variants Table of the GeneMarker HTS viewer.

1. Double-click the color swatch that is displayed for a variant status to open the Select Color dialog box, select a different color, and then click OK.

Figure 1-31: Select Color dialog box



The Select Color dialog box closes and you return to the Colors tab of the Application Settings dialog box. The newly selected color is displayed for the selected coverage or variant type.

2. If you are done specifying your settings, then click OK to close the Application Settings dialog box and return to the GeneMarker HTS main window; otherwise, continue to add and/or edit the information on this tab or any of the other tabs as needed.

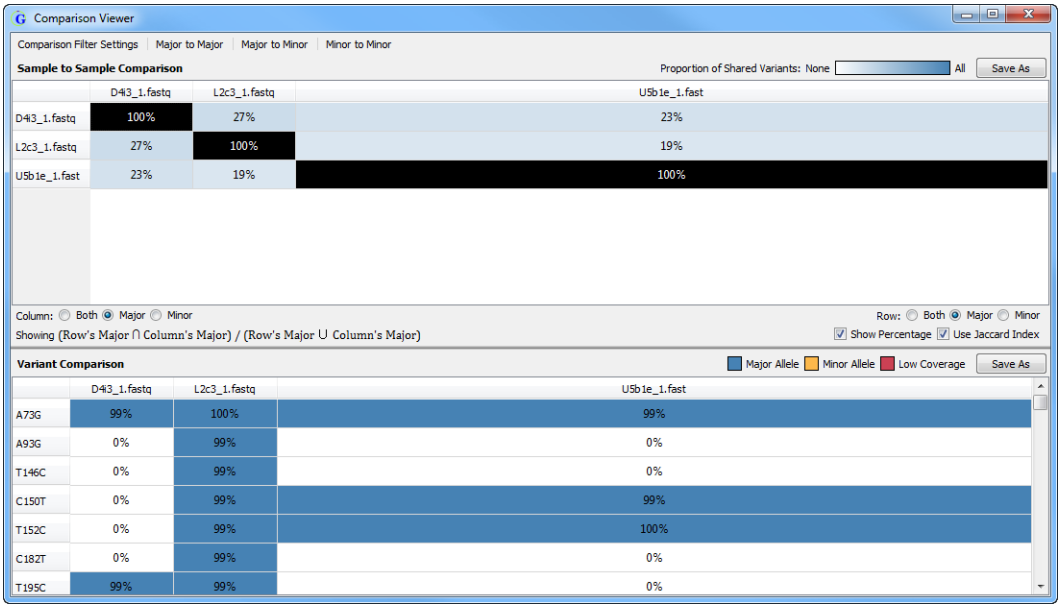
To edit the Comparison Viewer highlighting colors

The Comparison Viewer is divided into two panes: the Sample to Sample Comparison (top) pane and the Variant Comparison (bottom) pane.

- In the Sample to Sample Comparison (top) pane, diagonal cells that compare a sample to itself are highlighted by default in black. All other cells in the correlation table compare one sample to another, and by default, are highlighted in a blue color gradient from 0 to 100%, where 0% cells are blank/white.
- In the Variant Comparison (bottom) pane, variants that can be classified as a major allele or as a minor allele are highlighted, by default, in blue or gold, respectively. If a variant was determined to be a Low Coverage variant, then by default, the variant is highlighted in red.

See Figure 1-32 on page 33.

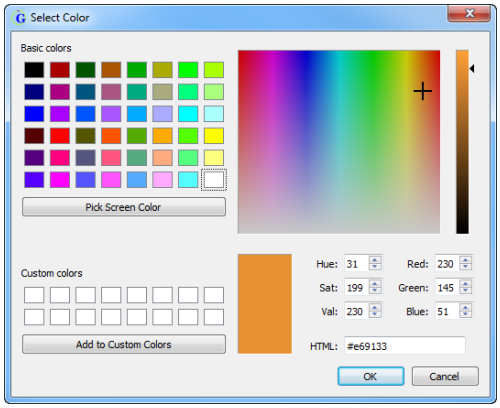
Figure 1-32: Comparison Viewer



1. Change one or more highlighting colors as needed. To do so, double-click the color swatch that is displayed for a nucleotide or variant type to open the Select Color dialog box, select a different color, and then click OK.

   The Select Color dialog box closes and you return to the Colors tab of the Application Settings dialog box. The newly selected color is displayed for the selected nucleotide or variant type.

Figure 1-33: Select Color dialog box



2. If you are done specifying your settings, then click OK to close the Application Settings dialog box and return to the GeneMarker HTS main window; otherwise, continue to add and/or edit the information on this tab or any of the other tabs as needed.

To specify the performance settings

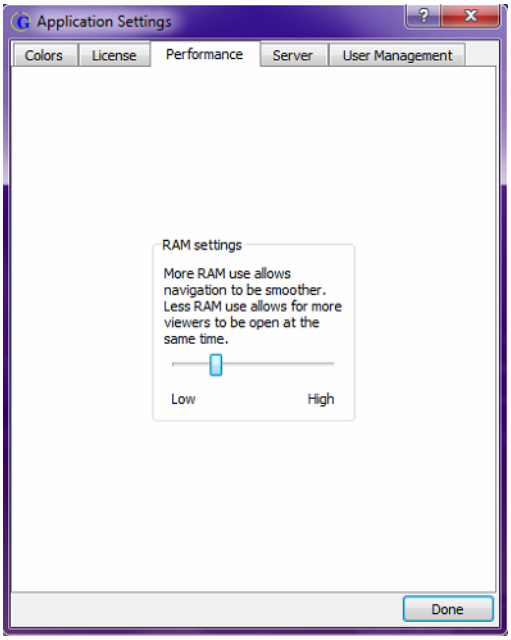
Performance settings determine the speed at which a sample file opens in the Mito Alignment Viewer as well as the speed at which you can navigate the viewer display.

1. On the Application Settings dialog box, open the Performance tab.

   The tab displays the current RAM use on a sliding scale, where:

   - More RAM use allows for fewer Mito Alignment Viewers to be open at the same time, but navigation in these viewers is much "smoother."
   - Less RAM use allows for multiple Mito Alignment Viewers to be open at the same time.

Figure 1-34: Application Settings dialog box, Performance tab



2. Optionally, use the sliding scale to adjust the amount of RAM that your GeneMarker HTS instance uses. (Five levels between Low and High are available.)
3. If you are done specifying your settings, then click OK to close the Application Settings dialog box and return to the GeneMarker HTS main window; otherwise, continue to add and/or edit the information on any of the other tabs as needed.

Chapter 2
GeneMarker HTS Project Setup and Management

GeneMarker HTS is ideal for the analysis of mtDNA data from leading NGS sequencing platforms including Illumina and Ion Torrent. The New Project dialog box provides all the options for adding a new project for mtDNA data analysis. You can also open and work with existing projects in GeneMarker HTS. This chapter details all the options and functions that are available for setting up and managing your GeneMarker HTS projects.

This chapter covers the following topics:

- "Overview of the New Project Dialog Box" on page 37.
- "Working with Projects in GeneMarker HTS" on page 49.
- "GeneMarker HTS Project Output Files" on page 57.

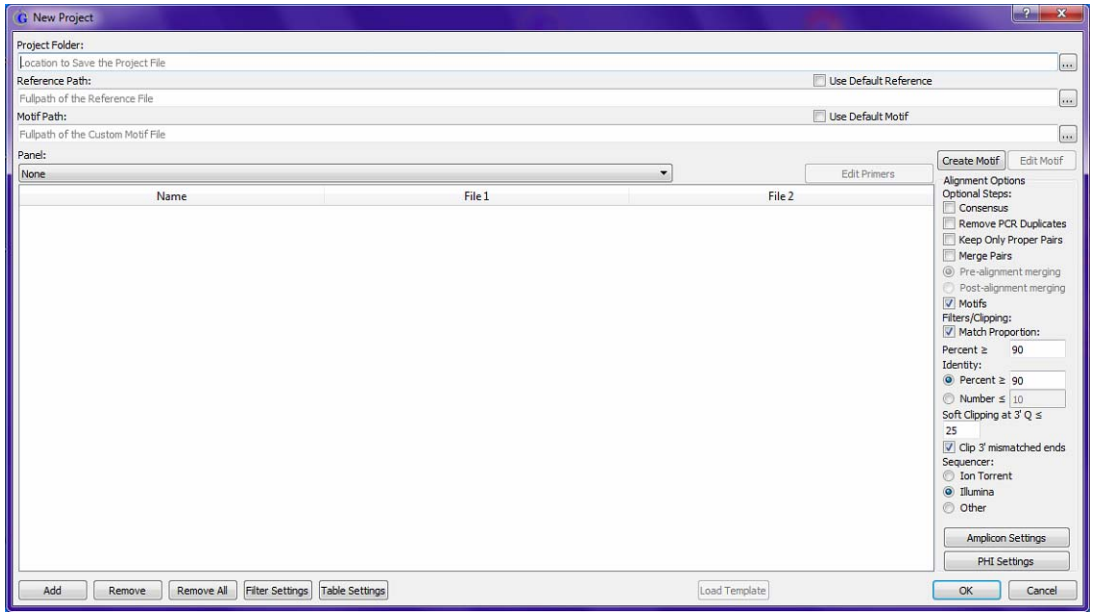
Chapter 2
GeneMarker HTS Project Setup and Management

Overview of the New Project Dialog Box

For detailed information about working with an existing GeneMarker HTS project see, "Working with Projects in GeneMarker HTS" on page 49.

You use the New Project dialog box to set up a new project for analyzing your data. You select the data that is to be analyzed for a project in the dialog box, and you also specify the alignment settings for the project in the dialog box. To open the New Project dialog box, at the top of the GeneMarker HTS main window, click New.

Figure 2-1: New Project dialog box



The following table describes the settings on the New Project dialog box the first time that the dialog box is opened. Any selections that you make are persistent within your GeneMarker HTS instance and across instances.

Table 2-1: New Project dialog box settings

Setting	Description
Project Folder	Location in which to save the project file. The default location is the folder in which the project sample files are saved. For example, if the project sample files are saved in C:\GM_HTS\Project_A_SampleFiles, then the project file is also saved in \GM_HTS\Project_A_SampleFiles. To select a different location, next to the Project Folder field, click the Browse button to open the Project Save Location dialog box, and then browse to and select a different folder.

Table 2-1: New Project dialog box settings (Continued)

Setting	Description
Reference Path Use Default Reference	Do one of the following: • To use the revised Cambridge reference sequence, or rCRS, that is supplied with GeneMarker HTS, select Use Default Reference. • To use a different reference, clear Use Default Reference, and then next to the Reference Path field, click the Browse button to open the Open Reference dialog box, and then browse to and select the appropriate reference. **Warning:** If you select a different reference, then it must in either FASTA or GBK format.
Note: Motif options are available only if Motifs is selected for an Alignment option. (See "Motifs" on page 39.) If Motifs is not selected as an alignment option, then all Motif options are unavailable. SoftGenetics recommends the use of a motif for mtDNA analysis.	
Motif Path Use Default Motif	You use a motif (a .txt file) to adjust the alignment to meet the expectations of your analysis. See "Using a motif" on page 41. Do one of the following: • To the motif data that is supplied with GeneMarker HTS, select Use Default Motif. • To use a different motif, clear Use Default Motif and then next to the Motif Path field, click the Browse button to open the Open Motif dialog box, and then browse to and select the appropriate custom motif. After you select a motif, the full directory path, including the filename, for the selected custom motif is displayed in the Motif Path field.
Panel	Required for STR data. Optional for mtDNA data. Leave Panel set to the default value of None or, on the Panel dropdown list, select from one of the pre-supplied panels: • Promega_Powerseq_AutoMitoY • Promega_Powerseq_AutoY • Promega_Powerseq_FullMito • Promega_Powerseq_MitoControlRegion **Note:** The available panels are for the simultaneous analysis of autosomal- and Y-STRs along with mitochondrial data from one sample. For a complete description of the available panels, see https://www.promega.com/resources/profiles-in-dna/2016/analyzing-data-from-next-generation-sequencers-using-the-powerseq-automitoy-system/.
Edit Primers	Under development.
Create Motif	Opens the Motif Editor dialog box, which provides the options for creating the custom motif that is to be used for your mtDNA analysis. If you select this option, then the built-in (default) motif is used as the starting point for creating the custom motif. See "To create or edit a custom motif" on page 42.
Edit Motif	Available only after you select a custom motif that is to be used for your mtDNA analysis. Opens the Motif Editor dialog box, which makes the selected custom motif available for editing. See "To create or edit a custom motif" on page 42.

Table 2-1: New Project dialog box settings (Continued)

Setting	Description
Alignment Options	
Optional Steps: (You can select one or more, or all of the following five optional steps.)	
Consensus	Do local indel realignment based on the consensus sequence. After variant calls are made based on the initial alignment, any reads that overlap indels but do not contain the indels are adjusted to include the indels if the new alignment is at least as good as the initial alignment.
Remove PCR Duplicates	On or Off. Reads are identified as PCR duplicates based on their start positions. If you are using paired-end data, then each pair of reads is considered together. Select to remove PCR duplicates, or leave blank and the PCR duplicates are not removed. **Note:** If you are using targeted sequencing, then you typically do not select this value. If you are using whole genome sequencing, then you typically do select this value.
Keep Only Proper Pairs	Ignore any reads that were not marked as being part of a proper pair, where a proper pair is defined as the following: • Both reads from the pair must be aligned and the alignment must be on opposite strands. • The 5’ end of the reverse read must be to the right of the 5’ end of the forward read.
Merge Pairs	Applicable only for mtDNA data that does *not* use a panel. If the reads are expected to fully overlap, then Merge Pairs is appropriate for mtDNA data. If you select Merge Pairs, then the overlapped portion of pairs is combined, which reduces coverage; however, because double counting is prevented, accuracy might be improved. If you select Merge Pairs, then the following options are enabled, and you can select one: • Pre-alignment merging - Results in slower data analysis and is more prone to error, especially in repetitive regions, but for certain situations such as those that involve high quality data and/or sequences that are not highly repetitive, then selecting this option can improve alignment. • Post-alignment merging - Because the reads might already be correctly overlapped after alignment, selecting this option results in faster data analysis. **Tip:** If you are not sure as to which option to select, analyze the data using both options, compare the results, and then make your selection.
Motifs	If selected, then the alignments of reads that cross motif regions with expected sequences are adjusted to match the list of variants. See “Using a motif” on page 41. **Note:** This is a default setting for all the Sequencer options (Ion Torrent, Illumina, and Other). If you do not want to use motifs, then you must manually clear this option.

Table 2-1: New Project dialog box settings (Continued)

Setting	Description
Filters/Clipping	
Match Proportion	Selected by default. Include the read if the proportion of the read that is aligned and not soft-clipped is greater than or equal to the indicated value. • Percent ≥: Default value is 90.
Identify	Reads that are less similar to the reference than indicated are not aligned. Select the way to identify the number of reads: • Percent ≥: Default value is 90. If the percentage of reads for a sample that do not match the reference are greater than or equal to the indicated value, then reads are not aligned. • Number ≤: Default value is 10. If the total number of reads for a sample that align to reference are less than or equal to the indicated value, then the reads are not aligned.
Soft Clipping at 3bp Q ≤	When the base call quality is less than or equal to the indicated quality score, then the 3' ends of reads are trimmed accordingly. See "Soft clipping based on quality scores" on page 44.
Clip 3' mismatched ends	Selected by default. If mismatched bases are found near the end of the alignment, then soft clipping is also carried out on the 3' ends of reads. See "Soft clipping based on quality scores" on page 44.
Note: GeneMarker HTS automatically adjusts the default values for both Identity and Soft Clipping based on the selected sequencer. The default settings are based on using an Illumina or Ion Torrent sequencer. If you change the value for either one or both of the settings, then GeneMarker HTS automatically selects "Other" for Sequencer.	
Sequencer - You can select only one value.	
Ion Torrent	Commonly used settings for Ion Torrent data.
Illumina	Selected by default. Commonly used settings for IIIlumina data.
Other	Select only if you must make changes to the common settings for your data.
Amplicon Settings	Opens the Amplicon Settings dialog box, which contains the settings for specifying which amplicon-sequenced regions were used for the mtDNA alignment project. See "Specifying amplicon regions" on page 45. **Note:** If this option is selected for an mtDNA alignment project, then these amplicon regions are displayed in an Amplicons pane in the GeneMarker HTS Viewer. See "Mito Alignment Viewer Layout" on page 61.
PHI Settings	Opens the PHI Settings dialog box, which contains the settings for specifying which personal health information to mask for a subject. You can specify which PHI to mask for a subject *only before* you carry out the corresponding mtDNA alignment project. See "Masking personal health information (PHI) for a subject" on page 46.
Filter Settings	Opens the Filter Settings dialog box, which contains options for customizing the display of the mtDNA and STR Viewers. See "Mito Alignment Viewer Filter Settings" on page 75 and/or "STR Analysis Viewer Filter Settings" on page 90.

