applied biosystems

HID STR Genotyper Plugin

USER GUIDE

v5.2.2

for use with:
Precision ID GlobalFiler™ NGS STR Panel
Precision ID IonCode™ 1–96 Kit in 96 Well PCR Plate

Catalog Numbers A30939, A33586
Publication Number MAN0015879
Revision C.0

For Research, Forensic, or Paternity Use Only. Not for use in diagnostic procedures.

ThermoFisher SCIENTIFIC

Manufacturer: Multiple Life Technologies Corporation manufacturing sites are responsible for manufacturing the products associated with the workflow covered in this guide.

Corporate entity: Life Technologies Corporation | Carlsbad, CA 92008 USA | Toll Free in USA 1 800 955 6288

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

DISCLAIMER: TO THE EXTENT ALLOWED BY LAW, LIFE TECHNOLOGIES AND/OR ITS AFFILIATE(S) WILL NOT BE LIABLE FOR SPECIAL, INCIDENTAL, INDIRECT, PUNITIVE, MULTIPLE, OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH OR ARISING FROM THIS DOCUMENT, INCLUDING YOUR USE OF IT.

Validation notice: The following Applied Biosystems™ panel has been internally tested but has not been validated under SWGDAM guidelines: Precision ID GlobalFiler™ NGS STR Panel.

Revision history: Pub. No. MAN0015879

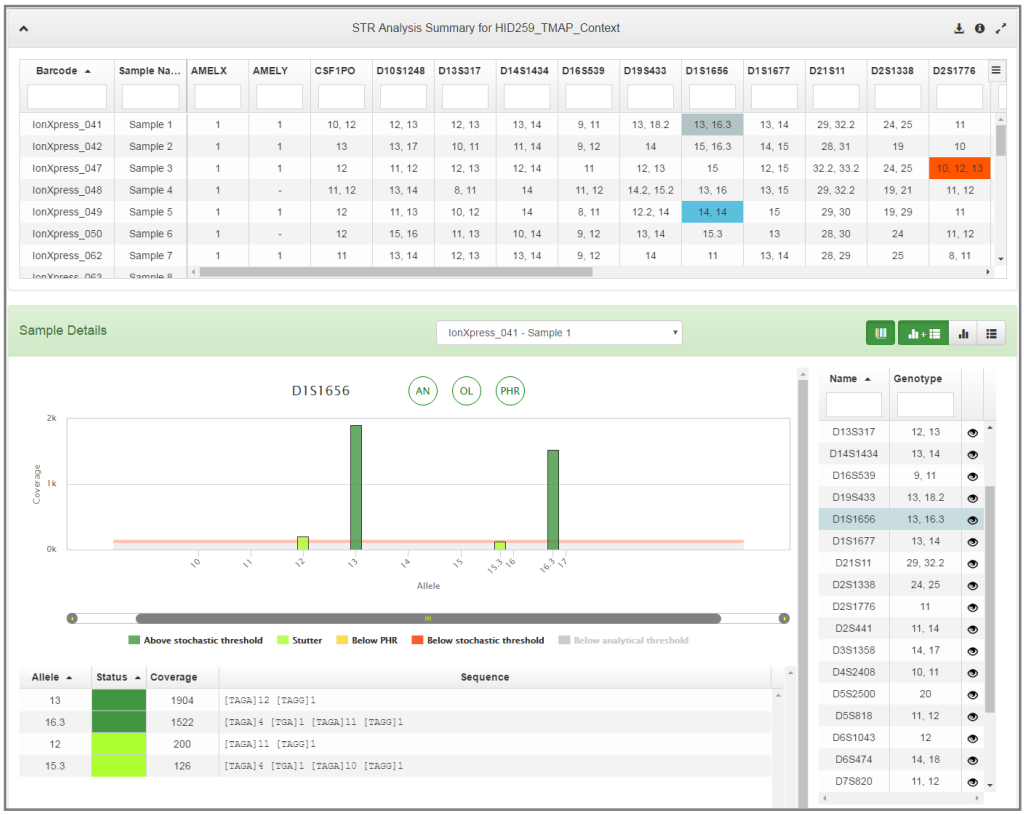
Revision	Date	Description
C.0	19 July 2017	• Updated for the 5.2.2 version of the plugin. • Instructions for installing and configuring the plugin updated. • Web links for downloading panel files and obtaining help updated.
B.0	18 May 2016	Regulatory statement updated.
A.0	12 May 2016	New document.

Contents

1 Plugin overview

Function of the plugin

The HID STR Genotyper Plugin is a software plugin that runs on the Torrent Server and is configured through the Torrent Browser. The application analyzes and interprets STR sequencing data obtained from the Ion PGM™ System.



Access Torrent Suite™ Software information

Visit **thermofisher.com/torrent-suite-software-docs.html** to access *Torrent Suite™ Software Help* and current Release Notes.

System requirements

Requirements to run the HID STR Genotyper Plugin:

- Ion PGM™ System—Torrent Suite™ Software 5.2.2
- Ion S5™ System—Torrent Suite™ Software 5.2.2

Requirements to access the data generated from the HID STR Genotyper Plugin:

- Google™ Chrome™ web browser

  Note: The HID STR Genotyper Plugin performs optimally using the Google™ Chrome™ web browser. This application is not recommended for use with other web browsers.

