illumina®

ForenSeq™ Universal Analysis Software Guide

FOR RESEARCH, FORENSIC, OR PATERNITY USE ONLY

ILLUMINA PROPRIETARY
Document # 15053876 v01
August 2016

Revision History

Document #	Date	Description of Change
Document # 15053876 v01	August 2016	Updated screen shots throughout the guide to reflect changes in the user interface for ForenSeq Universal Analysis Software v1.2. In the Supporting Information chapter, added directions for setting up an external data repository. Added a section on troubleshooting analysis errors. In the Run Metrics Tab section, revised the description of light orange in the sample representation tab graph to indicate that loci with low coverage might not have enough signal for all alleles to be distinguished from noise. Updated descriptions of icons for Allele Count and Interpretation Threshold. Updated descriptions of what causes interpretation threshold, allele count, and mixture detection quality indicators in STR genotype calling. Updated descriptions of what causes the interpretation threshold quality indicator in SNP genotype calling. Noted that the low coverage indicator can be seen in results from version 1.1 or earlier of the ForenSeq Universal Analysis Software. Updated the following thresholds for aSNPs in the Supporting Information section: • For rs1572018, changed the analytical threshold from > 5 to > 1.5, and the interpretation threshold from > 15 to > 4.5. • For rs735480, changed the analytical threshold from > 3.3 to > 1.5, and the interpretation threshold from > 10 to > 4.5. • For rs3916235, changed the interpretation threshold from > 10 to > 4.5. In the Analysis Procedures section for STR and SNP genotype calling, added quality considerations for the analytical threshold, and when signal is not detected at the locus. Added the icons Analytical Threshold and Not Detected. In the Reports section, added information on the new phenotype estimation report that is generated from the Project page. In the Reports section, added descriptions of the new coverage worksheets in the Project Level Report section. Updated descriptions of icons for Allele Count and Interpretation Threshold. In the section STR Genotype Calling, added a table with additional details regarding how gender is determined.

Document #	Date	Description of Change
Part # 15053876 Rev. C	June 2015	Updated software descriptions to ForenSeq Universal Analysis Software v1.1. • Updated screen shots throughout the guide to reflect changes in the user interface. • Added a section for modifying and updating the gender of a sample. • Added the section Define Loci Used in Population Studies for selecting loci for population calculations statistics, and noted that the software queries both population groups and loci defined by the administrator at the time of calculation. • Added information on selecting the 2p and 2p-p^2 population statistics calculation rule in the Defining the Statistics Calculation Method section. • Added instructions on applying the 2p rule to homozygous loci. • Noted in the STR Genotype Calling section that negative amplification control samples and reagent blank samples are not assessed for gender and by default are "inconclusive". • Added details on the system dashboard for the run progress bar, quick access to run metrics, analysis icon, and About information. • Added information on generating sample genotype reports from the Project page for multiple samples. • Updated the description of the Analysis page display to include the analysis version. • .Added information on how to indicate review status in the Viewing Run Quality Information section. • Added information on opening the Sample Details page from sample representation, positive amplification control, and negative amplification control tabs. • Noted that isometric alleles are now highlighted with a gray box. • Updated the name of the Locus Panel page to the Define Content page. • Updated the name of the Locus Thresholds page to the Analysis Values page.

Document #	Date	Description of Change
Part # 15053876 Rev. B	February 2015	Updated screenshots throughout the guide to reflect changes in the user interface. Added a System Settings chapter that details information on notifications, user management, changing locus thresholds, defining locus content, and population group settings. Added a section on Hardy-Weinberg expectations and linkage equilibrium. Removed the Analysis Settings section and moved threshold settings information to the System Settings chapter of the guide. Added a Creating a New Analysis section. Updated intensity and length graphs information to reflect a change in how the software splits and combines the display of STRs and SNPs on the graph. Updated information on the typed alleles indicator, which changed from true and false to a toggle switch, and can also be toggled on and off by clicking on the bar of the bar chart representing an allele for an STR, or pie chart for a SNP allele. Updated random match probability information to distinguish between the generation of population statistics and the random match probability calculation method. Updated report generation information to include Sample Summary and Sample Detail reports. Divided the Additional Analyses chapter to create a chapter for population statistics, sample comparison, and phenotype and bio-geographical ancestry estimation. Added a troubleshooting section for population group file upload. Changed the Settings section to Locus Level Settings. Updated the explanation of loci typed in the General Locus Information section to include selected locus content. Updated the population statistics calculations information to include population group selection.
Part # 15053876 Rev. A	December 2014	Initial Release

Table of Contents

Revision History iii
Table of Contents vii
Chapter 1 Getting Started 1
Introduction 2
Software Functions 3
ForenSeq Universal Analysis Software Concepts 4
Viewing the ForenSeq Universal Analysis Software 6
Logging in to the Software 7
Viewing the System Dashboard 8
Example Workflow 9
Creating a New Run 10
Viewing Run Details 13
Changing the Name or Description of a Run 15
Creating a New Run Version 16
Chapter 2 Sample and Run Results 17
Introduction 18
Viewing the Project Page 19
Viewing Run Quality Information 20
Creating a New Analysis 31
Viewing Sample Details Page 33
Chapter 3 Population Statistics 43
Introduction 44
Population Statistics Calculations 45
Generating Population Statistics 46
Population Statistics Results 46
Chapter 4 Sample Comparison 49
Introduction 50
Generating a Sample Comparison 51
Sample Comparison Results 52
Generating Sample Compare Population Statistics 54
Comparisons Tab 55
Chapter 5 Phenotype and Biogeographical Ancestry Estimation 57
Introduction 58
Phenotype and Biogeographical Ancestry Estimation 58
Generating pSNP and aSNP Information 59
Phenotype Estimation SNP Sample Details Table 60
Generating Phenotype Estimations 62
Phenotypes Tab 64
Chapter 6 Reports 65
Introduction 66
Project Level Report 67
Sample Genotype Report 70
Phenotype Estimation Report 75
Chapter 7 System Settings 79
Introduction 80
Notifications 82
User Management 83
Changing Locus Thresholds 85
Defining Locus Content for Analysis 89
Define Loci for Population Studies 91
Population Group Settings 92

Chapter 8 Analysis Metrics and Procedures 97
Introduction 98
Analysis Metrics 99
Analysis Procedures 100
Chapter 9 Troubleshooting 105
Introduction 106
Troubleshooting ForenSeq Runs 107
Troubleshooting Analysis Errors 112
Troubleshooting Population Group File Uploads 114
Chapter A Supporting Information 115
Human Sequencing Control Loci 116
Autosomal, Y, and X STR Filters and Thresholds 117
aSNP Thresholds 120
iSNP Thresholds 122
pSNP Thresholds 126
Set Up an External Data Repository 127
Index 133
Technical Assistance 135

Getting Started

Chapter 1

Introduction 2
Software Functions 3
ForenSeq Universal Analysis Software Concepts 4
Viewing the ForenSeq Universal Analysis Software 6
Logging in to the Software 7
Viewing the System Dashboard 8
Example Workflow 9
Creating a New Run 10
Viewing Run Details 13
Changing the Name or Description of a Run 15
Creating a New Run Version 16

Introduction

ForenSeq™ Universal Analysis Software is a complete DNA-to-data forensic software solution at the center of the MiSeq FGx™ Forensic Genomics System. The software performs analysis of sequenced genetic sample information for human identification and works in combination with the ForenSeq DNA Signature Prep Kit and the MiSeq FGx Instrument.

The ForenSeq Universal Analysis Software is on a standalone, dedicated server with a user interface specific for forensic genomics that enables run setup, sample management, analysis, and report generation. ForenSeq Universal Analysis Software provides population statistics and automated sample comparison, as well as an optional feature for estimating biogeographical ancestry, hair color, and eye color.

NOTE

Perform internal validation studies when developing protocols and interpretation guidelines for casework and DNA databasing using the ForenSeq Universal Analysis Software. It is possible that some features and recommendations detailed in the guide are not part of your analysis requirements, analysis routine, or necessary for high-quality analyses.

Software Functions

The following is a workflow outline of functions administered by the software:

1. Create an account with a user name and password for access to the ForenSeq Universal Analysis Software and the MiSeq FGx.
2. Input sample index combinations, sample types, and DNA primer mix selection for sequencing on the MiSeq FGx. Information is entered manually or by importing a text (*.txt) file.
3. The MiSeq FGx uses 2 software applications in tandem to produce images of clusters on the flow cell, perform image analysis, and call bases.
    a. During the run, MiSeq FGx Control Software captures images of clusters on the flow cell for image analysis, as well as operates the flow cell stage, gives commands to dispense reagents, and changes temperatures of the flow cell.
    b. Real-Time Analysis (RTA) software performs image analysis and base calling, and assigns a quality score to each base for each cycle as the run progresses. The completion of analysis by RTA and transfer of files initiates analysis on the ForenSeq Universal Analysis Software.
4. Using data from RTA, the ForenSeq Universal Analysis Software aligns reads to make allele and genotype calls. If interpretation or troubleshooting for a particular allele might be considered, the ForenSeq Universal Analysis Software provides quality indicators.
5. Review run analyses and generate reports on the ForenSeq Universal Analysis Software.

ForenSeq Universal Analysis Software Concepts

The following concepts and terms apply to the ForenSeq Universal Analysis Software:

Concept	Description
Amplicon	The result of PCR amplification of a targeted region of interest from input gDNA template.
Analysis Version	Analysis versions are a traceable, flexible way to reanalyze samples with new analysis settings. See *Creating a New Analysis* on page 31.
Analytical Threshold	This value (%) represents the limit of detection, and is a matter of internal laboratory policy-making.
Clusters	A clonal grouping of template DNA bound to the surface of a flow cell. Each cluster is seeded by a single, template DNA strand and is clonally amplified through bridge amplification until the cluster has roughly 1000 copies. Each cluster on the flow cell produces a single sequencing read. For example, 10,000 clusters on the flow cell would produce 10,000 single reads and 20,000 paired-end reads.
DNA Primer Mix	A set of PCR primers. Primers are tagged oligos targeting DNA sequences. In the ForenSeq DNA Signature Prep Kit, DNA Primer Mix A contains primers for STRs and identity SNPs (iSNPs). DNA Primer Mix B contains primers for STRs, iSNPs, phenotypic SNPs (pSNPs), and biogeographical ancestry SNPs (aSNPs).
Genotype	Alleles at a locus, which are typically heterozygous, homozygous, or hemizygous.
Human Sequencing Control	Pool of human DNA STR loci used as a control in the ForenSeq DNA Signature Prep Kit. Human Sequencing Control (HSC) helps run completion, and to highlight possible sequencing issues.
Index	DNA tags that are attached to target DNA sequences, enabling multiple samples to be pooled together for sequencing and demultiplexed post-run. See *Demultiplexing* on page 100.
Interlocus balance	The balance of read counts between loci in a sample. This balance is measured as the % CV of the read counts across all of the loci in the multiplex. See *General Locus Information* on page 34.
Intralocus balance	The balance of read counts between typed alleles at a locus. This balance is measured as the intensity of the minimum intensity typed allele divided by the intensity of the maximum intensity typed allele. If both are identical in intensity, the intralocus balance is 100%. See *Locus Detail Box Indicators* on page 35.
Interpretation Threshold	Sometimes referred to as a stochastic threshold. This value (%) can help in assessing both alleles of a heterozygote, and is a matter of internal laboratory policy-making.
Project	A project is a collection of analyzed results in the ForenSeq Universal Analysis Software and can be used to logically group and organize data. If desired, each sample in a run can be independently assigned to a different project, so that a run can have samples from more than 1 project. See *Viewing the Project Page* on page 19.

Concept	Description
Read	A sequence read refers to the data string of A, T, C, and G bases corresponding to the sample DNA. Millions of reads can be generated in a sequencing run.
Run Version	Run versions are a traceable, flexible way to recombine information to execute a new analysis. See *Viewing Run Details* on page 13 and *Creating a New Run Version* on page 16.
Sample Information	Information to set up and analyze a sequencing run, such as a list of samples and their index combinations, sample types and DNA primer mix.
Sample Type	The type or function of a sample. Sample types include sample, positive amplification control, negative amplification control, and reagent blank.
SNP	Single Nucleotide Polymorphism- Variation of a single nucleotide base within a DNA sequence, relative to a known DNA reference sequence.
STR	A DNA sequence containing a variable number (typically ≤ 50) of tandemly repeated short (2–6 bp) sequence motifs, such as (GATA)n.
Stutter	Polymerase slippage that can occur during PCR amplification of repetitive DNA sequences, the library preparation process, cluster generation, or when performing sequencing, and can create DNA amplification that is less than or greater than the size of a parent allele.

Viewing the ForenSeq Universal Analysis Software

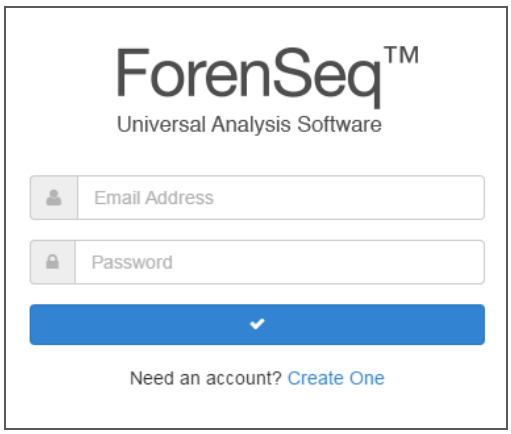
The ForenSeq Universal Analysis Software interface is viewed through a web browser. To view the interface, open a web browser on a computer with access to the network used by the ForenSeq Universal Analysis Software and the Instrument. Set screen resolution to a minimum of 992 pixels for optimal display.

NOTE
The application is optimized for use on the Google Chrome Browser.

To access the software, enter the server address in the web browser address bar. If the server is not listed in the Domain Name System (DNS), enter the server IP address of the network.

NOTE
No internet connection or internet access is required to view the software interface.

Logging in to the Software



1 Access the network server through a web browser.

2 Enter a user account name and password.
If you need an account, see *Creating an Account*.

3 Click the blue button with the checkmark.

4 To log out, click the drop-down list on your account email address at the top of the page and select **Log Out**.

NOTE
To change your password, see *Changing an Account Password* on page 84.

Creating an Account

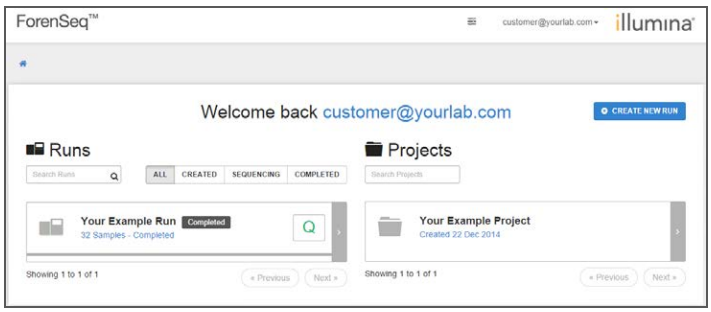
1 Access the network server through a web browser.

2 When the ForenSeq Universal Analysis Software page opens, click **Create One** at the bottom of the dialog box.

3 Enter a valid user account name and password.

4 Click the blue button with the checkmark.
After you create an account, a user with administrator-level system access must approve your account before you can access the system. See *Enabling and Disabling Accounts* on page 83.

Viewing the System Dashboard



A successful login opens the system dashboard, and shows the following information and functionality:

- List of runs, arranged by run creation date, with the most recent run appearing at the top of the list
- List of projects, arranged by project creation date, with the most recent project appearing at the top of the list
- Search tools for runs and projects
- Run filtering for the run states of All, Created, Sequencing, and Completed
- Analysis icon to access processing queue information, the last 3 analyses completed, and the status of the last analysis submitted
- Name of the logged in user with a drop-down list that includes access for log out, About information, and user profile settings
- Quality icon next to the run name that links to run quality metrics. See *Run Metrics Tab on page 27* for more information.
- Create a new run

Example Workflow

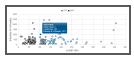
Create an account in the ForenSeq Universal Analysis Software. See *Logging in to the Software* on page 7.

Create a new run and save it. See *Creating a New Run* on page 10.

Log in to the Instrument with the account information created in the ForenSeq Universal Analysis Software. See the *MiSeq FGx Instrument Reference Guide (part # 15050524).*

On the Instrument, select and sequence the run you created and saved in the ForenSeq Universal Analysis Software. See the *MiSeq FGx Instrument Reference Guide (part # 15050524).*

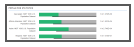
Review run quality information in the ForenSeq Universal Analysis Software. See *Viewing Run Quality Information* on page 20.



Review sample details. See *Viewing Sample Details Page* on page 33.

Optional Steps

Create a new run version or create a new analysis. See *Creating a New Run Version* on page 16 and *Creating a New Analysis* on page 31.



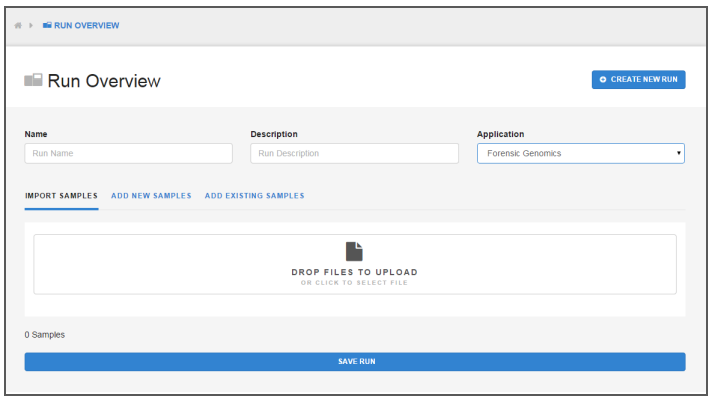
Generate population statistics. See *Population Statistics* on page 43.

Compare samples. See *Sample Comparison* on page 49.

Generate phenotype or biogeographical ancestry information. See *Phenotype and Biogeographical Ancestry Estimation* on page 58.

Generate project-level and sample genotype reports. See *Reports* on page 65.

Creating a New Run



Before performing a sequencing run on the Instrument, enter information in ForenSeq Universal Analysis Software about the run and the samples in the run.

After logging in, the software opens to the system dashboard.

1 Click **Create New Run**.

2 Enter the following information for sequencing and analysis.

   a Name—Enter a run name.
   b [Optional] Description—Enter a run description.
   c Application—Select Forensic Genomics.
   d Add sample information using 1 of the 3 methods described in Table 1, or a combination of the methods.

   Before you save a run, you can edit any of the run details and sample details, or remove a sample. To make changes, click in any field on the Create A Run page. To remove a sample, click the X on the right side of the sample information.

3 Click **Save Run**.

NOTE
The run version that is created for the initial analysis and is sequenced is version 1.0. For example, if the name of the run is Example Run, the run version is 1.0 - Example Run, and appears on the Project page. If you create additional versions of the run, the runs are numbered in sequence. See *Creating a New Run Version* on page 16.

Table 1 Create a Run

Method	Description	Steps
Import Samples	If you created a tab-delimited file (*.txt) with sample sheet information, import the sample sheet. See *Entering Sample Information* on page 12.	▸ Click the **Import Samples** tab. ▸ Click **Drop Files to Upload or Click to Select File**. ▸ Navigate to the (*.txt) file and click **Open**. ▸ Click the blue **Import Samples** button.
Add New Samples	Add new sample information.	▸ Click the **Add New Samples** tab. ▸ Enter a **Sample Name**. ▸ Enter a **Project Name**. If the project already exists, use the search field to find the project name. ▸ [Optional] Enter a **Sample Description**. ▸ Select an **i7 Index**. ▸ Select an **i5 Index**. ▸ Select a **Sample Type** from the drop-down list. ▸ Select a **Mix Type** from the drop-down list. ▸ When sample information is complete, click **Add New Sample**.
Add Existing Samples	Add samples already created.	▸ Click the **Add Existing Samples** tab. ▸ Enter an existing **Sample Name**. Use the search field to find the sample name. ▸ Enter a **Project Name**. If the project already exists, use the auto complete drop-down list to find the project name. ▸ [Optional] Enter a **Sample Description**. ▸ Select an **i7 Index**. ▸ Select an **i5 Index**. ▸ Select a **Sample Type** from the drop-down list. ▸ Select a **Mix Type** from the drop-down list. ▸ When sample information is complete, click **Add New Sample**.

Entering Sample Information

Sample information can be directly input on the Create a Run page. You can also create a tab delimited *.txt file, and import it into the software on the same page.

Creating a Text File with Sample Information

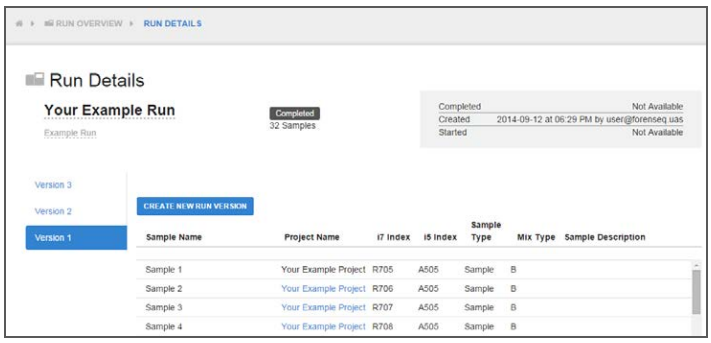
SampleName	Project	i7Index	i5Index	SampleType	SampleDescription	MixType
Sample1	Tst Prj	R701	A501	Sample	Sample 1	A
Sample2	Tst Prj	R701	A502	Positive Amplification Control	PAC 1	B
Sample3	Tst Prj	R701	A503	Negative Amplification Control	NAC 1	A
Sample4	Tst Prj	R701	A504	Reagent Blank	RA 1	B

To create a *.txt file for a run, make columns of information, with these headers in the columns.

Table 2 Column Headings for Sample Information

Column Header	Field Description
SampleName	Name of the sample.
Project	Project associated with the sample. If desired, each sample can be independently assigned to a different project, so that a run can have samples from more than 1 project. See *Viewing the Project Page* on page 19.
i7Index	i7 index ID for the sample. The i7 index range is R701 to R712.
i5Index	i5 index ID for the sample. The i5 index range is A501 to A508.
SampleType	Options for the sample type include the following: • Sample • Reagent Blank • Negative Amplification Control • Positive Amplification Control
SampleDescription	Describe the sample here.
MixType	Enter DNA mix type A or B for the sample. DNA Primer Mix A contains primers for STRs and iSNPs. DNA Primer Mix B contains primers for STRs, iSNPs, pSNPs, and sSNPs.

Viewing Run Details

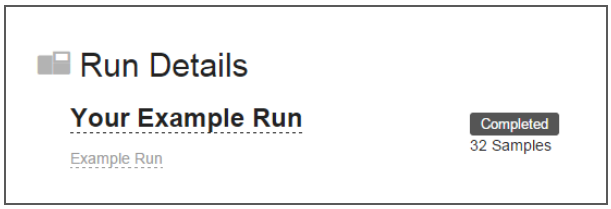


To view the Run Details page, click the name of the run on the system dashboard. There are several ways to find a run on the system dashboard:

- The system dashboard lists runs on the left.
- To narrow your search of the run list, click a buttons next to the Search Runs box to sort; All, Created, Sequencing, or Completed.
- On the system dashboard, use page navigation at the bottom of the page, under the runs column.
- Enter the run name in the Search Runs box.
- To return to the system dashboard, use the breadcrumb trail at the top of the page.

The Run Details page contains the following information:

- Name of the run
- Description of the run (if a description is entered when the run is created)
- Status of the run. See *Viewing Run Status* on page 14
- Number of samples in the run. If more than 1 run version, the number of samples in the most recent version of the run



- Time and date the run is created in the ForenSeq Universal Analysis Software
- Time and date the sequencing of the run began on the Instrument
- Time and date the run completed sequencing on the Instrument
- Run version
- A list of samples in that version of the run

Viewing Run Status

After you create and save a run, a box next to the name of the run indicates the current run status.

Table 3 Run Status Descriptions

Run Status	Description
Created	Run setup information is saved in the ForenSeq Universal Analysis Software.
Sequencing	Sequencing of the run on the Instrument is in progress.
Sequencing Paused	Sequencing of the run is on pause on the Instrument. To pause a run, see the *MiSeq FGx Instrument Reference Guide (part # 15050524)*.
Completed	Sequencing of the run on the Instrument is complete.
Completed-Error	The run did not finish sequencing on the Instrument because it was manually stopped, or because of an error.

Changing the Name or Description of a Run

1. Navigate to the Run Details page.
2. Click the name or description of the run.
3. Enter the new name in the name field, or the new description in the description field.
4. Click the blue checkmark box to save the change, or click the X box if you do not want the save the change.

Creating a New Run Version

The Create New Run Version feature is a traceable, flexible way to create alternate analyses of your samples. In a new run version, you can use existing samples or add new samples. You can modify run index combinations, sample types, project assignment, and DNA primer mix types. When the new run version is set up, you can execute an analysis with the new information. New run versions are available in the Instrument for sequencing. Create A New Run Version is available after a run completes sequencing.

1 Open the Run Details page. See *Viewing Run Details* on page 13.

2 Click the **Create New Run Version** button.

3 Select **Edit Samples** to revise sample information. Add samples or modify existing sample information to describe the samples in a run.

4 Click **Save Changes**.

5 To execute an analysis on the new version, select **Execute Analysis**.

NOTE
After you execute analysis on a run version, that run version cannot be deleted.

6 To delete the run version before you execute analysis, select **Delete Version**.

NOTE
Version 1.0 of a run is the run version that is created for the initial analysis and is sequenced. If you create additional versions of the run, the runs are numbered in sequence. For example, if you create a new version of the run 1.0 - Example Run 1.0, it becomes 2.0 - Example Run after sequencing. If you create a third version to sequence of the same run, it becomes 3.0 - Example Run, and so forth.

Part # 15053876 v01

Sample and Run Results

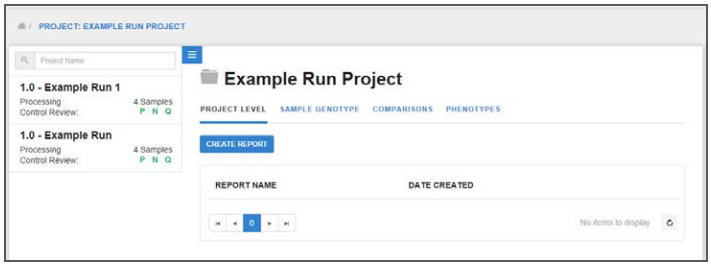
Chapter 2

Introduction 18
Viewing the Project Page 19
Viewing Run Quality Information 20
Creating a New Analysis 31
Viewing Sample Details Page 33

Introduction

Information and results for a project are opened from the Project page. Quality metrics, sample representation, as well as positive and negative control results are accessible for each analysis. Analysis settings can be viewed, modified, or saved using a template, and renamed. Modified settings can then be used to reanalyze samples in a project. The software also features charts and tables of STR and SNP locus results.