Table 2-1: New Project dialog box settings (Continued)

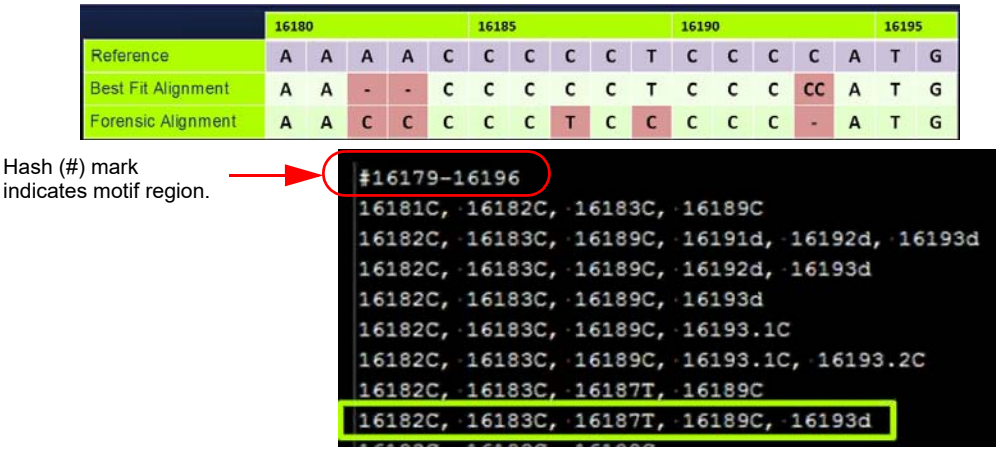
Setting	Description
Table Settings	Opens the Table Settings dialog box, which displays a list of all the data fields that are to be displayed in the Variants table in the Mito Alignment Viewer. See "Mito Alignment Viewer Filter Settings" on page 75.
Note: SoftGenetics has set default values that it has determined to be the most appropriate for mtDNA analysis for both the Filter Settings and Table Settings, but you can always edit any or all of these values. See "Mito Alignment Viewer Filter Settings" on page 75.	

Using a motif

Traditionally, alignment is focused on minimizing the number of mismatches between the reference and a sample read, or a "best fit" alignment; however, because forensic analysis has established conventions for the position of many common variants, this traditional approach to alignment is not always preferred when analyzing mtDNA. As a result, GeneMarker HTS provides the option of using a motif to better meet your analysis requirements. A *motif* is a list of variant calls within a specified region that are translated into an expected sequence. You define this list of variant calls in a *motif file*, which is a text (.txt) file that has the following format:

- The first line for each motif definition in the file is preceded with a hash tag/pound symbol (#) and this line identifies the start and end points for the region.
- The variant calls in the motif file are listed with standard forensic nomenclature, for example, Deletion: 16191**d**, Insertion: 16193.1**C**, and SNP: 1184**A.**

Figure 2-2: Example of Best Fit alignment compared to Forensic alignment



When you are creating a motif file, note the following:

- Spacing of the variants is not absolute, but you must use commas to separate the variant calls.
- Regions cannot overlap.
- No two motifs in the same region can resolve to the same sequence.

You can use the Check For Errors option that is available as part of the built-in Motif Editor to confirm that no two motifs in the same region resolve to the same sequence.

If you elect to use motif alignment, then the alignments of reads that cross motif regions with expected sequences are adjusted to match the list of variants. Reads that start or end in motif regions are soft-clipped. You can use the built-in motif that is supplied with GeneMarker HTS, or you can the Motif Editor to create and/or edit your own custom motifs. See "To create or edit a custom motif" below.

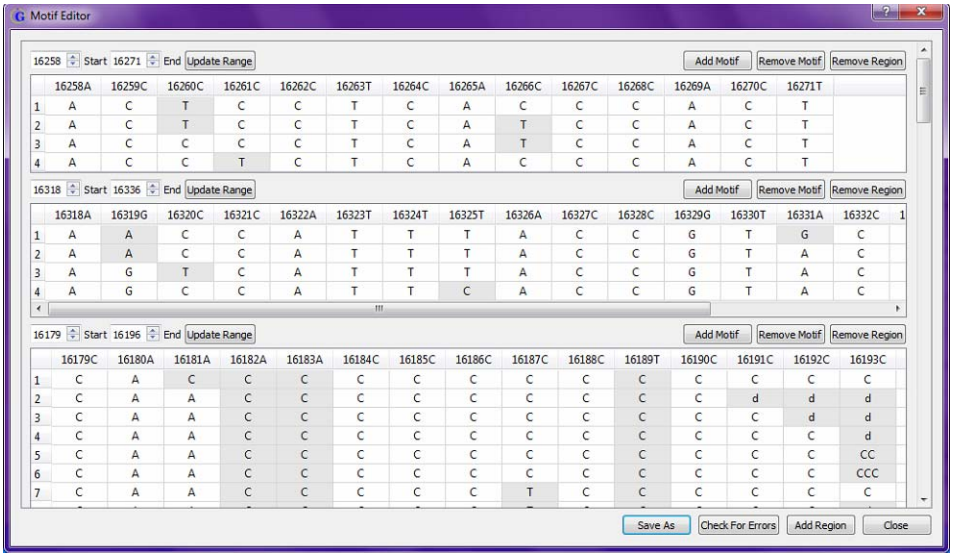
To create or edit a custom motif

Although you can manually edit a motif file, SoftGenetics strongly recommends that you use the built-in Motif Editor instead.

1. On the New Project dialog box, do one of the following:

Option	Description
Create a custom motif	Under Alignment Options, select Motifs, and then click Create Motif. The Motif Editor dialog box opens. The built-in (default) motif is loaded in the dialog box. Expected variant calls are highlighted in gray.
Edit an existing custom motif	1. Next to the Motif Path field, click the Browse button to open the Open Motif dialog box, and then browse to and select the appropriate custom motif. 2. Click Edit Motif. The Motif Editor dialog box opens. The selected custom motif is loaded in the dialog box. Expected variant calls are highlighted in gray.

Figure 2-3: Motif Editor dialog box



2. For *each* motif region, do any or all of the following as needed:

Option	Description
Update motif range	Adjust the Start and/or End values for a motif region (manually edit the values, or use the Up Arrow/Down Arrow buttons to adjust the value one base at a time), and then click Update Range. **Tip:** Because motifs require exact matches, shorter motifs are better for alignment.
Add Motif	Adds a new motif as the last in the list of displayed motifs for the region. The default sequence for a new motif is always set to the reference sequence.
Remove Motif	Select a motif, and then click to remove the selected motif from the motif region.
Remove Region	Click to remove a region entirely from the motif.

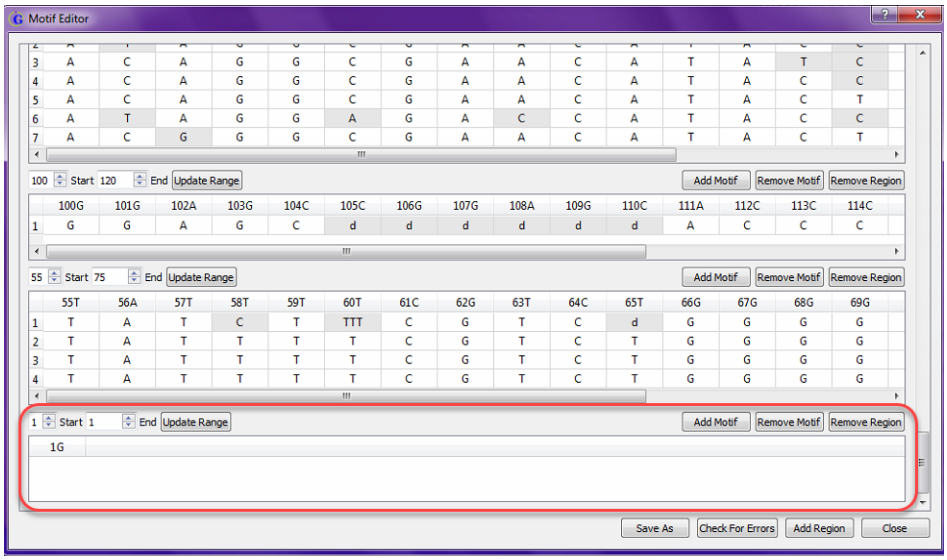
3. If you are done with customizing the motif, then continue to Step 4; otherwise, optionally, to add one or more regions to the motif, do the following:

   a. Click Add Region.

      A blank new region placeholder is displayed at the bottom of the Motif Editor dialog box.

Use the scroll bar on the right side of the dialog box to scroll to the bottom of the dialog box and view the placeholder region.

Figure 2-4: Motif Editor dialog box with a new region placeholder



   b. Specify the start and end positions for the new region, and then click Update Range.

   c. Create the needed motifs for the new range. (See Step 2.)

4. Optionally, to check for errors before saving the customized motif, click Check For Errors.

   If errors exist in the customized motif, then a Motif File Errors dialog box that details the errors in the motif file opens; otherwise, the Motif File Errors dialog box displays the message “No errors.”

5. Click OK to close the Motif File Errors dialog box, and if the file did not contain any errors continue to Step 6; otherwise, iteratively correct the file, and then click Check For Errors until you receive the message “No Errors,” and then continue to Step 6.

6. To save the new motif file, click Save As.

   The Save Motif dialog box opens. The file format is set to .txt and you cannot edit this.

7. Specify the name for the motif file, and then browse to and select the location in which to save the file, and then click Save.

   The Save Motif dialog box closes. The Motif Editor dialog box remains open.

8. Click “X” to close the Motif Editor dialog box.

   You return to the New Project dialog box. The full path to the customized motif file, including the file name, is displayed in the Motif Path field. The motif can now be used for mtDNA analysis.

Soft clipping based on quality scores

If the base call quality is less than or equal to the quality score that is specified in the Alignment Settings for a project, then the 3’ ends of reads are soft-clipped. If mismatched bases are found near the end of the alignment, then soft clipping is also carried out on the 3’ ends of reads according to the following:

1. Starting at the end of the alignment sequence, move towards the middle of the alignment sequence, with a quality score for each nucleotide in the sequence calculated as follows.

   - Add 1 for a matched base; otherwise, subtract 3 for a mismatched base. A quality score < -6 is not allowed.

2. Continue calculating a quality score for each base in the alignment sequence until a nucleotide with a quality score of 6 is found.
3. Move back from this position towards the end of the read until a mismatch is found.
4. Soft clip from this mismatch through the end of the read.

For example:

- The alignment results in a CIGAR string of 100=2X10=1X3=.
- A score is calculated going back from the end: 0, 1, 2, 3, 0, 1, 2, 3, 4, 5, 6.
- Moving back to the end, soft clipping is started at the first mismatch through the end, resulting in four bases total being soft-clipped: 100=2X10=4S.

Specifying amplicon regions

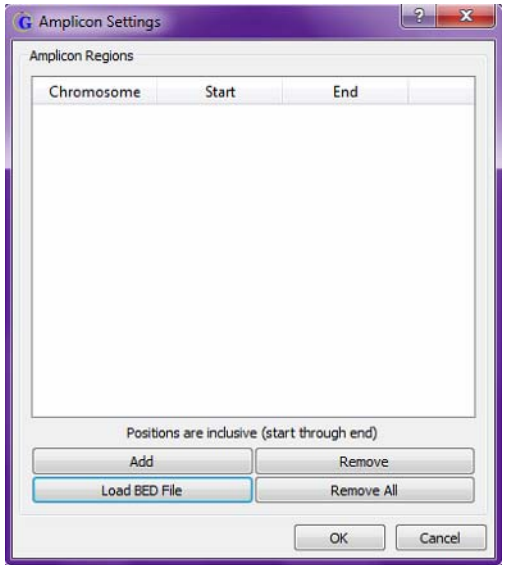
The Amplicon Settings dialog box contains the settings for specifying which amplicon-sequenced regions were used for the mtDNA alignment project. If you select Amplicon Settings for an mtDNA alignment project, then you must do one of the following:

- You must load a BED file that includes the amplicon regions.
- You must manually specify the amplicon regions.

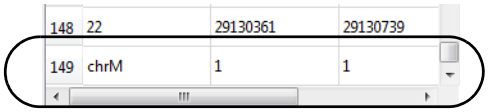
1. On the New Project dialog box, click Amplicon Settings.

   The Amplicon Settings dialog box opens.

Figure 2-5: Amplicon Settings dialog box



2. Do one of the following:

Option	Description
Load an amplicon BED file	Click Load BED File to open the Open BED File dialog box, and then browse to and select the appropriate BED file.
Note: For detailed information about the format for a BED file, see "BED file" on page 97.	
Manually add amplicon regions **Note:** You can manually add as many regions as needed.	Click Add. A new region is added as the last region in the list of displayed regions. By default, the chromosome is set to M, and both the Start and End positions are set to "1," but you can edit any or all of these values. *Figure 2-6: New amplicon region* 

Option	Description
Remove a region **Note:** You can remove only a single region at a time.	Select the region, and then click Remove. **Warning:** No confirmation message is displayed before you remove a region, so be cautious when doing do.
Remove all regions in a single step	Click Remove All.

3. Click OK.

   The Amplicon Settings dialog box closes. The New Project dialog box remains open.

4. Continue with setting up the new project as appropriate.

Masking personal health information (PHI) for a subject

When carrying out a forensic analysis, a lab technician must remain aware of the potential of exposing personal health information (PHI) about a subject that is not relevant to the investigation. For example, the alignment results could reveal that a subject is a carrier for a an inherited disease such as Leigh's syndrome. To prevent this, you can hide regions from the alignment that could reveal PHI. You can elect to load default hidden regions, which hides the PHI regions that MITOMAP has identified, and/or hide the PHI regions that you explicitly define in a BED file.

For detailed information about the format for a BED file, see "BED file" on page 97.

You can specify which personal health information to mask for a subject only before carrying out the corresponding mtDNA alignment project.

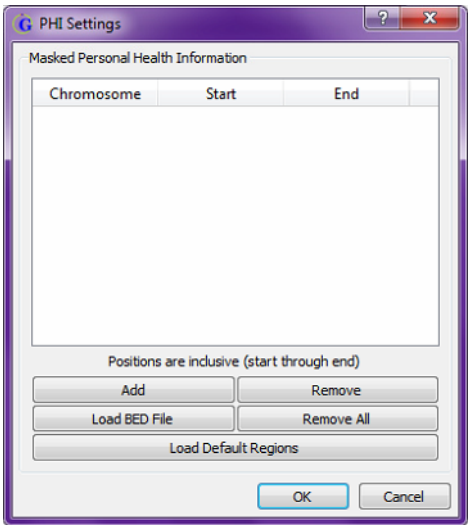
For a default or custom hidden region, the following is applicable:

- Nucleotides in the region are not displayed in the Pile-Up pane in the GeneMarker HTS Viewer.
- The CIGAR string is hidden in the tooltip for a read in the Pile-Up pane.
- You cannot copy sequences for reads that cross the region.
- Allele counts are hidden from all reports.
- No variants are called within the region.

1. On the New Project dialog box, click PHI Settings.

   The PHI Settings dialog box opens. See Figure 2-7 on page 47.

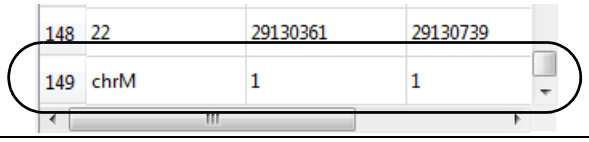
Figure 2-7: PHI Settings dialog box



2. Do one or both of the following:

Option	Description
Hide the PHI regions that MITOMAP has identified	Click Load Default Hidden Regions. **Note:** For a list of the MITOMAP locations that are the built in to PHI, see the following: • http://www.mitomap.org/foswiki/bin/view/MITOMAP/MutationsRNA • http://www.mitomap.org/foswiki/bin/view/MITOMAP/MutationsCodingControl
Hide the PHI regions that you have identified	Click Load BED File to open the Open BED File dialog box, and then browse to and select the appropriate BED file.

3. Optionally, after you have loaded the default and/or custom regions that are to be hidden, do any or all of the following as needed:

Option	Description
Manually add a region **Note:** You can manually add as many regions as needed.	Click Add. A new region is added as the last region in the list of displayed regions. By default, the chromosome is set to M, and the Start and End positions are both set to "1," but you can edit any or all of these values as needed. *Figure 2-8: Manually added PHI region* 

Option	Description
Remove a region **Note:** You can remove only a single region at a time.	Select the region, and then click Remove. **Warning:** No confirmation message is displayed before you remove a region, so be cautious when doing do.
Remove all regions in a single step	Click Remove All.

4. Click OK.

   The PHI Settings dialog box closes. The New Project dialog box remains open. The selected regions will be hidden from alignment.

5. Continue with setting up the new project.

Working with Projects in GeneMarker HTS

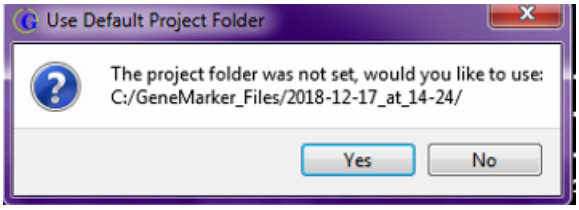
The New Project dialog box provides all the options for adding a new project in GeneMarker HTS. You can also open and work with existing projects in GeneMarker HTS.

To add a new project

Before you add a new project in GeneMarker HTS, it is important to note the following:

- If you do not specify a folder in which to save the project, then GeneMarker HTS informs you that the project folder was not set, and asks you if you want to use the default project folder, which is the folder that contains the project sample files.