2 General procedures

Install the plugin

1. To install the plugin, sign in to the Torrent Server via the Torrent Browser.
2. Select ⚙ ▾, then click **Plugins**.
3. Click **Install or Upgrade Plugin**.
4. In the **Install or Upgrade Plugin** dialog, click **Select File**, navigate to and select the HID_STR_Genotyper_v5_2_2 file, then click **Upload and Install**.
5. When the upload completes, refresh the web page to ensure that the upload was successful.
6. In the HID_STR_Genotyper row, select the checkbox to enable the plugin.
   See *Torrent Suite™ Software Help* for further information.

Download and install hotspots and targets files

Download the configuration files

Precision ID GlobalFiler™ NGS STR Panel files can be downloaded from Thermo Fisher Cloud soon. Until the panel BED and analysis parameter files are available on Thermo Fisher Cloud, contact your local Thermo Fisher Scientific FAS or Technical Support for the files.

Panel-specific target BED files and analysis parameters:

File type	File name
Panel BED	PrecisionID_GlobalFiler_NGS_STR_Panel_targets.bed
Analysis parameters	PrecisionID_GlobalFiler_NGS_STR_Panel_AnalysisParams.json

When the download is complete, install the panel *.bed file.

Install the configuration files in the Torrent Suite™ Software

Note: Do not install the *.json file; it is installed in a later setup stage.

1. Sign in to the Torrent Server via the Torrent Browser.
2. Select ⚙ ▼, then click **References**.
3. In the left navigation pane, select **Target Regions**, then click **Add Target Regions**.
4. Click **Select File**, then navigate to and select the Precision_ID_GlobalFiler_NGS_STR_Panel_Targets_vX.bed file.
5. From the **Reference** dropdown list, select **hg19 - Homo sapiens hg19**.
6. (*Optional*) Enter a description and notes.
7. Click **Upload Target Regions File**.
8. Refresh the web page to ensure the upload was successful.
9. In the left navigation pane, select **Hotspots**, then click **Add Hotspots**.
10. Click **Select File**, then navigate to and select the Precision_ID_GlobalFiler_NGS_STR_Panel_Hotspot_vX.bed file.
11. From the **Reference** dropdown list, select **hg19 - Homo sapiens hg19**.
12. (*Optional*) Enter a description and notes.
13. Click **Upload Hotspots File**.
14. Refresh the web page to ensure the upload was successful.

Run the plugin

Two types of runs are available for the HID STR Genotyper Plugin:

- **Planned Run** (page 9)—Use this type of run when configuring the plugin through the Planned Run Wizard.
- **Reanalysis Run** (page 12)—Use this type of run to reanalyze the data with different parameters.

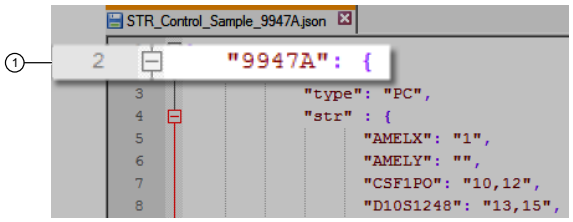
Create a Planned Run

Use this procedure to configure the plugin during Planned Run creation so that the plugin runs automatically after a sequencing run.

1. Sign in to the Torrent Server via the Torrent Browser.
2. Select **Plan ▸ Templates**.
3. In the navigation pane, click **Human Identification**, then click the **Applied Biosystems Precision ID Globalfiler STR Panel** template from the **Template Name** list
4. In the **Plan** tab, enter or select the following options.

   Note: Sample name information is transferred to the Converge™ Software with the results.

Option	Action
Run Plan Name (required)	Enter a name for the Planned Run. Do not use spaces or special characters.
Analysis Parameters	Select **Default (Recommended)**
Reference Library	Select **hg19**
Target Regions	Select **PrecisionID_Globalfiler_NGS_STR_Panel_targets.bed**
Hotspot Regions	Select **PrecisionID_Globalfiler_NGS_STR_Panel_hotspot.bed**
Use same reference & BED filles for all barcodes	Leave the option selected.
Number of barcodes	Select the appropriate number of barcodes.
Sample Tube Label	Leave blank (recommended); information is auto-populated during the Ion Chef™ template preparation run.
Chip ID	Leave blank Information is auto-populated after the Ion Chef™ template preparation run).
Barcode	Select the appropriate information from the dropdown list.
Sample (required)	Enter a name for each sample. Do not use spaces or special characters.
Control Type	Leave blank.

Option	Action
Sample ID	• **If not including a control**—**Sample ID** is optional. • **If including a control**—Enter a **Sample ID** using this naming convention: <Name of your choice>--<`Control Type`>-<`Control Name`> , where – `Control Type` is PC or NTC (positive or negative control) – The `Control Name` is provided in the control profile file.  ① In this example, 9947A is the `Control Name`. **Example positive control name:** Sample1234--PC-9947A **Example negative control name:** Sample5678--NTC-NegCtrl
Sample Description	(*Optional*) If you have already created a case for this sample in the Converge™ Software, enter the Case ID.
Reference	Leave blank.
Bead Loading (%)	Use default (**30**)
Key Signal (1-100)	Use default (**30**)
Usable Sequence (%)	Use default (**30**)

5. Click the **Ion Reporter** tab, confirm the following options, then click **Next**.
   - **Ion Reporter Account—None** (*default*)
   - **Sample Grouping—Self**

6. In the **Application** tab, confirm the following options, then click **Next**.
   - **Application—Human Identification** (*default*)
   - **Target Technique—AmpliSeq DNA** (*default*)