Viewing the Project Page



When you add samples to a run, you associate every sample to a project. Because each sample can be independently assigned to a project, a run can have samples with more than 1 project. Information and results for a project appear in a series of pages in the software interface from the Project page. To view project-level results in a report, see *Project Level Report* on page 67.

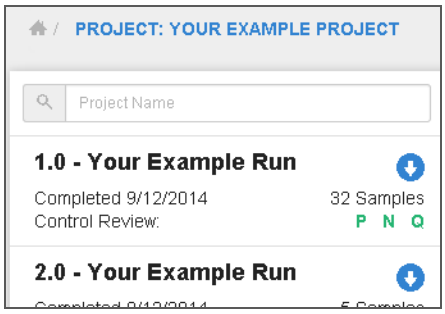
Navigating to the Project Page

There are several ways to find a Project page.

- Access a Project page from the Run Details page of your completed sequencing run.
- Access the Project page from the system dashboard, which lists projects on the right.
- On the system dashboard, use page navigation at the bottom of the page, under the projects column to find the relevant project. Click the name of a project to access the Project page.
- Enter the project name in the Search Projects field. Click the name of a project to access the Project page.
- On the Project page, use the Project Name search field to navigate to other projects you want to view. Click the name of a project to access the Project page.
- To return to the system dashboard, use the breadcrumb trail at the top of the page.

Viewing Run Quality Information

The analyses in a project, along with the associated version number, are listed on the left panel of the Project page. Click the name of an analysis to view run quality metrics and sample metrics.

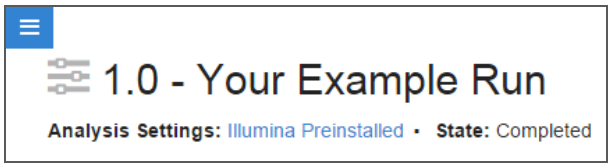


The following basic run information shows in the left panel below the analysis name:

Analysis Info	Description
Completed Date	Date the run completed sequencing, or, if the analysis is not complete, the word "Processing"
Samples	Number of samples in the analysis for the project
Control Review	Snapshot of the overall quality indication for the following: • Positive control samples (P) • Negative control samples (N) • Quality metrics for the run (Q) Green indicates that all the metrics were within an acceptable range. Orange indicates that at least 1 metric was not within the predefined range for acceptability and requires further investigation. Black indicates that there are no metrics assessed with the analysis.

After you click an analysis of a project on the left panel of the Project page, an Analysis page opens. At the top middle of the Analysis page, analysis settings, the analysis version, and the status (state) of the analysis for the run shows below the analysis name. Analysis Settings indicate if Illumina preinstalled settings are in use. If the analysis settings are from an analysis template, the name of the template is displayed. See *Changing Locus Thresholds* on page 85.

Reviewers can indicate a review status for each analysis and each sample. The default review state is Primary Review, with the option to select Technical Review, and then Review Complete.



The following table lists possible analysis states.

Table 4 State of Analysis

State	Description
New	Analysis by the software did not begin yet.
Queued	Analysis is in line to begin processing.
In Progress	Analysis by the software is in progress.
Completed	Analysis by the software completed successfully.
Errored	Analysis did not complete successfully. Analysis stopped with no results because of an error.

After selecting an analysis name on the left panel, a box with the following color-coded run quality indicators appears on the page, from left to right. Click the icon for further details.

Table 5 Run Quality and Sample Control Icons

Icon	Description
P	Positive control quality indicator for positive amplification controls and HSC metrics. See *Viewing Positive Control Metrics*.
N	Negative control quality indicator for reagent blank samples and negative amplification control samples. See *Viewing Negative Control Metrics* on page 24.
Q	Displays overall quality metrics for the run. See *Viewing Quality Metrics* on page 26.

Viewing Positive Control Metrics

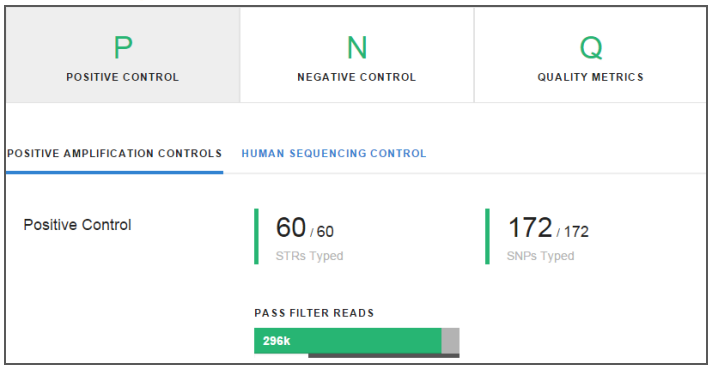
The analyses in a project, along with the analysis version number, are listed on the left panel of the Project page. To view results for positive control samples, click the analysis name on the Project page to open the Analysis page, and then on the Positive Control icon (P). Click the tabs to display positive amplification control and HSC.

Positive Control metrics are available after the analysis is complete. Color indicators on the positive control icon (P) show the outcome of control metrics.

Table 6 Positive Control Icon Color Indicator

Color	Positive Amplification Control Samples	HSC Samples
Green	All samples designated positive amplification control in the run, assigned to the analysis, have a green status. All STRs and SNPs are typed and all alleles are concordant with the known Control DNA 2800M.	Overall the sample has sufficient intensity coverage, and the genotype for each locus is concordant.
Orange	Orange status indicates that at least 1 STR or SNP is not typed, is typed but discordant with Control DNA 2800M, or the number of reads in the sample is less than 85,000.	The sample might not have sufficient read coverage, or the genotype for a locus is not concordant. Lists all effected loci.

Positive Amplification Control Tab



The Positive Amplification Control tab lists all samples identified as positive amplification control. Each sample is compared with the known Control DNA 2800M multi-locus genotype for each of the STRs and SNPs in the DNA primer mix that is assigned to the sample. Next to each sample name is a short vertical color bar with the number of STRs and SNPs typed out of the total possible number of STR loci and SNP loci.

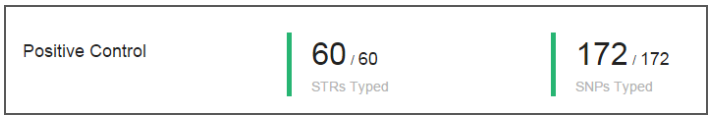


Table 7 Positive Amplification Control Vertical Bar Color Indicators for STRs

Color	Indication
Green	All STRs in the sample are typed and concordant with the known Control DNA 2800M.
Orange	Indicates that ≥ 1 STR in the sample is not typed, or is typed yet discordant with the known Control DNA 2800M muli-locus genotype.

Table 8 Positive Amplification Control Vertical Bar Color Indicators for SNPs

Color	Indication
Green	All SNPs in the sample are typed and concordant with the known Control DNA 2800M.
Orange	Indicates that ≥ 1 SNP in the sample is not typed, or is typed yet discordant with the known Control DNA 2800M muli-locus genotype.

For each positive amplification control sample, the following information is displayed:

- The name of each positive amplification control sample, which can be clicked to open the Sample Details page
- The number of typed STRs and SNPs out of the total possible number of STRs and SNPs in the ForenSeq primer mix
- A list of discordant loci compared to Control DNA 2800M
- The number of reads

NOTE
Control DNA 2800M is a control for single-source male gDNA and is the positive amplification control DNA provided with the ForenSeq DNA Signature Prep Kit.

If warranted, a discordance table shows all discordant STRs and SNPs compared to Control DNA 2800M and includes the following information:

- Locus name
- Genotype observed in the sample
- Genotype expected from the Control DNA 2800M positive amplification control

A locus is determined to be discordant for the following reasons:

- 1 or more loci in the positive amplification control is not called
- The genotype of 1 or more positive amplification control loci differs from the known Control DNA 2800M genotype

Each positive amplification control sample has a horizontal bar that indicates the number of reads. The dark gray shadow bar below the bar indicates a guideline for a sufficient number of reads of 85,000 and above. To view sample scores on a more exact point in the range of values, mouse over the graphs.

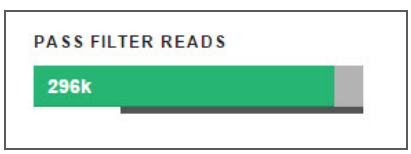
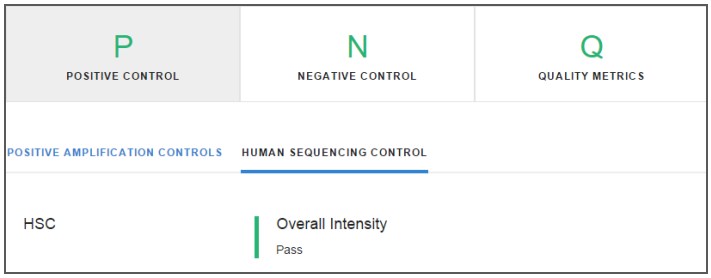


Table 9 Positive Amplification Control Number of Reads Bar Color Indicators

Color	Indication
Green	The total number of reads is within the expected range.
Orange	The total number of reads is not within the expected range. Orange does not necessarily mean that the run fails. It is possible to use these data, however Illumina recommends verifying the sample data. Look at the data in more detail on the Sample Details page and review the quality results on the Quality Metrics page. See *Viewing Sample Details Page* on page 33, and *Viewing Quality Metrics* on page 26.

Human Sequencing Control Tab



The Human Sequencing Control tab indicates whether HSC meets criteria for intensity, and genotype concordance.

The loci in the sequencing control are assessed for an expected minimum intensity level and the correct genotype call. If the overall intensity status fails, then a list of relevant loci and the locus lengths is displayed.

Next to the sample name HSC is a short vertical color bar.

Table 10 Human Sequencing Control Vertical Bar Color Indicators

Color	Indication
Green	Indicates that the sample meets or exceeds minimum intensity level criterion and genotype concordance.
Orange	Indicates that the sample does not meet 1 or more criteria. If the bar is orange, then it is possible that the overall intensity for HSC is lower than expected. If any loci are listed on the page, then those loci do not meet intensity or genotype concordance criteria, and their locus name and length in base pairs (bp) are listed. For a complete list of HSC loci, see *Human Sequencing Control Loci* on page 116.

Viewing Negative Control Metrics

The analyses in a project, along with the analysis version number, are listed on the left side of the navigation bar of the Project page. To view results for negative control samples, click the analysis name on the Project page to open the Analysis page, and then on the Negative Control icon (N). Click the tabs to display reagent blanks and negative amplification control metrics.

Negative Control metrics are available after analysis is complete. A color indicator on the negative control icon (N) shows the outcome of control metrics.

Table 11 Negative Control Icon Color Indicators

Color	Indication
Green	All samples designated reagent blank or negative amplification control in the run, assigned to the analysis, have a green status. Green status indicates that in each control sample, no SNPs, or STRs are typed.
Orange	At least 1 sample designated Reagent Blank or Negative Amplification Control in the analysis has an orange status. Orange status indicates that ≥ 1 STR or SNP is typed.
Grey	No samples are designated reagent blank, nor negative amplification control, in the analysis assigned to the project.

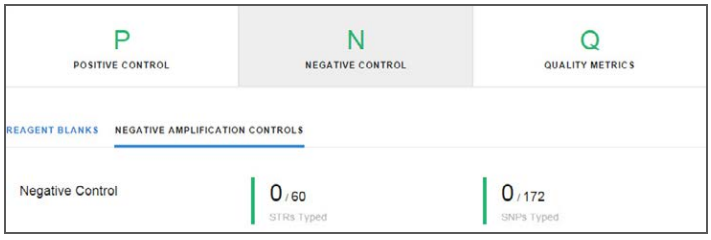
Reagent Blanks Tab

The Reagent Blanks tab lists all samples identified as reagent blank. Each sample is compared with STRs and SNPs in the DNA primer mix assigned to the sample. Next to each sample name, the tab shows how many STRs and SNPs are typed out of the total possible number of target loci. You can click the sample name to open the Sample Details page.

Table 12 Reagent Blanks Tab Vertical Bar Color Indicators

Color	Indication
Green	Indicates that no STR or SNP locus is typed.
Orange	Indicates that ≥ 1 STR or SNP locus is typed.

Negative Amplification Controls Tab



All samples identified as negative amplification control are listed in the Negative Amplification Controls tab. You can click the sample name to open the Sample Details page. Each negative amplification control sample is compared with the STRs and SNPs in the DNA primer mix assigned to that sample. Next to each sample name, the tab shows how many STRs and SNPs are typed out of the total possible number of target loci.

Table 13 Negative Amplifications Controls Tab Vertical Bar Color Indicators

Color	Indication
Green	Indicates that no STR or SNP locus is typed.
Orange	Indicates that ≥ 1 STR or SNP locus is typed.

Viewing Quality Metrics

The analyses in a project, along with the analysis version number, are listed on the left side of the Project page. To view overall run quality metrics and sample representation, click the name of an analysis on the Project page to open the Analysis page, and then on the Quality Metrics icon (Q). Click the tabs to display Run Metrics and Sample Representation. Run metrics information can also be accessed on the system dashboard by hovering on the Q next to the run name.

Quality Metrics for the run shows intensity (number of reads) of each sample in the analysis, and a number of cycles completed indicator. This page enables remote monitoring of the run, and all the run metrics available during sequencing on the Instrument.

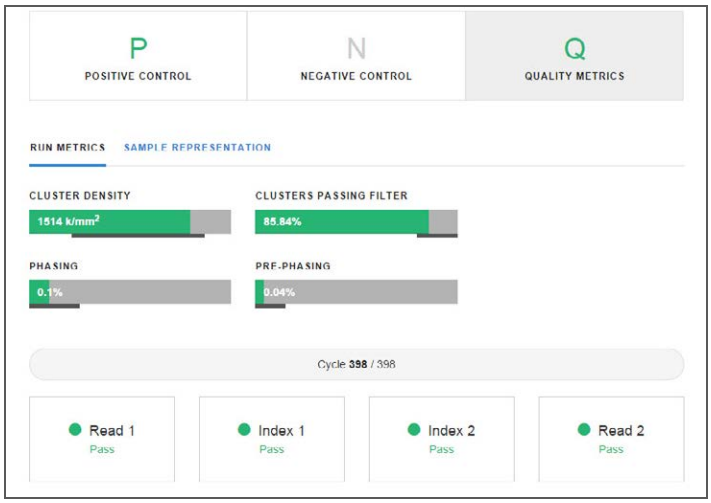
As the run progresses, a gray run progress bar is displayed on the system dashboard under the run name, and the Instrument updates the ForenSeq Universal Analysis Software with details of run progress. The metrics presented mirror the information that is displayed on the MiSeq FGx during sequencing. After sequencing of the run completes, the information for the run is preserved for run quality assessment, as this Quality Metrics page remains static so that results from the run are available.

A color indicator on the quality metrics icon (Q) shows the outcome of quality metrics.

Table 14 Quality Metrics Icon Color Indicators

Color	Indication
Green	All quality metrics are in the recommended range for a run.
Orange	1 or more quality metric is not in the recommended range for a run.

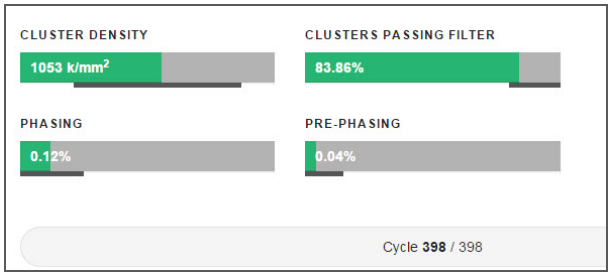
Run Metrics Tab



To view overall run quality, click the Run Metrics tab. Horizontal bars show the acceptable range of values for cluster density, clusters passing filter, phasing, and prephasing. A color dot indicates Read and Index quality.

Table 15 Run Metrics Horizontal Bar Color Indicators

Color	Indication
Green	Green indicates that values are within the acceptable range.
Orange	Orange indicates the possibility that further evaluation of the quality of the run is recommended.
Dark Gray	The dark gray shadow bar below the result bar indicates the passing range.



- **Cluster Density (K/mm²)**—Shows the number of clusters per square millimeter for the run. For ForenSeq runs, Illumina recommends a target cluster density range of 400–1650 K/mm². Cluster density values outside of the target range can still produce results that are sufficient to use for analysis. Values that deviate substantially from the target range can negatively impact other quality metrics, and decrease the quantity of valuable data from the run.
- **Clusters Passing Filter (%)**—Shows the percentage of clusters passing filter based on the Illumina chastity filter, which measures quality. The filter can detect low quality base calls. Data appears only after cycle 25.

NOTE
The chastity of a base call is the ratio of the intensity of the greatest signal divided by the sum of the 2 greatest signals. If more 1 one base call has a chastity value of less than 0.6 in the first 25 cycles, reads do not pass the quality filter.

For ForenSeq runs, Illumina recommends a target clusters passing filter value of ≥ 80%. Clusters passing filter values that are outside of the target range can still produce results that are sufficient to use for analysis. Values that deviate substantially from the target range can negatively impact other quality metrics, and decrease the quantity of data from the run.

- **Phasing (%)**— Shows the percentage of molecules in a cluster that fall behind the current cycle within Read 1 and Read 2. Low percentages indicate good run statistics. For ForenSeq runs, Illumina recommends a phasing value of ≤ 0.25%. Phasing values outside of the target range can still produce results that are sufficient to use for analysis. See *Phasing and Prephasing* on page 99.
- **Prephasing (%)**— Shows the percentage of molecules in a cluster that run ahead of the current cycle within Read 1 and Read 2. Low percentages indicate good run statistics. For ForenSeq runs, Illumina recommends a prephasing value of ≤ 0.15%. Prephasing values outside of the target range can still produce results that are sufficient to use for analysis. See *Phasing and Prephasing* on page 99.
- **Cycle**— Shows the number of sequencing cycles completed. 1 cycle includes the chemical addition and imaging of 1 base for each cluster on a flow cell. A total of 398 cycles are performed on a ForenSeq run.

Table 16 Read and Index Metrics Icon Color Indicators

Color	Indication
Green	The average quality for assessed reads is within the recommended range.
Orange	The average quality for assessed reads is not within the recommended range.
Gray	The read or index did not occur yet in the sequencing run.

- **Read 1**— Read 1 follows the Read 1 sequencing protocol using Sequencing by Synthesis (SBS) reagents. The Read 1 sequencing primer is annealed to the template strand during the cluster generation step. The RTA software evaluates the first 50 cycles of the Read 1 segment of the run for quality. Read 1 quality metrics are displayed only after cycle 50 is complete.
- **Index 1**— The Read 1 product is removed, and the Index 1 (i7) sequencing primer is annealed to the same template strand as in Read 1. Following Index Read preparation, the Index 1 (i7) Read performs 8 cycles of sequencing. The RTA software evaluates all 8 cycles of the Index 1 segment of the run for quality. Quality results are displayed only after cycle 359 is complete.
- **Index 2**— The Index 1 (i7) Read product is removed, and the template anneals to the grafted P5 primer on the surface of the flow cell. The run proceeds through an additional 7 chemistry-only cycles in which no imaging occurs, followed by 8 cycles of

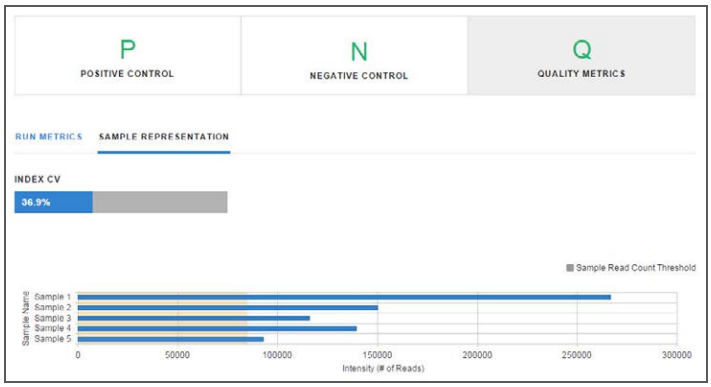
sequencing. The RTA software evaluates all 8 cycles of the Index 2 (i5) segment of the run for quality. Quality results are displayed only after cycle 367 is complete.

- **Read 2**— The Index Read 2 product is removed, and the original template strand is used to regenerate the complementary strand. Then, the original template strand is removed to allow hybridization of the Read 2 sequencing primer. Read 2 follows the standard paired-end sequencing protocol using SBS reagents. The RTA software evaluates all cycles of the Read 2 segment of the run for quality.

Table 17 Cycle Increments and Corresponding Sequencing Phases

Cycle Number	Sequencing Phase
1 to 351	Read 1
352 to 359	Index 1
360 to 367	Index 2
368 to 398	Read 2

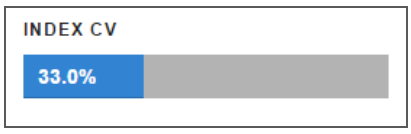
Sample Representation Tab



Quality Metrics for the run show intensity (number of reads) of each sample in the analysis, and a number of cycles completed indicator. This page enables remote monitoring of the run, and all the run metrics available during sequencing on the Instrument. Click the tabs to display Run Metrics and Sample Representation.

The quality Sample Representation tab shows an overall view of the number of reads, and read distribution for each sample in the analysis.

- **Index CV**— Displayed as a percentage, the number of reads that are assigned to each sample. CV is the coefficient of variation for the number of read counts across all indexes. Index CV represents the distribution of read counts of the samples in the run.



Below the Index CV bar is a graph that lists each sample, with a bar showing the number of reads. To see the exact number of reads for a sample, mouse over the bar. You can click the bar or sample name to open the Sample Details page.

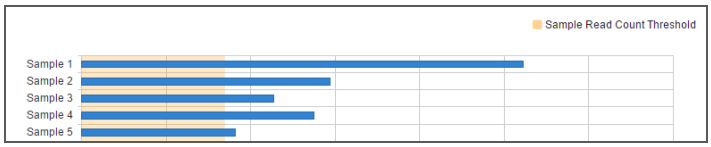
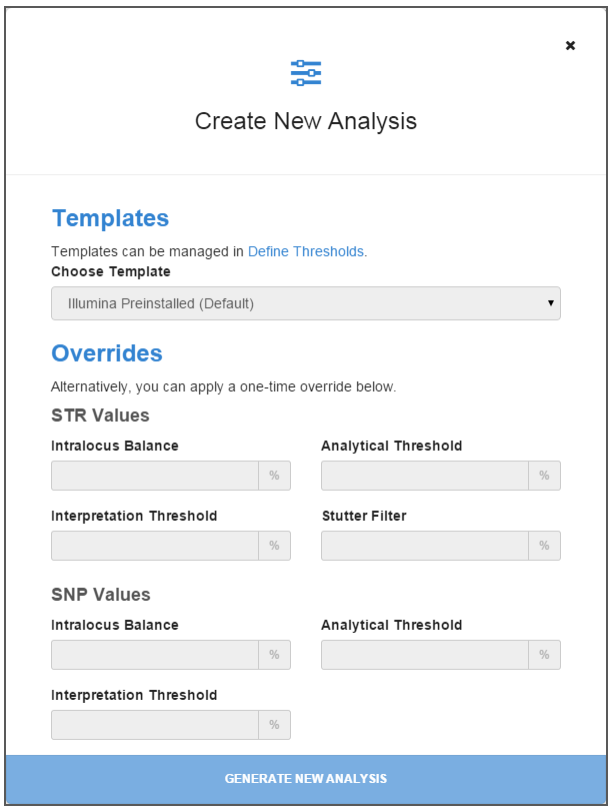


Table 18 Color Indicators of the Sample Representation Tab Graph

Color	Indication
Light Orange	0 to 85,000 reads are highlighted in the graph in light orange. Interpret samples below 85,000 reads with caution, as loci with low coverage might not have enough signal for alleles to be distinguished from noise.
Blue	A sample with a blue bar in the graph indicates a sample that meets signal intensity requirements.
Dark Orange	A sample with a dark orange bar in the graph indicates a sample that does not meet signal intensity guidelines.

Creating a New Analysis



The Create New Analysis feature is a traceable, flexible way to create alternate versions of major analyses. In a new analysis, you can select a different system settings template or apply a 1-time override to system settings. The creation of a template requires administrator access. See *Creating a Locus Threshold Template* on page 85. Create New Analysis is available after a run completes sequencing.

If you apply a 1-time override, you have the option of changing intralocus balance, analytical threshold, and interpretation threshold settings for STRs and SNPs. You can also change the stutter filter setting for STRs.

The Create New Analysis command is accessed on the Analysis page. See *Viewing Run Quality Information* on page 20.

1. Select the name of a major analysis on the left panel.
   Major analyses version numbers end in 0. For example, an analysis with the name *3.0 - Your Example Run* is a major analysis. An analysis with the name *3.1 - Your Example Run* is not a major analysis, and no new analysis can be created from it.
2. Click the **Create New Analysis** button.
3. Select from the Choose Template drop-down list or enter 1-time override values in the Overrides section.
4. Click **Generate New Analysis**.

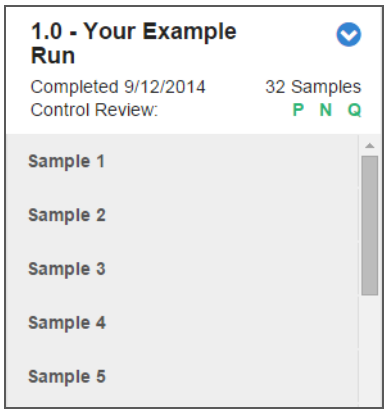
NOTE
After you execute analysis, the analysis version cannot be deleted.

NOTE

The run version created for the initial analysis and sequenced is version 1.0, and the initial analysis of that run is analysis version 1.0. If you create additional analyses for a run, the analyses are numbered in sequence. For example, if you create a new analysis of 1.0 - Example Run, the analysis name is 1.1 - Example Run. If you create an analysis of the run a second time, the analysis name is 1.2 - Example Run. If you create a new analysis of the run a third time, the analysis name is 1.3 - Example Run, and so forth. Create New Analysis can be applied to major analyses only, such as 1.0, 2.0, 3.0, etc.

Viewing Sample Details Page

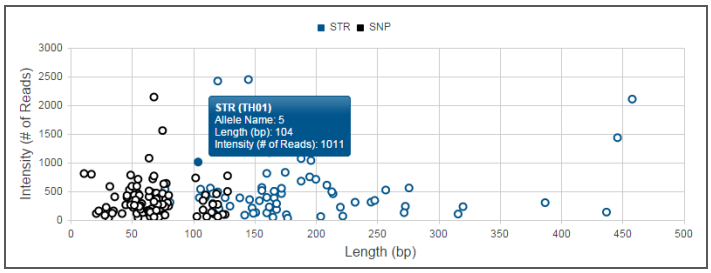
To view sample locus results, click the Project page. See *Viewing the Project Page* on page 19. Locate an analysis on the left pane of the page, and click the blue arrow to open the list of samples in the run. A blue down arrow indicates that the sample list for the analysis is open. To close the list of samples, click the blue arrow. A blue arrow that points left indicates that the sample list is closed.



To view the Sample Details page with STR and iSNP results, click a sample in the list. The sample name, index numbers, gender, sample type, DNA primer mix type, and description are displayed at the top of the page.