Figure 2-9: Create Path prompt



- Every GeneMarker HTS project is named project. pjt, and the project file as well as all the project output files are saved in a folder that, by default, is named as <project current date>_@_<project run time>, with the date in the following format: year_month_day and the time in 24-hour time. For example, a project folder named 2018_12_12_@14-28 indicates a project that was run on December 12th, 2018 at 2:18 pm in the afternoon. This naming convention allows you to save multiple projects in the same folder without overwriting an existing project.
- The default reference that is supplied in GeneMarker HTS is the revised Cambridge Reference Sequence (rCRS). If you select a different reference, then it must in either FASTA or GBK format.
- When you add a new project in GeneMarker HTS, the sample files can be in one of two formats – compressed FASTQ files (fastq.qz) or uncompressed FASTQ files (.fastq).
- The sample file names for a project are automatically generated from the names of the loaded files, but you can always edit a sample file name.
- GeneMarker HTS automatically groups paired reads in a sample, but you can split a group of paired reads into individual samples.
- GeneMarker HTS automatically adjusts the default values for both Identity and Soft Clipping based on the Sequencer that you select. If you change the value for either one or both settings, then GeneMarker HTS automatically selects “Other” for Sequencer, and you cannot select a different value.
- Filter settings and Table Settings control the criteria for variant calling and for displaying project information in the Variants table in the Mito Alignment Viewer. Filter settings determine the alleles that are called for a project in the STR Analysis Viewer. You can

adjust Filter Settings and Table Settings before or after alignment of the samples in a project.

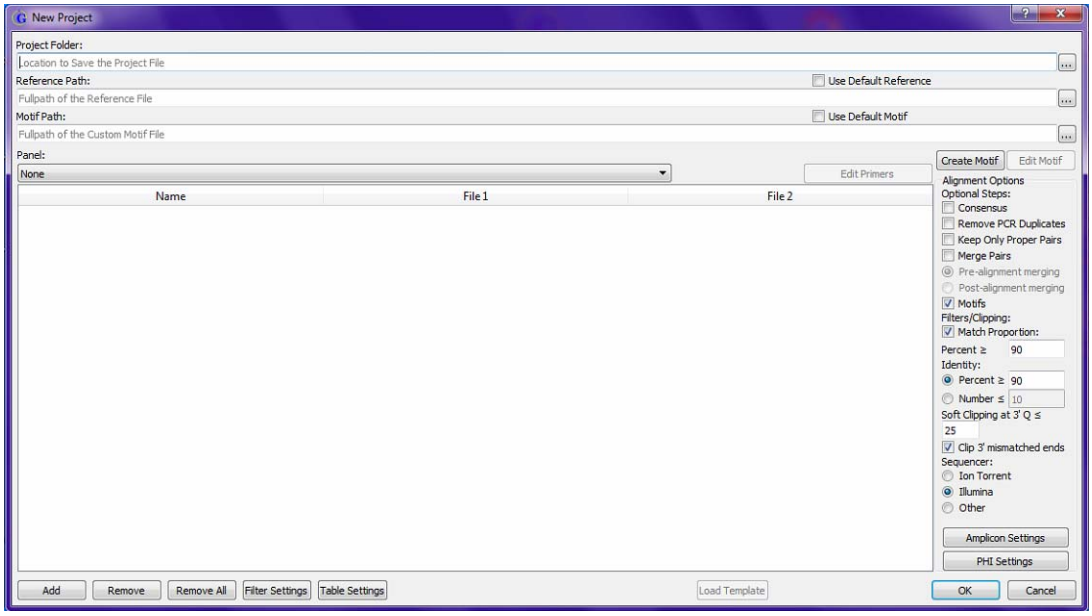
- Amplicon Settings define the amplicon-sequenced regions that were used for the mtDNA alignment project. You can adjust Amplicon Settings only *before* the alignment of the samples in a project.
- PHI Settings control which personal health information to mask for a subject. You can adjust PHI Settings only *before* the alignment of the samples in a project.

To add a new project in GeneMarker HTS, do the following:

1. On the GeneMarker HTS main window, click New.

   The New Project dialog box opens.

Figure 2-10: New Project dialog box



2. Do one of the following to specify the location in which to save the project:
   - Leave the Project Folder blank, and then, by default, after you select the project sample files, the project folder is set to the same folder as the project sample files.
   - Next to the Project Folder field, click the Browse button to open the Project Save Location dialog box, and then browse to and select where the project (*.pjt) file is to be stored.
   - In the Project Folder field, manually enter the full path to the folder in which the project is to be saved.

If the specified folder does not exist, then GeneMarker HTS will prompt you to create the path before the project is created. See Step 9.

3. If the project is to use the default reference that is supplied with GeneMarker HTS, then select Use Default Reference; otherwise, leave Use Default Reference blank, and then next to the Reference Filename field, click the Browse button to open the Open Reference dialog box, and then browse to and select the reference that the project is to use.

The default reference is the revised Cambridge reference sequence (rCRS). If you select a different reference, then it must be in either FASTA or GBK format.

4. If the project is to use the built-in motif that is supplied with GeneMarker HTS, then under Alignment Options, select Motifs, and then select Use Default Motif; otherwise, under Alignment Options, select Motifs, and then do *one* of the following:

   - Select Create Motif, and then create a custom motif.
   - Next to the Motif Path field, click the Browse button to open the Open Motif dialog box and browse to and select an existing motif.

If needed, you can edit the existing motif before using it in the project. For information about creating and editing a motif, see "To create or edit a custom motif" on page 42.

5. Specify all the options for the project, including the sequencer. See Table 2-1 on page 37.

GeneMarker HTS automatically adjusts the default values for both Identity and Soft Clipping based on the Sequencer that you select. The default values are based on using an Illumina or Ion Torrent sequencer. If you change the value for either one or both settings, then GeneMarker HTS automatically selects "Other" for Sequencer, and you cannot select a different value.

6. To add the samples for the project, click Add to open the Add Samples dialog box, and then select the sample files for the project.

The sample files can be in one of two formats – compressed FASTQ files (fastq.qz) or uncompressed FASTQ files (.fastq).

   The sample files are added to the project and are displayed in the Sample Files pane of the New Project dialog box. The sample file names are automatically generated from the names of the loaded files.

7. Optionally, you can do any or all of the following for loaded sample files:

   - To split paired read sample files into separate sample entries, right-click any one of the paired read sample files, and on the context menu that opens, click Split.

To combine split paired read files, right-click any of the split files, and on the context menu that opens, click Combine with <filename>.

   - To change the name of a sample file, double-click the file to select it, and then

change the name of the file.

- To remove one or more sample files from a project, select the sample file (CTRL-click to select multiple files), and then click Remove.
- To remove all sample files in a single step, click Remove All.

8. Optionally, do any or all of the following:
   - To adjust the settings that affect variant calling and therefore, the variants that are displayed in the Variants table of the Mito Alignment Viewer, click Filter Settings to open the Filter Settings dialog box, and then edit any or all of the values on the Mito Variant tab and/or the Mito Alignment tab.
   - To adjust the settings that affect the information that is displayed in the Variants table of the Mito Alignment Viewer and therefore, project report generation, click Table Settings to open the Table Settings dialog box, and then indicate which information is to be displayed in the Variants table and which is not.

For detailed information about Filter Settings and Table Settings for an mtDNA alignment project, see "Mito Alignment Viewer Filter Settings" on page 75.

   - To adjust the settings that determine the alleles that are called for a project in the STR Analysis Viewer, click Filter Settings to open the Filter Settings dialog box, and then edit any or all of the values on the STR Amplicon tab and/or the SNP Amplicon tab.

For detailed information about Filter Settings for an STR analysis project, see "STR Analysis Viewer Filter Settings" on page 90.

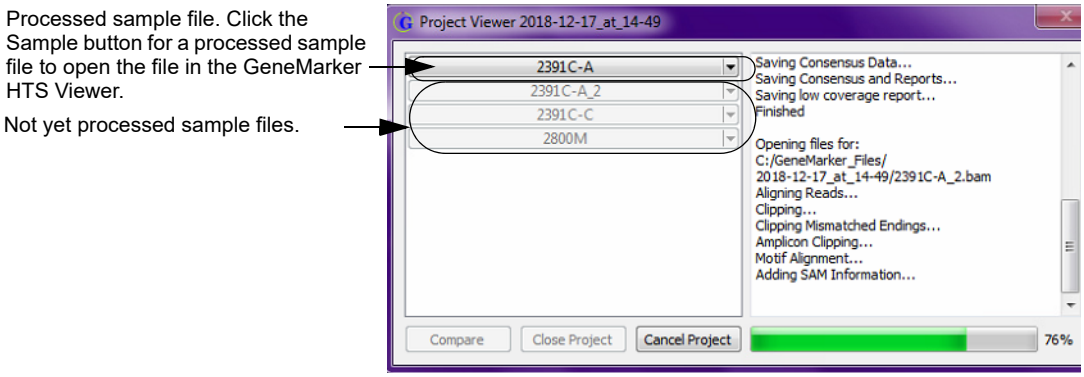
   - To adjust the amplicon regions used for the alignment, click Amplicon Settings, and then add or remove regions as appropriate. See "Specifying amplicon regions" on page 45.
   - To adjust the settings that mask the personal health information (PHI) for the subject, click PHI Settings to open the PHI Settings dialog box, and then add and/or remove default and/or custom hidden regions as appropriate. See "Masking personal health information (PHI) for a subject" on page 46.

9. Click OK.

   The Project Viewer dialog box opens. The samples that are being processed for the project are displayed on the left side of the dialog box in the form of "Sample buttons." The progress of processing a sample file is displayed on the right side of the dialog box. After a sample file is successfully processed for a project, the Sample button for the project is enabled. You can click the Sample button for any completed sample file to view its results in either the Mito Alignment Viewer (for an mtDNA only project) or the STR Analysis Viewer (for an STR analysis project), even if other sample files are being processed or are waiting to be processed.

Depending on your project setup, other messages might open before the Project Viewer dialog box opens, such as a message about creating a non-existent project folder. Answer all the prompts as needed to resolve any issues and open the Project Viewer dialog box.

Figure 2-11: Project Viewer dialog box with processed and unprocessed Sample files



For a detailed list of the output files for a GeneMarker HTS project, see "GeneMarker HTS Project Output Files" on page 57.

10. Continue to Chapter 3, "GeneMarker HTS Viewers," on page 59.

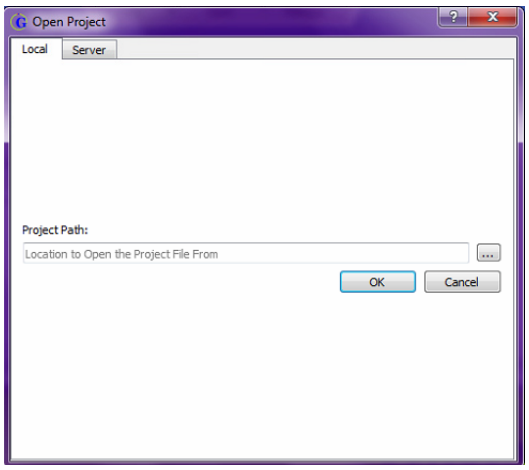
To work with an existing project

You can load a project that has been saved locally to the GeneMarker HTS client.

1. On the GeneMarker HTS main window, click Open.

   The Open Project dialog box opens. The Local tab is the open tab.

Figure 2-12: Open Project dialog box, Local tab



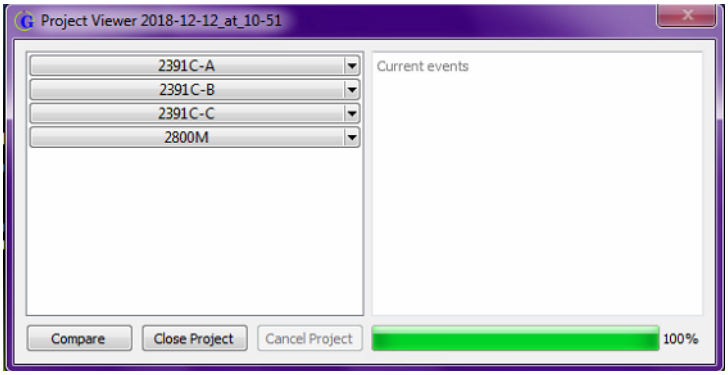
2. Next to the Project Path field, click the Browse button to open a secondary Open Project dialog box, and then browse to and select the project (*.pjt) file.

   The secondary Open Project dialog box closes. You return to the primary Open Project dialog box. The full path to the folder in which the project was saved, including the project name (project. pjt) is displayed in the Project Path field.

3. Click OK.

   The primary Open Project dialog box closes and the Project Viewer dialog box opens. The left side of the dialog box displays all the samples for the project. The project status is shown as 100% completed.

Figure 2-13: Project Viewer dialog box for a completed project



4. Do one the following:

   - Double-click a Sample button.
     - If a panel was not used for the project, then the Mito Alignment Viewer opens, displaying the mtDNA alignment results for the sample. You can open multiple individual sample files at the same time. See Figure 2-14 on page 55.See
     - If a panel was used for the project, then the STR Analysis Viewer opens, displaying the STR analysis results for the sample. You can open multiple individual sample files at the same time. See Figure 2-15 on page 56.

You can then open the Mito Alignment Viewer directly from the STR Analysis Viewer and view the mtDNA alignment results for the project.

   - To compare the results for all the samples in the project in a single Comparison Viewer window, click Compare. See Figure 2-16 on page 56.

When you open a sample file for a project in the Mito Alignment Viewer, the filter settings and table settings that were in place at the time that the project was saved determine the variants that are called and the information that is displayed in the Variants table. You can always edit these settings as needed to better suit your working needs: you can click Filter Settings to open the Filter Settings dialog box and edit the settings on the Mito Variant and/or Mito Alignment tabs, or you can click Load to select a saved project settings file. You can also request to open the viewer with alternate alignment filters. For detailed information about the Mito Alignment Viewer display, see "Mito Alignment Viewer Filter Settings" on page 75.

Figure 2-14: Mito Alignment Viewer

When you open a sample file for a project in the STR Analysis Viewer, the filter settings that were in place at the time that the project was saved determine the alleles that are called for the project. You can always edit these settings as needed to better suit your working needs: you can click Filter Settings to open the Filter Settings dialog box and edit the settings on the STR Amplicon and/or SNP Amplicon tabs, or you can click Load to select a saved project settings file. For detailed information about the STR Analysis Viewer display, see "STR Analysis Viewer Filter Settings" on page 90.

Figure 2-15: STR Analysis Viewer

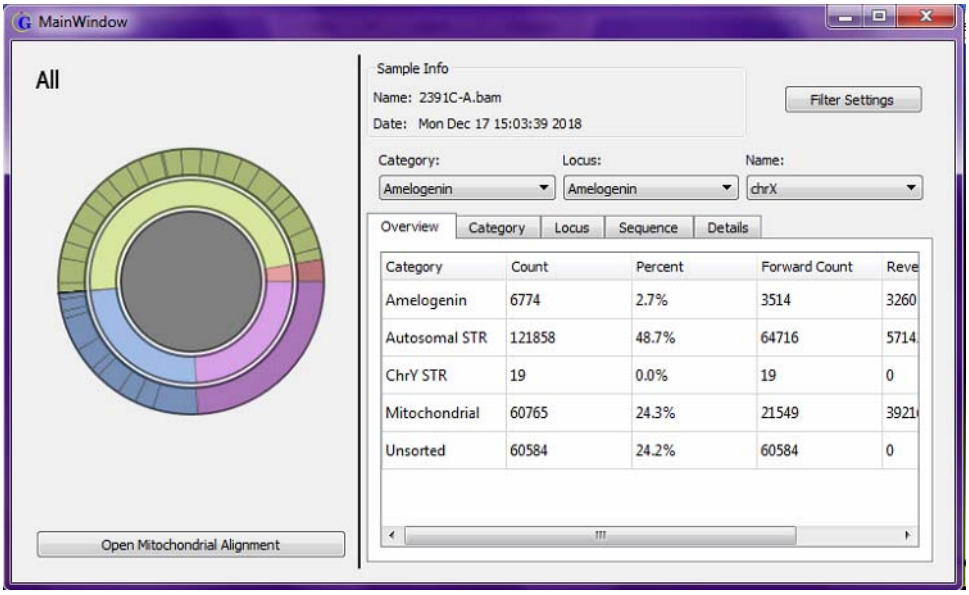
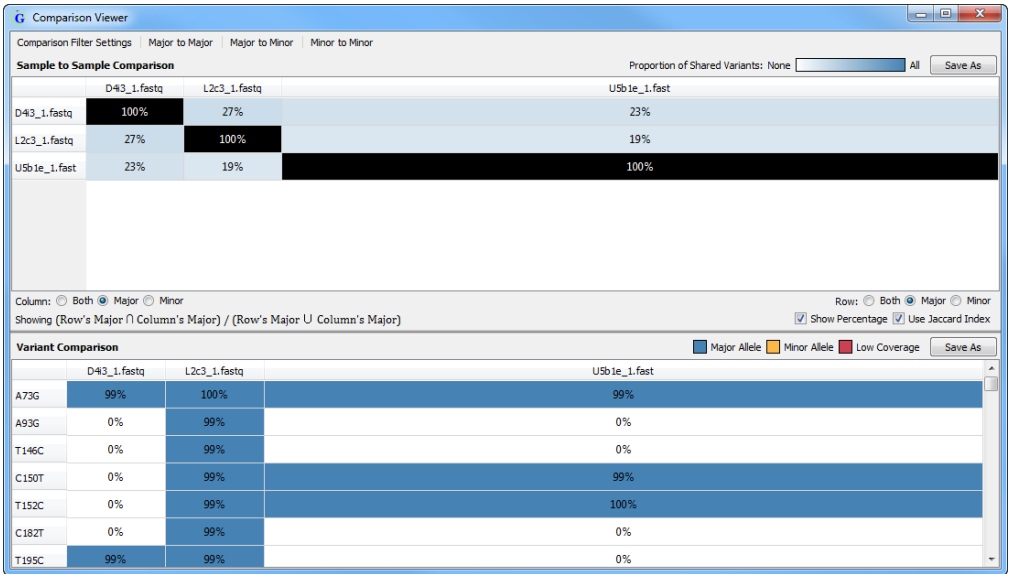


Figure 2-16: Comparison Viewer



GeneMarker HTS Project Output Files

The table below lists the files and reports that are generated for each sample in an mtDNA alignment project.