7. In the **Kits** tab, select the appropriate options according to the following table, then click **Next**.

Option	Action
Sample Preparation Kit	Leave blank.
Library Kit Type	Select **Ion AmpliSeq 2.0 Library Kit**. **IMPORTANT!** For HID applications, **Ion AmpliSeq 2.0 Library Kit** is the appropriate selection for use with the Precision ID Library Kit.
Template Kit	Select the appropriate template kit. For example: • Ion OneTouch™ 2 System: Select **OneTouch**, then select **Ion PGM Hi-Q OT2 Kit - 200** from the dropdown list. • Ion Chef™ System: Select **IonChef**, then select **Ion PGM Hi-Q Chef Kit** from the dropdown list.
Sequencing Kit	Select **Ion PGM Hi-Q Sequencing Kit**.
Flows	Enter **850**.
Control Sequence	Leave blank.
Chip Type	Select the appropriate chip type.
Barcode Set	Select**IonCode**.
Mark as Duplicates Reads	Leave the option deselected.
Base Calibration Mode	Select **Default Calibration**.
Enable Realignment	Leave the option deselected.

8. In the **Plugins** tab, click **Clear Selections**, select **HID_STR_Genotyper_v5_2_2**, then click the blue **Configure** link.

9. In the **Plugin Configuration** dialog:
   a. Select **hg19 (Homo sapiens)** from the **Select reference** dropdown list.
   b. Select **PrecisionID_GlobalFiler_NGS_STR_Panel_targets.bed** from the **Select BED file** dropdown list.

10. In the **Analysis Settings** pane of the **Plugin Configuration** dialog, upload the PrecisionID_GlobalFiler_NGS_STR_Panel_AnalysisParams.json file:
    a. Click ☰ in the top-right corner of the **Analysis Settings** pane.
    b. Click **Browse...**, then follow the instructions on the screen to upload the *.json file.

11. Click **Save Changes** to close the dialog box, then click **Next**.

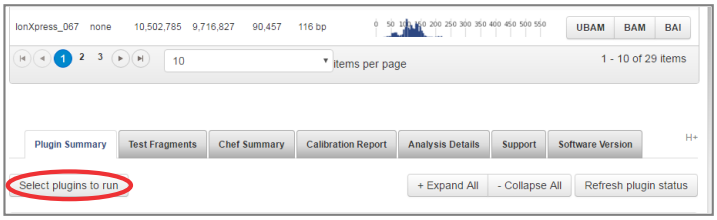
12. (*Optional*) In the **Projects** tab, assign a project, then click **Next**.
    - Select an existing project from the list. –*or*–
    - Click **Add Project**, then enter one or more project names.
13. In the **Plan** tab, enter the barcode and sample information.
14. Click **Plan Run**, then continue with sequencing.

When the run is complete, go to Chapter 3, "Review the data".

Reanalyze

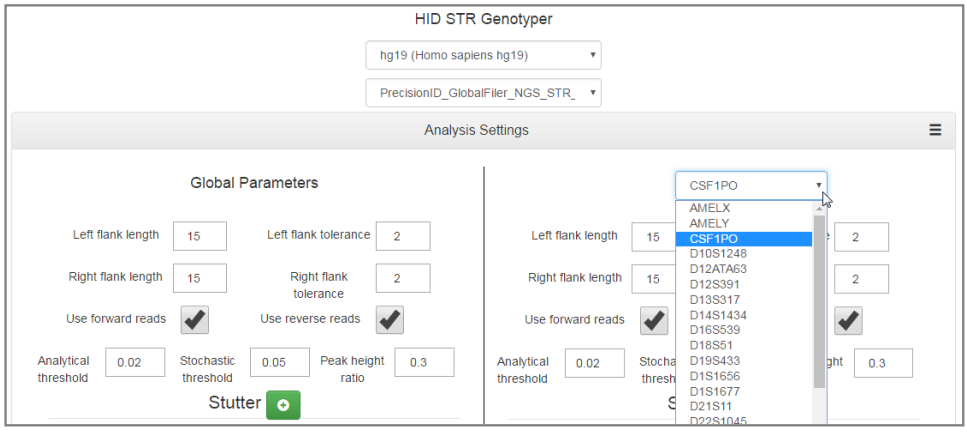
Use this type of run to reanalyze data, for example, if the experiment requires a different BED file or parameters.

1. Sign in to the Torrent Server via the Torrent Browser.
2. Select **Data** ▸ **Completed Runs & Results**, then click a **Report Name** to view the run summary.
3. In the analysis report page, click **Plugin Summary** to scroll to the bottom of the page, then click **Select Plugins to Run**.



4. In the dialog, select **HID_STR_Genotyper v5_2_2** from the list.
5. In the HID STR Genotyper Plugin configuration page:
    a. Select **hg19 (Homo sapiens hg19)** from the **Select reference** dropdown list.
    b. Select **PrecisionID_GlobalFiler_NGS_STR_Panel_targets.bed** from the **Select BED file** dropdown list.

6. In the **Analysis Settings** pane, select various loci from the top-right dropdown list to review their analysis parameters.



See Chapter 4, "Configure the input" for detailed information on panel BED files and plugin analysis parameters.

7. For each locus, modify the parameters as needed.

   Note: Any modified parameter appears in boldface, while default parameter settings are gray.

8. Save the changes to a panel BED file for future import.
   a. Click ☰ in the top-right corner of the **Analysis Settings** pane.
   b. Click **Save to file**, then follow the instructions on the screen.