Detailed iSNP and STR locus information is formatted in tables below the intensity (number of reads) and length (bp) information graphs. A table with STR locus information is on the left, and a table with SNP information is on the right. There is a quality control indicator icon legend above the tables.

Intensity and Length Graphs

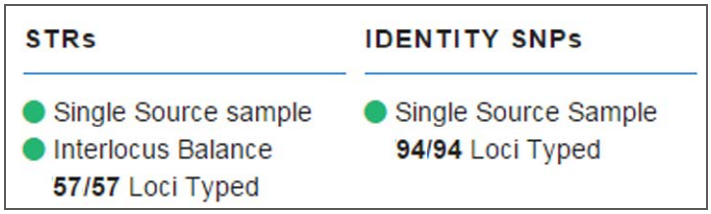


To see sample intensity and length information for STRs graphed separately from SNPs, click the blue STR or black SNP squares above the graph to turn the display off and on. Mouse over the squares to highlight the display of STRs or SNPs. Mouse over a dot in the graph to view the exact length in base pairs, and precise intensity level for an allele. The display for STRs includes the allele name.

- **Intensity (# of Reads)**—For each typed STR or iSNP allele in a sample, the graph shows signal intensity reported as the number of reads.

- **Locus Length (bp)**—For each typed allele in a sample, the graph shows the length of the amplicon as the number of base pairs between the PCR primers for the allele (STR) or locus (iSNP).

General Locus Information



The following STR and iSNP information shows above sample locus detail tables:

- **Single-Source Sample**—The single-source indicator is based on the results from the ForenSeq Universal Analysis Software algorithms for mixture detection. This approximation is performed separately for STRs and SNPs. For STRs, the number of loci with allele count quality control indicators must be > 5. For SNPs, the number of imbalance quality control indicators for all SNP loci must be > 10. See *STR Genotype Calling* on page 101 and *SNP Genotype Calling* on page 102. The calculation of the single-source indicator is updated if manual updates are made to the genotypes. See *Updating and Modifying Typed Results* on page 42.

NOTE
Evaluate and verify single-source sample results based on your internal guidelines.

Table 19 Single-Source Color Indicators

Color	Indication
Green	Indicates that the sample might be a single-source sample.
Orange	Indicates that the sample might contain multiple DNA contributors.

- **Interlocus Balance**—For STRs, the balance of the read counts between STR loci in a sample. This balance is measured as the % coefficient of variation (CV) of the read counts across all of the STR loci in the ForenSeq multiplex. The % CV is calculated as the standard deviation of the total read counts of each locus divided by the mean of locus read counts.

Table 20 Interlocus Balance Color Indicators

Color	Indication
Green	The balance is in the recommended range.
Orange	The interlocus balance metric falls outside of the optimum range.

- **Loci Typed**— Number of loci typed (by the analysis software or updated manually) out of all STR loci. Any loci that are deselected at the time of analysis creation are not included in the number of typed loci. See *Selecting Locus Content* on page 89. For SNPs, the number of loci typed represents the iSNPs only. Phenotypic SNP (pSNP) and biogeographical ancestry (aSNP) information is available in the *Phenotype Estimation*

SNP Sample Details Table on page 60. For more information on DNA primer mix types, see *Entering Sample Information* on page 12.

Locus Detail Box Indicators

Each box in the STR and iSNP sample locus details table has a color border that indicates whether the locus is within or outside of quality metric guidelines. If applicable, quality control indicator icons are displayed in the box.

Table 21 Locus Detail Box Color Indicators

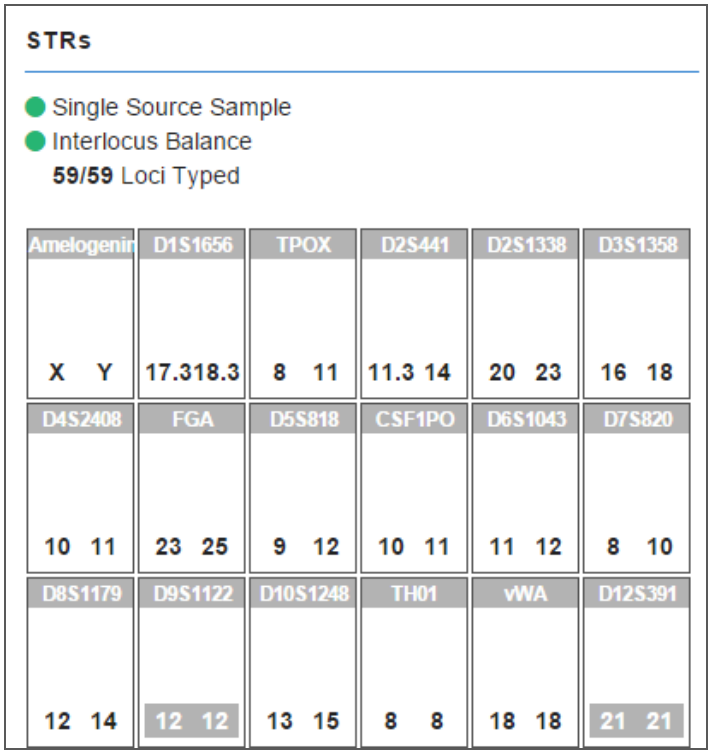
Color	Indication
Gray	Quality control indicators are not triggered for any locus.
Orange	1 or more quality control indicators are triggered for the locus.

If applicable, a quality control indicator icon shows in an STR or iSNP locus detail box. For more information about quality indicator icons, see *STR Genotype Calling* on page 101 and *SNP Genotype Calling* on page 102.

Table 22 Sample Details Quality Control Indicator Icons

Symbol	Symbol Name	Indication
	Stutter	For STRs, the stutter filter percentage is exceeded by 1 or more sequences at a stutter position of a possible parent allele. See *STR Genotype Calling* on page 101.
	Allele Count	For STRs, depending on the assigned gender of the sample, there are more alleles above the analytical threshold than expected, or that can be attributed to stutter. Serves as a potential tool for mixture detection and resolution. See *Allele Counting* on page 101.
	Imbalanced	Read count ratio falls below the user-defined intralocus balance in the analysis settings. For more information on imbalance, see *STR Genotype Calling* on page 101 and *SNP Genotype Calling* on page 102.
	Low Coverage	Signal above the interpretation threshold was not detected as defined in the analysis settings. See *STR Genotype Calling* on page 101 and *SNP Genotype Calling* on page 102. The low coverage indicator can be seen in results from version 1.1 or earlier of the ForenSeq Universal Analysis Software.
	Interpretation Threshold	Indicates that there is at least one allele that is above the analytical threshold, but below the interpretation threshold, that is not attributed to stutter.
	User Modified	At least 1 allele was manually edited as typed (True) or not typed (False). See *Updating and Modifying Typed Results* on page 42.
	Analytical Threshold	Indicates that the locus has signal below the analysis threshold, and no alleles above the interpretation threshold.
	Not Detected	No signal was detected for the locus.

STR Sample Details Table



The STR Sample Detail table consists of boxes for each target locus in the multiplex. All boxes are labeled with the locus name at the top. Click a box to see locus details. To view sample locus details in a report, see *Sample Genotype Report* on page 70.

A box for amelogenin is in the upper left corner. All other boxes display the STR alleles.

For autosomal STRs in single-source samples, 2 different numbers in a box indicate a heterozygous genotype. Two identical numbers in a box indicate a homozygous genotype, based on length alone. Review the locus details for further information and potential sequence variation.

For autosomal STRs, if more than 2 alleles are typed, a plus sign is displayed with the number of additional alleles. STRs with isometric alleles are highlighted with a gray box.

Click a box to open a locus details table to view the signal intensity and the sequence for the alleles at the locus. The display is color-coded, with rows in the table matching the colors of the bars in the chart.

The rows in the locus details table can be sorted by allele number, or by intensity, which is the number of reads. Click the Allele column heading to sort the data according to the allele call. Click the Intensity column heading to sort rows according to the number of reads (intensity).

In the locus details table, a toggle switch in the Typed column that displays a blue background and is in position on the right side indicates that the allele is contributing to the genotype for the locus. A toggle switch in the Typed column that displays a gray background and is in position on the left side indicates that the allele is not contributing to the genotype of the locus. The table also lists potential stutter and DNA repeat sequences.

The locus details bar chart shows potential alleles along the X axis. The analytical threshold is in dark gray, and interpretation threshold for intensity is in light gray. Hover over the information icon to the left of the legend to see analytical and interpretation threshold values. The precise signal intensity of a potential allele is displayed by mousing over the bar chart, and is listed in the locus details table. When STR alleles of the same length but of different a sequence are detected, they are displayed in the bar chart as a single bar with a horizontal line delineating each sequence variant (allele). Letters highlighted in white indicates isometric SNPs in the sequence.

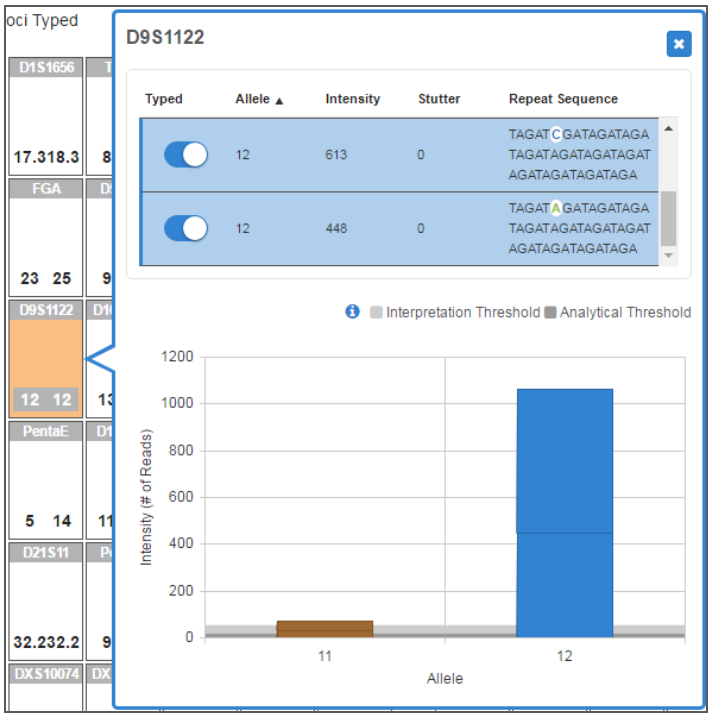
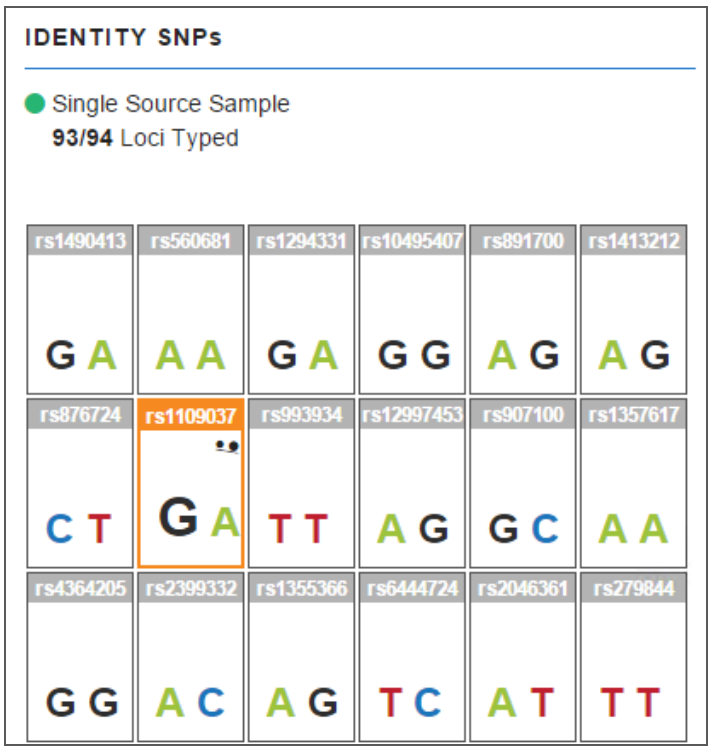


Table 23 STR Locus Details Table and Bar Chart Color Indicators

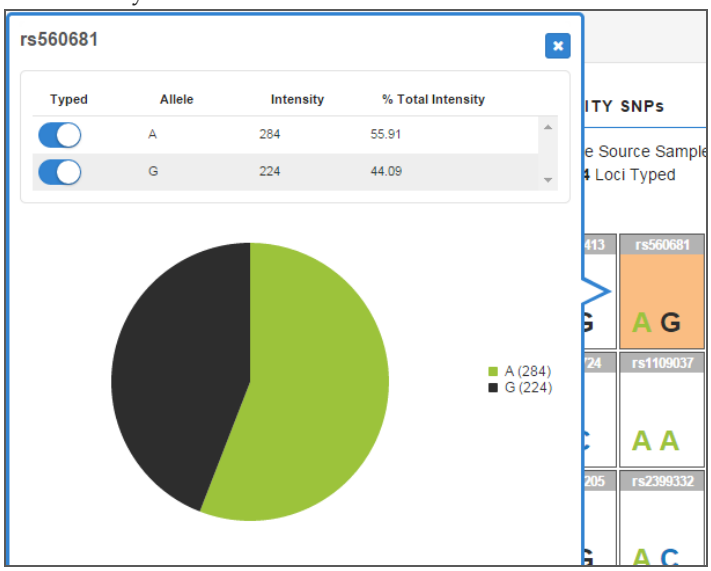
Color	Indication
Blue	Typed allele.
Brown	DNA sequence that is not typed, as it is above the analytical threshold, but below the stutter filter and is potentially stutter, or is the same length as a typed allele.
Pink	DNA sequence that is not typed, and is above the analytical threshold and stutter filter %. It meets analytical threshold and stutter filter limits. However, intensity falls below the interpretation threshold, or there are no other typed alleles in the position of the parent allele relative to potential stutter. For more information on stutter filtering, see *STR Genotype Calling* on page 101.
Red	DNA sequence that is not typed and is above the analytical threshold, interpretation threshold, and stutter filter %, and is not the same repeat length as a typed allele. The signal does not fall below the stutter filter, relative to any typed allele. For more information on stutter filtering, see *STR Genotype Calling* on page 101.
Grey	Signal not typed. Falls below the analytical threshold.

SNP Sample Details Table



The SNP Sample Details table consists of boxes for each target locus in the multiplex. All boxes are labeled with the locus name at the top. Click a box to see locus details. To view sample locus details in a report, see *Sample Genotype Report* on page 70.

In each box, nucleotide abbreviations refer to the genotype of the locus. Signal imbalance is indicated by the font size of 1 letter larger than the other. The larger of the 2 letters has a greater intensity.



Click a box to open a locus details table and pie chart. The display is color-coded, with the nucleotide letters in the locus details table matching the colors of the pie chart.

The rows in the locus detail box can be sorted by Typed, Intensity, or % Total Intensity by clicking their respective column headings.

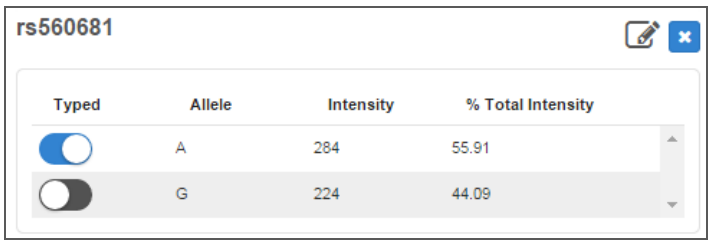
The table and pie chart display the intensity and relative percentage of the signal at the SNP locus position. A toggle switch in the Typed column that displays a blue background and is in position on the right side indicates that analysis settings instructed the ForenSeq Universal Analysis Software to call a particular allele. A toggle switch in the Typed column that displays a gray background and is in position on the left side indicates that analysis settings that drive allele calling are not met.

The legend to the right of the pie chart, and the Intensity column in the table above display the intensity of possible SNP alleles. Mouse over the pie chart segments, or look in the % Total Intensity column of the table to see the percentage of total intensity at the locus.

Table 24 SNP Locus Details Table and Pie Chart Nucleotide Color Indicators

Color	Nucleotide
Green	A
Red	T
Blue	C
Black	G

Updating and Modifying Typed Results



You can change the typed status of a sample in a locus detail box in the STR or SNP Sample Details table.

1 In the Sample details table, click the box to open the table and chart.
A toggle switch in the Typed column that displays a blue background and is in position on the right side indicates that the allele is typed. A toggle switch in the Typed column that displays a gray background and is in position on the left side indicates that the allele is not typed.

2 To change the allele from typed to untyped, click the toggle switch so that the switch displays a gray background and is in position on the left.

3 To change the allele from untyped to typed, click the toggle switch so that the switch displays a blue background and is in position on the right.

The allele can also be changed from typed to untyped, or from untyped to typed, by clicking on the bar (in the case of an STR) or the part of a pie chart (in the case part of a SNP) that represents the allele.

To apply the 2p rule to a homozygous locus, click the toggle switch so that the switch displays a blue background and is in position on the right. For more information, see *Defining the Statistics Calculation Method* on page 92 and *Applying the 2p Rule to a Homozygous Locus* on page 93.

A green message box opens near the bottom of the page that indicates your changes are saved. A user modified icon that looks like a pencil and paper appears in the box in the sample details table and the open details box. If you return the allele to its original typed status, the user modified icon no longer appears. The typed status of a sample cannot be modified while a report is generating that includes the sample.

Updating and Modifying Gender Results

You can change the gender of a sample on the Sample Details page. See *Viewing Sample Details Page* on page 33. The gender of the sample is displayed at the top of the page. When you change the gender of a sample, the change only applies to that particular analysis.

1 To change the gender of the sample, click **Gender**.
The Modify Gender Call box opens.

2 Select a gender.

3 Click **Submit**.

Population Statistics

Introduction 44
Population Statistics Calculations 45
Generating Population Statistics 46
Population Statistics Results 46

Introduction

ForenSeq Universal Analysis Software features the ability to calculate population statistics relative to an allele frequency database. After STR and iSNP multilocus genotyping is complete, population statistics can be calculated automatically. Population groups are installed with the software and their use can be excluded or included in calculations. Customized population groups can also be added for statistical calculation. See *Population Group Settings* on page 92.

Population Statistics Calculations

The ForenSeq Universal Analysis Software performs probability calculations for multilocus autosomal STR profiles. The genotype for each locus must either meet the 2n copy number expected for autosomal loci, or must be uncalled. Calculations follow Scientific Working Group for DNA Analysis Methods (SWGDAM) guidelines. The model assumes that populations are in Hardy-Weinberg equilibrium, that loci are in linkage equilibrium with one another, and a basic correction is included for population substructure.

NOTE

If the copy number is inconsistent, such as 3 alleles typed for an autosomal STR, update and modify the locus calls in order successfully calculate population statistics. See *Updating and Modifying Typed Results* on page 42.

NOTE

Population statistics calculations follow SWGDAM guideline 5.2.1.2. The formula for homozygote genotypes, in accordance with NRCII 4.4a., is p2 + p(1-p)theta, where theta is defined with the population group. For more information on theta correction, see http://www.nfstc.org/pdi/Subject07/pdi_s07_m02_03.htm.

Population statistics calculations use the following information:

- Single-source autosomal STR genotype calls
- Selected population groups, or National Institute of Standards and Technology (NIST) population allele frequency estimates[1]
- Minimum allele frequency, as defined by selected population groups
- Population substructure correction factor θ, when applicable to selected population groups

The ForenSeq Universal Analysis Software queries the population groups and loci defined by the administrator at the time of calculation.

Resources

1. Hill CR, Duewer DL, Kline MC, Coble MD, Butler JM. U.S. population data for 29 autosomal STR loci. *Forensic Sci Int Genet.* 2013;7(3):e82–e83.

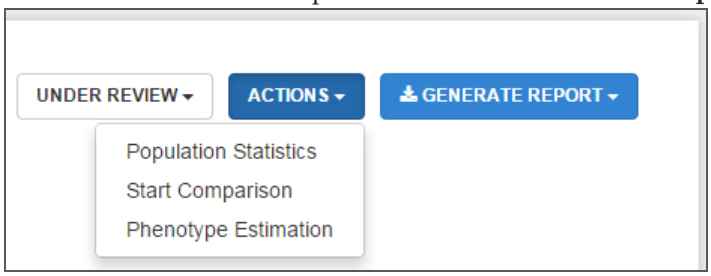
Generating Population Statistics

Follow these instructions to calculate population statistics.

1 From the Project page, locate an analysis on the left pane, and click the blue arrow to open a list of samples. See *Viewing the Project Page* on page 19.

2 Select a sample from the project.

3 Click the blue Actions drop-down list button and select **Population Statistics**.



4 Select Population Groups to include in the random match probability calculation.

5 Click **Generate**.

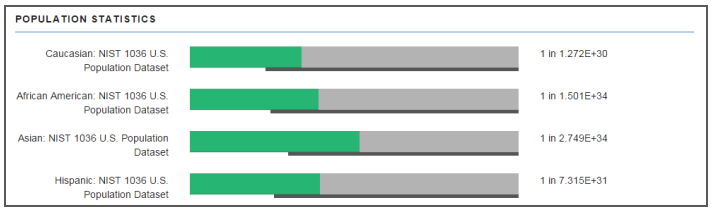
6 A population statistics processing bar shows calculation progress.

7 Population statistics results are displayed in horizontal bars.

NOTE
Population statistics generation can take several minutes. While waiting for population statistics to process, you can navigate to other pages, and perform other functions in the software.

8 To view population statistics results later, return to the Project page, and the Comparisons tab, and select the relevant analysis.

Population Statistics Results



Population group names that are used in the calculation are displayed on the left. Rarity of probability is represented with a horizontal green bar. The horizontal result bar is scaled to the range of possible frequencies for a population group. The range represents all values possible with the population group and the loci used, and is represented on log10 scale. The gray bar represents the source attribution threshold set for the system. To modify source attribution threshold values, see *Changing the Source Attribution Threshold* on page 93.

Hardy-Weinberg Expectations and Linkage Equilibrium

Results of preliminary data analyses of the autosomal STRs and iSNPs in the ForenSeq DNA Signature Prep Kit are consistent with previous experience and publications with genetic identity DNA markers. The autosomal genetic identity markers in the ForenSeq kit generally meet expectations of independence at the population level. The recommendations

of the NRC Report II (1996) should be followed for estimating the rarity of a multilocus DNA profile.[1] For kinship analyses, physically close markers may not be inherited independently. There are 2 options recommended for the use of genetic data for kinship analyses: (1) Incorporate the recombination rate and maximum likelihood estimates of haplotype frequencies for the pair of loci. Note that if the data do not involve double heterozygotes for any individual, then the recombination rate is not needed in the computations. LE assumption allows haplotype frequency estimates from the product of allele frequencies; or (2) use only 1 of the 2 loci in a kinship analysis, which would be the more informative of the 2 in a specific case. (Dr. Bruce Budowle, Ph.D, Executive Director of the University of North Texas Institute of Applied Genetics, and Dr. Ranajit Chakraborty, Ph.D, Director of the Center for Computational Genomics of the University of North Texas Institute of Applied Genetics, personal communication, February 3, 2015.)

Resources

1. Committee on DNA Forensic Science: An Update, National Research Council. The evaluation of DNA forensic evidence. Washington D.C.: National Academies Press; 1996.

Sample Comparison

Introduction 50
Generating a Sample Comparison 51
Sample Comparison Results 52
Generating Sample Compare Population Statistics 54
Comparisons Tab 55

Introduction

The ForenSeq Universal Analysis Software sample comparison feature evaluates the genotyping results between 2 samples using autosomal, X and Y STRs and iSNPs. You can compare samples from the same project, or different projects. Sample comparison compares typed STRs and typed iSNPs from the 2 samples. If DNA Primer Mix B is used, aSNPs and pSNPs can be used in phenotype estimation, but are not included in the comparison function. A typed STR and iSNP locus is designated as discordant, within the context of a sample comparison, if the genotype is not shared by both samples.

Generating a Sample Comparison

Follow these instructions to compare 2 single-source samples.

1 From the Project page, locate an analysis on the left pane, and click the blue arrow to show a list of samples. See *Viewing the Project Page* on page 19.

2 Select a sample from the project. In the sample comparison, this sample is called Sample A.

3 Click the blue Actions drop-down list button and select **Start Comparison**.

4 Click the blue **Choose Sample** button.
A Sample Compare Selection box opens.

5 In the Sample Compare Selection box, select a project from the Project Name drop-down list.

6 A list of samples opens that are associated with the project. Use the search field or scroll through the list of samples from the project to find the second sample you want to compare. In the sample comparison, this sample is called Sample B.

7 Sample comparison starts automatically when you click the blue **Select Sample** button.

8 A sample comparison progress bar indicates how much time is left until the comparison is complete.
The sample comparison page opens.

9 When the comparison is complete, results are displayed.

NOTE
Sample comparison generation can take several minutes. While waiting for sample comparison to process, you can navigate to other pages, and perform other functions in the software.

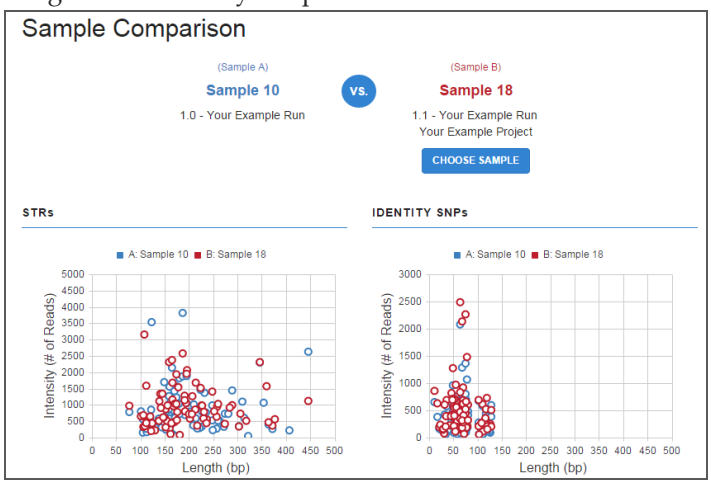
To view sample comparison results later, return to the Project page and click the Comparisons tab. A link to the results page is presented next to a description of the compared samples and the date.

After the sample comparison is complete, you can select another sample to compare by clicking the blue Choose Sample button showing beneath the name of the second sample on the Sample Comparison page.

Sample Comparison Results

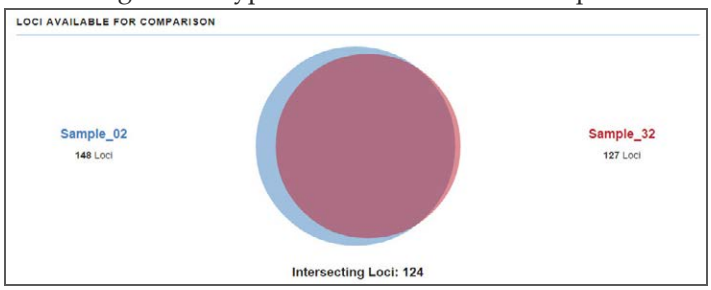
The results for sample comparison are color coded. Blue is for sample A, and red is for sample B. Sample comparison results are illustrated in 3 ways:

- Length and Intensity Graph



Sample intensity (number of reads) and length (bp) information for typed STRs and iSNPs are graphed separately, with STRs on the left, and iSNPs on the right. This graph behaves similarly to the length and intensity graphs on the Sample Details page. For more information, see *Intensity and Length Graphs* on page 33. On the Sample Comparison page, the comparison samples are overlaid on top of each other.