File	Description
BAM/BAI	Contains the fully detailed alignment results.
Alignment Statistic Log	Contains a summary of the alignment results.
Consensus Sequence	Details the consensus sequence with the major alleles in FASTA format. **Note:** A major allele is the allele with the highest frequency at a given position. All major alleles are reported unless the allele is the same as the reference. For detailed information about viewing and reporting on major and minor alleles for a sample alignment, see "Mito Alignment Viewer Reports" on page 83.
Consensus Statistics	Lists the number of reads with each allele at each position in the sample.
Major Allele Report	Variant report for all major alleles. **Note:** You can copy and paste a major allele string into EMPOP.
Minor Allele Report	Variant report for all minor alleles. **Note:** A minor allele is any other allele with a frequency that is less than the allele with the highest frequency at a given position For detailed information about viewing and reporting on major and minor alleles for a sample alignment, see "Mito Alignment Viewer Reports" on page 83.
Low Coverage Report	Lists all the regions with coverages that fell below the Total Coverage threshold that was specified for the project.
Project Settings	The settings (alignment, motif selection, and so on) for the project. **Note:** You can load a project settings file that was saved for one project and use this file in another project.
Variants	Combination of the Major Allele Report and Minor Allele Report.

Chapter 3
GeneMarker HTS Viewers

Three viewers are available in GeneMarker HTS application for viewing the results of a project: the Mito Alignment Viewer, the STR Analysis Viewer, and the Comparison Viewer.

- The Mito Alignment Viewer is a viewing and editing tool that you use to view and edit the results of your mtDNA alignment projects and produce a variety of interactive reports that summarize the alignment and other analysis information.
- The STR Analysis Viewer is an interactive view-only tool that you use to view the results of a simultaneous analysis of autosomal- and Y-STRs along with mitochondrial data for your project samples.
- The Comparison Viewer is a viewing tool that you use to view a sample to sample comparison as well as a variant comparison of all the samples in the project at the same time.

This chapter details these viewers as well as the reporting options that are available for your mtDNA alignment project results.

This chapter covers the following topics:

- "Mito Alignment Viewer Layout" on page 61.
- "Mito Alignment Viewer Navigation and Options" on page 69.
- "Mito Alignment Viewer Filter Settings" on page 75.
- "Mito Alignment Viewer Reports" on page 83.
- "STR Analysis Viewer Layout" on page 86.
- "STR Analysis Viewer Filter Settings" on page 90.
- "Comparison Viewer" on page 93.

Chapter 3
GeneMarker HTS Viewers

Mito Alignment Viewer Layout

You use the Mito Alignment Viewer to view the results of your mtDNA alignment projects. You can edit the display settings for the viewer and you can use viewer options to edit the results of the alignment project. You can load one or more sample files in their own Mito Alignment Viewer windows while a project is being run, or you can load one or more sample files for a saved project. (See "Working with Projects in GeneMarker HTS" on page 49.) When you open a sample file for any project in the Mito Alignment Viewer, the Filter Settings and Table Settings that were in place at the time that the project was saved determine the variants that were called and the information that is displayed in the Variants Table in the viewer, but you can always edit these settings as needed to better suit your working needs, or you can request to open the Mito Alignment Viewer with alternate alignment filters. (See "Mito Alignment Viewer Filter Settings" on page 75.)

- If you load multiple sample files from *different* projects, then each project sample opens in its own separate Mito Alignment Viewer window, and the viewer windows are *not* linked. You can independently navigate and edit the display settings for each sample file in its own viewer window.
- If you load multiple sample files from the *same* project, then each project sample opens in its own separate Mito Alignment Viewer window, but the viewer windows *are* linked. All the viewer windows zoom and scroll *horizontally* across the reference in unison. Any changes that you make to the Filter Settings for one viewer window are applied to all opened viewer windows for the project's sample files.

Figure 3-1: Mito Alignment Viewer

From top to bottom, the Mito Alignment Viewer has six major components:

- The title bar. See "Title bar" below.
- The main menu. See "Main menu" below.
- The Global View pane. See "Global Viewer pane" on page 63.
- The Pane displays. See "Pane displays" on page 64.
- The Variants table. See "Variants table" on page 67.
- The Status bar. See "Status Bar" on page 68.

Title bar

The Mito Alignment Viewer title bar displays the name for the alignment project file that is being analyzed.

Figure 3-2: Mito Alignment Viewer Title bar

Main menu

The Mito Alignment Viewer main menu is set up in a standard Windows menu format with menu commands grouped into menus across the menu bar.

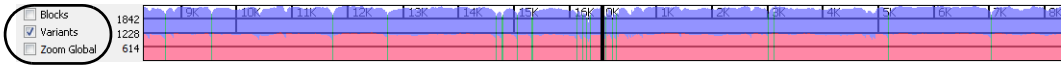
Menu Option	Description
Reports	Generate one of three reports for the project: the Coverage report, the Template Length report, or the Read Length report. See "Mito Alignment Viewer Reports" on page 83.
Filter Settings	Opens the Filter Settings dialog box, which contains options for changing the viewer display. See "Mito Alignment Viewer Filter Settings" on page 75.
Load Report	Opens the Load Report dialog box, in which you can select a sample report for display in the viewer. See "To save/load reports for an mtDNA project" on page 73.
Save BAM Edits	If you have made any edits to the BAM file, then you can save the edited BAM file as a new BAM file.
Save Report	Opens the Save Report dialog box, in which you specify a folder for saving the information that is currently displayed in the Variants table as a report. See "To save/load reports for an mtDNA project" on page 73.
Table Settings	Opens the Table Settings dialog box, which displays a list of all the columns that are currently displayed in the Variants table of the Mito Alignment Viewer. By default, all available columns are initially displayed in the Variants table. To remove a column from the Variants table display, clear its selection. To add a column to the display, select it. See "Mito Alignment Viewer Filter Settings" on page 75.
Toggle Table	Toggles the display (On or Off) of the Variants table in the Mito Alignment Viewer. See "Variants table" on page 67.

Global Viewer pane

The Global Viewer pane shows the depth of coverage for the sample. By default, the forward read coverage is displayed in blue and the reverse read coverage in red. A solid black line/ hollow black rectangle indicates the range of the pile-up. By default, green lines indicate the locations of major variants and orange lines indicate the locations of minor variants.

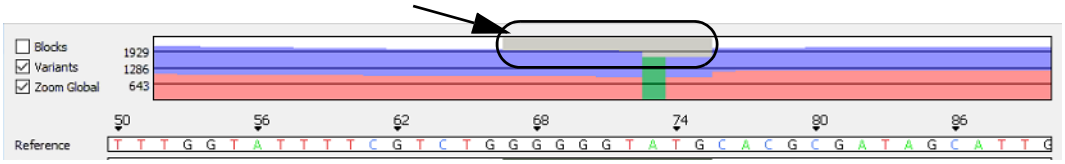
The zoom level for the Pile-Up pane determines the appearance of a line/rectangle indicator in the Global View pane. You can change all the default display colors for the pane on the Colors tab of the Application Settings dialog box. See "Specifying GeneMarker HTS Application Settings" on page 27.

Figure 3-3: Global Viewer pane



By default, gray shading indicates areas of coverage that fall below the Total Coverage threshold that was specified for the Filter Settings for the project.

Figure 3-4: Shading alerts for low coverage regions in the Global View pane



For detailed information about the Total Coverage threshold, see "To specify the filter settings for an mtDNA project display" on page 75. "N" is displayed for the consensus sequence in these areas of low coverage in the Consensus Sequence pane. See "Pane displays" on page 64.

To zoom the display of the Global Viewer pane to only that for the pile-up region, select Zoom Global.

For information about the Blocks option and the Variants option, see "Pane displays" on page 64.

Figure 3-5: Zoom Global selected

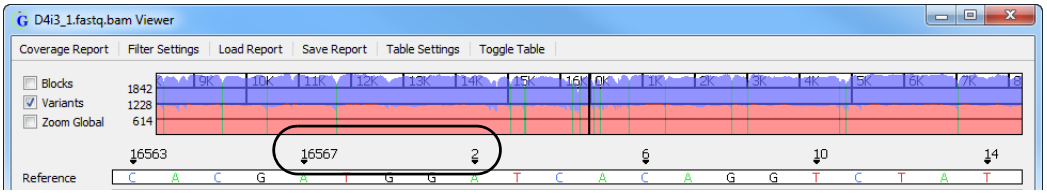
To save the image that is displayed in the Global Viewer pane as image with forward coverage stacked on top of reverse coverage, right-click in the Global Viewer pane, and on the context menu that opens, click Save Global Coverage Image. By default, the image is saved in PNG/JPEG format in the project folder, but you can always select a different location. You must name the image when you save it.

Pane displays

The Pane displays consists of up to four separate panes. Reference Position numbers are displayed at the top of the Panes displays. By definition, an mtDNA reference is circular, and this is reflected in the reference position numbers. As you scroll horizontally through a project display, the position numbers increase up to the end position of the reference, and then the reference position numbering starts over again.

For information about horizontally scrolling the display in the Mito Alignment Viewer, see "Mito Alignment Viewer Navigation and Options" on page 69.

Figure 3-6: Circular reference reflected in reference position numbering in Panes display



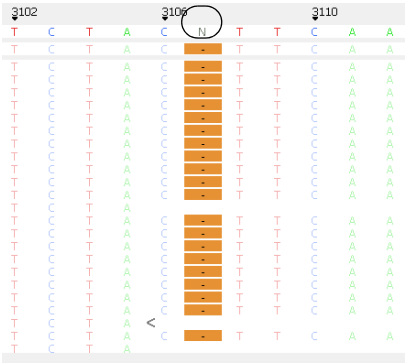
From top to bottom, the four panes in the Panes display are the following:

- The Reference pane. See "Reference pane" below.
- The Consensus Sequence pane. See "Consensus Sequence pane" on page 65.
- The Amplicons pane. See "Amplicons pane" on page 65.
- The Pile-Up pane. See "Pile-Up pane" on page 65.

Reference pane

The Reference pane displays the sequence of the reference to which the samples were aligned. To maintain position numbering in an mtDNA alignment project, an "N" is used instead of a nucleotide read in the Reference sequence. Variant calls are not reported for reference positions with an "N."

Figure 3-7: "N" for a reference position to maintain position numbering

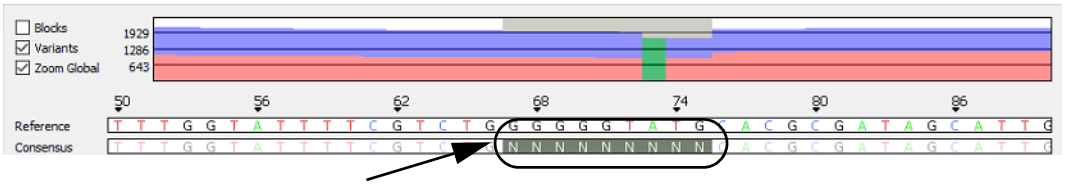


Position 3107 in the rCRS has an "N" to maintain position numbering. Bases that are aligned at this position are ignored.

Consensus Sequence pane

The Consensus Sequence pane displays the consensus sequence for the major alleles, where a major allele is defined as the allele with the highest frequency at a given position. If any areas of coverage fall below the Total Coverage threshold that was specified for the Filter Settings for the project, then the consensus sequence is set to “N” for these regions in the Consensus Sequence pane.

Figure 3-8: Low coverage regions in the Consensus Sequence pane

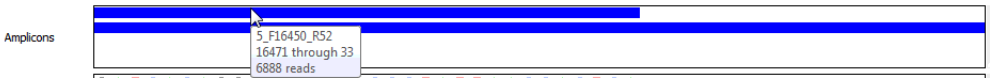


For detailed information about the Total Coverage threshold, see “To specify the filter settings for an mtDNA project display” on page 75. The corresponding areas of low coverage are highlighted in gray in the Global Viewer pane. See “Global Viewer pane” on page 63.

Amplicons pane

The Amplicons pane is displayed only if Amplicon Settings were specified for the alignment project. The pane displays the amplicon-sequenced regions that were used for the mtDNA alignment project. If you hold your mouse pointer over an amplicon region, a tooltip opens that displays the following information: the chromosome, the chromosome start and end positions, and the number of reads that were assigned to the region.

Figure 3-9: Amplicons pane

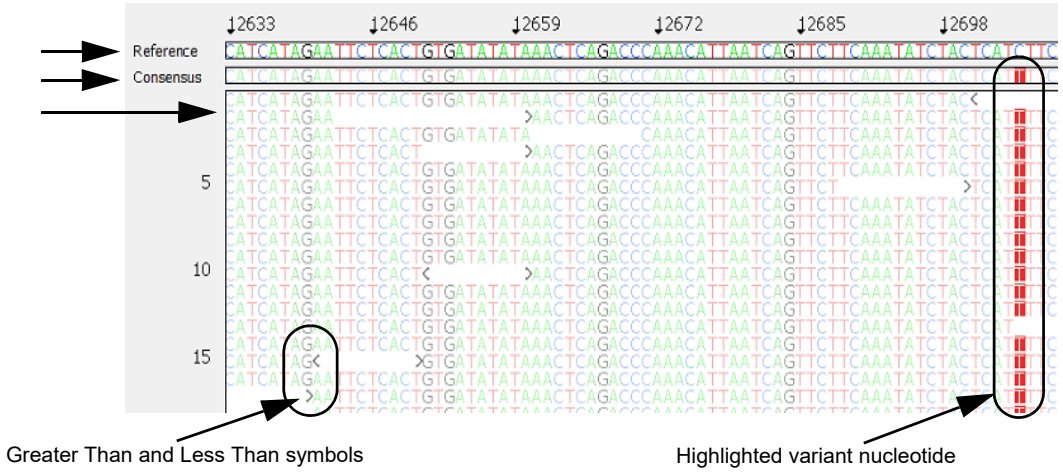


Pile-Up pane

The Pile-Up pane displays a view of all the reads for the sample as they align to the reference sequence. Greater Than and Less Than symbols point to the 5’ end of the read, where the Greater Than (>) symbol is used for forward reads and the Less Than (<) symbol is used for reverse reads. See Figure 3-10 on page 66.

Unlike the other panes in the display (Reference, Consensus Sequence, and Amplicons), the Pile-Up pane is not explicitly identified by name in the viewer; however, it is always shown in the same location: if Amplicon Settings were not specified for the alignment project, then after the Consensus Sequence pane, or if Amplicon Settings were specified for the alignment project, then after the Amplicons pane.

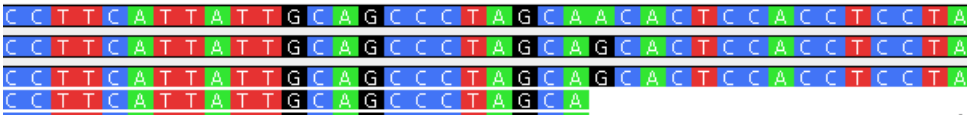
Figure 3-10: Variants selected for Pane displays



Two options displayed next to the Global Viewer pane are available for changing the Pane displays:

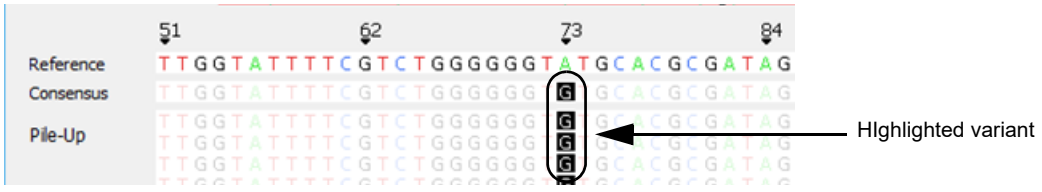
- Blocks – If Blocks is selected, then the background for each nucleotide is a color and the nucleotide (A, T, C, and G) is white.

Figure 3-11: Blocks selected for Pane displays



- Variants – Selected by default. If Variants is selected, then for all non-variant nucleotides, the background is white and the nucleotide (A, T, C, and G) is displayed in color. For variant nucleotides, the background is colored, and the nucleotide is displayed in white, which is referred to as a *highlighted* variant.