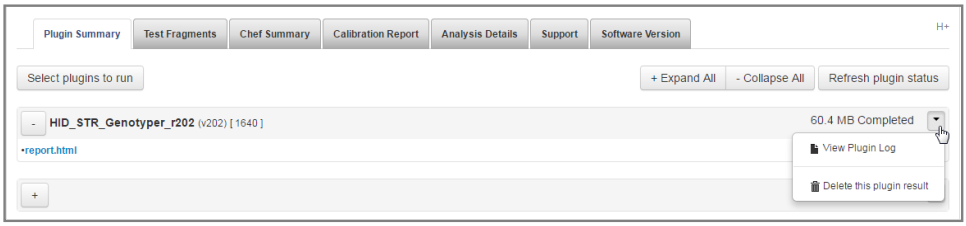
9. Click **Submit Query** to launch the analysis.

When the run is complete, go to Chapter 3, "Review the data".

3 Review the data

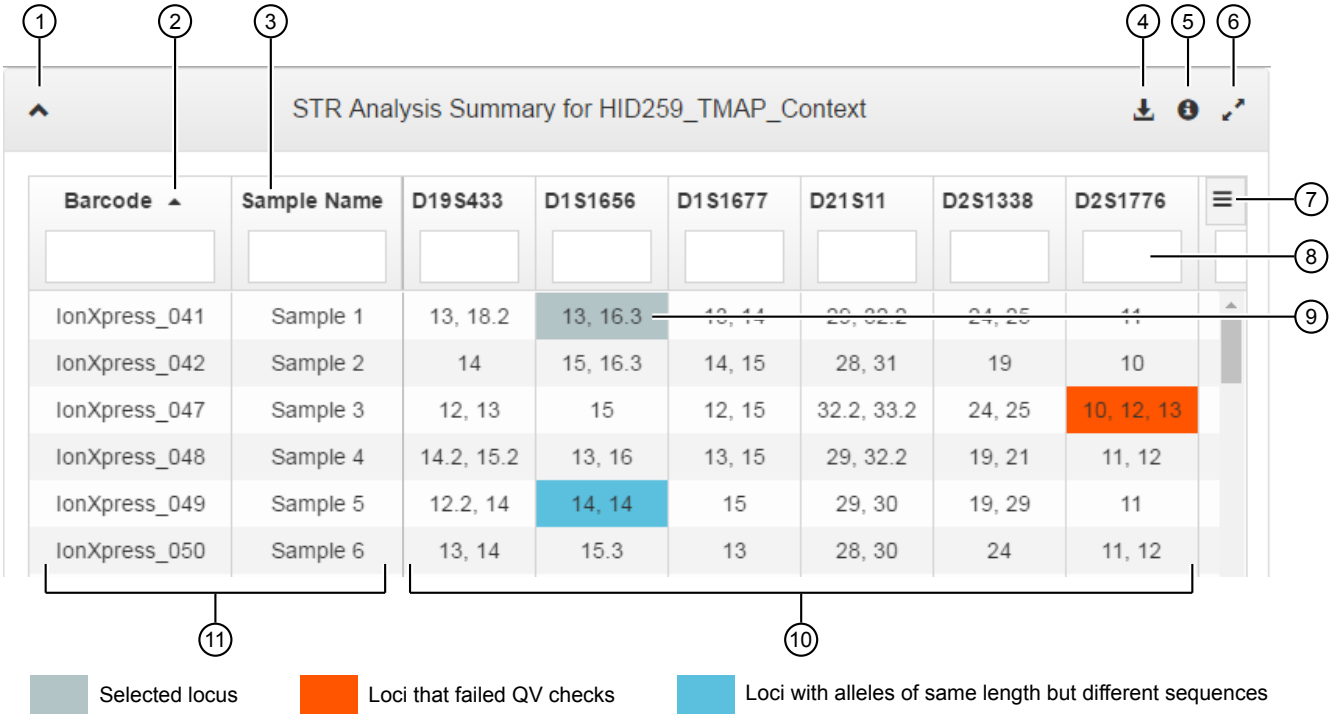
The progress of the analysis can be monitored by viewing the plugin log file.

1. (*Optional*) During analysis, click , then select **View Plugin Log**.



2. When at least one sample has been analyzed, click **report.html** to view the results.

View analysis summary table



① Show or hide the summary table
② Sort table by barcode
③ Sort table by sample name
④ Display the download options
⑤ Display run and analysis details
⑥ Expand summary table
⑦ Export table as .csv file
⑧ Search for rows containing entered text
⑨ Select cell to display coverage and genotype details
⑩ Available actions: scroll, select, and search
⑪ Available actions: pin, search, and sort