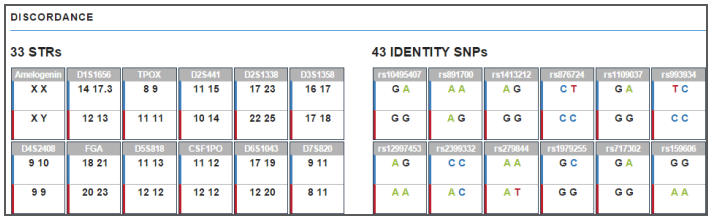
- Venn Diagram of Typed Loci Available for Comparison



A Venn diagram shows the intersection of typed loci from each sample. The total number of typed loci in each sample is represented by a circle. On each side of the diagram is the name of the sample and the total number of typed STR loci and typed iSNPs.

The number of intersecting loci is displayed below the Venn diagram circles. These are the loci that are called in both samples, and are visually represented by the overlap of the sample circles.

- Table of Discordance

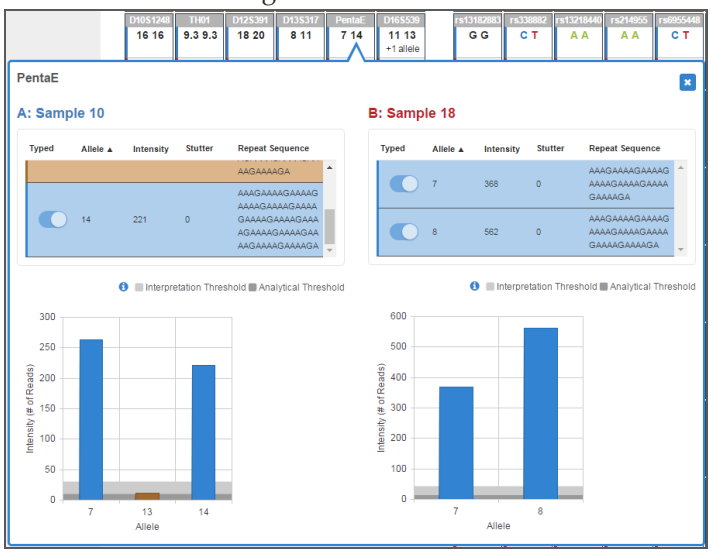


Amelogenin	D1S1656	TPOX	D2S441	D2S1338	D3S1358
X X	14 17.3	8 9	11 15	17 23	16 17
X Y	12 13	11 11	10 14	22 25	17 18
D4S2408	FGA	D5S818	CSF1PO	D6S1043	D7S820
9 10	18 21	11 13	11 12	17 19	9 11
9 9	20 23	12 12	12 12	12 20	8 11

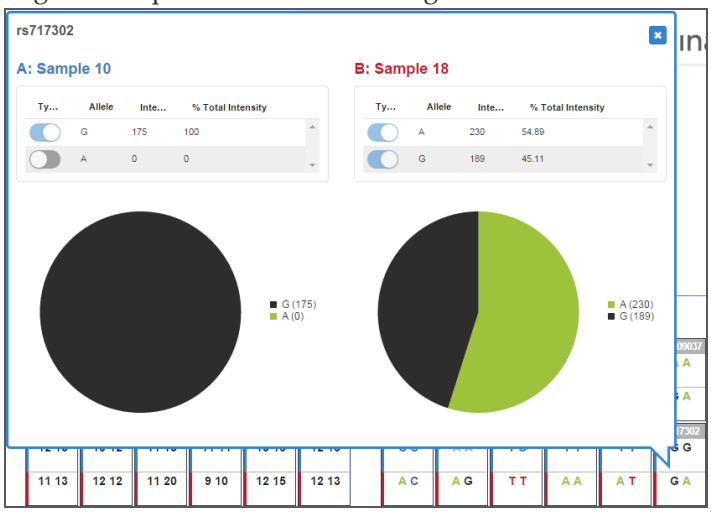
rs10495407	rs891700	rs1413212	rs876724	rs1109037	rs993934
G A	A A	A G	C T	G A	T C
G G	A G	G G	C C	G G	C C
rs12997453	rs2399332	rs276944	rs1979255	rs717302	rs159606
A G	C C	A A	G C	G A	G G
A A	A C	A T	G G	G G	A A

If there are typed loci that do not have the same genotype in both samples, they are displayed in a table showing discordance. Selecting a discordant locus presents the results for both samples at the locus for further comparison.

Tables in the STR detail boxes can be sorted by allele and intensity by clicking their respective column headings.



Tables in the SNP detail boxes can be sorted by allele, intensity, and % total intensity by clicking their respective column headings.



Generating Sample Compare Population Statistics

After a sample comparison is complete, you can generate population statistics from the Sample Compare page for the intersection loci typed in Sample A. All of the loci that overlap between the samples are used in the calculation, regardless of discordance. To generate population statistics for a subset of loci, update and modify locus typing on the Sample Details page. See *Updating and Modifying Typed Results* on page 42.

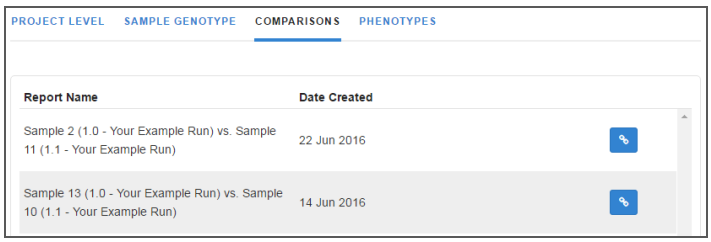
NOTE
To enable generating a comparison, the software requires that each genotype reflects the locus copy number.

1. When a Sample Comparison is complete, scroll to the bottom of the Sample Comparison results page to the Population Statistics section and select the blue **Generate Population Statistics** button.
2. Select Population Groups to include in population statistics calculation.
3. Click **Generate**.
4. A population statistics progress bar indicates how much time is left until calculations are complete.
5. Population statistics results are displayed in horizontal bars.

NOTE
Population statistics generation can take several minutes. While waiting for population statistics to process, you can navigate to other pages, and perform other functions in the software.

To view population statistics results later, return to the Project page and click the Comparisons tab.

Comparisons Tab



The Comparisons tab lists all sample comparisons performed on samples in a project. A blue link button connects to sample comparison results, which are identical to the results that open when a sample comparison is complete. See *Sample Comparison Results* on page 52 and *Generating Sample Compare Population Statistics* on page 54.

1 On the Project page, select the **Comparisons** tab. See *Viewing the Project Page* on page 19.

2 Click the blue **Link** button in the list to display sample comparison results for a sample in the project.

NOTE
If you performed a sample comparison on samples in 2 different projects, access the sample comparison report from the Project page of sample A.

Phenotype and Biogeographical Ancestry Estimation

Chapter 5

Introduction 58
Phenotype and Biogeographical Ancestry Estimation 58
Generating pSNP and aSNP Information 59
Phenotype Estimation SNP Sample Details Table 60
Generating Phenotype Estimations 62
Phenotypes Tab 64

Introduction

The phenotype estimation feature analyzes pSNPs and aSNPs to display genotype results for SNPs that are indicative of hair color, eye color, and biogeographical ancestry relative to major population groups. The use of this feature is optional.

Phenotype and Biogeographical Ancestry Estimation

Probabilities for hair and eye color estimation are obtained from the HIrisPlex model, a multinomial logistic regression model.[1,2,3] Biogeographical ancestry estimation is obtained by principal component analysis (PCA). The model was trained on the 1000 Genomes data.[4] The unknown sample is projected onto the pre-trained first 2 components based on its aSNP genotype calls.

NOTE
For information on 1000 Genomes populations, see 1000genomes.org.

Resources

1. Walsh S, Chaitanya L, Clarisse L, et al. Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. *Forensic Sci Int Genet.* 2014 Mar;9:150–161.

2. Liu F, van Duijn K, Vingerling JR, et al. Eye color and the prediction of complex phenotypes from genotypes. *Curr Biol.* 2009 Mar 10;19(5):R192–193.

3. Walsh S, Lui F, Ballantyne KN, van Oven M, Lao O, Kayser M. IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet.* 2011 June;5(3):170–180.

4. Abecasis GR, Altshuler D, Auton A, et al. A map of human genome variation from population-scale sequencing. *Nature.* 2010 October;28;467(7319):1061–1073.

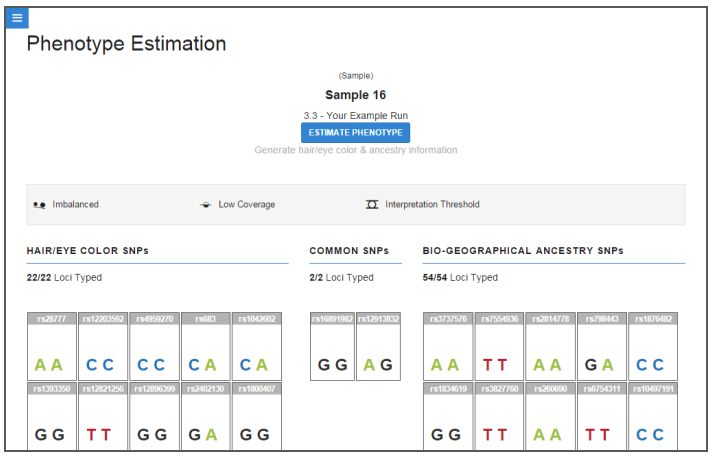
Generating pSNP and aSNP Information

Perform the following steps to generate phenotype estimation SNP information:

1. From the Project page, locate an analysis on the left pane, and click the blue arrow to open a list of samples. See *Viewing the Project Page* on page 19.
2. Select a sample from the project.
3. Click the blue **Actions** drop-down list button.
4. Select **Phenotype Estimation**.
5. The Phenotype Estimation table opens.

NOTE
The phenotype estimation feature is available for samples analyzed with DNA Primer Mix B.

Phenotype Estimation SNP Sample Details Table



The Phenotype Estimation SNP Sample Details table shows data for pSNPs and aSNPs. To generate a phenotype estimation from the SNP information, see *Generating Phenotype Estimations* on page 62. Three columns of SNPs are displayed in the Phenotype Estimation sample details table:

- Hair and eye color (pSNPs and iSNPs)
- SNPs for estimating hair color, eye color, and biogeographical ancestry, listed as Common SNPs
- Biogeographical ancestry (aSNPs)

NOTE

Genotype results for hair, eye color, and common SNPs are required to produce a result for hair and eye color estimation. If 1 of the SNPs is not typed, no result is generated. In contrast, only 1 biogeographical ancestry SNP is required to be typed to produce results for biogeographical ancestry. The absence of SNPs in the biogeographical ancestry estimation diminishes the accuracy of the estimation. The biogeographical ancestry estimation model is retrained on each execution using only the aSNPs with a multilocus genotype.

The following information is displayed above sample locus detail tables:

- **Loci Typed**— The number of pSNP and aSNP target loci identified in the run, out of all possible target loci in the DNA Primer Mix B multiplex.

Each box in the Phenotype Estimation sample details table has a color border that indicates whether the locus is within or outside of quality metrics guidelines. If applicable, quality control indicator icons are displayed in the box. Click on a locus box to display genotyping table and circle plot. For information on the table and pie charts in the SNP locus detail table, see *SNP Sample Details Table* on page 40.

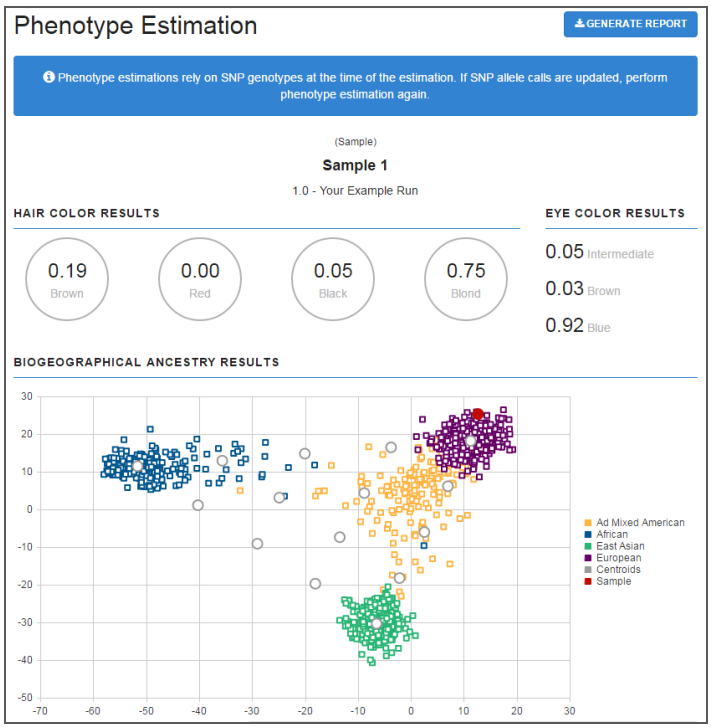
Table 25 Locus Typed Box Color Indicators

Color	Indication
Gray	Quality control indicators are not triggered for any locus.
Orange	One or more quality control indicators are triggered for the locus.

Table 26 Sample Details Quality Control Indicator Icons

Symbol	Symbol Name	Indication
	Imbalanced	Allele balance percentage falls below the defined ratio for intralocus balance in the analysis settings. For more information on imbalance, see *SNP Genotype Calling* on page 102.
	Low Coverage	Signal above the interpretation threshold was not detected as defined in the analysis settings. The low coverage indicator can be seen in results from version 1.1 or earlier of the ForenSeq Universal Analysis Software. See *SNP Genotype Calling* on page 102.
	Interpretation Threshold	Indicates that the locus has at least 1 typed allele, and at least 1 allele is above the analytical threshold but below the interpretation threshold.
	Analytical Threshold	Indicates that the locus has signal below the analysis threshold, and no alleles above the interpretation threshold.
	Not Detected	No signal was detected for the locus.

Generating Phenotype Estimations



Perform the following steps to generate a phenotype estimation:

1 From the Phenotype Estimation page, click the blue **Estimate Phenotype** button. For information on how to access the Phenotype Estimation page, see *Phenotype and Biogeographical Ancestry Estimation* on page 58.

2 A phenotype estimation progress bar indicates how much time is left until the estimation is complete.

3 The Phenotype Estimation chart opens.

NOTE
If SNP allele calls are updated after you generate a phenotype estimation, perform phenotype estimation again.

- **Hair Color Results**— Hair phenotype percentages based in the HIrisPlex model described by Wash et al.[1]
- **Eye Color Results**— Eye color phenotype percentages based in the HIrisPlex model described by Wash et al.
- **Biogeographical Ancestry Results**— Graph of the result from the principal component analysis of the sample data set. A training data set of 1000 Genomes populations is used to create the principal component analysis model. Then the results from several 1000 Genomes populations and the unknown sample are plotted with the results of the model.

NOTE
For information on 1000 Genomes populations, see 1000genomes.org.

Centroids give perspective to the results of the estimation and provide logical groupings of populations. The centroids are represented by the gray centers of the clusters. There are 3 centroids for 3 major ancestries (AFR, ASN, EUR in the middle of the blue, purple, and

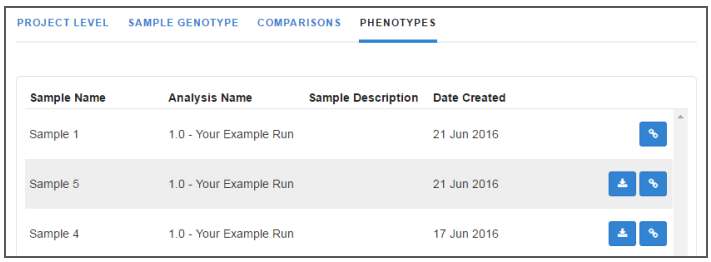
green clusters), and then orientational centroids at the one-quarter intervals between these 3 large groups.

- **Sample**— Displayed on the chart in red. Your unknown sample.
- **Distance to Nearest Centroid**— Proximity of the sample to the nearest centroid. This measurement gives an indication of how related the sample is to the general grouping for the centroid. For comparison, the distance is provided for the 1000 Genome samples contributing to the centroid.
- **SNP Panel Used**— Number of typed aSNPs applied to the biogeographical ancestry model, over all SNPs contained in DNA Primer Mix B.
- **1000 Genomes populations with samples in centroid with sample**— Click the black arrow to access drop-down list below the graph.
- **Reference samples in centroid with sample**— Click the black arrow to access the drop-down list below the graph.

Resources

1. Walsh S, Chaitanya L, Clarisse L, et al. Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. *Forensic Sci Int Genet.* 2014 Mar;9:150-161.

Phenotypes Tab



The Phenotypes tab is a list of all phenotype estimations performed on samples in a project. See *Phenotype and Biogeographical Ancestry Estimation* on page 58. A blue link button connects to phenotype estimation results, which are identical to the results that open when a phenotype estimation is complete. See *Generating Phenotype Estimations* on page 62.

1 On the Project page, select the **Phenotypes** tab. See *Viewing the Project Page* on page 19.

2 Click the blue **Link** button in the list to display phenotype estimation results for a sample in the project.

Reports

Introduction 66
Project Level Report 67
Sample Genotype Report 70
Phenotype Estimation Report 75

Introduction

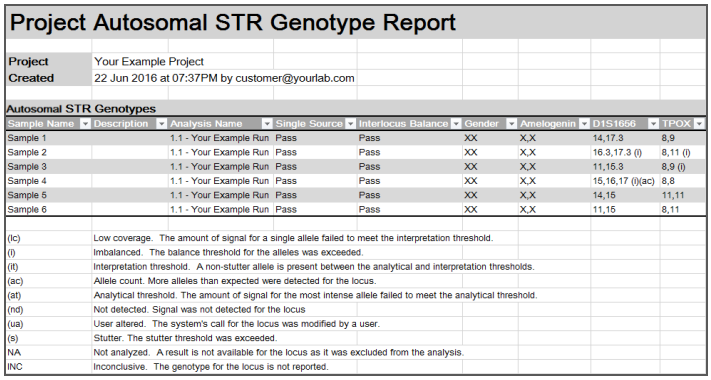
The ForenSeq Universal Analysis Software features the ability to generate and view the following reports:

- Project-level genotype
- Sample-level genotype
- Sample-level phenotype estimation

After report generation, links to analyses and reports are available in separate tabs on the Project page for later reference.

To create a Project Level genotype report, click the blue Create Report button on the Project page. The Phenotype Estimation report is generated on the Phenotype Estimation page. The Sample Genotype report can be created from the Project page or the Sample Details page.

Project Level Report



Project Autosomal STR Genotype Report

Project	Your Example Project
Created	22 Jun 2016 at 07:37PM by customer@yourlab.com

Autosomal STR Genotypes

Sample Name	Description	Analysis Name	Single Source	Interlocus Balance	Gender	Amelogenin	D1S1656	TPOX
Sample 1		1.1 - Your Example Run	Pass	Pass	XX	X,X	14,17.3	8,9
Sample 2		1.1 - Your Example Run	Pass	Pass	XX	X,X	16.3,17.3 (i)	8,11 (i)
Sample 3		1.1 - Your Example Run	Pass	Pass	XX	X,X	11,15.3	8,9 (i)
Sample 4		1.1 - Your Example Run	Pass	Pass	XX	X,X	15,16,17 (i)(ac)	8,8
Sample 5		1.1 - Your Example Run	Pass	Pass	XX	X,X	14,15	11,11
Sample 6		1.1 - Your Example Run	Pass	Pass	XX	X,X	11,15	8,11

(lc)	Low coverage. The amount of signal for a single allele failed to meet the interpretation threshold.
(i)	Imbalanced. The balance threshold for the alleles was exceeded.
(it)	Interpretation threshold. A non-stutter allele is present between the analytical and interpretation thresholds.
(ac)	Allele count. More alleles than expected were detected for the locus.
(at)	Analytical threshold. The amount of signal for the most intense allele failed to meet the analytical threshold.
(nd)	Not detected. Signal was not detected for the locus
(ua)	User altered. The system's call for the locus was modified by a user.
(s)	Stutter. The stutter threshold was exceeded.
NA	Not analyzed. A result is not available for the locus as it was excluded from the analysis.
INC	Inconclusive. The genotype for the locus is not reported.

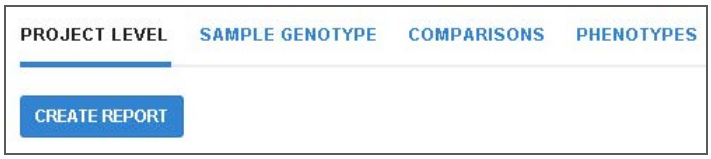
The Project Level report is a genotype report of results and supporting information for the STR loci and iSNPs across all samples and analyses in a project. The report is a Microsoft Excel workbook file that you can download and customize by using sorting filters and drop-down list. After Project Level reports are created, they are available for download from the Project page.

The Project Level genotype report workbook lists every sample name and its description, and includes single-source, interlocus balance, and gender results. The report shows the typed alleles for each locus of each sample of each analysis in a project. The worksheets with coverage information present the read counts and alleles for each typed allele from every sample in each analysis that is in a project.

If multiple analyses of the samples in the project exist, the order of the samples in the workbook is arranged by analysis name, with the most recent analysis appearing at the top of the list. The workbook is subdivided into 8 worksheets:

- Autosomal STRs
- Autosomal STR coverage
- Y-STRs
- Y-STR coverage
- X-STRS
- X-STR coverage
- iSNPs
- iSNP coverage

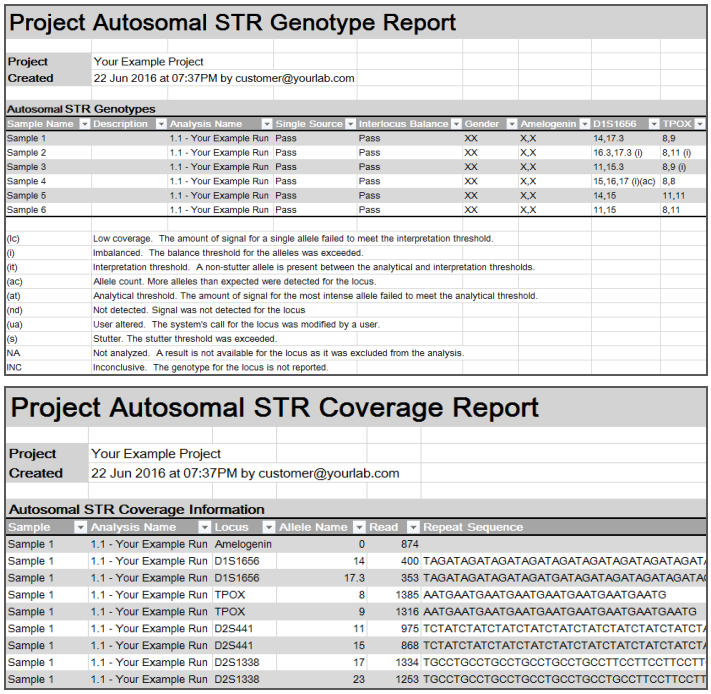
1. On the Project page, click the blue **Create Report** button. See *Viewing the Project Page* on page 19.
   On the right side of the report list, a Pending button appears for each report generated. When the report is complete, the Pending button becomes a Download button.



2. Click the blue **Download** button to access the report.

Autosomal STR Worksheets for Project Level Reports

The Autosomal STR worksheets in the Project Level report feature a genotype report and a coverage report. The reports show locus genotypes, quality control indicators, number of reads, and for STRs the nucleotide sequence of typed alleles. Alleles of the same length with different sequences at the locus are highlighted with a different font size. STRs with isometric alleles have sequence differences highlighted by font size. For information on the meaning of column headings and quality control indicators in the body of the worksheet, see *General Locus Information* on page 34 and *STR Sample Details Table* on page 37.



Project Autosomal STR Genotype Report

Project	Your Example Project
Created	22 Jun 2016 at 07:37PM by customer@yourlab.com

Autosomal STR Genotypes

Sample Name	Description	Analysis Name	Single Source	Interlocus Balance	Gender	Amelogenin	D1S1656	TPOX
Sample 1		1.1 - Your Example Run	Pass	Pass	XX	X,X	14,17.3	8,9
Sample 2		1.1 - Your Example Run	Pass	Pass	XX	X,X	16.3,17.3 (i)	8,11 (i)
Sample 3		1.1 - Your Example Run	Pass	Pass	XX	X,X	11,15.3	8,9 (i)
Sample 4		1.1 - Your Example Run	Pass	Pass	XX	X,X	15,16,17 (i)(ac)	8,8
Sample 5		1.1 - Your Example Run	Pass	Pass	XX	X,X	14,15	11,11
Sample 6		1.1 - Your Example Run	Pass	Pass	XX	X,X	11,15	8,11

(lc)	Low coverage. The amount of signal for a single allele failed to meet the interpretation threshold.
(i)	Imbalanced. The balance threshold for the alleles was exceeded.
(it)	Interpretation threshold. A non-stutter allele is present between the analytical and interpretation thresholds.
(ac)	Allele count. More alleles than expected were detected for the locus.
(at)	Analytical threshold. The amount of signal for the most intense allele failed to meet the analytical threshold.
(nd)	Not detected. Signal was not detected for the locus
(ua)	User altered. The system's call for the locus was modified by a user.
(s)	Stutter. The stutter threshold was exceeded.
NA	Not analyzed. A result is not available for the locus as it was excluded from the analysis.
INC	Inconclusive. The genotype for the locus is not reported.

Project Autosomal STR Coverage Report

Project	Your Example Project
Created	22 Jun 2016 at 07:37PM by customer@yourlab.com

Autosomal STR Coverage Information

Sample	Analysis Name	Locus	Allele Name	Read	Repeat Sequence
Sample 1	1.1 - Your Example Run	Amelogenin	0	874	
Sample 1	1.1 - Your Example Run	D1S1656	14	400	TAGATAGATAGATAGATAGATAGATAGATAGATAGAT
Sample 1	1.1 - Your Example Run	D1S1656	17.3	353	TAGATAGATAGATAGATGATAGATAGATAGATAGATA
Sample 1	1.1 - Your Example Run	TPOX	8	1385	AATGAATGAATGAATGAATGAATGAATG
Sample 1	1.1 - Your Example Run	TPOX	9	1316	AATGAATGAATGAATGAATGAATGAATGAATG
Sample 1	1.1 - Your Example Run	D2S441	11	975	TCTATCTATCTATCTATCTATCTATCTATCTATCTATCTA
Sample 1	1.1 - Your Example Run	D2S441	15	868	TCTATCTATCTATCTATCTATCTATCTATCTATCTATCTA
Sample 1	1.1 - Your Example Run	D2S1338	17	1334	TGCCTGCCTGCCTGCCTGCCTGCCTTCCTTCCTTCCTT
Sample 1	1.1 - Your Example Run	D2S1338	23	1253	TGCCTGCCTGCCTGCCTGCCTGCCTGCCTTCCTTCCTT

Y-STR and X-STR Worksheets for Project Level Reports

The Y-STR and X-STR worksheets in the Project Level report feature a genotype report and a coverage report. The reports show locus genotypes, quality control indicators, number of reads, and the nucleotide sequence of typed alleles for all samples and all analyses in a project. Alleles of the same length with different sequences at the locus are highlighted with a different font size. For information on the meaning of column headings and quality control indicators in the body of the worksheet, see *General Locus Information* on page 34 and *STR Sample Details Table* on page 37.