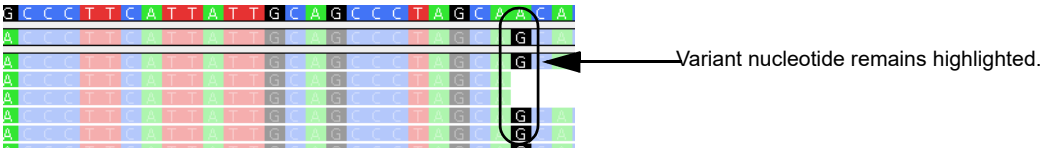
Figure 3-12: Variants selected for Pane displays



To change the background and nucleotide colors for the Pane displays, see *"Specifying GeneMarker HTS Application Settings" on page 27.*

You can select both Blocks and Variants for the display. In this case, all nucleotides are still displayed in white on a colored background; however, all nucleotides other than variant nucleotides are dimmed and all variant nucleotides are highlighted to visually differentiate them from non-variant nucleotides. See Figure 3-13 below. You can adjust the visibility of the dimmed variants. See "Specifying GeneMarker HTS Application Settings" on page 27.

Figure 3-13: Blocks and Variants selected for pane display in the Mito Alignment Viewer

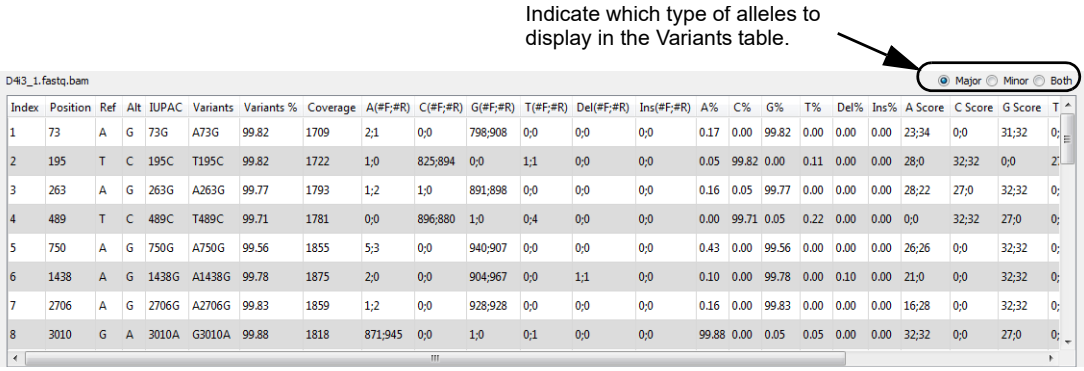


Variants table

By default, the Variants table lists each variant call in order of their sequence position, where:

- A *major* allele is defined as the allele with the highest frequency at a given position.
- All major alleles are reported unless the allele is the same as the reference.
- *Minor* alleles are any other allele at a given position with a frequency that is greater than the Minimum Threshold.
- Insertions are reported as major alleles if they occur at a frequency that is greater than or equal to 50% and as minor alleles if they occur at a frequency that is less than 50%.
- Because reference alleles can represent a secondary mutation of a major allele, reference alleles are reported as minor alleles if their frequencies occur at a rate that is greater than the minimum threshold for Allele Frequency.

Figure 3-14: Variants table



Index	Position	Ref	Alt	IUPAC	Variants	Variants %	Coverage	A(#F;#R)	C(#F;#R)	G(#F;#R)	T(#F;#R)	Del(#F;#R)	Ins(#F;#R)	A%	C%	G%	T%	Del%	Ins%	A Score	C Score	G Score	T
1	73	A	G	73G	A73G	99.82	1709	2;1	0;0	798;908	0;0	0;0	0;0	0.17	0.00	99.82	0.00	0.00	0.00	23;34	0;0	31;32	0;
2	195	T	C	195C	T195C	99.82	1722	1;0	825;894	0;0	1;1	0;0	0;0	0.05	99.82	0.00	0.11	0.00	0.00	28;0	32;32	0;0	2
3	263	A	G	263G	A263G	99.77	1793	1;2	1;0	891;898	0;0	0;0	0;0	0.16	0.05	99.77	0.00	0.00	0.00	28;22	27;0	32;32	0;
4	489	T	C	489C	T489C	99.71	1781	0;0	896;880	1;0	0;4	0;0	0;0	0.00	99.71	0.05	0.22	0.00	0.00	0;0	32;32	27;0	0;
5	750	A	G	750G	A750G	99.56	1855	5;3	0;0	940;907	0;0	0;0	0;0	0.43	0.00	99.56	0.00	0.00	0.00	26;26	0;0	32;32	0;
6	1438	A	G	1438G	A1438G	99.78	1875	2;0	0;0	904;967	0;0	1;1	0;0	0.10	0.00	99.78	0.00	0.10	0.00	21;0	0;0	32;32	0;
7	2706	A	G	2706G	A2706G	99.83	1859	1;2	0;0	928;928	0;0	0;0	0;0	0.16	0.00	99.83	0.00	0.00	0.00	16;28	0;0	32;32	0;
8	3010	G	A	3010A	G3010A	99.88	1818	871;945	0;0	1;0	0;1	0;0	0;0	99.88	0.00	0.05	0.05	0.00	0.00	32;32	0;0	27;0	0;

By default, the Variants table shows both (major and minor) allele types, but you can choose to show only major alleles or only minor alleles. The specific alleles and the information that is displayed for the alleles in the Variants table are determined by the Filter Settings and Table Settings that were in place at the time that the project was saved. See "Mito Alignment

Viewer Filter Settings" on page 75. You can click a column header to change the sort order of the data. An icon in the column header indicates the sort direction. An Up arrow indicates an ascending sort and a Down arrow indicates a descending sort.

Status Bar

After the Mito Alignment Viewer opens, a temporary status bar is displayed in the lower left corner of the viewer below the Variants table. The status bar displays the progress of loading the project in the viewer. After the project is fully loaded, the status bar briefly displays 100%, and then it is closed and is no longer displayed in the viewer.

Figure 3-15: Status bar

You must wait until after the status for loading a project is 100% and the status bar closes before you can work with the project in the Mito Alignment Viewer.

Mito Alignment Viewer Navigation and Options

You can easily navigate the Mito Alignment Viewer using your mouse as well as some keyboard hotkeys. In various areas of the viewer, you can use the mouse as well as some keyboard hotkeys to work with and edit the variant information in the viewer, and in some areas of the viewer, you can also right-click to open a context menu that contains a list of options for working with and editing the variant information.

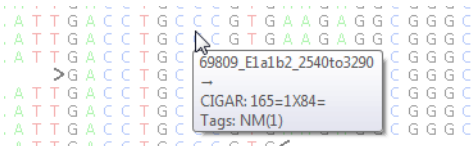
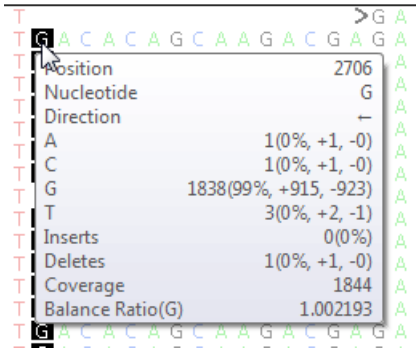
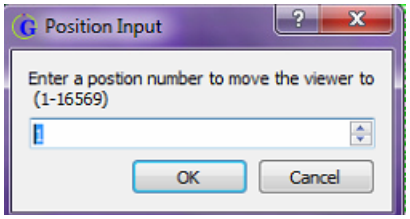
Mito Alignment Viewer navigation

You can easily navigate the Mito Alignment Viewer using your mouse as well as some keyboard hotkeys in the Global Viewer pane, the Pane displays, and the Variants table.

Option	Description
Global Viewer pane	
Double-click any region in the Global Viewer pane to move the Pane displays to the same area.	
Pane displays	
Zoom In	Hold down the left mouse button and draw a box from the upper left hand corner of the pane towards the lower right hand corner. A box is formed around the area that being reduced for viewing.
Zoom Out	Hold down the left mouse button and draw a box from the lower right hand corner of the pane towards the upper left hand corner. **Note:** The magnification for zooming out is always 100%.
Horizontal Scroll	Hold down the right mouse button and then drag the cursor across the display from left to right, or right to left.
Vertical Scroll	Use the Vertical Scroll bar to the right of the Panes display, or use the Scroll wheel on your mouse.
Moving to each variant position in the Pile-Up pane	When the Mito Alignment Viewer first opens, the starting position for the Pile-Up pane is the first reference nucleotide position (Pos #1). You can use the following keyboard shortcuts to move the display to each *variant* position: • CTRL + f moves the display forward one variant position at a time. • CTRL +b moves the display backward one variant position at a time.
Move to a closest variant position in the Variants table	Right-click any position in the Pile-Up pane and on the context menu that opens, click Go to closest position in table. The variant that is closest to the selected position is highlighted in the Variants table.
Variants table	
Double-click any entry in the Variants table to move the Pane displays to the corresponding variant position.	

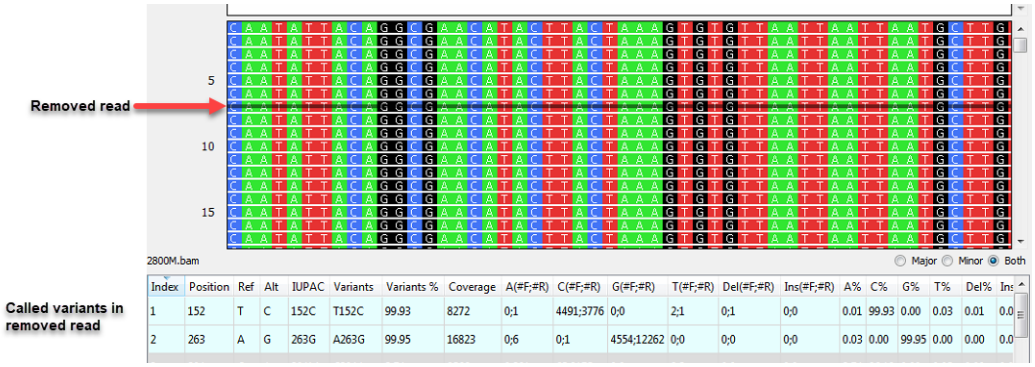
Mito Alignment Viewer options

In various areas of the viewer, you can use your mouse as well as some keyboard hotkeys to work with and edit the variant information in the viewer. In some areas, you can also right-click to open a context menu that contains a list of options for viewing and editing the variant information.

Option	Description
Pile-Up pane	
Open a tooltip with the following information for a specific read (variant or non-variant): • Read name • Read direction • Read CIGAR string • Read Tags	SHIFT + mouse scroll wheel click or CTRL +SHIFT, and then left mouse click. *Figure 3-16: Read tooltip* 
Open a tooltip with the following information for a specific variant: • The position number for the allele • The major allele at the position • The read direction for the allele • The counts for the different alleles at the position • The total coverage for the position • The Allele Balance Ratio • The Allele Balance Score	Click the mouse scroll wheel or CTRL + left mouse click. *Figure 3-17: Variant tooltip* 
Adding a variant	See "To add a variant at a given position" on page 72.
Jump to position	Right-click any position, and in the Position Input dialog box that opens, enter the position number to which to move the viewer. The position is displayed in the middle of the viewer. *Figure 3-18: Position Input dialog box* 
Go to closest position in table	Right-click any position, and on the context menu that opens, select this option to highlight the entry in the Variants table for the variant that is closest to this selected position.

Option	Description
Copy read's sequence	To copy the sequence from any read to your clipboard, right-click any read, and on the context menu that opens, click Copy read's sequence. You can then use standard Paste commands to paste the copied information into a third-party product such as Microsoft Word or Notepad.
Remove read/Undo Remove read	Right-click any position, and on the context menu that opens, select this option to remove the read from any consideration when analyzing or graphing the project results. A solid horizontal black line is displayed across the read to indicate that it has been removed. To add a removed read back, right-click the read, and on the context menu that opens, click Undo remove read. **Note:** If the removed read contains any called variants, then these variants are highlighted in light blue in the Variants table. See "Variants table" on page 67.
Variants table	
Copy selected	Click and hold the left mouse button and then draw a box (square or rectangle) around the cells that are to be copied. Right-click any of the copied cells, and on the context menu that opens, select Copy selected to copy the selected table cells to your clipboard. Use standard keyboard commands or menu commands to paste the copied cells into a third-party application such as Microsoft Word or Excel.
Deleting a variant	See "To delete a variant at a given position" on page 73.
Saving/Loading reports	See "To save/load reports for an mtDNA project" on page 73.

Figure 3-19: Called variants in a removed read



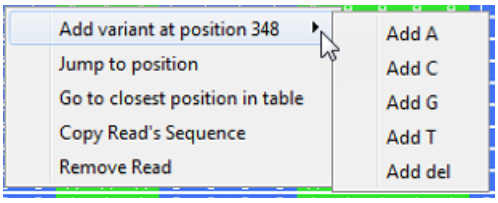
To add a variant at a given position

You can manually add a variant (insertion or deletion) at a position of your choice in the Pile-Up pane.

1. In the Pile-Up pane, right-click the position at which you are adding the variant.

   A context menu opens with various options, including options for adding a variant.

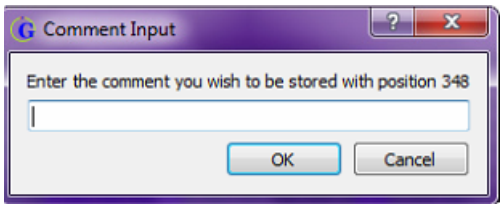
Figure 3-20: Pile-Up pane context menu



2. Click Add Variant at position <position #>, and then on the context menu that opens, select the appropriate option, such as Add A or Add del.

   The context menu closes and Comment Input dialog box opens.

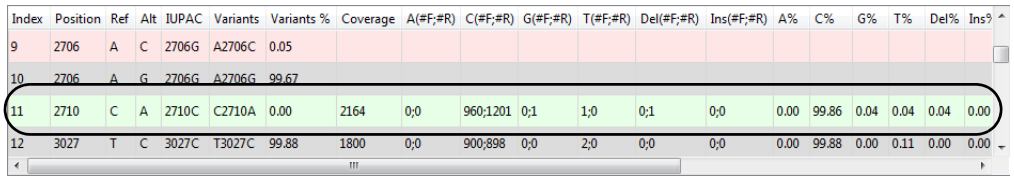
Figure 3-21: Comment Input dialog box



3. Optionally, enter the comment that is to be stored for the indicated position, and then click OK.

   The Comment Input dialog box closes. A new entry that is highlighted in green is displayed for the added variant in the Variants table.

Figure 3-22: New variant entry in Variants table



Index	Position	Ref	Alt	IUPAC	Variants	Variants %	Coverage	A(#F;#R)	C(#F;#R)	G(#F;#R)	T(#F;#R)	Del(#F;#R)	Ins(#F;#R)	A%	C%	G%	T%	Del%	Ins%
9	2706	A	C	2706G	A2706C	0.05													
10	2706	A	G	2706G	A2706G	99.67													
11	2710	C	A	2710C	C2710A	0.00	2164	0;0	960;1201	0;1	1;0	0;1	0;0	0.00	99.86	0.04	0.04	0.04	0.00
12	3027	T	C	3027C	T3027C	99.88	1800	0;0	900;898	0;0	2;0	0;0	0;0	0.00	99.88	0.00	0.11	0.00	0.00

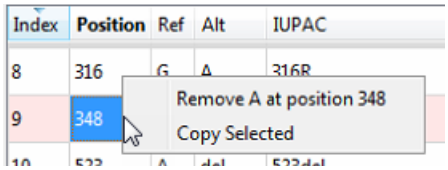
To delete a variant at a given position

You can manually delete a variant (insertion or deletion) from a project in the Variants table.

1. In the Variants table, right-click the variant that you are deleting.

   A context menu opens with two options: Remove the variant at the selected position and Copy Selected.

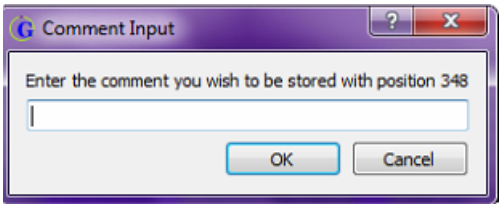
Figure 3-23: Variants table context menu



2. Select Remove <variant> at position < >.

   The context menu closes and Comment Input dialog box opens.

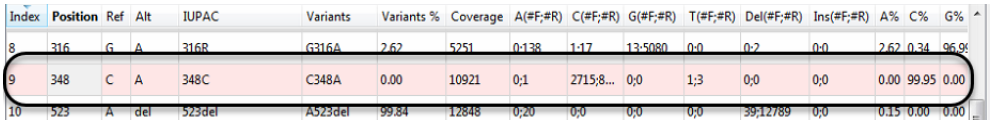
Figure 3-24: Comment Input dialog box



3. Optionally, enter the comment that is to be stored for the indicated position, and then click OK.

   The Comment Input dialog box closes. A new entry that is highlighted in red is displayed for the deleted variant in the Variants table.