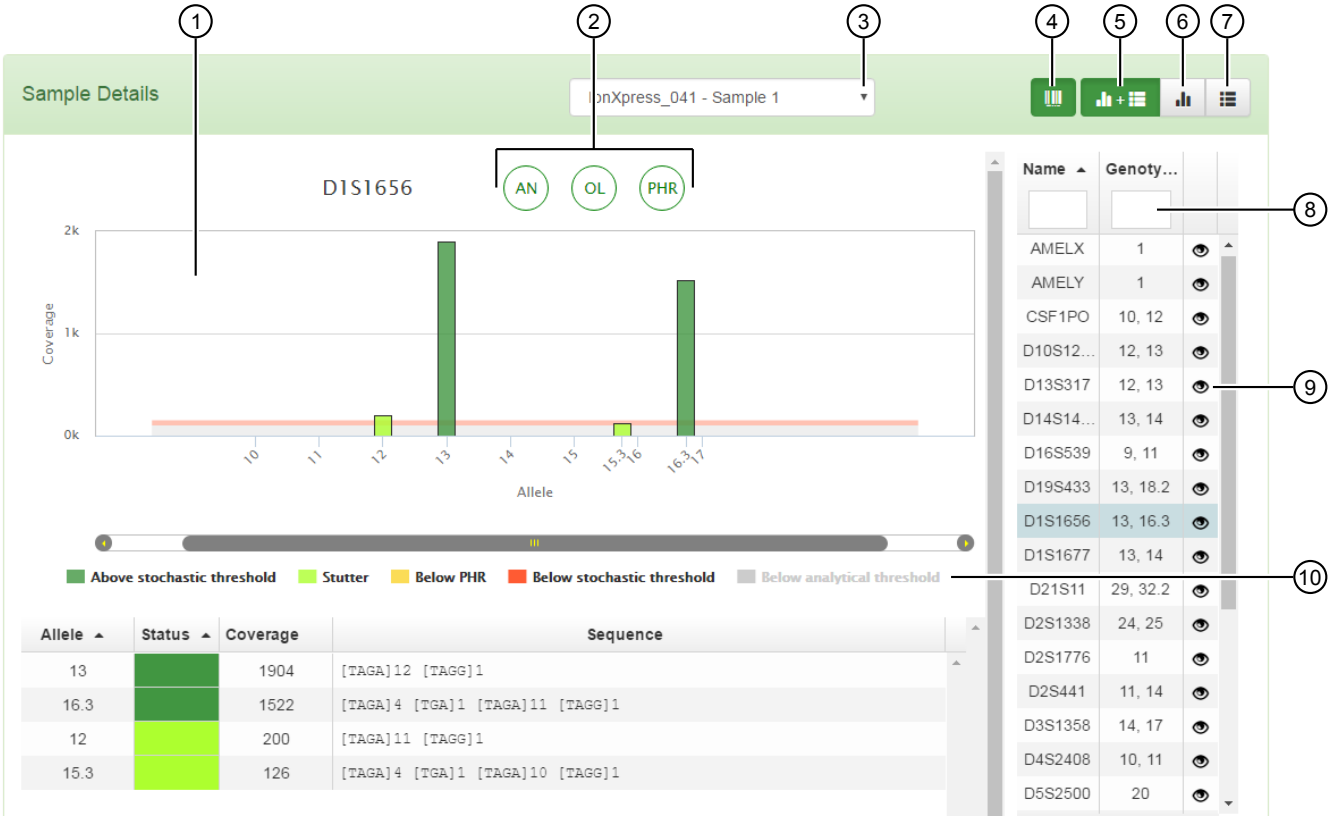
Download options for report summary table

Note: Click ⬇ to display the following download options.

- **PDF**—Contains sample genotype table, coverage plots for each sample, run details, and analysis parameters.
- **CSV**—Composite CSV file; contains genotype details for all samples in the run.
- **Panel**—BED file; used for analysis.
- **Parameters**—Contains analysis parameters used for analysis.
- **Archive**—Zip file; contains all plugin output files (does *not* include BAM files).

  Note: Your technical support representative may use this file for troubleshooting.