Project X STR Genotype Report

Project	Your Example Project
Created	22 Jun 2016 at 07:37PM by customer@yourlab.com

X STR Genotypes

Sample Name	Description	Analysis Name	Single Source	Interlocus Balance	Gender	DXS10135	DXS8378
Sample 1		1.1 - Your Example Run	Pass	Pass	XX	19.1,21	10,10
Sample 2		1.1 - Your Example Run	Pass	Pass	XX	24,25	10,10
Sample 3		1.1 - Your Example Run	Pass	Pass	XX	24,25,35,36 (i)(ac)	11,14
Sample 4		1.1 - Your Example Run	Pass	Pass	XX	17,28.1 (i)	10,12 (i)
Sample 5		1.1 - Your Example Run	Pass	Pass	XX	26,27 (i)(ac)(it)(s)	11,12
Sample 6		1.1 - Your Example Run	Pass	Pass	XX	25,29	11,12
Sample 7		1.1 - Your Example Run	Pass	Pass	XX	18,28	10,11

(lc)	Low coverage. The amount of signal for a single allele failed to meet the interpretation threshold.
(i)	Imbalanced. The balance threshold for the alleles was exceeded.
(it)	Interpretation threshold. A non-stutter allele is present between the analytical and interpretation thresholds.
(ac)	Allele count. More alleles than expected were detected for the locus.
(at)	Analytical threshold. The amount of signal for the most intense allele failed to meet the analytical threshold.
(nd)	Not detected. Signal was not detected for the locus
(ua)	User altered. The system's call for the locus was modified by a user.
(s)	Stutter. The stutter threshold was exceeded.
NA	Not analyzed. A result is not available for the locus as it was excluded from the analysis.
INC	Inconclusive. The genotype for the locus is not reported.

Project X STR Coverage Report

Project	Your Example Project
Created	22 Jun 2016 at 07:37PM by customer@yourlab.com

X STR Coverage Information

Sample	Analysis Name	Locus	Allele Name	Read	Repeat Sequence
Sample 1	1.1 - Your Example Run	DXS10135	19.1	1043	AAGAAAGAAAGAGAAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG
Sample 1	1.1 - Your Example Run	DXS10135	21	1103	AAGAAAGAAAGAGAAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG
Sample 1	1.1 - Your Example Run	DXS8378	10	6508	CTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT
Sample 1	1.1 - Your Example Run	DXS7132	13	1312	TAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGA
Sample 1	1.1 - Your Example Run	DXS7132	14	748	TAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG
Sample 1	1.1 - Your Example Run	DXS10074	15	2194	AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGGAA
Sample 1	1.1 - Your Example Run	DXS10074	16	1768	AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAA
Sample 1	1.1 - Your Example Run	DXS10103	17	87	TAGATTGACAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG

iSNP Worksheets for Project Level Reports

The iSNP worksheets in the Project Level report feature a genotype report and a coverage report. The reports show genotypes, quality control indicators, and the number of reads for typed alleles. For information on the meaning of column headings in the body of the worksheet, see *General Locus Information* on page 34 and *SNP Sample Details Table* on page 40.

Project iSNP Genotype Report

Project	Your Example Project
Created	22 Jun 2016 at 07:37PM by customer@yourlab.com

iSNP Genotypes

Sample Name	Description	Analysis Name	Single Source	Gender	rs1490413	rs560681	rs1294331
Sample 1		1.1 - Your Example Run	Pass	XX	A,A	A,G	A,A (it)
Sample 2		1.1 - Your Example Run	Pass	XX	G,A	A,A	G,G (it)
Sample 3		1.1 - Your Example Run	Fail	XX	G,A (i)	A,A	G,G
Sample 4		1.1 - Your Example Run	Fail	XX	G,G	A,G	G,G
Sample 5		1.1 - Your Example Run	Pass	XX	G,A	A,A	INC (it)
Sample 6		1.1 - Your Example Run	Pass	XX	G,A	A,G	G,A (i)
Sample 7		1.1 - Your Example Run	Pass	XX	G,G	A,A	INC (nd)

(lc)	Low coverage. The amount of signal for a single allele failed to meet the interpretation threshold.
(i)	Imbalanced. The balance threshold for the alleles was exceeded.
(it)	Interpretation threshold. A non-stutter allele is present between the analytical and interpretation thresholds.
(at)	Analytical threshold. The amount of signal for the most intense allele failed to meet the analytical threshold.
(nd)	Not detected. Signal was not detected for the locus
(ua)	User altered. The system's call for the locus was modified by a user.
NA	Not analyzed. A result is not available for the locus as it was excluded from the analysis.
INC	Inconclusive. The genotype for the locus is not reported.

Coverage Information

Locus	Allele Name	Typed Allele?	Reads
rs3737576	A	Yes	53
rs3737576	G	Yes	47
rs7554936	C	Yes	682
rs7554936	T	No	0
rs2814778	A	Yes	3870
rs2814778	G	No	0
rs798443	A	Yes	596
rs798443	G	No	0
rs1876482	C	Yes	1233
rs1876482	T	No	0

Sample Genotype Report

The Sample Genotype report is composed of data and charts of results and supporting information for the STR and iSNP loci of a single sample. The report provides the same information as in the sample locus details table. See *STR Sample Details Table* on page 37 and *SNP Sample Details Table* on page 40.

The report is a Microsoft Excel workbook file that you can download and customize by using sorting filters, and drop-down lists. After Sample Genotype reports are created, they are available for download from the Project page on the Sample Genotype tab.

The workbook is subdivided into 9 worksheets:

- Autosomal STRs
- Autosomal STR figure
- Y-STRs
- Y-STR figure
- X-STRs
- X-STR figure
- iSNPs
- iSNP figure
- Settings

1. On the Project page, select the **Sample Genotype** tab.
2. Click the blue **Generate Report** button.
3. Select **Sample Details** or **Sample Summary**.
   The Sample Details report includes typed and untyped alleles. The Sample Summary report includes typed alleles only.
4. In the Generate Reports box, select an analysis name from the drop-down list and select samples.
5. Click the blue **Generate** button.
   On the right side of the report list, a Pending button appears for each report generated. When the report is complete, the Pending button becomes a Download button.
6. Click the blue **Download** button to access the report.

NOTE
If Excel gives you a Protected View warning when you open a Sample Genotype report, click the Enable Editing button to view all the features of the workbook.

Sample Genotype reports for single samples can also be generated from the Sample Details page.

1. On the Project page, locate an analysis on the left pane of the page. See *Viewing the Project Page* on page 19.
2. Click the blue arrow to open a list of samples.
3. Click a sample in the list to open the Sample Details page.
4. Click the blue **Generate Report** button.
5. Select **Sample Details** or **Sample Summary**.

6 Click the blue **Generate** button.

Autosomal STR Worksheets for Sample Level Reports

The Autosomal STR worksheets in the Sample Genotype report show locus genotypes, quality control indicators, number of reads, and for STRs the nucleotide sequence of typed alleles. Alleles of the same length with different sequences at the locus are highlighted with a different font size. For information on the meaning of column headings and quality control indicators in the body of the worksheet, see *General Locus Information* on page 34 and *STR Sample Details Table* on page 37.

Sample Autosomal STR Report

Sample	Sample 12
Project	Your Example Project
Analysis	1.0 - Your Example Run
Run	Your Example Run
Called Gender	XY
Created	2014/10/30 1:25:47 AM by customer@yourlab.com

Autosomal STR Locus Information

Loci Typed	28 / 28
Single Source	Pass
Interlocus Balance	Pass

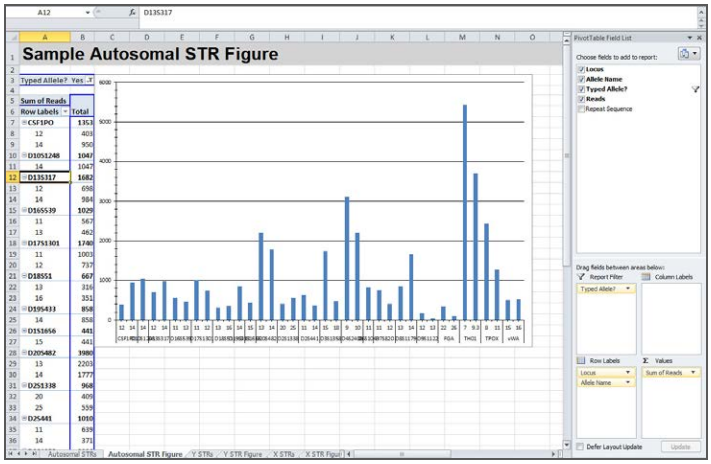
Locus	Genotype	Flags
Amelogenin	X,Y	
CSF1PO	12,13	
D10S1248	14,15	imbalanced
D12S391	19,19.3	
D13S317	12,13	
D16S539	12,13	
D17S1301	11,12	
D18S51	14,17	
D19S433	12,12	
D1S1656	12,13	

Coverage Information

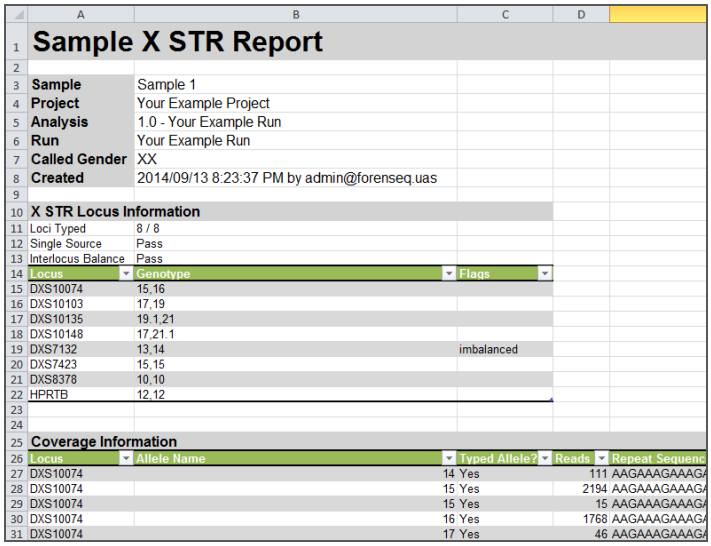
Locus	Allele Name	Typed Allele?	Reads	Repeat Sequence
Amelogenin	0	Yes	874	
CSF1PO	10	Yes	23	AGATAGATAGATAGATAGATAGATAGATAGATAGA
CSF1PO	11	Yes	694	AGATAGATAGATAGATAGATAGATAGATAGATAGA
CSF1PO	12	Yes	693	AGATAGATAGATAGATAGATAGATAGATAGATAGA
D10S1248	13	Yes	81	GGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
D10S1248	14	Yes	820	GGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
D10S1248	16	Yes	79	GGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
D10S1248	17	Yes	586	GGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
D12S391	16	Yes	16	AGATAGATAGATAGATAGATAGATAGATAGATAGA
D12S391	17	Yes	11	AGATAGATAGATAGATAGATAGATAGATAGATAGA
D12S391	17	Yes	119	AGATAGATAGATAGATAGATAGATAGATAGATAGA
D12S391	18	Yes	26	AGATAGATAGATAGATAGATAGATAGATAGATAGA
D12S391	18	Yes	874	AGATAGATAGATAGATAGATAGATAGATAGATAGA

For STRs, the figure worksheet shows a pivot table and pivot chart. The data for the table and chart in the figure worksheet are populated from the autosomal STR worksheet. The pivot chart illustrates the typed alleles for all the autosomal STRs. The pivot table is a summation of the reads for the typed alleles at each locus. The Typed Allele column default setting shows the value of typed alleles (Typed Allele = Yes). However you can modify the setting by clicking the filter icon to the right of the column heading.

Figure 1 Sample Genotype Report Autosomal STR Figure Worksheet



Y-STR and X-STR Worksheets for Sample Level Reports



Sample X STR Report

Sample	Sample 1
Project	Your Example Project
Analysis	1.0 - Your Example Run
Run	Your Example Run
Called Gender	XX
Created	2014/09/13 8:23:37 PM by admin@forenseq.uas

X STR Locus Information

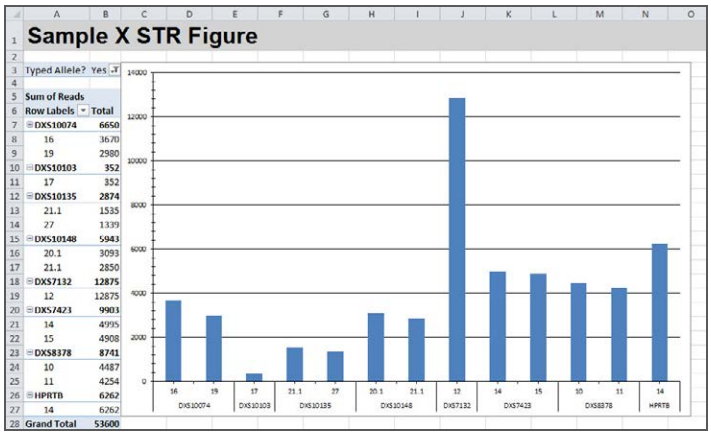
Loci Typed	8 / 8
Single Source	Pass
Interlocus Balance	Pass

Locus	Genotype	Flags
DXS10074	15,16	
DXS10103	17,19	
DXS10135	19.1,21	
DXS10148	17,21.1	
DXS7132	13,14	imbalanced
DXS7423	15,15	
DXS8378	10,10	
HPRTB	12,12	

Coverage Information

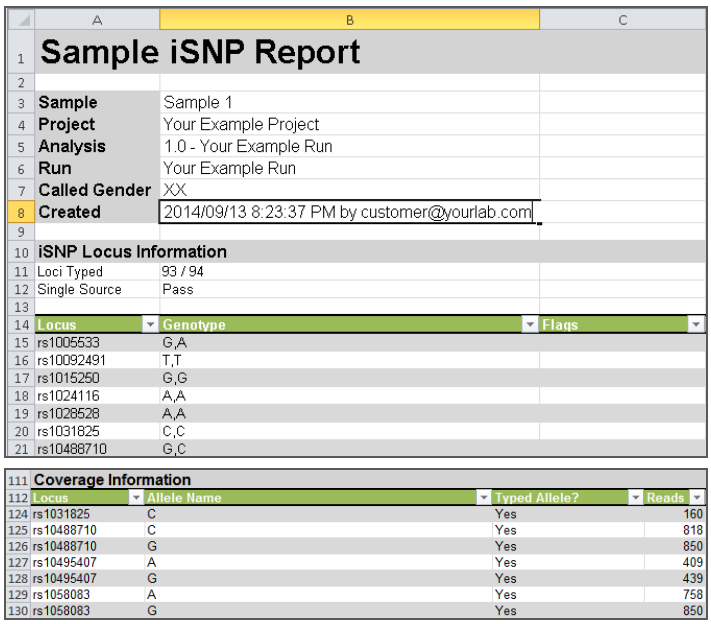
Locus	Allele Name	Typed Allele?	Reads	Repeat Sequenc
DXS10074	14	Yes	111	AAGAAAGAAAGA
DXS10074	15	Yes	2194	AAGAAAGAAAGA
DXS10074	15	Yes	15	AAGAAAGAAAGA
DXS10074	16	Yes	1768	AAGAAAGAAAGA
DXS10074	17	Yes	46	AAGAAAGAAAGA

The Y-STR and X-STR worksheets in the Sample Genotype report show locus genotypes, quality control indicators, number of reads, and the nucleotide sequence of typed alleles. Alleles of the same length with different sequences at the locus are highlighted with a different font size. For information on the meaning of column headings and quality control indicators in the body of the worksheet, see *General Locus Information* on page 34 and *STR Sample Details Table* on page 37.



The figure worksheets show a pivot table and pivot chart for Y-STRs and X-STRs. The data for the table and chart in the figure worksheets are populated from their respective worksheets. The pivot chart illustrates the information in the table on the left of the read sums per locus. The pivot table is a summation of the reads for an allele at each locus.

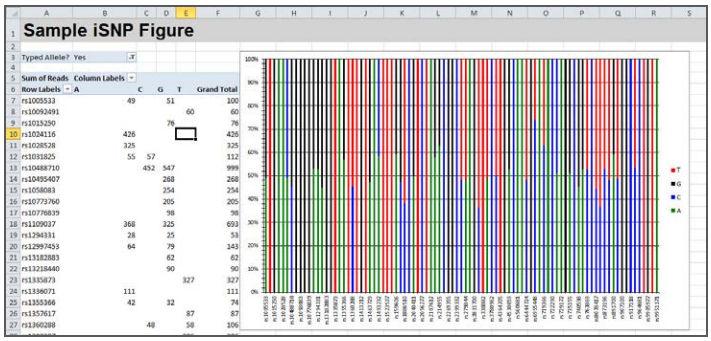
iSNP Worksheet for Sample Level Reports



	A	B	C
1	**Sample iSNP Report**		
2			
3	**Sample**	Sample 1	
4	**Project**	Your Example Project	
5	**Analysis**	1.0 - Your Example Run	
6	**Run**	Your Example Run	
7	**Called Gender**	XX	
8	**Created**	2014/09/13 8:23:37 PM by customer@yourlab.com	
9			
10	**iSNP Locus Information**		
11	Loci Typed	93 / 94	
12	Single Source	Pass	
13			
14	**Locus**	**Genotype**	**Flags**
15	rs1005533	G,A	
16	rs10092491	T,T	
17	rs1015250	G,G	
18	rs1024116	A,A	
19	rs1028528	A,A	
20	rs1031825	C,C	
21	rs10488710	G,C	

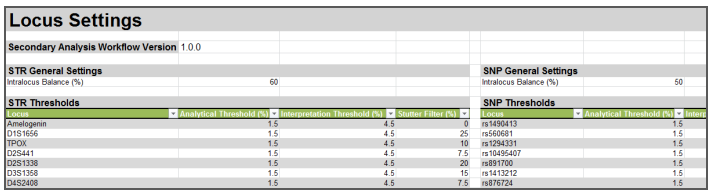
111	**Coverage Information**			
112	**Locus**	**Allele Name**	**Typed Allele?**	**Reads**
124	rs1031825	C	Yes	160
125	rs10488710	C	Yes	818
126	rs10488710	G	Yes	850
127	rs10495407	A	Yes	409
128	rs10495407	G	Yes	439
129	rs1058083	A	Yes	758
130	rs1058083	G	Yes	850

The iSNP worksheets in the Sample Genotype report show genotypes, quality control indicators, and the number of reads for typed alleles. For information on the meaning of column headings in the body of the worksheet, see *General Locus Information* on page 34 and *SNP Sample Details Table* on page 40.



The figure worksheet shows a pivot table and pivot chart for iSNPs. The data for the table and chart in the figure worksheet are populated from the worksheet. The pivot chart illustrates the information in the table on the left of the read sums per locus. In the pivot table, each of the iSNP loci are listed. The first column is the locus name, identified as an rs number. The second, third, fourth, and fifth columns list the possible alleles (ACGT) and the number of reads. The number of reads for each sample is listed by typed SNP allele in the Grand Total column. The chart is a 100% stacked bar chart representation of the percentage each allele contributes to a genotype.

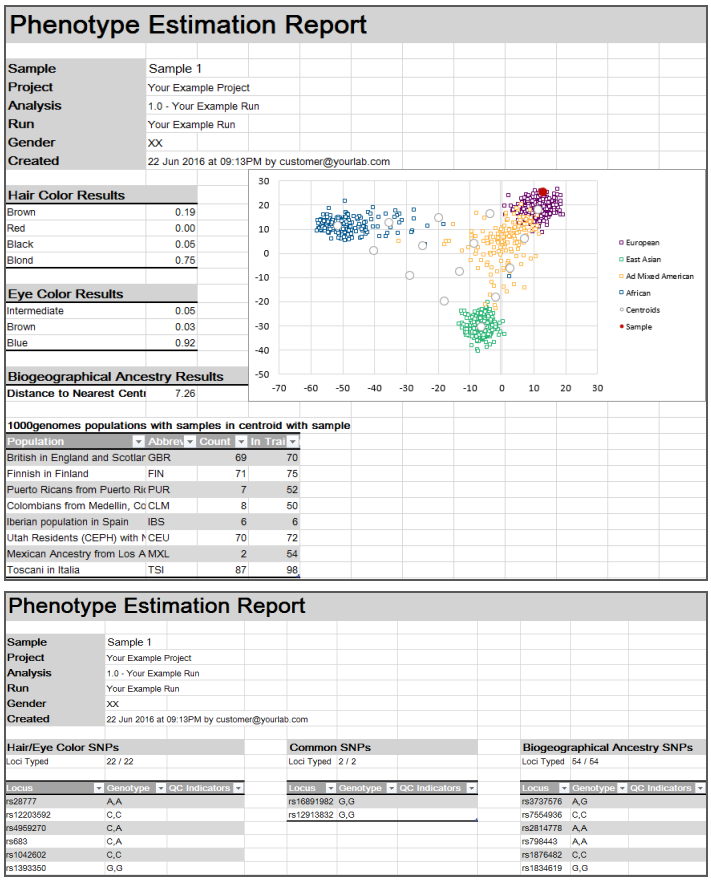
Locus Level Settings for Sample Level Reports



The Settings worksheet displays the analytical and interpretation thresholds and the stutter filters in use for the analysis of the sample.

Phenotype Estimation Report

The Phenotype Estimation report is composed of data and charts of results and supporting information for a single sample. The report provides the same information as a phenotype estimation, with the exception of reference samples in sample. See *Generating Phenotype Estimations* on page 62 and *Phenotype Estimation SNP Sample Details Table* on page 60.



Phenotype Estimation Report

Sample	Sample 1
Project	Your Example Project
Analysis	1.0 - Your Example Run
Run	Your Example Run
Gender	XX
Created	22 Jun 2016 at 09:13PM by customer@yourlab.com

Hair Color Results	
Brown	0.19
Red	0.00
Black	0.05
Blond	0.75

Eye Color Results	
Intermediate	0.05
Brown	0.03
Blue	0.92

Biogeographical Ancestry Results	
Distance to Nearest Centr	7.26

1000genomes populations with samples in centroid with sample

Population	Abbrev	Count	In Tra
British in England and Scotlar	GBR	69	70
Finnish in Finland	FIN	71	75
Puerto Ricans from Puerto Ri	PUR	7	52
Colombians from Medellin, Co	CLM	8	50
Iberian population in Spain	IBS	6	6
Utah Residents (CEPH) with N	CEU	70	72
Mexican Ancestry from Los A	MXL	2	54
Toscani in Italia	TSI	87	98

Phenotype Estimation Report

Sample	Sample 1
Project	Your Example Project
Analysis	1.0 - Your Example Run
Run	Your Example Run
Gender	XX
Created	22 Jun 2016 at 09:13PM by customer@yourlab.com

Hair/Eye Color SNPs	
Loci Typed	22 / 22

Locus	Genotype	QC Indicators
rs28777	A,A	
rs12203592	C,C	
rs4959270	C,A	
rs683	C,A	
rs1042602	C,C	
rs1393350	G,G	

Common SNPs	
Loci Typed	2 / 2

Locus	Genotype	QC Indicators
rs16891982	G,G	
rs12913832	G,G	

Biogeographical Ancestry SNPs	
Loci Typed	54 / 54

Locus	Genotype	QC Indicators
rs3737576	A,G	
rs7554936	C,C	
rs2814778	A,A	
rs798443	A,A	
rs1876482	C,C	
rs1834619	G,G	

Figure 2 Phenotype Estimation Report SNP Data Coverage

Coverage Information

Locus	Allele Name	Typed Allele?	Reads
rs3737576	A	Yes	53
rs3737576	G	Yes	47
rs7554936	C	Yes	682
rs7554936	T	No	0
rs2814778	A	Yes	3870
rs2814778	G	No	0
rs798443	A	Yes	596
rs798443	G	No	0
rs1876482	C	Yes	1233
rs1876482	T	No	0

The report is a Microsoft Excel workbook file that you can download and customize by using sorting filters, and drop-down lists. After a phenotype estimation for a sample is performed, a report can be created from the estimation. They are available for download from the Project page on the Phenotype tab.

The workbook is subdivided into 4 worksheets:

- Estimation
- SNP Data

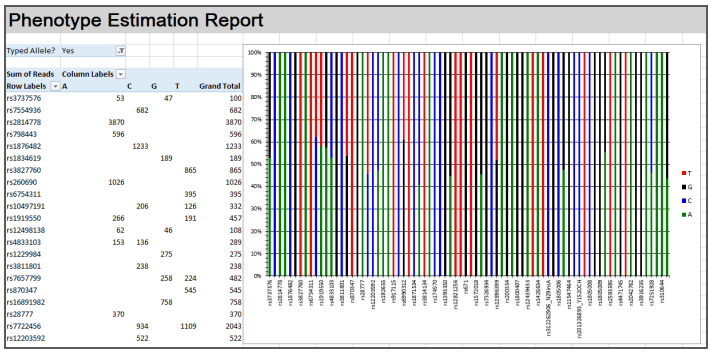
- SNP Balance Figure
- Settings

1. On the Project page, select the **Phenotypes** tab.
2. Click the blue link icon for a sample in the list.
   The samples on the list are ones that have a phenotype estimation.
3. Click the blue **Generate Report** button.
   When the report is complete, the Download icon displays.
4. Click the blue **Download** button to access the report.

Phenotype Estimation reports for single samples can also be generated from the Phenotype Estimation page after the estimation is complete. See *Generating Phenotype Estimations* on page 62.

NOTE
Typed alleles in the report reflect current data. If SNP allele calls are updated after you generate a phenotype estimation, perform phenotype estimation again and generate another report.