Figure 3-25: New variant entry in Variants table



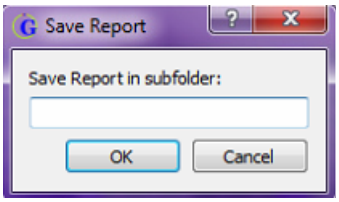
To save/load reports for an mtDNA project

After you generate an mtDNA project for the first time, the project settings (Filter Settings and Table Settings) are saved as Initial_Report at the Project Folder level. After you make any changes to the project settings and/or any manual edits to the project (for example, adding or deleting a variant), you can save these changes as another project report. Each report that you save is saved in a separate sub-folder under the primary project folder with a Change Log file that details the changes made to the project settings and project results. With a project opened in the Mito Alignment Viewer, you can load any report that has been saved for the project to dynamically update the information that is displayed for the project in the viewer.

1. Do any or all of the following as needed:
   - On the Mito Alignment Viewer main menu, click Filter Settings to open the Filter Settings dialog box, make any needed edits to the settings, and then click OK to save the edited settings. See "To specify the filter settings for an mtDNA project display" on page 75.
   - On the Mito Alignment Viewer main menu, click Table Settings to open the Table Settings dialog box, make any needed edits to the settings, and then click OK to save the edited settings. See "To specify the table settings for an mtDNA project display" on page 79.
   - In the Pile-Up pane, remove a read/undo a removed read, or add any variants as needed. See "Remove read/Undo Remove read" on page 71 or "To add a variant at a given position" on page 72.
   - In the Variants table, delete any variants as needed. See "To delete a variant at a given position" on page 73.
2. On the Mito Alignment Viewer main menu, click Save Report.

   The Save Report dialog box opens.

Figure 3-26: Save Report dialog box

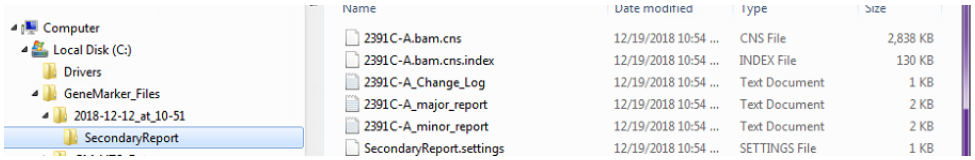


3. Enter a name for the subfolder in which the report is to be saved, and then click OK.

The subfolder name cannot contain any spaces or special characters. Underscores (_) are allowed.

The Save Report dialog box closes. The report is saved in a separate sub-folder under the primary project folder with a Change Log file that details the changes made to the project results. You can now load this report for a project to dynamically update the information that is displayed for the project in the viewer.

Figure 3-27: Saved Project report and folder structure



Mito Alignment Viewer Filter Settings

Two types of settings determine the information that is displayed for a project sample in the Mito Alignment Viewer: filter settings and table settings. When you open a sample for any mtDNA project in the Mito Alignment Viewer, the filter settings and table settings that were in place at the time that the project was saved determine the information that is displayed in the viewer, but you can always edit these settings as needed to better suit your working needs. Optionally, you can specify alternate alignment filters and open the Mito Alignment Viewer based on these filters. These alternate filters provide a way for you to edit the alignment without having to rerun the project.

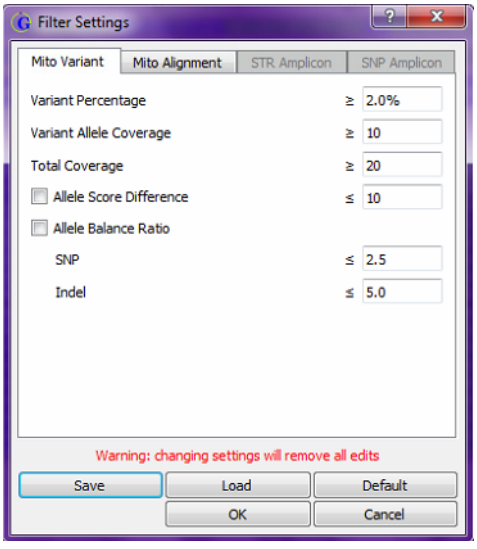
To specify the filter settings for an mtDNA project display

The Mito Variant tab and the Mito Alignment tab on the Filter Settings dialog box provide the options for adjusting the variant calling settings for an mtDNA project, which determine the variants that are displayed in the Variants table in the Mito Alignment Viewer. You can edit the default values for the filter settings, or you can load a project.settings file for a previously run project, and then apply the settings as-is, or optionally, you can edit any of the values, and then apply the settings.

1. On the Mito Alignment Viewer main menu, click Filter Settings.

   The Filter Settings dialog box opens. The Mito Variant tab is the open tab.

Figure 3-28: Filter Settings dialog box, Mito Variant tab



For Step2, note the following:

- *If you edit the values for any Filter Settings, then to return all settings to their default values (the factory values set by SoftGenetics) at the same time, click Default.*
- *If you edit any of the Filter Settings, then before you click OK to apply the settings, you can click Save to save the Filter Settings to a Settings(.settings) file.*

2. Do one of the following:
   - Edit the values for any of the Filter Settings values.
   - Click Load to open the Load Settings dialog box, and then browse to and select the appropriate project settings (.settings) file.

     The Filter Settings dialog box is populated with the values for the loaded project settings file. You can leave the values as-is, or you can edit any of the values.

Table 3-1: Filter Settings dialog box, Mito Variant tab

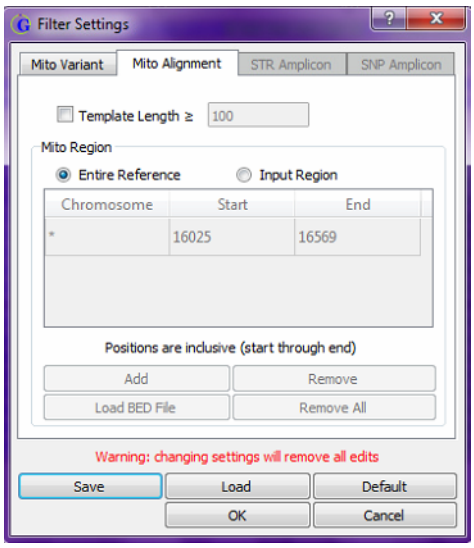
Option	Description
Variant Percentage	The percent of reads for which the variant was called at the indicated position.
Variant Allele Coverage	The number of reads that contained the variant at the indicated position.
Total Coverage	The total number of reads for the indicated position.
Allele Score Difference	A value that is calculated for each minor allele to determine if the allele is a true variant and therefore, is to be included in the analysis results, or if the allele is a false positive and therefore, is to be filtered from the results. See “Allele Score Difference” on page 78.
Allele Balance Ratio	A value that is calculated for each allele to determine if the allele is a true variant and therefore, is to be included in the analysis results, or if the allele is a false positive and therefore, is to be filtered from the results. See “Allele Balance Ratio” below.
SNP	Default value is ≤ 2.5.
Indel	Default value is ≤ 5.0.

3. If you need to change the reference region for the analysis, then go to Step 4; otherwise, click OK to close the Filter Settings dialog box and immediately apply the settings.

   The Mito Alignment Viewer display is dynamically updated based on the settings.

4. Open the Mito Alignment tab.

Figure 3-29: Filter Settings dialog box, Region Filters tab



5. Optionally, to filter by read length for paired data, select Template Length. You can leave the default value set to 100, or you can edit the value.

   The paired read data is kept if TLEN >= size or TLEN <= - size, where TLEN is Template Length.

If the data is non-paired, then read length is used for filtering.

6. Do one of the following:

   - To call variants in any region, select Entire Reference. (The default value.)
   - To limit variant calls to a single specific region, select Input Region, and then adjust the Start and End values.

By default, the Start and End values are set to define the region (16025 through 16569, inclusive) that is the most commonly specified for targeted sequencing.

   - To limit variant calls to multiple specific regions, select Input Region, select Load BED File, and then in the Load BED Settings dialog box, browse to and select the BED file that defines these multiple regions.

For detailed information about the format for a BED file, see "BED file" on page 97. After you load a BED file, you can use the Add and Remove options to add and remove regions for the file. If you add a region, you must specify the Start and End positions for it.

7. Click OK to close the Filter Settings dialog box and immediately apply the settings.

   The Mito Alignment Viewer display is dynamically updated based on the settings.

Allele Score Difference

In GeneMarker HTS, a quality score is calculated for every SNP or insertion allele as the median of the base quality score from the original data. Because base quality scores are not possible for a deletion, the quality scores for the bases that surround a deletion are used instead. The Allele Score Difference filter, which is based on the base quality score for an allele, is applied individually *for each direction* as follows:

- The total number of reads for the *minor* allele is recorded.
- The base quality score for each minor allele is subtracted from the base quality score for the major allele to calculate the Allele Score Difference.

If the Allele Score Difference is greater than its set threshold and the total number of reads for the minor allele is greater than or equal to five, then the variant is assumed to be a false positive and the minor allele is filtered from the alignment results; otherwise, if any of these criteria are not met, then the allele is assumed to be a true allele and is not filtered from the results.

Allele Balance Ratio

The Allele Balance Ratio is another filter that you can use to determine if an allele is a true allele or a false positive. The Allele Balance ratio is defined as the maximum value for *one* of the following:

- Allele Forward %/Total Forward % *or* Total Forward %/Allele Forward %
- Allele Reverse %/Total Reverse % *or* Total Reverse %/Allele Reverse %

with different default thresholds specified for SNPs (2.5) and indels (5.0).

Consider the following scenario, where:

- Two forward reads and 79 reverse reads with the allele (a SNP), for 81 total *allele* reads.
- 1758 total forward reads and 6055 total reverse reads, for 7812 total reads.

This results in the following:

Value	Calculation
Total Forward %	1758/7813 = 0.22501
Total Reverse %	6055/7813 = 0.77499
Allele Forward %	2/81 = 0.02469
Allele Reverse %	79/81 = 0.97531
Allele Forward %/Total Forward %	0.02469/0.22501 = 0.9113
Total Forward %/Allele Forward %	0.22501/0.02469 = **9.1134**
Allele Reverse %/Total Reverse %	0.97531/0.77499 = 1.2585

Value	Calculation
Total Reverse %/Allele Reverse %	0.77499/0.97531 = 0.7946
Because 9.1134 is > 2.5, this allele is filtered from the results.	

Although SoftGenetics has supplied the default values for what they have determined to be the most appropriate for the calculation of the Allele Balance Ratio, you can always edit one or both of these thresholds as needed. If you edit a value, you must try to maintain a balance of sensitivity versus specificity. For example, if you are concerned that variants are being called that are actually false positives and the variants also appear to be unbalanced, then you might tighten the restrictions on variant calling and decrease the appropriate threshold value. Conversely, if some expected variants are not being called, then you might loosen the restrictions on variant calling and increase the appropriate threshold value.

If an allele is determined to be one directional, then an Allele Balance Ratio is not calculated. Instead, the allele is not filtered from the results only if it is a major allele.

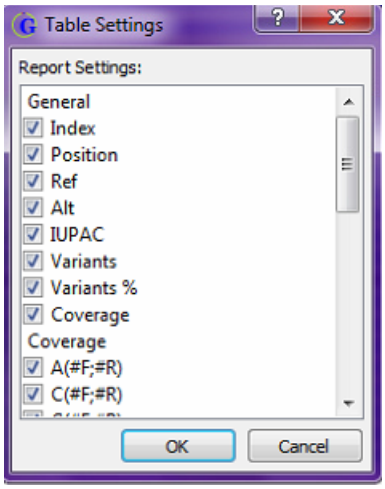
To specify the table settings for an mtDNA project display

The table settings determine the data columns (information) that are displayed in the Variants table for each called allele in a project. By default, all available data columns are displayed for an allele in the Variants table, but you can always specify which columns are to be displayed, and which are not. The information that is displayed for an allele is categorized as one of the following: General, Coverage, Coverage Percent, Score, or Modification.

1. On the Mito Alignment Viewer main menu, click Table Settings.

   The Table Settings dialog box opens. By default, all columns are selected for inclusion in the Variants table.

Figure 3-30: Table Settings dialog box



2. Leave the columns that are to be included in the Variants table selected; otherwise, to hide a column in the Variants table, clear its selection.

Column	Description
General	
Index	Unique system-generated identifier for the variant.
Position	The nucleotide position in the chromosome where the variant occurs.
Ref	The reference nucleotide at the variant position.
Alt	The alternate (variant) nucleotide at the variant position.
IUPAC	The IUPAC nucleotide code for the allele. If multiple alleles are called at a position, then the appropriate IUPAC ambiguity code is reported. **Note:** Mixtures of a deletion and other alleles use a lowercase letter.
Variants	The standard variant call.
Variant %	The variant frequency at the indicated position.
Coverage	The total number of reads (total coverage) at the variant position.
Coverage	
A #(F,R), C #(F,R), G #(F,R), T#(F,R)	The number of reads that show the indicated nucleotide at the variant location in the forward direction and the number of reads that show the indicated nucleotide at the variant location in the reverse direction.
Del #(F,R)	The number of reads that show a deletion at the variant location in the forward direction and the number of reads that show a deletion at the variant location in the reverse direction.
Ins #(F,R)	The number of reads that show an insertion at the variant location in the forward direction and the number of reads that show an insertion in the reverse direction at the variant location.
Coverage Percent	
A(%), C(%), G(%), T(%)	The percentage of reads that show the indicated nucleotide at the variant location.
Del(%)	The percentage of reads that show a deletion at the variant location.
Ins(%)	The percentage of reads that show an insertion at the variant location.
Score	
A Score, C Score, G Score, T Score	The median nucleotide call quality for the indicated allele (SNP) in both the forward and reverse directions.
Modifications	
Category	Indicates the variant status for the project in the Variants table - Called, Removed, Added,
Comment	Any comment that was entered for the added or deleted variant.

3. Click OK to close the Table Settings dialog box and immediately update the display of the Variants table.

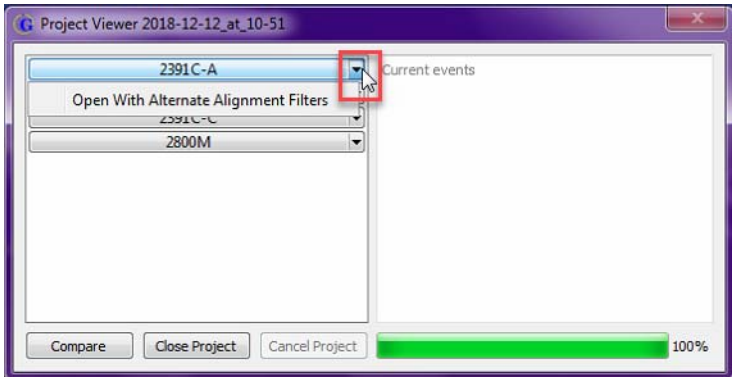
To specify alternate alignment filters

When you are setting up a new project, you must specify alignment filters as part of the project settings. Before you open an mtDNA project sample in the Mito Alignment Viewer, you can specify different values for some of these alignment filters and open the project in the Mito Alignment Viewer based on these edited filters. These alternate filters provide a way for you to edit the alignment without having to rerun the project.

1. For the appropriate sample on the Project Viewer dialog box, click the Dropdown arrow that is displayed at the far right of the Sample button.

   The option Open with Alternate Alignment Filters is displayed.

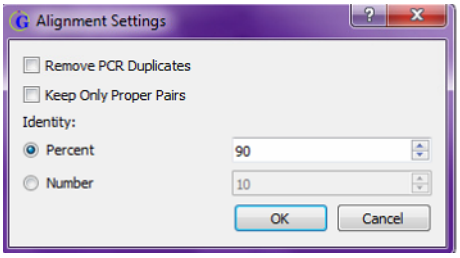
Figure 3-31: Accessing alternate alignment filters



2. Click Open with Alternate Alignment Filters.

   The Alignment Settings dialog box opens.

Figure 3-32: Alignment Settings dialog box



3. Specify the alternate values for the alignment filters.

Table 3-2: Alignment Settings

Option	Description
Remove PCR Duplicates	On or Off. Reads are identified as PCR duplicates based on their start positions. If you are using paired-end data, then each pair of reads is considered together. Select to remove PCR duplicates, or leave blank and the PCR duplicates are not removed. **Note:** If you are using targeted sequencing, then you typically do not select this value. If you are using whole genome sequencing, then you typically do select this value.

Table 3-2: Alignment Settings (Continued)

Option	Description
Keep Only Proper Pairs	Ignore any reads that were not marked as being part of a proper pair, where a proper pair is defined as the following: • Both reads from the pair must be aligned and the alignment must be on opposite strands. • The 5' end of the reverse read must be to the right of the 5' end of the forward read.
Identify	Reads that are less similar to the reference than indicated are not aligned. Select the way to identify the number of reads: • Percent ≥: Default value is 90. If the percentage of reads for a sample that do not match the reference are greater than or equal to the indicated value, then reads are not aligned. • Number ≤: Default value is 10. If the total number of reads for a sample that align to reference are less than or equal to the indicated value, then the reads are not aligned.