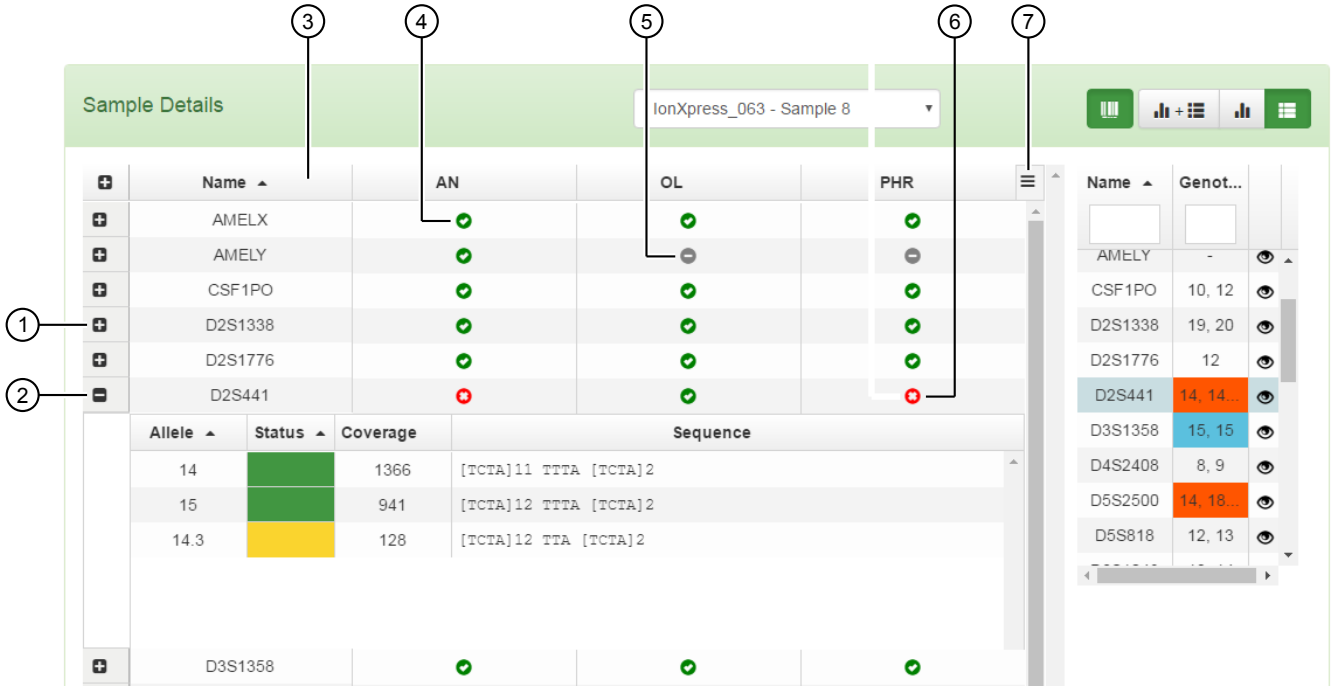
View report details



① Click and drag within chart to zoom
② Hover over to view locus QV
③ Select sample to view
④ Display details by sample or by locus
⑤ Display both coverage chart and genotype table
⑥ Display coverage chart for all loci
⑦ Display genotype table with locus QV for all loci
⑧ Search for rows containing entered text
⑨ View reads in IGV
⑩ Toggle displayed alleles with chart legend

Note: In the IGV view, the alignment to hg19 may not be accurate. Coverage details shown in IGV may not be the same as shown in the charts and tables.

Genotype table view



① Expand a locus marker
② Collapse a locus marker
③ Click a header to sort the genotype table
④ Status: passed all QV checks
⑤ Status: undefined; no alleles detected
⑥ Status: failed one or more QV checks
⑦ Export table as .csv file

Configure the input

Configuration overview

The plugin is designed to provide flexibility in terms of kit configuration and analysis parameter tuning. You can define the set of loci to be analyzed, then define analysis parameters for each locus and even for specific alleles.

There are two ways to configure the inputs for the plugin:

- **Panel BED file**—Properties that are intrinsic to an STR locus are defined in a panel BED file. These parameters do not change if the kit configuration changes, and can include locus coordinates in the reference, number of bases in the repeat unit, repeat motif, expected (ladder) alleles, etc.
- **Analysis parameters**—Properties that are specific to a kit or analysis run are defined in the plugin configuration dialog (see "Reanalyze" on page 12), and can include coverage thresholds, stutter ratios, etc. You can save these parameters to a file for use in future analyses.

Panel BED files

A BED file is a tab-delimited text file. It defines the list of loci to be analyzed along with various loci properties.

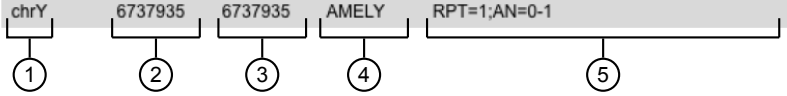
Upload a BED file as a **Target Regions** file to the **Torrent Server** via the Torrent Browser. See the *Torrent Suite™ Software User and Admin Guides* for detailed information on uploading BED files.

The following example shows the various fields of a BED file.

- The first line is required and must begin with `track`.
- The value of the **type** field must be `bedDetail`.
- The chromosome numbers and coordinates should match those in the reference against which the file is uploaded.
- The far-right column defines various attributes of the loci. Attributes are key-value pairs separated by a semi-colon (`;`).

  Note: Attributes are optional. If absent, the plugin assumes a default value.

```
track type=bedDetail name="STR_Early_Access" description="STR_Early_Access with motif structure"
chr1    230905362   230905429   D1S1656    MOTIF=[TAGA][TGA][TAGG][TG];REV=TRUE;ALLELES=10-20,10.3-20.3
chr2    1493425     1493456     TPOX
chr11   2192319     2192346     TH01       REV=TRUE;ALLELES=6-10,9.3
chr19   30417142    30417197    D19S433    MOTIF=[AAGG](AAGGTAGG);REV=TRUE
chrX    11315050    11315050    AMELX      RPT=1;AN=1
chrY    6737935     6737935     AMELY      RPT=1;AN=0-1
```



① Chromosome number (*required*)
② Start coordinate (*required*)
③ End coordinate (*required*)
④ Locus name (*required*)
⑤ Details (*optional*)

RPT attribute

The RPT attribute is the number of bases in a repeat unit.

This attribute impacts the allele name. The sequence length is divided by this value to obtain the allele name.

Value type	Default value	Example
Integer	4	• `RPT=3` (value for D12ATA63) • `RPT=4` (value for D2S1338)

ALLELES attribute

The ALLELES attribute is the expected list of alleles in a locus. It is similar to GeneMapper™ *ID-X* Software control alleles.

This attribute impacts OL (Off-Ladder) QV. If the sample has an allele in this locus that is not present in the list, the OL flag is triggered.

Value type	Default value	Example
• Numeric range (closed interval; for example, both ends included) • Individual values	No default value. All alleles are accepted and OL is not flagged.	• `ALLELES=10-12,10.3-12.3` (means 10, 11, 12, 10.3, 11.3, 12.3) • `ALLELES=6-10,9.3` (means 6, 7, 8, 9, 10, 9.3)

AN attribute

The AN attribute is the expected number of alleles in a locus.

This attribute impacts the AN (Allele Number) QV. If the number of non-noise alleles in the sample for this locus is outside the range, the AN flag is triggered.

Value type	Default value	Example
• Integer range • Max value	• 1-2 • Min=1 (if only max is given)	• `AN=0-1` (means at most one allele is expected in the locus) This is typically the case for Y markers. • `AN=2` The default min value is 1 so expect either one or two alleles.