SNP Balance Worksheet for Phenotype Estimation Reports


Phenotype Estimation Report

Typed Allele? Yes

Sum of Reads	Column Labels				
Row Labels	A	C	G	T	Grand Total
rs3737576		53		47	100
rs7554936			682		682
rs2814778		3870			3870
rs798443		596			596
rs1876482			1233		1233
rs1834619				189	189
rs3827760				865	865
rs260690		1026			1026
rs6754311				395	395
rs10497191			206	126	332
rs1919550		266		191	457
rs12498138		62	46		108
rs4833103		153	136		289
rs1229984			275		275
rs3811801			238		238
rs7657799				258	224	482
rs870347				545	545
rs16891982				758	758
rs28777		370			370
rs7722456			934	1109	2043
rs12203592			522		522

The SNP balance figure worksheet shows a pivot table and pivot chart for SNPs in use for the phenotype estimation report. The data for the table and chart in the SNP balance figure worksheet are populated from the worksheet. The pivot chart illustrates the information in the table on the left of the read sums per locus. In the pivot table, each of the SNP loci are listed. The first column is the locus name, identified as an rs number. The second, third, fourth, and fifth columns list the possible alleles (ACGT) and the number of reads. The number of reads for each sample is listed by typed SNP allele in the Grand Total column. The chart is a 100% stacked bar chart representation of the percentage each allele contributes to a genotype

Locus Settings for Phenotype Estimation Reports

Phenotype Estimation Locus Settings

Secondary Analysis Module Version	1.0.14351	
SNP General Settings		
Intralocus Balance (%)	50	
SNP Thresholds		
Locus	**Analytical Threshold (%)**	**Interpretation Threshold (%)**
rs3737576	1.5	4.5
rs7554936	1.5	4.5
rs2814778	1.5	4.5
rs798443	1.5	4.5
rs1876482	1.5	4.5
rs1834619	1.5	4.5
rs3827760	1.5	4.5

The Settings worksheet displays the analytical and interpretation thresholds for loci in the phenotype estimation of the sample. The worksheet also displays the intralocus balance setting for SNPs in the analysis.

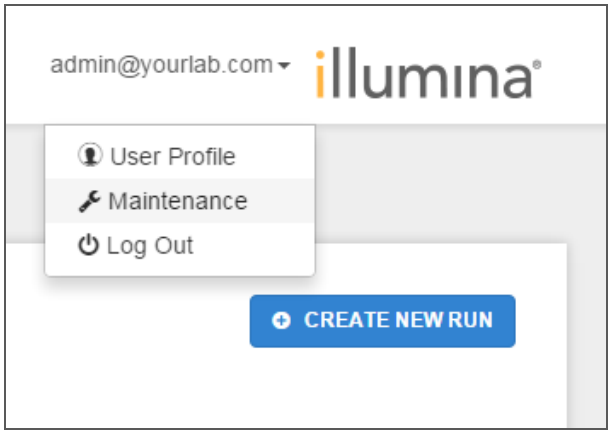
System Settings

Chapter 7

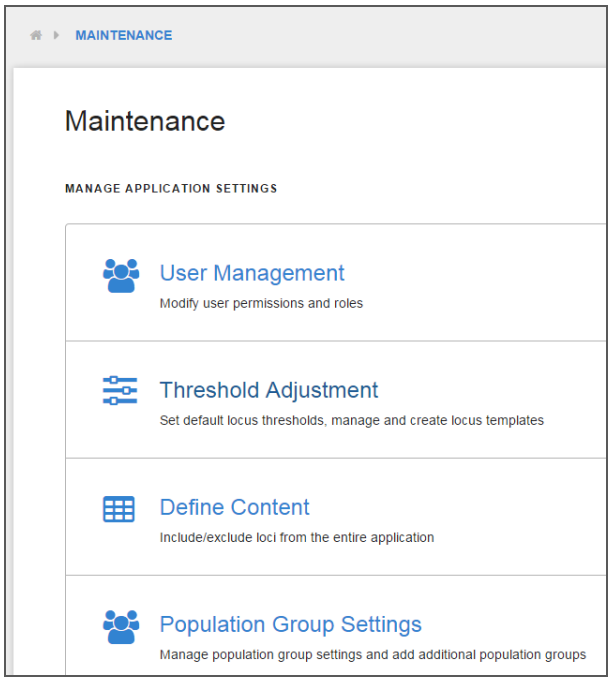
Introduction 80
Notifications 82
User Management 83
Changing Locus Thresholds 85
Defining Locus Content for Analysis 89
Define Loci for Population Studies 91
Population Group Settings 92

Introduction

System settings are available from the Maintenance page. The Maintenance page and system settings features are only visible and available to users with administrator account privileges.



To access the Maintenance page, click the arrow next to the user account drop-down list and select **Maintenance**.

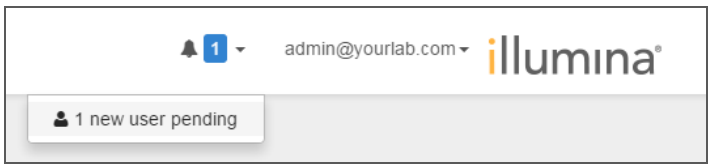


The following system settings features are accessed from the Maintenance page.

Table 27 System Settings Features

Settings Name	Description
User Management	Activate or deactivate user accounts. Modify user permissions and roles, including the assignment of administrator privileges.
Threshold Adjustment	Set default analysis thresholds. Create alternate analysis settings templates. Create locus templates.
Define Content	Control the loci that are displayed in the software and are analyzed by the software.
Loci Used in Population Studies	Control the loci that are included in population statistics calculations.
Population Group Settings	Select 1 or more population databases for the calculation of population statistics. You can select population databases included with the software, or upload your own.

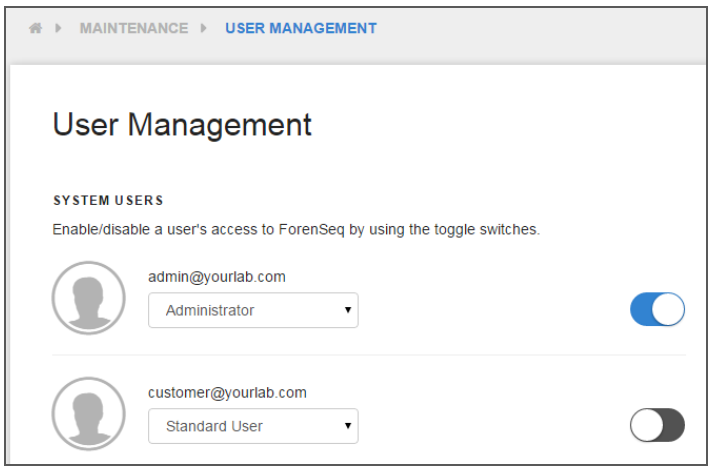
Notifications



When a user is logged on as an administrator, pending notifications that require attention appears in a blue box in the top right corner of the page next to a bell. To view pending notifications, click the down arrow next to the number of notifications.

When you click a notification information alert, a page related to the notification opens. For example, if the notification is to approve a new user, the User Management page opens. After you click a notification information alert, the reminder does not appear again on the page.

User Management



Access to the system is controlled by account management on the User Management page. The User Management page is launched from the Maintenance page. Users with administrator privileges enable and disable system accounts and assign account access levels.

When a user creates an account on the log in page of the software, an administrator approves the account on the User Management page before the user can log in and access the system. An enabled account has access to the ForenSeq Universal Analysis Software and the Instrument.

All accounts are listed on the User Management page. The access level for each account is displayed in a drop-down list below the account ID. A toggle switch next to the account indicates if the account is enabled or disabled. An account with a gray toggle switch is disabled, and the user cannot log in.

Enabling and Disabling Accounts

A user creates an account on the login page of the software. An administrator enables the account before the user can log in and access the system. When the user creates an account, a message is displayed for the administrator user that the account is pending creation. Sign-in is possible when the administrator approves the account.

After an administrator enables the account, which is controlled on the User Management page, the user can log in. If the user attempts to log in again before the account is approved, a message is displayed that the account was not approved for login, and to contact the administrator.

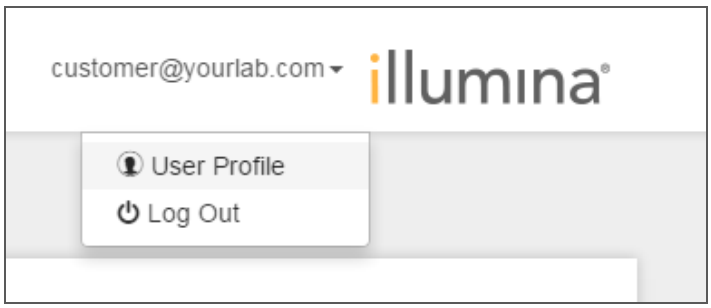
To enable an account, click the toggle switch next to the account name so that the switch displays a blue background and is in position on the right side.

To disable an account, click the toggle switch next to the account name so that the switch displays a gray background and is in position on the left side.

Assigning Account Access Levels

Each user is assigned to a group that determines access privileges for the system. A user with administrator privileges has all access privileges of a regular user, as well as the ability to configure system settings that are on pages launched from the Maintenance page. Administrator account privileges can be assigned or removed by other administrator users.

Changing an Account Password

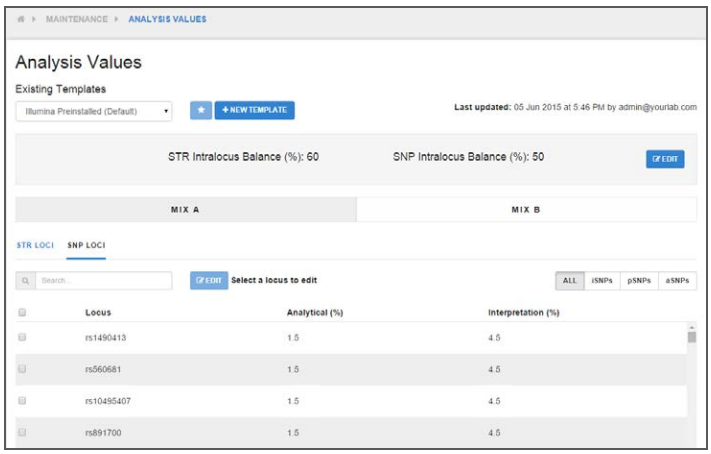


The current password associated with an account can be changed when you are logged in with the current password. A password change takes effect the next time you are required to log in.

To change a user password, perform the following steps:

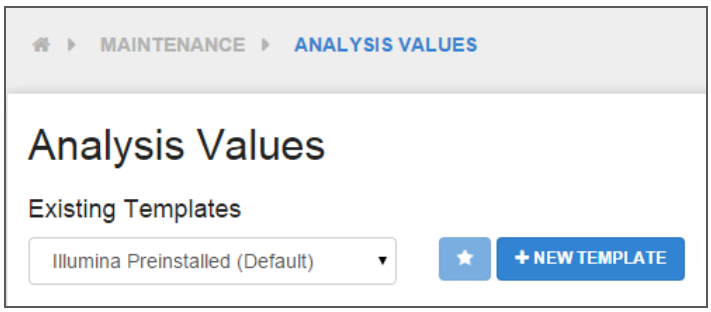
1. Select **User Profile** from the account drop-down list.
2. Enter the current password.
3. Enter the new password. The password must meet system requirements for a valid password.
4. In the Confirm New Password field, enter the new password again.
5. Select **Save**.

Changing Locus Thresholds



Locus thresholds that impact genotype calling and the setting of quality control indicators can be changed on the Analysis Values page. The Analysis Values page is launched from the Maintenance Page. See *Introduction* on page 80. Users with administrator privileges can change current thresholds or create new threshold templates. Changes to locus thresholds can then be applied to new analyses.

Creating a Locus Threshold Template

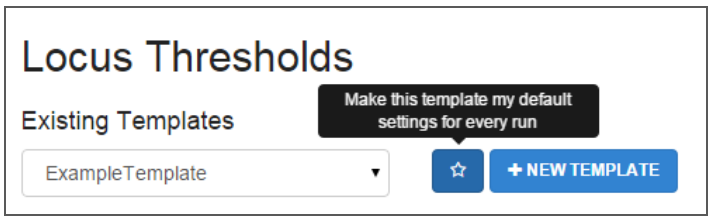


A collection of locus threshold settings is saved under a single name as a template for easy access and reference within the system. When you create a new template, it is based on the settings for the template open at the time of creation. A locus threshold template can be accessed on the Analysis Settings button.

To create a new template, access the Analysis Values page from the Maintenance page.

1. From the Existing Templates drop-down list, select a template on which to base the new template.
2. Click the **+New Template** button
3. Enter a name for the template.
4. Click **Save**.

Designating a Default Locus Threshold Template



Any locus threshold template can be designated as the system default template. The purpose of a default template is to define locus thresholds for major version analyses. A major version analysis is the first analysis conducted on a run version or at run completion.

To designate a template as the default template, access the Analysis Values page from the Maintenance page.

1 Select the template from the Existing Templates drop-down list.

2 Click the blue star icon.

Changing Analysis Level Settings

STR intralocus balance settings and SNP intralocus balance settings in a template are used for all loci across an analysis. These analysis level settings apply to all DNA primer mix types and loci.

To change the analysis level settings for a template, access the Analysis Values page from the Maintenance page.

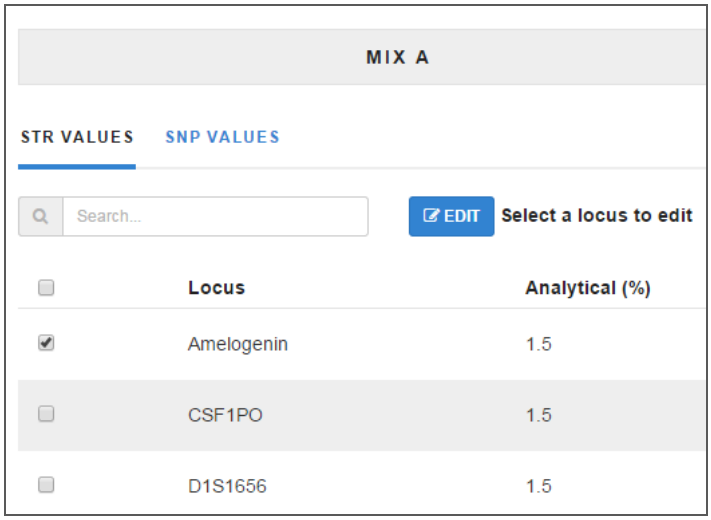
1 Select the template from the Existing Templates drop-down list.

2 Enter new STR Intralocus Balance or SNP Intralocus Balance values.
If the values were changed before, click the blue **Edit** button on the row that displays STR Intralocus Balance and SNP Intralocus Balance values. Then enter new STR Intralocus Balance or SNP Intralocus Balance values.

3 Click **Save**.

When you click save, the new settings are applied to the template. The Last Updated time stamp in the upper right corner of the page is refreshed with the new time and date, confirming when the change in analysis settings values is made to the template.

Changing Locus Level Settings

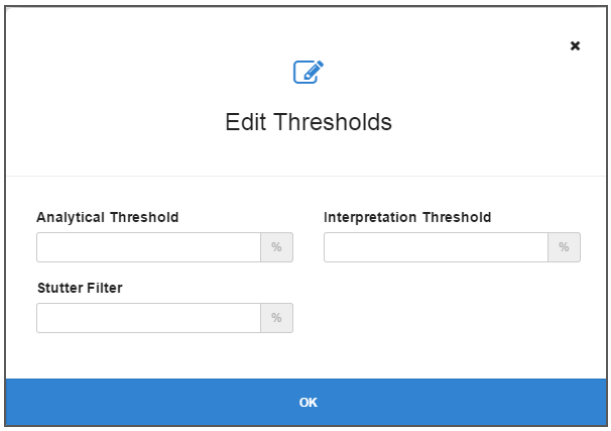


If you change the thresholds for an individual locus, it is important to consider whether you want to apply the change to both DNA Primer Mix A and DNA Primer Mix B.

To change the locus level settings for a template, access the Analysis Values page from the Maintenance page.

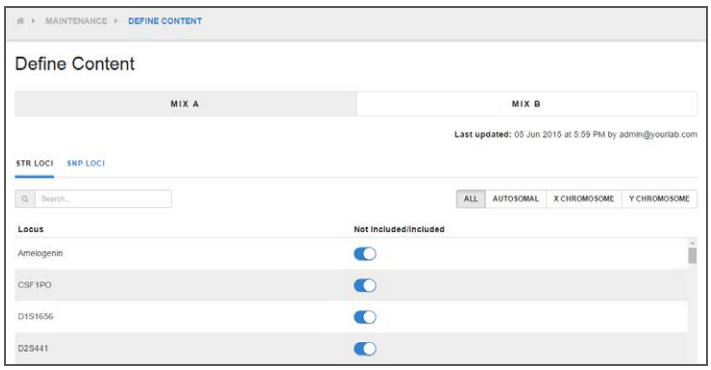
1. Select the template from the Existing Templates drop-down list.
2. Select DNA Primer Mix A or DNA Primer Mix B tab.
3. Select the STR or the SNP loci tab.
   To find a particular locus, you can use the Search feature, or separate the loci by selecting a filter button above the list.
   - For STRs, the list can be filtered for all, autosomal, X chromosome, or Y chromosome
   - For SNPs, the list can be filtered for all, iSNPs, aSNPs, or pSNPs

1. To select loci for locus level changes, click the small box to the left of the locus name. To select all loci in the tab, click the small box at the top of the list on the left side.
2. Click the blue **Edit** button.
3. An Edit Threshold box opens.
4. Enter new values for locus level settings.



5 Click **OK**.
When you click OK, the new settings are applied to the template. The Last Updated time stamp in the upper right corner of the page is refreshed with the new time and date, confirming when the change in locus level settings values is made to the template.

Defining Locus Content for Analysis



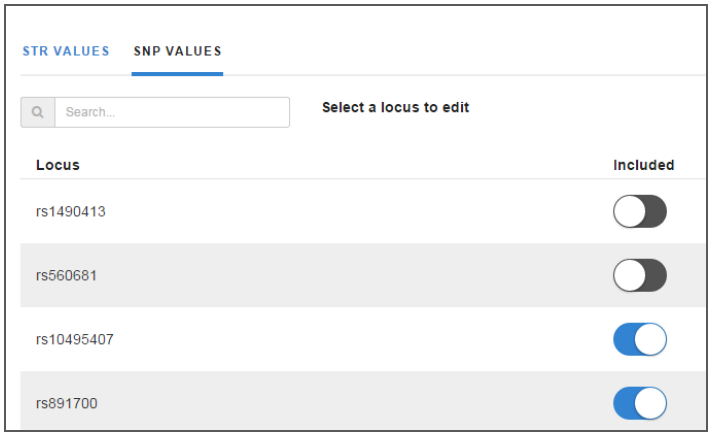
The ForenSeq Universal Analysis Software is equipped to analyze all loci that are included in ForenSeq kits. While the software can analyze the entire range of loci, locus content can be defined by an administrator. Deselected loci are not analyzed, and results for the loci are not generated.

The selection of loci included for analysis is controlled on the Define Content page. By default, the software includes all kit loci in analysis. The Define Content page is launched from the Maintenance Page. See *Introduction* on page 80.

NOTE
The gender that is assigned to a sample is based on the presence of signal at a sufficient number of Y loci. When modifying the loci included for analysis, at least 3 X loci or 3 Y loci are required in order to determine gender. If not enough X or Y loci are active, gender cannot be determined for the samples, and all samples are assigned the status of *Inconclusive*. The reason that gender cannot be called in this instance is because the maximum copy number for a locus is the maximum copy number for an assigned gender call. For example, the X loci will have a copy number of 2.

NOTE
The deactivation of ≥ 1 pSNP prevents hair and eye color estimation if you execute phenotype estimation. The complete set of ForenSeq pSNPs are required.

Selecting Locus Content



By default, the software includes all kit loci in analysis. To define loci in analyses, access the Define Content page from the Maintenance page.

1 Select DNA Primer Mix A or DNA Primer Mix B tab.

2 Select the STR or the SNP loci tab

To find a particular locus, you can use the Search feature, or separate the loci by selecting a filter button above the list.

- For STRs, the list can be filtered for all, autosomal, X chromosome, or Y chromosome
- For SNPs, the list can be filtered for all, iSNPs, aSNPs, or pSNPs

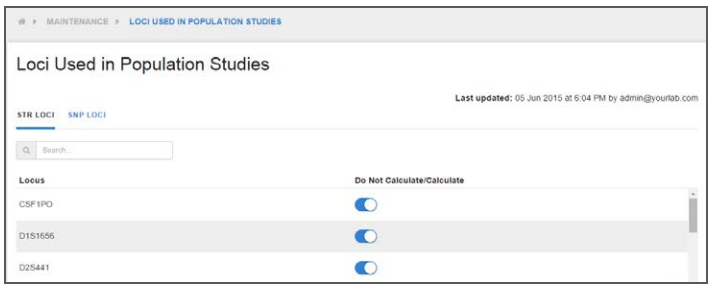
1 To deselect a locus, click the toggle switch next to the locus name so that the switch displays a gray background and is in position on the left side.

NOTE

Deselecting a locus is specific to the DNA primer mix type. If a locus is present in more than 1 mix type, you might want to consider adjusting both DNA primer mixes.

2 To select a locus to include for analysis, click the toggle switch next to the locus name so that the switch displays a blue background and is in position on the right side.

Define Loci for Population Studies



By default, the ForenSeq Universal Analysis Software uses all loci that are included in ForenSeq kits in calculating population statistics. Locus content for calculating population statistics can be defined by an administrator. Deselected loci are not in use.

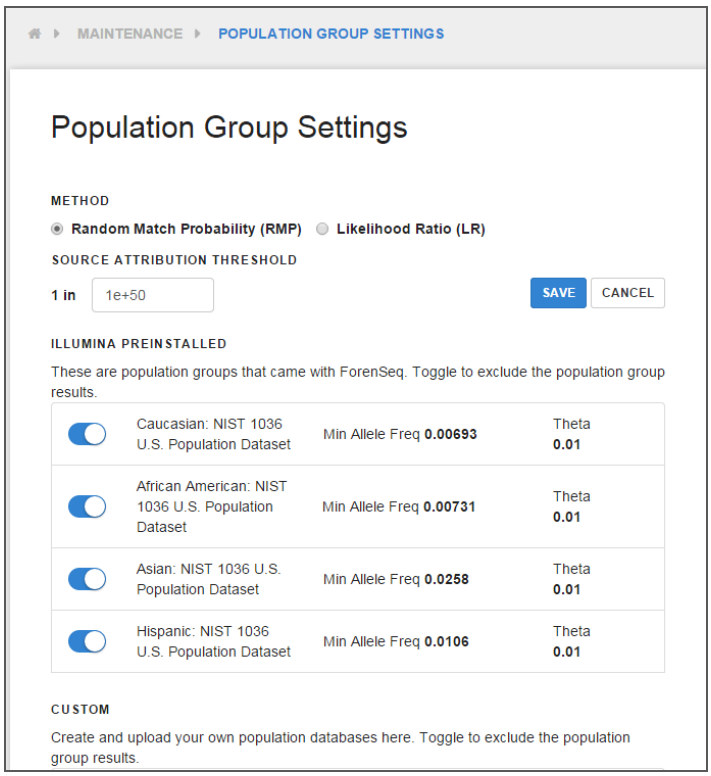
The selection of loci included for population statistics calculations is controlled on the Locus Used in Population Studies page.

Selecting Locus Content for Population Studies

By default, the software includes all kit loci in analysis. To define locus in analysis, access the Define Content page from the Maintenance page.

1. Select the STR or the SNP loci tab.
   To find a particular locus, you can use the Search feature, or separate the loci by selecting a filter button above the list.
   - For STRs, the list can be filtered for all, autosomal, X chromosome, or Y chromosome
   - For SNPs, the list can be filtered for all, iSNPs, aSNPs, or pSNPs
2. To deselect a locus, click the toggle switch next to the locus name so that the switch displays a gray background and is in position on the left side.
3. To select a locus to include, click the toggle switch next to the locus name so that the switch displays a blue background and is in position on the right side.

Population Group Settings



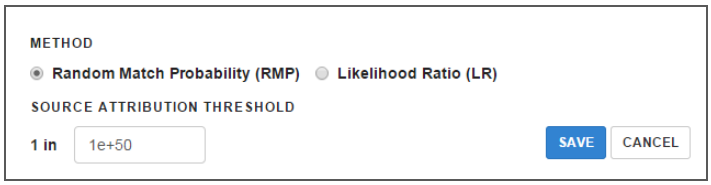
Population statistics calculations by the software require 2 components: at least 1 population group database to reference, and a method of calculation. The software comes installed with several population group databases that are ready for use. You can use population databases included with the software, or upload your own. Although random match probability is the default calculation method, you can select likelihood ratio instead. The software also supports the ability to define source attribution thresholds. Source attributions are relative to guidelines from Budowle et al.[1]

Population group database selection, method of population statistics calculation, and source attribution thresholds are controlled on the Population Group Settings page. The Population Group Settings page is launched from the Maintenance Page. See *Introduction* on page 80.

Resources

1. Budowle B, Chakraborty R, Carmody G, Monson KL. Source Attribution of a Forensic DNA Profile. *Forensic Science Communications*. 2000;2(3).

Defining the Statistics Calculation Method



The ForenSeq Universal Analysis Software supports the calculation of population statistics by either random match probability or likelihood ratio methods, with use of the 2p or 2p - p^2 rule. Use of the likelihood ratio method creates the inverse of the result retrieved from that of the random match probability method.

By default, the software uses random match probability and uses the 2p - p^2 setting for population statistics calculations. A change to the method used for calculation can be selected on the Population Group Settings page. If you select a different method of calculation, this method is used for all new calculations in the system going forward. Existing results remain unchanged.

NOTE
If you select Likelihood Ratio, the software automatically inverts the source attribution threshold setting. See *Changing the Source Attribution Threshold* on page 93.

The selection of calculation methods is controlled on the Population Group Settings page. To change the calculation method or calculation rule for population statistics, access the Population Groups Settings page from the Maintenance page.

1. Select **Random Match Probability (RMP)** or **Likelihood Ratio (LR)** in the Method area of the page.
2. Select **Use 2p** or **Use 2p - p^2** in the Population Statistics Calculation area of the page.
3. Click **Save**.

Applying the 2p Rule to a Homozygous Locus

When the genotype of a locus is homozygous (diploid), the formula for population statistics calculation is $P(A)^2 + P(A) \times (1 - P(A)) \times \theta$ where P(A) is the effective frequency of allele A in use for the reference population, and θ is the population structure correction factor.

The method of calculation can be changed for a homozygous locus by applying the 2p rule in the locus detail box. For directions, see *Updating and Modifying Typed Results* on page 42.

The results of applying the 2p rule to a homozygous locus in the locus detail box is dependent on the population statistics calculation setting on the Population Group Settings page. When the setting is Use 2p, and the 2p rule is applied in the locus detail box, the calculation formula becomes 2 x P(A). In contrast, when the population statistics calculation setting is Use 2p - p^2, and the 2p rule is applied in the locus detail box, the calculation formula becomes 2 x P(A) - P(A)^2.

Changing the Source Attribution Threshold

The source attribution threshold is defined to facilitate visualization of population group statistics in a display bar. Population statistics that are less than the frequency of the source attribution threshold are colored on the results bar in orange. The threshold is visible on the display as a gray bar below the result.

There are 2 source attribution threshold settings, each corresponding with a particular method of population group statistics calculation. Calculation of population statistics is performed by the software as either random match probability or likelihood ratio methods. Use of the likelihood ratio method creates the inverse of the result retrieved from that of the random match probability method. By default, as corresponds with the random match probability method, the source attribution threshold is set to > 1.

NOTE
If you select Random Match Probability, the software automatically inverts the source attribution threshold setting. See *Changing the Source Attribution Threshold* on page 93.