4. Click OK.

   The Alignment Settings dialog box and the Project Viewer dialog box close. The Mito Alignment Viewer opens for the sample. The values for these alternate filters determine the information that is displayed for the project sample in the Mito Alignment Viewer:

Mito Alignment Viewer Reports

Three reports are available for generation from the Mito Alignment Viewer:

- Coverage report - The Coverage report provides a high-level overview of the fraction of bases (the entire reference range, or the range of bases that was specified for the Filter Settings for the project) that have some level of coverage that is greater than or equal to the Total Coverage threshold that was specified for the project. The report displays the coverage distribution of sample reads without directional information (a solid black line) and individually for both the forward and reverse directions (blue and red lines, respectively). The report also details low coverage regions in a separate table below a Coverage Curve graph. The report is useful for identifying regions that were not adequately sequenced because of low coverage. See Figure 3-33 below.

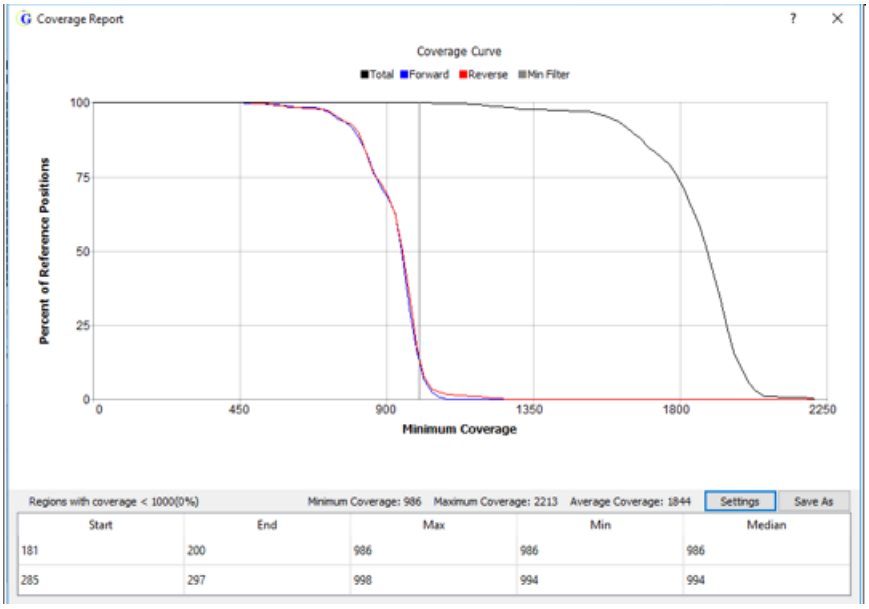
  Optionally, you can do one or both of the following for the Coverage report:

- Template Length Report - The Template Length report shows a length distribution based on TLEN values in the BAM file for paired-end data. TLEN is a positive value in the BAM file for forward read data and the same value, but stored as a negative value, for the reverse read data. See Figure 3-35 on page 84.

- Read Length report - The Read Length report consists of two graphs. The Read Length graph shows the length of the sequences in the BAM file. The Aligned Read Length shows the length of the sequences, ignoring any soft-clipping in the CIGAR string. See Figure 3-36 on page 85.

All three reports are available from the Reports option on the Mito Alignment Viewer main menu. After a report is generated, it is displayed in a separate window.

Figure 3-33: Coverage report



Optionally, you can do one or both of the following for the Coverage report:

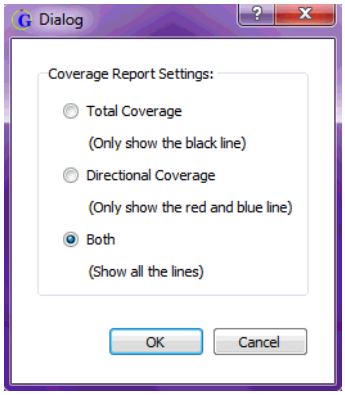
Option	Description
Change the Coverage Curve Report Display	1. Click Settings. The Coverage Curve Report Settings dialog box opens. The dialog box shows the options for changing the display of the report. By default, Both is selected, which results in all lines (non-directional (black), forward (blue), and red (reverse)) being displayed on the report. *Figure 3-34: Coverage Curve Report Settings dialog box*  2. Do one of the following, and then click OK to close the dialog box and immediately apply the changes. • To display only the total coverage for all reads (non-directional black line), select Total Coverage. • To display only the forward and reverse reads (directional blue and red lines), click Directional Coverage.
Save the Low Coverage Regions table as a text file	Click Save As. You must provide a name for the report. By default, the Low Coverage Regions report is saved in the top level project folder, but you can always select a different location.

Figure 3-35: Template Length report

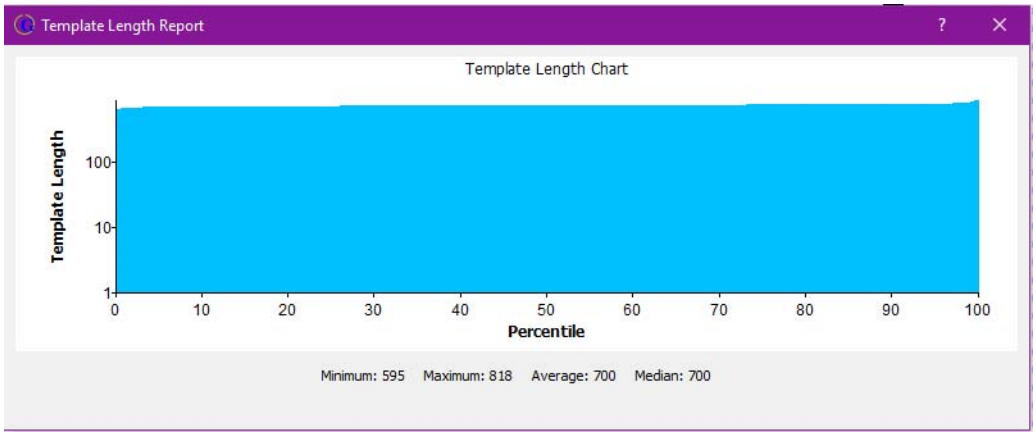
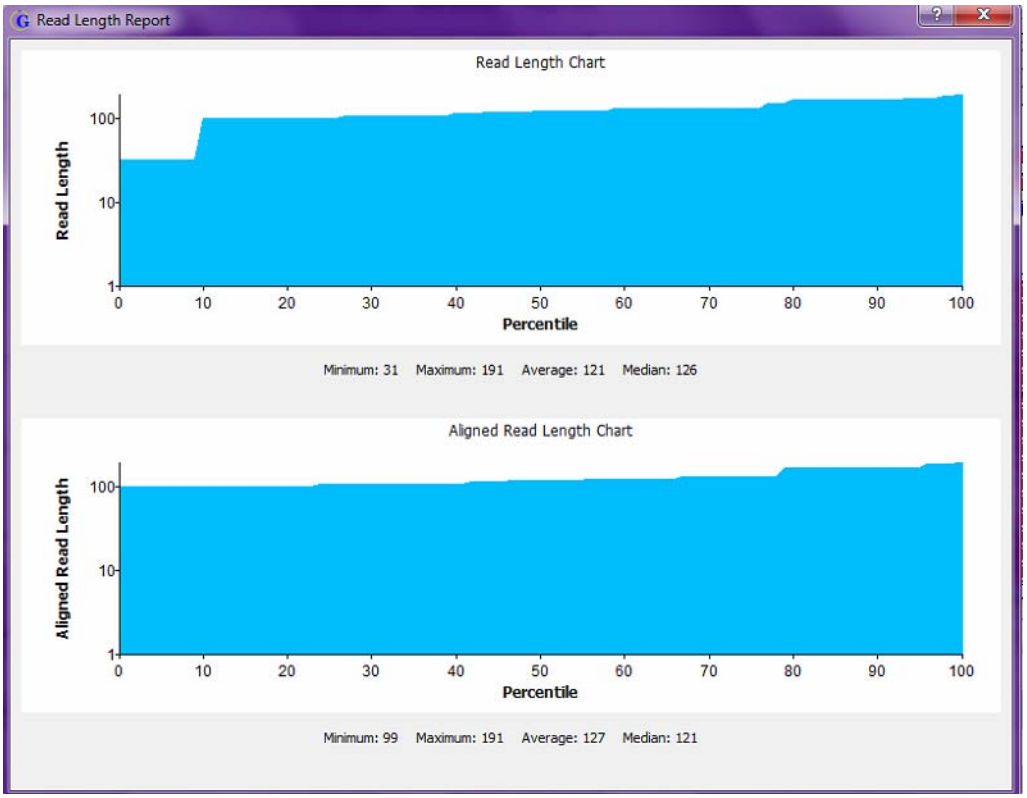


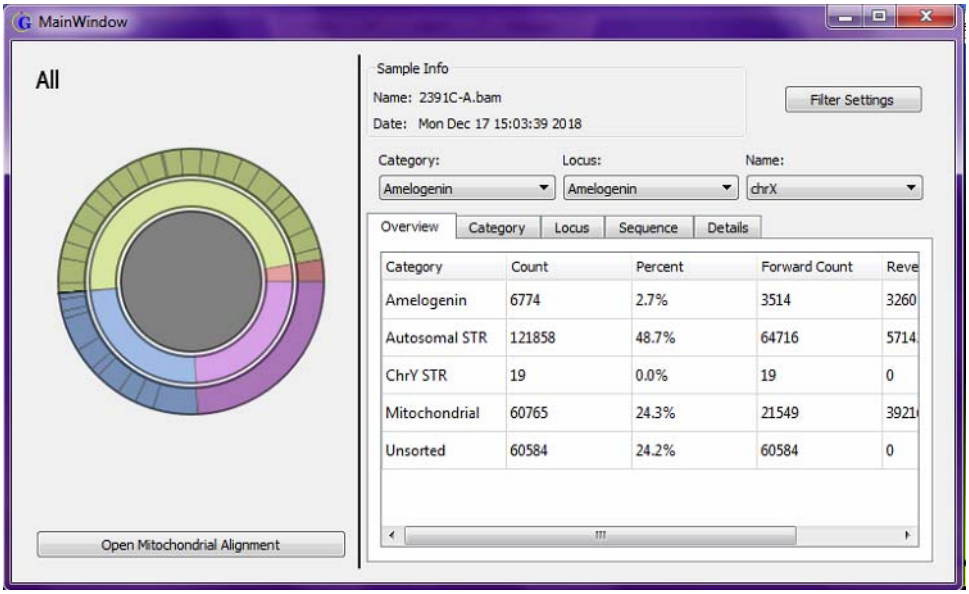
Figure 3-36: Read Length report



STR Analysis Viewer Layout

The STR Analysis Viewer is an interactive view-only tool that is available in the GeneMarker HTS application for viewing the results of a simultaneous analysis of autosomal- and Y-STRs along with mitochondrial data for your project samples. You use the STR Analysis Viewer to view a summary of the autosomal- and Y-STRs analysis results for your project. You can then open the Mito Alignment Viewer directly from the STR Analysis Viewer and view the mtDNA alignment results for the project. When you open a sample file for any project in the STR Analysis Viewer, the Filter Settings that were in place at the time that the project was saved determine the alleles that were called for the project, but you can always edit these settings as needed to better suit your working needs. See "STR Analysis Viewer Filter Settings" on page 90.

Figure 3-37: STR Analysis Viewer



The STR Analysis Viewer has two distinct but interactive panes: the Sunburst Plot (left) pane and the Sample (right) pane. Both panes provide a way to navigate project results with the Sunburst Plot pane providing an image of the project results and the Sample pane providing a tabulated display of the project results. The panes displays are linked.The information that is displayed in the Sample pane determines the part of the Sunburst Plot that is highlighted in gray. Conversely, clicking anywhere on the Sunburst plot pane updates the tabulated results in the Sample pane accordingly. See:

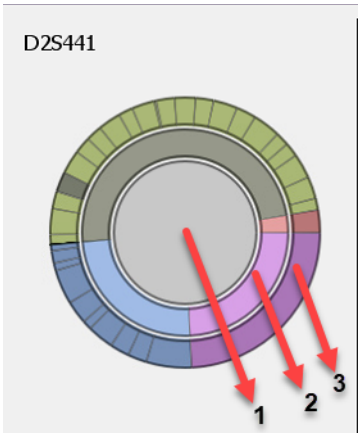
- "Sunburst Plot pane" on page 87.
- "Sample pane" on page 88.

Sunburst Plot pane

The Sunburst Plot consists of a series of three concentric circles, with the circles visually representing the following information:

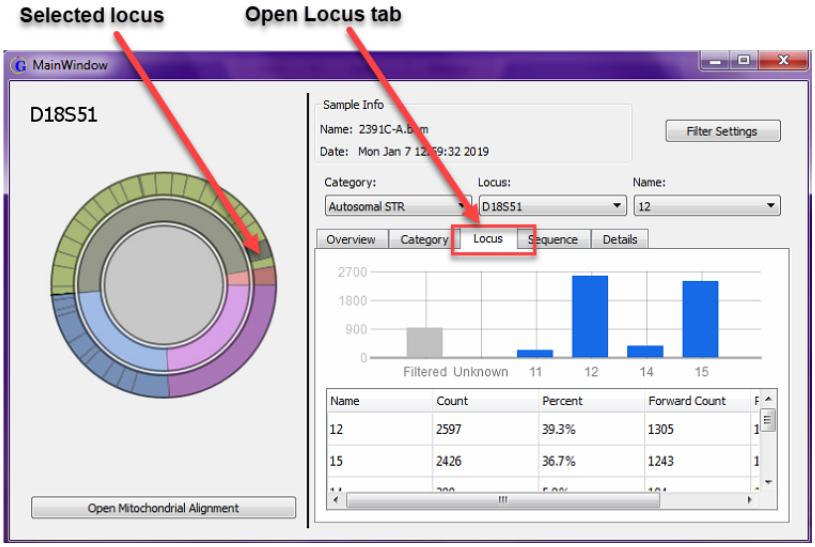
- Innermost circle (1)- The project overview.
- Middle circle (2) - The DNA category.
- Outer circle (3) - All the loci for each DNA category.

Figure 3-38: Sunburst Plot



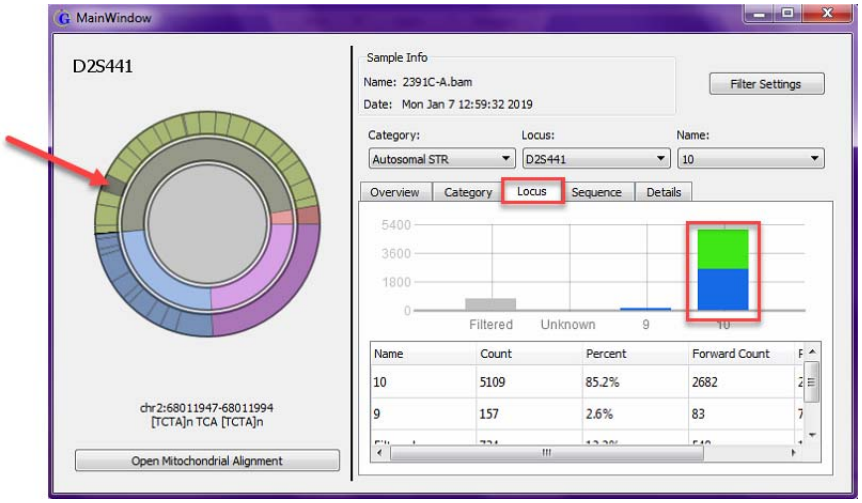
When you *click* (do not double click) any of the circles anywhere, the appropriate parts of the plot are highlighted in gray, and the Sample Analysis Results Table display is updated. For example, you can click a locus for the Autosomal STR category, and the Sample Analysis Results Table display is updated accordingly, with the Locus tab being the open tab.

Figure 3-39: Interactive and linked displays for the STR Analysis Viewer



The interactions between the Sunburst Plot and the Sample pane provide an option for viewing distinct sequence variations in iso-alleles. See Figure 3-40 below.

Figure 3-40: Viewing sequence variations in iso-alleles in the STR Analysis Viewer



Sample pane

The Sample pane provides both summary and detailed analysis results information for the selected sample. The sample name and the date that the sample was analyzed (the sample information) are displayed at the top of the Sample pane. The Sample Analysis Results table is displayed below the sample information and details the analysis results for the sample based on the selected Category, Locus, and (Chromosome) Name. The information is displayed across five tabs: Overview, Category, Locus, Sequence, and Details.

Table 3-1: Sample Analysis Results Table panes

Tab	Description
Overview	When the STR Analysis Viewer first opens for a sample, the Overview tab is the open tab. The tab displays a summary of the analysis results for the project. • Category - The DNA that was amplified for the sample. If the DNA could not be categorized based on the panel information, then the DNA is placed in the Unsorted category. • Percent - Total Number of Category Reads/Sum Total of Category Reads • Forward Count - Total number of forward reads for the sample. • Reverse Count - Total number of reverse reads for the sample.
Category	The loci that were analyzed for the DNA category.
Locus	The detailed information for the locus that is selected on the Locus dropdown list.
Sequence	The sequence for the selected locus.

Table 3-1: Sample Analysis Results Table panes

Tab	Description
Details	Details information for either the selected locus or selected sequence.

You can select a different Category, Locus, and/or Name to update the table display accordingly, or optionally, the entire table is interactive for navigating the project results. For example, you can double-click a specific locus on the Category tab not only to update all the tabs on the table but also to open and view the Details tab for the locus, or you can double-click an entry on the Sequence tab to open and view the Details tab for the sequence. Whenever the Sample Analysis Results Table display is updated, the location of the gray highlighting on the Sunburst Plot is also updated.