REV attribute

For some loci, the forensic community uses the reverse strand to denote STR sequences. For example, conventional notation for D5S818 is `[AGAT]` whereas the hg19 reference strand is `[TCTA]`. To display the compressed sequence in the conventional form, set the REV attribute field to TRUE.

This attribute impacts the compressed allele sequence.

Value type	Default value	Example
• TRUE • FALSE	FALSE	• `REV=TRUE` (value for D5S818, CSF1PO, etc.)

MOTIF attribute

The MOTIF attribute is the repeat structure.

This attribute impacts the compressed allele sequence. It is used to find the number of repeats for each part.

Value type	Default value	Example
Nucleotide sequence enclosed in [ ], { }, ( ). • Specify each repeat unit only once. • Order is not important.	First **RPT** bases after the start coordinate.	• `MOTIF=[AAGG](AAGGTAGG)` • `MOTIF=[AGAT][GATA]` • `MOTIF=[TCTG]{TA}{TCA}`

Define repeat structure

The MOTIF attribute defines the repeat structure of a locus. To support the various ways in which compound repeats can be represented, the plugin provides three formats for denoting a repeat unit.

- **Bases within [ ]**
  Form a repeat unit and count towards final allele length.

Motif	Sequence	Compressed	Allele
`[AGAT][GATA]`	`AGAT AGAT GATA GATA`	`[AGAT]2 [GATA]2`	4

- **Bases within { }**
  Form a constant unit and do *not* count towards final allele length.
  Constant units are required and their absence reduces the final allele length by as many bases.

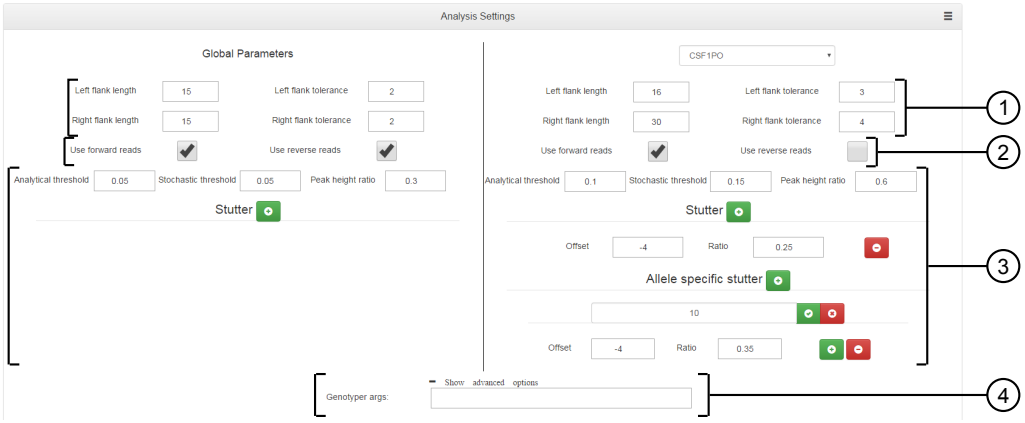
Motif	Sequence	Compressed	Allele
`[TCTA]{TA}`	`TCTA TCTA TA TCTA`	`[TCTA]2 TA [TCTA]1`	3
`[TCTA]{TA}`	`TCTA TCTA TCTA TCTA`	`[TCTA]4`	3.2

- **Bases within ( )**
  Form a constant unit and count towards final allele length.
  Constant units are required and their absence reduces the final allele length by as many bases.

Motif	Sequence	Compressed	Allele
`[AAGG](AAGGTAGG)`	`AAGG TAGG AAGG AAGG`	`AAGGTAGG [AAGG]2`	4
`[AAGG](AAGGTAGG)`	`GGTAGG AAGG AAGG`	`GGTAGG [AAGG]2`	3.2

Analysis parameters

There are four broad sections in the **Analysis Settings** window. To view a short description of a parameter, hover over its label.



① Read processing
② Read filtering
③ Coverage thresholding
④ Advanced

Parameter settings and files

Use optimized parameters

Use the custom parameter file (GlobalFiler_NGS_STR_Panel_AnalysisParams.json), for correct analysis of the Precision ID GlobalFiler™ NGS STR Panel. The custom parameter file is included in the set of downloaded panel files.

1. Click ☰ in the top-right corner of the **Analysis Settings** pane.
2. Click **Choose File**, then follow the instructions on the screen to upload the *.json file.

Set parameters manually

(*Optional*) Tune analysis parameters for each locus and for specific alleles.

Note: You can modify the settings in the plugin configuration dialog, then save the settings for future import.

Note: The values in the **Global Parameters** pane are default values if the corresponding value for a particular locus has not been specified. Any modified parameter appears in bold font, while default parameter settings are gray.

Read processing

Initially, the flank-matching algorithm processes reads from a sample. This algorithm searches for flanks of specific length on either side of the STR region. Reads are transferred to downstream analysis only if both flanks are present.

You can adjust algorithm settings for:

- The number of bases for which to search in each flank.
- The number of mismatches (SNPs/indels) that are allowed in the flank.