To change the source attribution threshold in use by the system, access the Population Group Settings page from the Maintenance page. Perform the following steps.

1 If random match probability is in use for population statistics calculations, enter a value > 1 in the Source Attribution Threshold field. If likelihood ratio is in use, enter a value < 1.

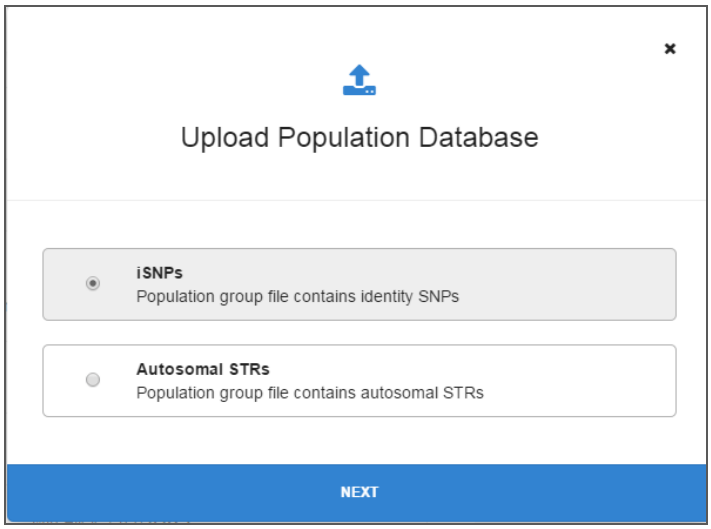
2 Click **Save**. Click Cancel to return to the source attribution threshold you changed.

Adding Population Groups

In addition to the population groups installed with the software, you can add custom population groups of your own to use in calculating population statistics. To add a population group, access the Population Group Settings page from the Maintenance page.

1 Select the **Add Population Group** button in the Custom area of the page.

2 Click **iSNPs** or **Autosomal STRs** to define the population group you are adding.



3 Click the **Next** button.

4 Select a population group file for upload.

5 Click **Upload Population Database**.
When you upload the file, results of the file validation are displayed. If you receive an error message, in many cases you can still proceed with the upload, defining the population group without the samples with errors. See *Troubleshooting Population Group File Uploads* on page 114. To abandon uploading the population group, close the Upload Population Database window.

6 Click **Next**.

7 Enter a name for the population group in the Database Name field.
A default value is displayed in the Minimum Allele Frequency field. The default value is calculated from the population size using the equation Minimum Allele Frequency = 5/2N. N is the number of subjects in the group.

8 Click **Finish**.

Custom Population Group Requirements

To add a custom population group successfully, the software requires that the file conform to a particular format.

NOTE
Types of loci that are not supported for custom population group files include X STRs, Y STRs, aSNPs, pSNPs, and loci with 3 alleles.

Table 28 Requirements for Custom Population Group Files

Type of Requirement	Details
File Type	File contents are text tab-delimited (*.txt).
Type of STR Loci Supported	Autosomal STRs with genotype entries that contain numbers
Type of SNP Loci Supported	iSNPs with genotype entries that contain A, C, T, or G. Each genotype must contain exactly 2 alleles.
Homozygotes	Entry contains the allele 2 times. For example, enter A A, or 12 12.
Number of Subjects in the Population Group	> 3 and ≤ 2500
Locus Names	Match the name of the locus in the software.
Table Rows	The first row of the table contains the loci being defined. In the rows that follow, each subject identifier is unique to a row and is not repeated in other rows.
Table Columns	The first column of the table contains the subject identifiers for each row. Each subject must contain a genotype for each locus in the file.

Custom Population Group File Examples

There are 2 options for the entry of genotypes in custom population files. Make sure that the file is consistent, with each sample in the file entered in the table in the same way.

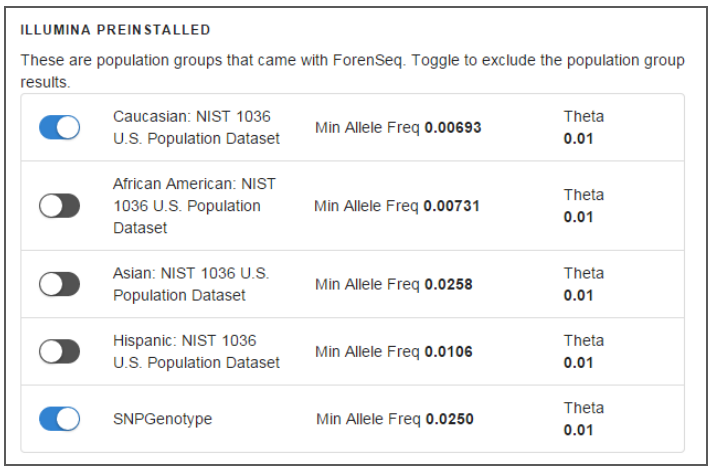
In the first table option, each allele is in an independent cell, so that each locus has 2 columns of data.

SampleCode	CFS1PO	CFS1PO	D10S1248	D10S1248	DS12S391	DS12S391
STRSample1	11	12	14	14	17	21
STRSample2	9	11	12	13	16	16
STRSample3	10	12	14	15	16	20
STRSample4	11	12	11	14	16	19
STRSample5	8	12	14	14	15	18

In the second table option, each locus is in a single column, with alleles separated by a comma.

SampleCode	rs1005533	rs10092491	rs1015250	rs1024116	rs1028528	rs1031825
SNPSample1	A,G	T,C	C,G	A,A	A,G	C,C
SNPSample2	G,A	T,T	G,C	A,A	G,A	A,C
SNPSample3	G,G	T,C	G,G	G,G	G,A	A,A
SNPSample4	G,A	C,T	C,G	A,G	A,G	C,C
SNPSample5	A,A	T,T	C,G	G,A	G,A	A,A

Selecting Population Groups



Population groups installed with the software and any population groups you defined for calculating population statistics are listed on the Population Group Settings page. The minimum allele frequency and Theta value is displayed next to each population group. By default, all population groups installed with the software are available for population statistics calculations.

To deselect a population group from use for population statistic calculations, click the toggle switch next to the population group name so that the switch displays a gray background and is in position on the left side.

To select a population group for use for population statistic calculations, click the toggle switch next to the population name so that the switch displays a gray background and is in position on the left side.

Analysis Metrics and Procedures

Chapter 8

Introduction 98
Analysis Metrics 99
Analysis Procedures 100

Introduction

ForenSeq Universal Analysis Software performs a series of analysis steps on data from the MiSeq FGx and RTA software to obtain STR and SNP information from samples sequenced on the instrument.

Analysis Metrics

During the sequencing run, RTA generates data files that include analysis metrics used by ForenSeq Universal Analysis Software. Metrics that appear in the ForenSeq Universal Analysis Software are clusters passing filters, base call quality scores, and phasing and prephasing values.

Clusters Passing Filter

During analysis, RTA filters raw data to remove any reads that do not meet the overall quality as measured by the Illumina chastity filter. The chastity of a base call is calculated as the ratio of the brightest intensity divided by the sum of the brightest and second brightest intensities.

Clusters pass filter (PF) when no more than 1 base call in the first 25 cycles has a chastity of < 0.6.

Quality Scores

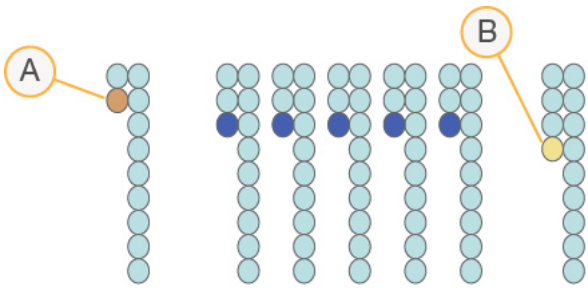
For each read, the quality scores are compared against benchmark values for the assay. If the average quality of the read is lower than the benchmark value, an orange indicator is provided.

During the sequencing run, base call quality scores are calculated during acquisition.

Phasing and Prephasing

During the sequencing reaction, each DNA strand in a cluster extends by 1 base per cycle. A small portion of strands might become out of phase with the current incorporation cycle, either falling a base behind (phasing) or jumping a base ahead (prephasing). Phasing and prephasing rates indicate an estimate of the fraction of molecules that became phased or prephased in each cycle.

Figure 3 Phasing and Prephasing



A Read with a base that is phasing
B Read with a base that is prephasing

The number of cycles performed in a read is 1 more cycle than the number of cycles analyzed.

Analysis Procedures

ForenSeq Universal Analysis Software performs analysis on sequencing data generated on the MiSeq FGx from sample libraries prepared with the ForenSeq DNA Signature Prep Kit. Analysis procedures include demultiplexing, sequence alignment, allele counting, and genotype calling. Quality control indicators address aspects such as DNA mixture detection and genotype call quality.

The ForenSeq Universal Analysis Software addresses STR and SNP loci amplified in either of the 2 DNA primer mixes in the ForenSeq DNA Signature Prep Kit. Analysis is performed on amplified and sequenced DNA for STR and SNP loci, including the following:

- Autosomal STRs
- X-STRs
- Y-STRs
- iSNPs
- pSNPs
- aSNPs

Analysis is geared toward single-source DNA samples as well as challenging forensic samples. The target type is determined to be a SNP or a STR for sequence alignment purposes. Each read is then aligned to the hg19 human reference sequences that correspond to the loci evaluated in the primer pool to determine the locus and the repeat length.

After alignment, potential alleles are counted from the numbers of reads. Counts are used to detect if the sample is a mixture, call genotypes, and assign quality control indicators. Quality control indicators can help in mixture detection and sample interpretation.

NOTE
For more information on hg19 sequences, see http://genome.ucsc.edu/cgi-bin/hgGateway.

Demultiplexing

Demultiplexing is the first step in post-sequencing analysis. Demultiplexing separates data among pooled samples based on short index sequences that tag samples from different libraries. Each Index Read sequence is compared to the index sequences specified in sample information entered on the Create a New Run page. The ForenSeq kit enables pooling of many i7 and i5 index pairs into the same run.

Alignment of STRs

Each read is evaluated against all the possible loci in the primer pool to determine its source. If a read can be aligned to multiple loci, a scoring mechanism is employed and the read is attributed to the highest scoring locus.

Evaluation of the read is done by identifying the start and stop of the repeat on both sides of the read as compared to the reference sequence. In the case of STRs, the read is then further evaluated to identify the repeated region of the read for additional allele calling refinement.

Alignment of SNPs

Each read is evaluated against all the possible loci in the primer pool to determine its source. If a read can be aligned to multiple loci, a scoring mechanism is employed and the

read is attributed to the highest scoring locus. Sequence-based variation in SNPs is not based on repetitive elements as they are for STRs.

Allele Counting

The software determines the length of the STR sequence between the forward and reverse PCR primers, and the number of repeats within that sequence. In addition to determining the length of the sequence, an analysis algorithm accounts for the potential that STR alleles of the same length can vary in sequence, therefore representing 2 different alleles.

STR Genotype Calling

Genotype calling of an STR occurs as the ForenSeq Universal Analysis Software considers signal at a target. The analysis algorithm uses allele counts and several parameters to call (type) an allele.

First, analytical and interpretation thresholds are determined through the application of the supplied threshold percentage values. The user-defined percentage value is multiplied by the sum of read counts at a locus, respectively. In cases of low coverage, a minimum number of reads is used (650) as the number of reads for the locus in the determination of thresholds.

After automated STR allele calling, the gender of the sample is determined based on the signal at the X and Y loci. The software uses gender designation at the Y STR loci and X STR loci to dictate the expected copy number at these loci. The number of alleles typed at a locus is important in setting genotype quality control indicators such as allele count. The number of alleles called also determines whether a locus with a single non-stutter allele above the interpretation threshold is a homozygote or hemizygote. By default, negative amplification control samples and reagent blank samples are not assessed for gender, displaying "Inconclusive" in place of XX or XY.

Table 29 Gender Determination

Inconclusive	XX	XY
Negative control samples are always assigned a gender of inconclusive. A sample is also assigned a gender of inconclusive if it meets either of the two conditions: • Of the X STR loci in the sample, less than 3 loci have signal above the analytical threshold. • Less than 3 Y STR loci are active in the analysis.	For a sample to receive the designation of gender XX, all of the following conditions must apply: • The sample is not a negative control. • Of the X STR loci in the sample, at least 3 loci have signal above the analytical threshold. • Less than 3 Y STR loci have signal greater than the analytical threshold, but at least 3 loci are active in the analysis.	For a sample to receive the designation of gender XY, all of the following conditions must apply: • The sample is not a negative control. • Of the X STR loci in the sample, at least 3 loci have signal above the analytical threshold. • Of the Y STR loci in the sample, at least 3 loci have signal above the analytical threshold.

After gender-calling, the software checks the read counts for the following quality considerations:

Low Coverage— If the allele with the highest reads does not meet the effective interpretation threshold, the locus is not processed further. The low coverage indicator can be seen in results from version 1.1 or earlier of the ForenSeq Universal Analysis Software.

Stutter Filtering— Stutter is computed as sequence stutter, where the software checks the repeat units in the STR sequence. This computation can enable the separation of a minor contributor allele from stutter when the sequences differ.

A sequence is considered stutter, with offset *k* repeat units if:

- The sequences differ in length by *k* times the reference repeat unit length
- The number of repeat units in the potential stutter are identical to an adjacent unit of length of the reference repeat unit
- The are no (other) mismatches between the sequences

In the case of tetra- and penta-nucleotide repeats, the reference repeat unit lengths are 4 and 5 respectively. The software assesses differences in the sequences for *k* = -2, -1 and +1 repeat units. That is, with n the length in base pairs of a given allele, the software assesses differences for sequences at the n-8 or n-10, n-4 or n-5, and n+4 or n+5 positions for tetra- and penta-nucleotide repeats, respectively.

The stutter filter setting depends on the differences in repeat units. The stutter filter setting is applied as is for the n-4 and n-5 positions. The stutter filter setting is squared it at the n-8 or n-10 and the n+4 or n+5 positions.

The expected stutter intensity is the product of the stutter filter setting and the called allele intensity at a repeat length. For example, a stutter filter setting of 10% (0.1) is squared (0.1 x 0.1 = 0.01 or 1%) before multiplying by the intensity of the called parent allele.

A stutter quality control indicator for elevated stutter is displayed when both of these conditions exist:

- Uncalled read intensity, with sequence of a potential stutter of a called allele
- Uncalled read intensity is greater than the maximum expected, user-defined stutter % of the potential parent allele

Interpretation Threshold— If the read count of any DNA sequence has a read count of a non-stutter DNA sequence between the analytical and interpretation threshold, an interpretation threshold quality indicator icon appears.

Allele Count— Indicated when the signal from possible non-stutter alleles is greater than the analytical threshold, and is greater than the copy number of the locus. The copy number at the locus is defined by the expected zygosity of the locus.

Analytical Threshold— When a locus has signal below the analytical threshold and no alleles exceed the interpretation threshold, an analytical threshold quality indicator icon appears.

Not Detected—When no signal is present at the locus, a not detected quality control indicator quality indicator icon appears.

Imbalance— If the calculated intralocus balance is below the intralocus balance threshold, and the locus is not homozygous, an imbalance quality control indicator is triggered.

Mixture Detection— A sample is indicated as a possible mixture when the total number of loci that are assigned the many alleles quality control indicator is > 5.

SNP Genotype Calling

Genotype calling of a SNP occurs as the ForenSeq Universal Analysis Software determines that DNA sequence at a targeted locus passes quality control thresholds. The analysis algorithm uses allele counts and several parameters to call (type) an allele.

Then the software checks the read counts for the following quality considerations:

Per Locus QC— A quality check for each locus is performed. Based on the assessment, a preliminary genotype is assigned based on the signal intensity for each allele.

Interpretation Threshold— If the read count of either allele present is between the analytical and interpretation thresholds, an interpretation quality indicator icon appears. The interpretation threshold percent is defined in the analysis settings.

Imbalance— If the count of the reference allele divided by the count of the alternate allele is less that the user-defined imbalance ratio, and the locus is not homozygous, then it is indicated as imbalanced.

Not Detected—When no signal is present at the locus, a not detected quality control indicator quality indicator icon appears.

Mixture Detection— A sample is indicated as a possible mixture when the total number of loci that are imbalanced is > 10.

Analytical Threshold— When a locus has signal below the analytical threshold and no alleles exceed the interpretation threshold, an analytical threshold quality indicator icon appears.

Troubleshooting

Introduction 106
Troubleshooting ForenSeq Runs 107
Troubleshooting Analysis Errors 112
Troubleshooting Population Group File Uploads 114

Introduction

This chapter features troubleshooting tips for ForenSeq sequencing runs, analysis errors, and population group file uploads. Most of the troubleshooting recommendations for ForenSeq sequencing runs include library prep steps, which are detailed in the *ForenSeq DNA Signature Prep Guide (part # 15049528)*.

Troubleshooting ForenSeq Runs

Problem	Resolution
Sample Intensity (Number of Reads) Low	If sample intensity for a sample is low, this condition can indicate a problem with sample quantity or sample quality. 1 Make sure that DNA concentrations are not below the recommended amount. 2 Depending on the source, purify the DNA or dilute, and then sequence the sample again. 3 Load fewer samples and sequence the samples again. Make sure to maintain a minimum of 8 samples in the run.

Problem	Resolution
Cluster Density High	Review the Run Metrics tab. If cluster density for the run is high, but the reads are above the sample read count threshold in the Sample Representation tab, proceed with analysis. If there is insufficient data for analysis, it can be due to 1 of the following reasons. ▸ Sample from the Purified Library Plate (PLP) or the Library Normalization Beads 1 (LNB1) reagent was over-pipetted during normalization. a If the sample has a high number of alleles with quality control indicators, then rerun the sample on a new sequencing run, and confirm that correct volumes and reagents are used for normalization. b If the volume of sample used is correct, consider further diluting the PCR product in the Pooled Normalization Libraries (PNL) tube. c If normalization volumes or reagents were not correct, redo the *Normalize Libraries* process and resequence. ▸ Too much adapter dimer in library prep. a Run 1–2 µl of the PCR product in the PLP plate on a Bioanalyzer or Fragment Analyzer and check the amount of primer dimer. Contact Illumina Technical Support with any questions. b If there is more than 5 % adapter dimer in the PCR product you check, redo the *Purify Libraries* process on the remaining PCR product in the PLP plate using the 30 µl remaining in each well. Adjust the *Purify Libraries* process to use 30 µl SPB. Add 32.5 µl Resuspension Buffer (RSB) and transfer 30 µl to a new 96-well PCR plate. Proceed with the *Normalize Libraries* process. c [Optional] If SPB are limited, purify the PCR product in the PLP plate using a standard column cleanup method. Use manufacturer instructions. ▸ Not enough Hybridization Buffer (HT1) added, or too much normalized library added to HT1 when preparing libraries. If the sample has a high number of alleles with quality control indicators, then rerun the sample, on a new sequencing run, making sure the normalized library is sufficiently diluted.

Problem	Resolution
Cluster Density Low	Review the Run Metrics tab. If cluster density for the run is low, but the reads are above the sample read count threshold in the Sample Representation tab, proceed with analysis. If the sample has an insufficient number of reads due to low density, it is likely due to 1 of the following reasons: ▸ During library prep, DNA input was too low or overly degraded, or the DNA dilution was incorrect. If possible, repeat library prep with more DNA. ▸ During library prep, a critical reagent was not added. Repeat library prep. ▸ LNB1 was not pipetted sufficiently during normalization. Consider diluting the PCR product in the PNL tube less. ▸ HT1 was over-pipetted when preparing libraries. Repeat sequencing using the correct volume. ▸ Excessive sample was lost during SPB purification. Repeat library prep using bead-handling best practices. ▸ Sample not denatured properly. a Make sure that the final library pool is denatured with HP3. b Make sure that the HP3 pH is above 12.5. c Make sure that the final library pool is heat denatured. ▸ Instrument blockage. a Perform a maintenance wash and repeat sequencing. b Make sure to refill the wash tray and bottle for every wash. c Make all Instrument wash solutions daily.
Percentage of Clusters Passing Filter Low	If the percentage of clusters passing filter is too low, it is likely due to 1 of the following reasons: ▸ Cluster density is too high. See troubleshooting for Cluster Density High. ▸ Phasing or prephasing is too high. a Perform a maintenance wash and repeat sequencing. b Make sure to refill the wash tray and bottle for every wash. c Make all Instrument wash solutions daily. ▸ If the problem is still not resolved, it could be that reagents are not performing as expected. Contact Illumina Technical Support.

Problem	Resolution
Phasing or Prephasing High	If phasing or prephasing is too high, it is likely due to 1 of the following reasons: ▸ Cluster density is too high. See troubleshooting for Cluster Density High. ▸ Instrument blockage. a Perform a maintenance wash and repeat sequencing. b Make sure to refill the wash tray and bottle for every wash. c Make all Instrument wash solutions daily. ▸ If the problem is still not resolved, it could be that reagents are not performing as expected. Contact Illumina Technical Support.
Run Does Not Complete	If the sequencing run stops before completion, it is likely due to 1 of the following reasons: ▸ When preparing libraries, HSC was not added. Add HSC according to the *Denature and Dilute Libraries* process and repeat the sequencing run. ▸ Cluster density is too high. See troubleshooting for Cluster Density High. ▸ If an error message is displayed, contact Illumina Technical Support.
Read 1 or Read 2 Quality Scores Low	If quality scores for Read 1 or Read 2 are too low, it is likely due to 1 of the following reasons: ▸ Phasing and prephasing are too high. See troubleshooting for Phasing and Prephasing High. ▸ Cluster density is too high. See troubleshooting for Cluster Density High. ▸ If the problem is still not resolved, it could be that reagents are not performing as expected. Contact Illumina Technical Support.
Index 1 or Index 2 Quality Scores Low	If quality scores for Index 1 or Index 2 are too low, it is likely due to 1 of the following reasons: ▸ Phasing and prephasing are too high. See troubleshooting for Phasing and Prephasing High. ▸ Cluster density is too high. See troubleshooting for Cluster Density High. ▸ There was low diversity in the Index Read because not enough samples were sequenced. Repeat sequencing with a minimum of 8 samples. ▸ If the problem is still not resolved, it could be that reagents are not performing as expected. Contact Illumina Technical Support.

Problem	Resolution
Reads Per Sample (Intensity) Low	Review the Sample Representation tab. If the samples are above the sample read count threshold, then proceed with analysis. If reads per sample are below the sample read count threshold, it is likely due to 1 of the following reasons: ▸ During library prep, DNA input was too low, or the DNA dilution was incorrect. If possible, repeat library prep with more DNA. ▸ During library prep, a critical reagent was not added. Repeat library prep. ▸ LNB1 was not pipetted sufficiently during normalization. a If the sample has a high number of alleles with quality control indicators, then rerun the sample on a new sequencing run. b If the sample has a low number of alleles with quality control indicators, consider diluting the PNL less and rerun the samples on a new sequencing run. c If the run does not pass, redo the *Normalize Libraries* process. ▸ HT1 was over-pipetted when preparing libraries. Redo the *Normalize Libraries* process. ▸ The PCR product was not heat-denatured when preparing libraries. Perform the denature heating step on the Diluted Normalized Libraries (DNL) tube and repeat the sequencing run.
Reads Per Sample (Intensity) for Human Sequencing Control Low	Review the Sample Representation tab. If the samples are above the sample read count threshold, then proceed with analysis. If reads per sample are below the sample read count threshold, it is likely due to 1 of the following reasons: ▸ When preparing libraries, HSC was not added. Add HSC according to the *Denature and Dilute Libraries* process and repeat the sequencing run. ▸ HT1 was over-pipetted when preparing libraries. Repeat the sequencing using the correct volume. ▸ The sample was not heat-denatured when preparing the sample for sequencing. Perform the denature heating step and repeat the sequencing run. ▸ HSC was not denatured with HP3 when preparing the sample for sequencing. Make sure that the HSC is denatured with HP3 before adding it to the library, and repeat the sequencing run.

Troubleshooting Analysis Errors

Problem	Resolution
Analysis Error Message	When you click Create New Analysis, and you receive an analysis error message, it is likely due to 1 of the following reasons: ▸ The Illumina ForenSeq Analysis service is not running. a On the ForenSeq server desktop, click the Windows icon in the lower left corner of the screen. b Enter "Services" to find the Services application. c Open the **Services** application. d Locate Illumina ForenSeq Analysis in the list of services. e If the status of the service is not "Running", right-click on the service and select **Start**. f After the service has started, try to the Create New Analysis command again. ▸ The Illumina ForenSeq Analysis service is running, but the system cannot locate the run data to perform an analysis. a On the ForenSeq server, open the log file for the analysis service at (C:\Illumina\Forenseq UAS\Analysis\logs\application.log). b Find the message, "EXCEPTION OCCURRED:System.Exception" in the log file to confirm the problem is that the system cannot locate the run data. c If the run repository is stored on a networked location, confirm the network is running. d Navigate to the run folder listed in the log file as the RunStoragePath to confirm that the data exists in the expected location. e Try the Create New Analysis command again.

Problem	Resolution
All Q Icons on Dashboard are Gray	When all Q icons on the dashboard are gray, and no information is available when you hover over the icons, this condition can indicate the following: ▸ The system cannot locate the run data. a Click the Windows icon in the lower left corner of the screen. b Enter "Services" to find the Services application. c Open the **Services** application. d Locate Illumina ForenSeq Analysis in the list of services. e If the status of the service is not "Running", right-click on the service and select **Start**. ▸ The run repository is stored on a networked location, and the network is not connected. a Navigate to the run folder described in the configuration file. b Confirm that data exists in the expected location. c Confirm that no permission changes were made to access the run folder location.

Troubleshooting Population Group File Uploads

Error Message	Item Affected	Interpretation and Possible Resolution
File format is unrecognized by the system	File	Unable to proceed because the system cannot recognize the provided data. Make sure that the file is text tab-delimited (*.txt).
Population group definition does not meet minimum or maximum size requirements	File	Unable to proceed because the population group is too small or too large for the system. Make sure that the population group is > 3 and ≤ 2500.
Locus does not match an expected locus name	Locus	The locus is not defined in the population group because system results cannot be correlated to the locus. Make sure that loci names are consistent with loci names in the software.
Number of alleles defined for the locus is not consistent with expectations	Locus	The locus is not defined in the population group because the number of alleles is not equal to 2. Review allele entries.
Invalid chromosome type	Locus	The locus is not defined in the population group because it is not an autosomal STR. Remove any X STRs and Y STRs from the population group.
Allele name is not recognized	Subject	The subject results are not included in the population group because the allele name is inconsistent with expectations. If your file contains STRs, make sure that entries contain repeat numbers and no letters. If your file contains SNPs, make sure that entries contain letters, and no repeat numbers.
Data has an incomplete profile	Subject	The subject results are not included in the population group because results are missing for at least 1 locus. Make sure that each subject has a genotype for each locus in the file.
Subject name is not unique to the population group	Subject	The subject results are not included in the population group because there are 2 or more rows with the same sample identifier. Remove duplicate sample identifiers.