STR Analysis Viewer Filter Settings

Filter settings determine the information that is displayed for a project sample in the STR Analysis Viewer. When you open a sample for any STR analysis project in the STR Analysis Viewer, the filter settings that were in place at the time that the project was saved determine the alleles that were called for the project, but you can always edit these settings as needed to better suit your working needs.

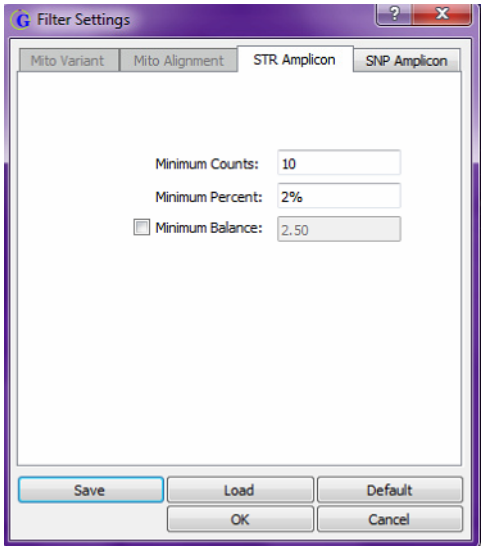
To specify the filter settings for an STR analysis project display

The STR Amplicon tab and the SNP Amplicon tab on the Filter Settings dialog box provide the options for adjusting the minimum count and minimum percent that are required to call an allele in an STR analysis project. You can edit the default values for the filter settings, or you can load previously saved filter settings, and then apply the settings as-is, or optionally, you can edit any of the values as needed, and then apply the settings.

1. On the STR Analysis Viewer, click Filter Settings.

   The Filter Settings dialog box opens. The STR Amplicon tab is the open tab.

Figure 3-41: Filter Settings dialog box, STR Amplicon tab



For Step2, note the following:

- *If you edit the values for any Filter Settings, then to return all settings to their default values (the factory values set by SoftGenetics) at the same time, click Default.*
- *If you edit any of the Filter Settings, then before you click OK to apply the settings, you can click Save to save the Filter Settings to a Settings(.settings) file.*

2. Do one of the following:
   - Edit the values for any of the Filter Settings values as needed.
   - Click Load to open the Load Settings dialog box, and then browse to and select the appropriate project settings (.settings) file.

     The Filter Settings dialog box is populated with the values for the loaded project settings file. You can leave the values as-is, or you can edit any of the values.

Table 3-2: Filter Settings dialog box, STR Amplicon tab

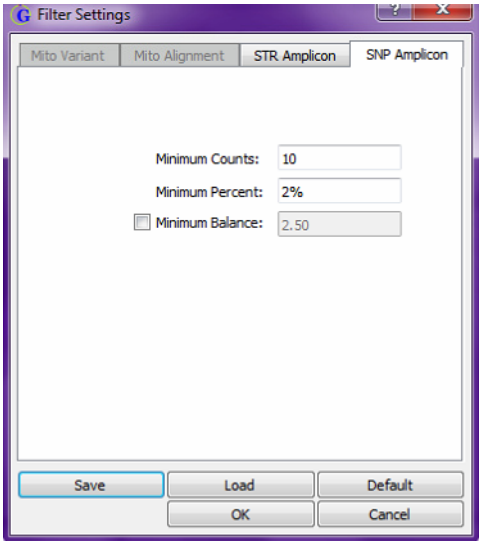
Option	Description
Minimum Counts	Default is 10.
Minimum Percent	Default value is 2%.
Minimum Balance	Used primarily as a research tool. The calculations are as follows: 1. Total Balance Ratio: min (total forward, total reverse)/(total forward + total reverse) 2. STR Group Balance Ratio: min (group forward, group reverse)/(group forward + group reverse) 3. Final Ratio Balance: (STR Group Balance Ratio)/(Total Balance Ratio) If the value of the Final Ratio Balance for an STR group is greater than the value that is specified for the Minimum Balance (default value of 2.50, but you can edit this), then the STR group is filtered from the results.

3. If you need to adjust the default values for SNP Amplicon filters, then go to Step 4; otherwise, click OK to close the Filter Settings dialog box and immediately apply the settings.

   The STR Analysis Viewer display is dynamically updated based on the settings.

4. Open the SNP Amplicon tab.

Figure 3-42: Filter Settings dialog box, SNP Amplicon tab



5. Edit the values for any of the Filter Settings values as needed.

Table 3-1: Filter Settings dialog box, SNP Amplicon tab

Option	Description
Minimum Counts	Default is 10.
Minimum Percent	Default value is 2%.
Minimum Balance	Similar to Minimum Balance on the STR Amplicon tab. You can leave the value set to the default value of 2.50, or edit the value as appropriate.

6. Click OK to close the Filter Settings dialog box and immediately apply the settings.

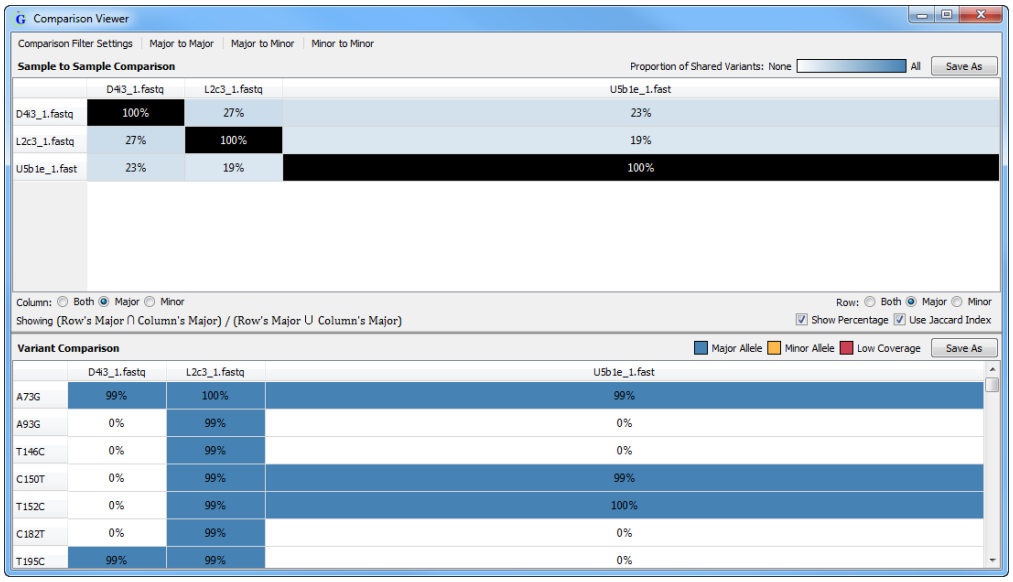
   The STR Analysis Viewer display is dynamically updated based on the settings.

Comparison Viewer

You can compare the results for all the samples in the project in a single Comparison Viewer window. To do so, open the appropriate project, and then in the Project Viewer dialog box, click Compare to open the Comparison Viewer window.

For detailed information about opening an existing project, see "To work with an existing project" on page 53.

Figure 3-43: Comparison Viewer



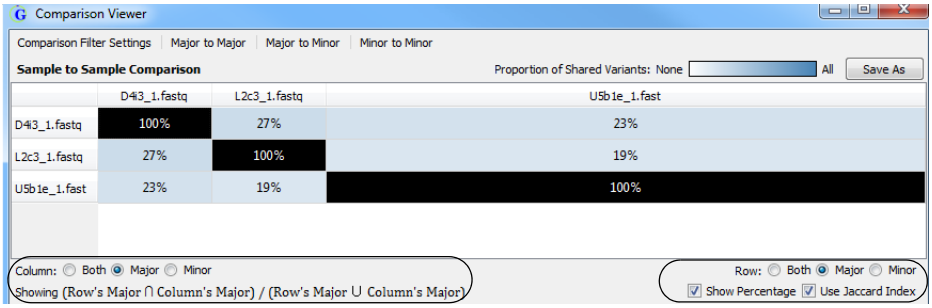
The Comparison Viewer is divided into two panes: the Sample to Sample Comparison (top) pane and the Variant Comparison (bottom) pane. Each pane displays unique information about a project. You can use the Save As option to save the information that is displayed in both panes as text (.txt) file with a filename of your choosing. By default, the text file is saved in the same directory as the project, but you can always select a different location.

Sample to Sample Comparison pane

The Sample to Sample Comparison pane for the Comparison Viewer is a correlation table, that, by default, displays the percentage of major alleles that are shared among all the project samples, with the Jaccard Index measurement used to calculate the comparison. Diagonal cells in the correlation table compare a sample to itself and by default, are highlighted in black. All other cells in the correlation table compare one sample to another, and by default, are highlighted in a blue color gradient from 0 to 100%, where 0% cells are blank/white. See Figure 3-44 on page 94.

You can change the default highlighting colors that are used in the Sample to Sample Comparison pane. See "Specifying GeneMarker HTS Application Settings" on page 27.

Figure 3-44: Comparison Viewer, default display



Options are available at the bottom of the Sample to Sample Comparison pane for the following:

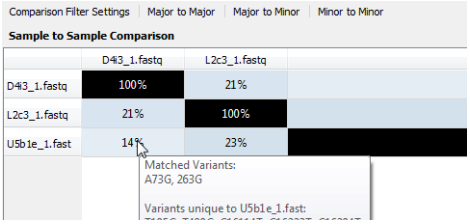
- Changing the variant types that are used for calculating the correlation (Major, Minor, or Both). You can select the same or different variant types for the column and the row.

You can also use the options that are available on the Comparison Viewer main menu (Major to Major, Major to Minor, and Minor to Minor) to quickly reset the variant types that are used for the calculation of the correlation.

- Changing the calculation that is used for the correlation. By default, the Jaccard Index measurement is used for the calculation, which results in the variant values in either the row or the column being used in the divisor. If you clear this option, then a Simple measurement is used for the calculation, which results in only the variant value in the column being used in the divisor.
- Showing the percentage of shared variants, or the absolute number of shared variants. By default, Show Percentage is selected, but you can clear this option to show absolute number.

As you change any of the options for the correlation calculation, both the correlation table display and the correlation calculation are updated accordingly. You can hold your mouse pointer over any entry in the correlation table to open a tooltip that displays the percentage or absolute number of matched variants, and if applicable, all the variants that are unique to each appropriate sample.

Figure 3-45: Correlation Table tooltip



Variant Comparison pane

The Variant Comparison pane lists the all the variants that were found in at least one sample for the project, where each column is a different sample and the variant percentage for the sample is displayed. The display provides a quick way of visually determining which variants were shared among samples and which were not. Variants that can be classified as a major allele or as a minor allele are highlighted, by default, in blue or gold, respectively. If a variant was determined to be a Low Coverage variant, (the Total Coverage for the variant falls below the Total Coverage threshold that was set for the project), then by default, the variant is highlighted in red.

You can change the default highlighting colors that are used in the Variant Comparison pane. See "Specifying GeneMarker HTS Application Settings" on page 27.

Figure 3-46: Variant Comparison pane

The specific variants that are initially displayed in the Variant Comparison pane are determined by the Filter Settings, including the Total Coverage threshold, that were in place at the time that the project was saved. To dynamically change the display, on the Comparison Viewer main menu, click Comparison Filter Settings to open the Filter Settings dialog box, and make any needed modifications. See "To specify the filter settings for an mtDNA project display" on page 75.

Chapter 3
GeneMarker HTS Viewers

Glossary

B

BED file

Also known as Region of Interest (*.bed file). A BED file is a tab-delimited text file. You can upload a BED file only if the reference sequence contains chromosome information. Each row in the file contains a region of the reference that is to be used for the analysis, and at a minimum, the file must contain the following information:

- Field #1 - Chromosome number or name for the region (for example, chrM)
- Field #2 - Chromosome start position (for example, 300)
- Field #3 - Chromosome end position (for example, 305)
- Field #4 - Optional description column (for example, Region 102)

The positions are 0-based and open-ended. For example, a start of 10 and end of 15 would include reference positions 11, 12, 13, 14, and 15.

C

Consensus alignment

An optional alignment step. Because each read is considered on its own, indels that are near the end of reads are sometimes not correctly aligned. If this alignment option is selected, then an extra step is added after the initial alignment. In this extra step, indel positions are re-examined to possibly improve the alignment of these reads based on the alignment of other reads that have the indel.

M

Motif alignment

An optional alignment step. When equivalent sequences can be aligned in multiple ways (motifs), an analyst might not prefer the alignment method that the software selects. This step ensures that the selected motif is the motif that is defined by the motif file.

P

PCR duplicates

PCR duplicates are a set of pairs (paired-end data) or reads (non-paired-end data) that have been generated from the same original fragment. The “Remove PCR Duplicates” optional

alignment step attempts to identify these reads and ignores all but one pair (or read) in each set.

Personal Health Information (PHI)

Personal health information can be inferred from some regions of whole mtDNA sequence. To maintain privacy during whole mtDNA analysis, a BED file can be used to specify the positions that must be hidden from the analyst.

Proper Pairs

A pair of aligned reads is referred to as a “proper pair” if the reads are aligned on opposite strands in the correct orientation. Non-proper pairs can result from misalignment, sequencing error, or off-target amplification.

Index

A

About button on the GeneMarker HTS main window 26

Allele Balance Ratio

defined 78

Allele Score Difference

defined 78

amplicon regions

specifying for a project 45

Amplicons pane in the Mito Aligment Viewer 65

application settings

specifying for GeneMarker HTS 27

Application Settings icon on the GeneMarker HTS main window ... 25

Assumptions for the manual 10

C

Comparison Viewer 93

Sample to Sample Comparison pane 93

Variant Comparison pane 95

Consensus Sequence pane in the Mito Aligment Viewer 65

context menu for the Mito Alignment Viewer 70

Conventions used in the manual ... 9

Coverage report 83

F

filter settings

specifying for a project in the Mito Alignment Viewer 75

specifying for a project in the STR Analysis Viewer 90

G

GeneMarker HTS

main window 25

About button 26

Application Settings icon 25

Online Help button 26

Project options 25

title bar 25

registering 18

registering offline 20

specifying the application settings for 27

starting 24

Global Viewer pane

Mito Alignment Viewer 63

L

layout of the Mito Alignment Viewer 61

Global Viewer pane 63

main menu 62

Pane displays 64

Amplicons pane 65

Consensus Sequence pane 65

Pile-up pane 65

Reference pane 64

status bar 68

title bar 62

Variants table 67

layout of the STR Analysis Viewer 86

Sample pane 88

Sunburst Plot pane 87

M

main menu

Mito Alignment Viewer 62

main window for GeneMarker HTS 25

Mito Alignment Viewer

adding a variant in the Pile-Up pane for 70

context menu for 70

deleting a variant for a project in the Variants table for 70

layout of 61

Global Viewer pane 63

main menu 62

Pane displays 64

Amplicons pane 65

Consensus Sequence pane 65

Pile-up pane 65

Reference pane 64

status bar 68

title bar 62

Variants table 67

navigation of 69

saving/loading reports for a project in 73

specifying the filter settings for a project in 75

specifying the table settings for a project in 79

Mito Alignment Viewer reports

Coverage report 83

Read Length report 83

Template Length report 83

motif
creating or modifying 42
defined 41
file format for 41
mtDNA alignment project reports
Coverage report 83
Read Length report 83
Template Length report 83

N

navigating the Mito Alignment Viewer 69
New Project dialog box
overview of 37

O

Online Help button on the GeneMarker HTS main window ... 26
Organization of the manual 10
output files for a project 57

P

Pane displays
Mito Alignment Viewer 64
Amplicons pane 65
Consensus Sequence pane 65
Pile-up pane 65
Reference pane 64
personal health information
defined 46
masking for a subject 46
PHI
see personal health information
Pile-up pane in the Mito Aligment Viewer 65
project
adding new in GeneMarker HTS 49
output files for 57
saving/loading reports for in the Mito Alignment Viewer 73
specifying which amplicon-sequenced regions were used for 45
working with an existing one in GeneMarker HTS 53
Project options on the GeneMarker HTS main window 25

R

Read Length report 83
Reference pane in the Mito Aligment Viewer 64
reports
saving/loading for a project in the Mito Alignment Viewer 73

S

Sample pane in the STR Analysis Viewer 88
Sample to Sample Comparison pane in the Comparison Viewer ... 93
soft-clipping
defined 44
SoftGenetics License Server
installing 16
Special information about the manual 9
status bar
Mito Alignment Viewer 68
STR Analysis Viewer
layout of 86
Sample pane 88
Sunburst Plot pane 87
specifying the filter settings for a project in 90
Sunburst Plot pane in the STR Analysis Viewer 87

T

table settings
specifying for a project in the Mito Alignment Viewer 79
Template Length report 83
title bar
GeneMarker HTS main window 25
Mito Alignment Viewer 62

U

Using the manual 9

V

variant
adding in the Pile-Up pane for the Mito Alignment Viewer 70
deleting from the Variants table in the Mito Alignment Viewer 70
Variant Comparison pane in the Comparison Viewer 95
Variants table in the Mito Alignment Viewer 67