The length of the flanks depends on the primer design. The default value for flank length is 15 and the tolerance is 2. For best results, set the flank length so that:

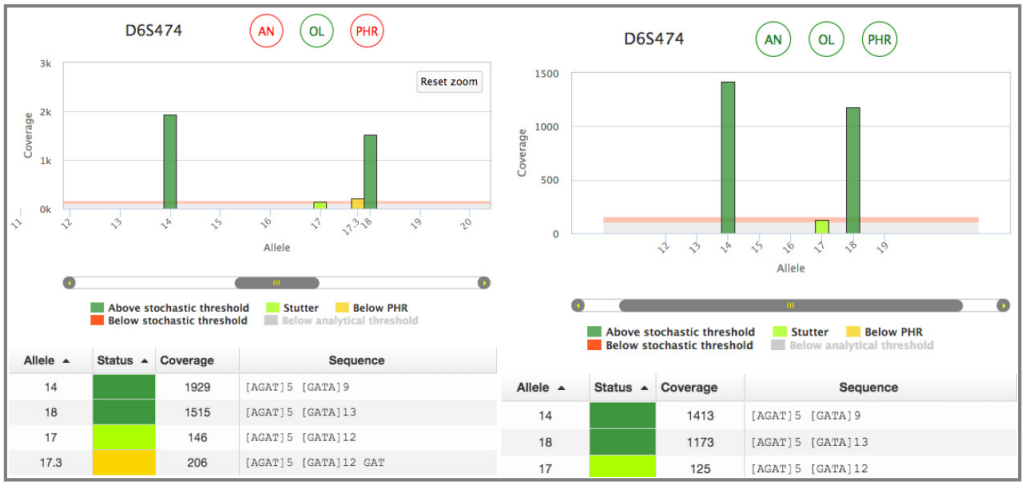
- There are at least as many bases sequenced beyond the STR region flank length.
- The length is sufficiently long to uniquely identify a locus.

Read filtering

Sometimes, reads on one strand tend to have systematic errors. Ignore reads from the failing strand by deselecting them.

Note: There will be less coverage for alleles because the deselected strand is ignored during analysis.

In the following example, an artifact has been removed by ignoring the forward strand. Notice the decrease in coverage and the absence of 17.3 (yellow peak) on the right-side chart.



Coverage thresholding

Alleles are classified into five categories based on their coverage.

Category	Description
Below analytical threshold	allele coverage < analytical threshold • These reads are considered as noise and are not considered for further analysis.
Below PHR	allele coverage < (PHR) × (max allele coverage at this locus)
Below stochastic threshold	allele coverage < stochastic threshold
Above stochastic threshold	allele coverage > analytical threshold, and allele coverage > stochastic threshold • Allele is not considered as a stutter.
Stutter	allele coverage > analytical threshold • Position is at stutter location • Allele coverage < (stutter ratio) × (primary allele coverage)

Specify stutters

Define a stutter by its offset from the primary allele (number of bases) and the maximum expected stutter ratio (percentage). Offset is negative if the stutter is shorter than the primary allele and positive otherwise.

In the **Analysis Settings** pane:

- Click ⊕ to add a stutter.
- Click ⊖ to delete a stutter.

Specify allele-specific stutters

Stutters can be defined for a particular allele. If an allele has a different stutter characteristic than others, you can explicitly specify its stutter properties.

In the **Analysis Settings** pane:

- Click ⊕ to add an allele-specific stutter. Enter the name of the allele, define one or more stutters, then click ✓ to save changes.
- Click ⊖ to delete a stutter.

Allele-specific stutters have higher precedence than locus-level stutters followed by global stutters.

In the following example for a tetrameric locus, there is a 20% stutter at the +1 position (4 bases to the right). However, the property is modified to 10% stutter for allele 12 and 15% for allele 14. In addition, for allele 14, there is a 20% stutter at the –1 position (4 bases to the left).

Advanced parameters

Use the following settings to accelerate data analysis and for advanced troubleshooting.

Parameter	Description	Default
-d	Down sampling rate. Increase this parameter to process fewer reads. Modifying this parameter will affect coverage. Example: "-d 5" will process 1 in every 5 reads.	1 (no down sampling)

A Kits and reagents

Precision ID GlobalFiler™ NGS STR Panel

For detailed information about this panel, see the *Precision ID Panels with Ion PGM™ System Application Guide* (Pub. No. MAN0015830).

Panle	Cat. No.	Storage
Precision ID GlobalFiler™ NGS STR Panel	A30939	–30°C to –10°C

Precision ID IonCode™ 1–96 Kit in 96 Well PCR Plate

The Precision ID IonCode™ 1–96 Kit in 96 Well PCR Plate (Cat. No. A33586) contains a set of 96 unique barcode adapters in a 96-well plate format for use in manual library preparation. When used in combination with the Precision ID Library Kit, this kit enables pooling of up to 96 libraries for multiplex sequence analysis.

Component	Quantity	No. of reactions	Storage
Precision ID IonCode™ 1–96 Kit in 96 Well PCR Plate	1 × 96-well plate (20 µL/well)	960 (10 reactions per barcode)	–30°C to –10°C

Documentation and support

Related documentation

Document	Publication number
Precision ID Panels with Ion PGM™ System Application Guide	MAN0015830
Torrent Suite™ Software Help	See thermofisher.com/torrent-suite-software-docs.html
IonCode™ Barcode Adapters 1–384 Kit Product Information Sheet	MAN0014640

Customer and technical support

For support:

- **In North America**—Send an email to **HIDTechSupport@thermofisher.com**, or call **888-821-4443 option 1**.
- **Outside North America**—Contact your local support office.
- For latest services and support information for all locations, go to **thermofisher.com/support**.

Limited product warranty

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