Supporting Information

Appendix A

Human Sequencing Control Loci 116
Autosomal, Y, and X STR Filters and Thresholds 117
aSNP Thresholds 120
iSNP Thresholds 122
pSNP Thresholds 126
Set Up an External Data Repository 127

Human Sequencing Control Loci

The following is a list of loci that are contained in human sequencing control with allele lengths.

Table 30 Human Sequencing Control Loci and Lengths

Locus	Allele Length
D3S1358	136
D5S818	68
D7S820	93
D8S1179	61
D13S317	118
D16S539	118
D18S51	131
D21S11	178
FGA	126
PentaE	340
TPOX	45
DYF387S1	185
DYS391	88
DYS392	297
Y-GATA-H4	120
DXS7423	145
DXS10074	193
DXS10103	129
PentaD	220
DYS448	311
DYS460	303

Autosomal, Y, and X STR Filters and Thresholds

The following is a list of autosomal, Y, and X STRs with values for stutter filter, analytical, and interpretation thresholds.

Table 31 STR Locus Thresholds

Loci	% Stutter	% Analytical	% Interpretation
Amelogenin	0	> 1.5	> 4.5
CSF1PO	< 10	> 1.5	> 4.5
D1S1656	< 25	> 1.5	> 4.5
D2S441	< 7.5	> 1.5	> 4.5
D2S1338	< 20	> 1.5	> 4.5
D3S1358	< 15	> 1.5	> 4.5
D4S2408	< 7.5	> 1.5	> 4.5
D5S818	< 12.5	> 1.5	> 4.5
D6S1043	< 12.5	> 1.5	> 4.5
D7S820	< 10	> 1.5	> 4.5
D8S1179	< 25	> 1.5	> 4.5
D9S1122	< 12.5	> 1.5	> 4.5
D10S1248	< 20	> 1.5	> 4.5
D12S391	< 33	> 1.5	> 4.5
D13S317	< 12.5	> 1.5	> 4.5
D16S539	< 20	> 1.5	> 4.5
D17S1301	< 20	> 1.5	> 4.5
D18S51	< 22	> 1.5	> 4.5
D19S433	< 12.5	> 1.5	> 4.5
D20S482	< 15	> 1.5	> 4.5
D21S11	< 10	> 1.5	> 4.5
D22S1045	< 20	> 1.5	> 4.5
FGA	< 25	> 1.5	> 4.5
PentaD	< 7.5	> 1.5	> 4.5
PentaE	< 10	> 1.5	> 4.5
TH01	< 10	> 1.5	> 4.5
TPOX	< 10	> 1.5	> 4.5

vWA	< 22	> 1.5	> 4.5
DYS19	< 15	> 1.5	> 4.5
DYS385a-b	< 20	> 1.5	> 4.5
DYF387S1	< 20	> 1.5	> 4.5
DYS389I	< 20	> 1.5	> 4.5
DYS389II	< 35	> 5	> 15
DYS390	< 15	> 1.5	> 4.5
DYS391	< 20	> 1.5	> 4.5
DYS392	< 30	> 1.5	> 4.5
DYS437	< 45	> 1.5	> 4.5
DYS438	< 15	> 1.5	> 4.5
DYS439	< 15	> 1.5	> 4.5
DYS448	< 15	> 3.3	> 10
DYS460	< 15	> 1.5	> 4.5
DYS481	< 50	> 1.5	> 4.5
DYS505	< 15	> 1.5	> 4.5
DYS522	< 15	> 1.5	> 4.5
DYS533	< 15	> 1.5	> 4.5
DYS549	< 22	> 1.5	> 4.5
DYS570	< 22	> 1.5	> 4.5
DYS576	< 15	> 1.5	> 4.5
DYS612[1]	< 35	> 1.5	> 4.5
DYS635	< 15	> 3.3	> 10
DYS643	< 20	> 1.5	> 4.5
Y-GATA-H4[2,3]	< 35	> 1.5	> 4.5
HPRTB	< 15	> 1.5	> 4.5
DXS7132	< 22	> 1.5	> 4.5
DXS7423	< 15	> 1.5	> 4.5
DXS8378	< 15	> 1.5	> 4.5
DXS10074	< 25	> 1.5	> 4.5
DXS10103	< 22	> 1.5	> 4.5
DXS10135	< 22	> 1.5	> 4.5

Resources

1. D'Amato ME, Ehrenreich L, Cloete K, Benjeddou M, Davison S. Characterization of the highly discriminatory loci DYS449, DYS481, DYS518, DYS612, DYS626, DYS644 and DYS710. Forensic Sci Int Genet. 2010 Feb;4(2):104-10.

2. Butler JM, Schoske R, Vallone PM, Kline MC, Redd AJ, Hammer MF. A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. Forensic Sci Int. 2002 Sep 10;129:10-24.

3. Short Tandem Repeat DNA Internet Database. SRM 2395-Human Y-Chromosome DNA Profiling Standard. National Institute of Standards and Technology (NIST). http://www.cstl.nist.gov/strbase/srm2395.htm. Updated December 15, 2009. Accessed June 9, 2015.

aSNP Thresholds

The following is a list of aSNPs with analytical, and interpretation thresholds.

Table 32 aSNP Thresholds

Loci	% Analytical	% Interpretation
rs3737576	> 1.5	> 4.5
rs7554936	> 1.5	> 4.5
rs2814778	> 1.5	> 4.5
rs798443	> 1.5	> 4.5
rs1876482	> 1.5	> 4.5
rs1834619	> 1.5	> 4.5
rs3827760	> 1.5	> 4.5
rs260690	> 1.5	> 4.5
rs6754311	> 1.5	> 4.5
rs10497191	> 1.5	> 4.5
rs1919550	> 1.5	> 4.5
rs12498138	> 1.5	> 4.5
rs4833103	> 1.5	> 4.5
rs1229984	> 1.5	> 4.5
rs3811801	> 1.5	> 4.5
rs7657799	> 1.5	> 4.5
rs870347	> 1.5	> 4.5
rs7722456	> 1.5	> 4.5
rs192655	> 1.5	> 4.5
rs3823159	> 1.5	> 4.5
rs917115	> 1.5	> 4.5
rs1462906	> 1.5	> 4.5
rs6990312	> 1.5	> 4.5
rs2196051	> 1.5	> 4.5
rs1871534	> 1.5	> 4.5
rs3814134	> 1.5	> 4.5
rs4918664	> 1.5	> 4.5
rs174570	> 1.5	> 4.5

rs1079597	> 1.5	> 4.5
rs2238151	> 1.5	> 4.5
rs671	> 1.5	> 4.5
rs7997709	> 1.5	> 4.5
rs1572018	> 1.5	> 4.5
rs2166624	> 1.5	> 4.5
rs7326934	> 1.5	> 4.5
rs9522149	> 1.5	> 4.5
rs200354	> 1.5	> 4.5
rs1800414	> 1.5	> 4.5
rs12439433	> 1.5	> 4.5
rs735480	> 1.5	> 4.5
rs1426654	> 1.5	> 4.5
rs459920	> 1.5	> 4.5
rs4411548	> 1.5	> 4.5
rs2593595	> 1.5	> 4.5
rs17642714	> 1.5	> 4.5
rs4471745	> 1.5	> 4.5
rs11652805	> 1.5	> 4.5
rs2042762	> 1.5	> 4.5
rs7226659	> 1.5	> 4.5
rs3916235	> 1.5	> 4.5
rs4891825	> 1.5	> 4.5
rs7251928	> 1.5	> 4.5
rs310644	> 1.5	> 4.5
rs2024566	> 1.5	> 4.5
rs1689198	> 1.5	> 4.5
rs1291383	> 1.5	> 4.5
rs16891982	> 1.5	> 4.5
rs12913832	> 1.5	> 4.5

iSNP Thresholds

The following is a list of iSNPs with analytical, and interpretation thresholds.

Table 33 iSNP Thresholds

Loci	% Analytical	% Interpretation
rs1490413	> 1.5	> 4.5
rs560681	> 1.5	> 4.5
rs1294331	> 1.5	> 4.5
rs10495407	> 1.5	> 4.5
rs891700	> 1.5	> 4.5
rs1413212	> 1.5	> 4.5
rs876724	> 1.5	> 4.5
rs1109037	> 1.5	> 4.5
rs993934	> 1.5	> 4.5
rs12997453	> 1.5	> 4.5
rs907100	> 1.5	> 4.5
rs1357617	> 1.5	> 4.5
rs4364205	> 1.5	> 4.5
rs2399332	> 1.5	> 4.5
rs1355366	> 1.5	> 4.5
rs6444724	> 1.5	> 4.5
rs2046361	> 1.5	> 4.5
rs279844	> 1.5	> 4.5
rs6811238	> 1.5	> 4.5
rs1979255	> 1.5	> 4.5
rs717302	> 1.5	> 4.5
rs159606	> 1.5	> 4.5
rs13182883	> 1.5	> 4.5
rs251934	> 1.5	> 4.5
rs338882	> 1.5	> 4.5
rs13218440	> 1.5	> 4.5
rs1336071	> 1.5	> 4.5
rs214955	> 1.5	> 4.5

rs727811	> 1.5	> 4.5
rs6955448	> 1.5	> 4.5
rs917118	> 1.5	> 4.5
rs321198	> 1.5	> 4.5
rs737681	> 1.5	> 4.5
rs763869	> 1.5	> 4.5
rs10092491	> 1.5	> 4.5
rs2056277	> 1.5	> 4.5
rs4606077	> 1.5	> 4.5
rs1015250	> 1.5	> 4.5
rs7041158	> 1.5	> 4.5
rs1463729	> 1.5	> 4.5
rs1360288	> 1.5	> 4.5
rs10776839	> 1.5	> 4.5
rs826472	> 1.5	> 4.5
rs735155	> 1.5	> 4.5
rs3780962	> 1.5	> 4.5
rs740598	> 1.5	> 4.5
rs964681	> 1.5	> 4.5
rs1498553	> 1.5	> 4.5
rs901398	> 1.5	> 4.5
rs10488710	> 1.5	> 4.5
rs2076848	> 1.5	> 4.5
rs2107612	> 1.5	> 4.5
rs2269355	> 1.5	> 4.5
rs2920816	> 1.5	> 4.5
rs2111980	> 1.5	> 4.5
rs10773760	> 1.5	> 4.5
rs1335873	> 1.5	> 4.5
rs1886510	> 1.5	> 4.5
rs1058083	> 1.5	> 4.5
rs354439	> 1.5	> 4.5

rs1454361	> 1.5	> 4.5
rs722290	> 1.5	> 4.5
rs873196	> 1.5	> 4.5
rs4530059	> 1.5	> 4.5
rs1821380	> 1.5	> 4.5
rs8037429	> 1.5	> 4.5
rs1528460	> 1.5	> 4.5
rs729172	> 1.5	> 4.5
rs2342747	> 1.5	> 4.5
rs430046	> 1.5	> 4.5
rs1382387	> 1.5	> 4.5
rs9905977	> 1.5	> 4.5
rs740910	> 1.5	> 4.5
rs938283	> 1.5	> 4.5
rs8078417	> 1.5	> 4.5
rs1493232	> 1.5	> 4.5
rs9951171	> 1.5	> 4.5
rs1736442	> 1.5	> 4.5
rs1024116	> 1.5	> 4.5
rs719366	> 1.5	> 4.5
rs576261	> 1.5	> 4.5
rs1031825	> 1.5	> 4.5
rs445251	> 1.5	> 4.5
rs1005533	> 1.5	> 4.5
rs1523537	> 1.5	> 4.5
rs722098	> 1.5	> 4.5
rs2830795	> 1.5	> 4.5
rs2831700	> 1.5	> 4.5
rs914165	> 1.5	> 4.5
rs221956	> 1.5	> 4.5
rs733164	> 1.5	> 4.5
rs987640	> 1.5	> 4.5

rs2040411	> 1.5	> 4.5
rs1028528	> 1.5	> 4.5

pSNP Thresholds

The following is a list of pSNPs with analytical, and interpretation thresholds.

Table 34 pSNP Thresholds

Loci	% Analytical	% Interpretation
rs28777	> 1.5	> 4.5
rs12203592	> 1.5	> 4.5
rs4959270	> 1.5	> 4.5
rs683	> 1.5	> 4.5
rs1042602	> 1.5	> 4.5
rs1393350	> 1.5	> 4.5
rs12821256	> 1.5	> 4.5
rs12896399	> 1.5	> 4.5
rs2402130	> 1.5	> 4.5
rs1800407	> 1.5	> 4.5
rs312262906_N29insA	> 1.5	> 4.5
rs1805005	> 1.5	> 4.5
rs1805006	> 1.5	> 4.5
rs2228479	> 1.5	> 4.5
rs11547464	> 1.5	> 4.5
rs1805007	> 1.5	> 4.5
rs201326893_Y152OCH	> 1.5	> 4.5
rs1110400	> 1.5	> 4.5
rs1805008	> 1.5	> 4.5
rs885479	> 1.5	> 4.5
rs1805009	> 1.5	> 4.5
rs2378249	> 1.5	> 4.5
rs16891982	> 1.5	> 4.5
rs12913832	> 1.5	> 4.5

Set Up an External Data Repository

This section describes how to set up an external data repository on your network or on an external data storage device.

Important Notes

- For the new location on the network, all services and files are accessed by the same user, typically, a dedicated service user or group. When set up is complete, verify that the service user or group can read and write files at the remote location and that all web and windows services are running as that user.
- In the following steps, domain and service_user are placeholders for the names of the network domain and the network account, respectively, linking the ForenSeq Universal Analysis Software server box.
- In the following steps, the parent to the Runs folder is the folder that contains the Runs folder in the external repository.
- During the entire move operation, do not create or start new runs or new analyses. Doing so results in an unstable system state.
- In the config files, add .<domain>.com at the end of the machine name in the new run storage location entry. For example, if the machine is Nas01, the domain is Illumina, and the runs are in UAS001\Runs on that machine. The config entry for the management service is \\nas01.illumina.com\UAS001\Runs.

  NOTE
  The new location should have a parent folder for the Runs folder. In the example, the parent folder is UAS001.

- The steps listed below occur on either the network or the ForenSeq Universal Analysis Software server. It is not necessary to do any steps on the MiSeq FGx.

Overview of Steps

1. Request a new network user.
2. Stop the ForenSeq Universal Analysis Software Management service.
3. Create the parent folder.
4. Copy the existing runs folder.
5. Validate the service_user's accessibility.
6. Configure the Forenseq Universal Analysis Software Management and Forenseq Universal Analysis Software Web services.
7. Set up the configuration file for the management service.
8. Configure the Analysis service.
9. Set up the configuration file for the Analysis service.
10. Start the Analysis service.
11. Start the ForenSeq Universal Analysis Software service.

Detailed Steps

Request a New Network User

1 Contact your technology support department to create a new network user with access to all services and file locations, including read/write permissions.

Stop the ForenSeq Universal Analysis Software Management Service

1 Open the Windows Start menu and enter IIS.

2 Select Internet Information Services (IIS) Manager in the search box.

3 In the Connections pane, expand **domain\service_user > Sites** to view ForenSeq Universal Analysis Software sites.

4 Select **ForenSeq UAS Management**.

5 In the Actions pane of IIS under Manage Website, select **Stop** to enable the Start button.

Create the Parent Folder

1 Create and name the parent folder to the Runs folder.

2 Right-click on the new folder and select Properties.

3 Select the Sharing tab.

4 Select **Share**.

5 Enter the service_user account in the Choose people on your network to share with dialog box.

6 Select **Add** and verify that the user name appears in the list box near the bottom of the screen.

7 Click the arrow under Permission Level for the service_user and check **Read/Write**.

8 Select the Share tab.

9 Select **Done** in the File Sharing dialog box.

10 Click **OK** in the Properties dialog.

11 Close the window.

NOTE
Do not create the Runs folder.

Copy the Existing Runs Folder

NOTE
Depending on the number of runs, this portion of the operation may take several hours. Consider scheduling this step to run overnight.

1 In File Explorer, browse to C:\Illumina\Forenseq UAS.

2 Right-click the **Runs** folder and Select **Copy**.

3 In File Explorer, navigate to the parent folder of the new Runs folder.

4 Open the **parent folder**, in the folder contents, right-click and select **Paste**.

5 Open some run files to verify that all of the content in the run files was copied to the new Runs folder.

6 On the ForenSeq Universal Analysis server, rename the Runs folder at C:\Illumina\Forenseq UAS\Runs to RunsOld.

CAUTION
Do not delete the old Runs folder from the ForenSeq Universal Analysis Software server.

Validate the Service_User's Accessibility

NOTE
Skip this step if installing the run files on a local drive.

1 In File Explorer, right-click **Network** and select **Map network drive**.

2 Record the drive letter displayed in the Drive field.

3 In the Address Bar, paste the path to the new network location including the Runs folder.
For example: \\nas01.illumina.com\UAS001\Runs

4 Uncheck **Reconnect at sign-in**.

5 Check **Connect using different credentials**.

6 Click **Finish**.

7 Enter the service_user name and password in the Windows Security dialog.
The domain is pre-selected and has the same name as the ForenSeq Universal Analysis Software machine domain.

8 Click **OK**.

9 Open several run files to verify that they have content.

10 From File Explorer, open any Run folder .

11 In the contents of the Runs folder, right-click and select **New > Text Document**.

12 Verify the new text document is created in the Runs folder.

13 Delete the new text document by right-clicking it and selecting **Delete**.

Configure the ForenSeq Universal Analysis Software Management Service

1 In IIS, select **Application Pools** in the Connections pane.

2 In the Application Pools pane, right-click **ForenSeq UAS Management**.

3 Select **Advanced Settings**.

4 In the Advanced Settings dialog box, scroll down to the Process Model section, select **Identity Entry**, and click the **ellipses (…)**.

5 Select **Custom Account**.

6 Click **Set**.

7 Enter the User name as domain\service_user in the Set Credentials dialog box.
- domain is the network domain
- service_user is the new user

8 Enter and confirm the Password for service_user

9 Click **OK**.

10 In the Advanced Settings dialog box, click **OK** .

11 In the Application Pools pane, right-click **ForenSeq UAS Management**.

12 Select **Recycle** or **Start**, whichever is enabled.

Configure the ForenSeq Universal Analysis Software Web Service

1 In IIS, select **Application Pools** in the Connections pane.

2 In the Application Pools pane, right-click **ForenSeq UAS Web**.

3 Select **Advanced Settings**.

4 Scroll down to the Process Model, select the **Identity Entry** click the **ellipses (…)**.

5 In the Application Pool Identity dialog box, select **Custom Account**.

6 Click **Set**.

7 Enter the User name as domain\service_user.
- domain is the network domain
- service_user is the new user

8 In the Set Credentials dialog box, enter and confirm the Password for service_user.

9 Click **OK**.

10 In the Advanced Settings dialog box, click **OK** .

11 In the Application Pools pane, right-click the **ForenSeq UAS Web**.

12 Select **Recycle** or **Start**, whichever is enabled.

Set Up the Configuration File for the Management Service

1 In File Explorer, browse to C:\Illumina\Forenseq UAS\Management.

2 Right-click the **Web** file and select **Edit** to open the file in Windows Notepad.
The Web file shows as Web.config if extensions are shown.
Write permissions are required to edit the file.

3 From the Edit menu, select **Find**.

4 In the Find what box, enter "RunFileRootPath" (including the quotes).

5 Click **Find Next**.

6 Update the entry with the path to the remote location including the Runs folder.
In this example, the updated entry is in **bold**:
<add key="RunFileRootPath" value="**\\nas01.illumina.com\UAS001\Runs**" />

7 Save and close the configuration file.

Configure the Analysis Service

1 Open the Windows Start Menu, enter Services, select **Services**, and scroll down to Illumina ForenSeq Analysis.

2 Right-click **Illumina ForenSeq Analysis** and select **Stop**.

3 Once successfully stopped, right-click **Illumina ForenSeq Analysis** again and select **Properties**.

4 Select the **Log On** tab.

5 In the Log on as section, select **This Account**.

6 In the This account box, enter the domain\service_user.
- domain is the network domain
- service_user is the new user

7 Enter the Password for service_user and click **Confirm**.

8 In the Properties dialog box, click **OK** to close it.
You should see a confirmation of the Log on as a Service permission grant.

Set Up the Configuration File for the Analysis Service

1 In File Explorer, browse to C:\Illumina\Forenseq UAS\Analysis.

2 Right-click **Matchbox.Analysis.ServiceHost.exe.config** and select **Edit** to open the file in Windows Notepad.
- The Web file shows as Web.config if extensions are shown.
- Write permissions are required to edit the file.

3 From the Edit menu, select **Find**.

4 In the Find what box, enter "RootDirectory" (including the quotes).

5 Click **Find**.

6 Update the RootDirectory to be the remote location to store the Runs folder. Do not include the Runs folder at the end of the path.
In this example, the updated entry is in **bold**:
<add key="RootDirectory" value="**\\nas01.illumina.com\UAS001**"/>

7 Save and close the configuration file.

Start the Analysis Service

1 From the Search field in the Windows Start Menu, enter Services, and select **Services**.

2 Scroll down the Services(Local) list to Illumina ForenSeq Analysis.

3 Right-click **Illumina ForenSeq Analysis** and select **Start**.

4 Verify in the Status column that Illumina ForenSeq Analysis is running.

Start the ForenSeq Universal Analysis Software Service

1 In the Connections pane of IIS, expand **domain\service_user > Sites** to view ForenSeq UAS sites.

2 Select **ForenSeq UAS Management**.

3 In the Actions pane, select **Start**.

4 Close the IIS window.

Index

A

account creation 7, 9
accounts
  access levels 83
  disabling 83
  enabling 83
  passwords 84
alignment 100-101
allele
  counting 100-101
  lengths 116
allele count quality indicator 35
amelogenin 37
amplicon 4
analysis
  execute 16
  generate new 31
  version 4, 22, 31
analysis settings 20, 74, 77, 85-86
analytical threshold 4, 37, 74, 77, 101-102, 117, 120, 122, 126

B

biogeographical ancestry 50, 58
biogeographical ancestry SNP (aSNP) 59-60, 62, 70, 75-76, 100, 120

C

cluster density 27, 106
clusters 4, 99
clusters passing filter 27, 99, 106
Control DNA 2800M 22
customer support 135
cycles 26-27, 99

D

define content 80
demultiplexing 100
DNA primer mix 12, 16, 26, 33, 50, 60, 62, 100
documentation 135

G

gender 35, 42, 67, 101
genotype
  calling 101-102
  definition 4

H

help, technical 135
Human Sequencing Control (HSC) 4, 22, 24, 106, 116

I

identity SNP (iSNP) 33-34, 50, 52, 60, 67, 69-70, 73, 100, 122
imbalance 35, 60, 102
index 4, 10, 12, 16, 27, 100
Index 1 27, 106
Index 2 106
index CV 29, 34
intensity 26, 29, 33, 37, 40, 52, 106
interlocus balance 4, 34, 37, 67
interpretation threshold 35, 37, 60, 74, 77, 101-102, 117, 120, 122, 126
intralocus balance 4, 35, 77, 86
IP address 6
isometric alleles 37, 68

L

likelihood ratio 93
locus
  detail box indicators 35
  excluding 89, 91
  including 89, 91
  length 33
  threshold template 85-86
  thresholds 85, 87
log in 9
log out 7
low coverage 35, 60, 101

M

maintenance 80
MiSeq FGx 3, 26, 98, 100
Mix Type 10, 12, 16, 87, 89, 91
mixture detection 34-35, 37, 100, 102

N

negative control 25-26
negative control icon 25
notifications 82
number of reads 29, 52

P

passing filter (PF) 27
passwords 84
phasing 27, 99, 106
phenotype 9, 50, 58, 64
phenotype estimation reports 75-77
phenotypic SNP (pSNP) 59-60, 62, 75-76, 100, 126
population group settings 44, 80, 92, 94-96
population statistics 44, 46, 54, 93
positive control 22
prephasing 27, 99, 106

primer 27, 33, 100-101
project 12
  analyses 20, 22
  assignment 16
  definition 4
  list 8
  page 20
  phenotype 64
  population statistics 46
  reports 66-68
  results 19
  sample comparison 50

Q

quality control indicator 33
quality metrics icon 26
quality scores 99

R

random match probability 45, 93
Read 1 27, 106
Read 2 27, 106
reads
  number 22, 26, 33
reagent blanks 25
reports
  phenotype estimation 75-77
  project 66-68
  sample 66, 68-74
  workflow 9
review indicators 20
RTA 3, 27, 98-99
run
  create new 8-10, 12
  delete version 16
  description 13
  description change 15
  filtering 8
  name 10, 12
  name change 15
  new version 16
  quality 20, 27
  remote monitoring 26, 89, 91
  samples 13
  save 10
  search 8
  troubleshooting 106
  version 9, 13
  workflow 9

S

sample
  add 10
  compare 9
  description 12
  edit 16
  import 10
  information 10
  locus results 33
  name 12
  reports 66, 69, 71-74
  representation 26
  single-source 34, 37, 51, 67, 102
  type 4, 10, 12
sample comparison 50-52, 54-55
sample representation 29
screen resolution 6
sequencing phase 27
sequencing run
  created 14
  details 13
  paused 14
  quality scores 20
  status 14
server address 6
settings 80
SNP
  alignment 101
  genotype calling 102
  reports 69, 73, 75-76
  sample details 40
  threshold 120, 122
source attribution threshold 46, 93
state of analysis 20
STR
  aligment 100
  genotype calling 101
  reports 68, 70-72
  sample details 37
  threshold 117
stutter 35, 101, 117
  definition 4
  filter 37, 74
system dashboard 8

T

technical assistance 135
template 20, 31, 85-86
threshold adjustment 46, 80, 93
typed 37, 40, 42

U

user management 80, 83
user modified
  indicator 35

V

viewing interface 6

W

web browser 6

Technical Assistance

For technical assistance, contact Illumina Technical Support.

Table 35 Illumina General Contact Information

Website	www.illumina.com
Email	techsupport@illumina.com

Table 36 Illumina Customer Support Telephone Numbers

Region	Contact Number	Region	Contact Number
North America	1.800.809.4566	Japan	0800.111.5011
Australia	1.800.775.688	Netherlands	0800.0223859
Austria	0800.296575	New Zealand	0800.451.650
Belgium	0800.81102	Norway	800.16836
China	400.635.9898	Singapore	1.800.579.2745
Denmark	80882346	Spain	900.812168
Finland	0800.918363	Sweden	020790181
France	0800.911850	Switzerland	0800.563118
Germany	0800.180.8994	Taiwan	00806651752
Hong Kong	800960230	United Kingdom	0800.917.0041
Ireland	1.800.812949	Other countries	+44.1799.534000
Italy	800.874909		

Safety data sheets (SDSs)—Available on the Illumina website at support.illumina.com/sds.html.

Product documentation—Available for download in PDF from the Illumina website. Go to support.illumina.com, select a product, then select **Documentation & Literature**.

Illumina
5200 Illumina Way
San Diego, California 92122 U.S.A.
+1.800.809.ILMN (4566)
+1.858.202.4566 (outside North America)
techsupport@illumina.com
www.illumina.com