A Sysmex Group Company

Interpret

NGS Analysis Software

CytoSure SureSeq

Interpret User Guide

for Interpret v3.5

OGT

Founded by Professor Sir Ed Southern, OGT supplies world-class genetics research solutions to leading clinical and academic research institutions.

- CytoSure ® — Class-leading products offering the complete array and NGS solution for clinical genetics research.

- CytoCell ® — High-quality fluorescence *in situ* hybridisation (FISH) probes for the detection of gene rearrangements related to inherited genetic disease and cancer.

- **SureSeq™** — Hybridisation-based NGS panels and library preparation reagents delivering flexibility and unparalleled coverage uniformity for clinical research applications.

For more information, visit www.ogt.com.

Contents

User Guide for Software Version 3.6

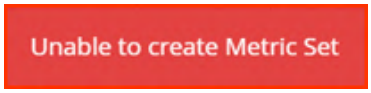
Released November 2023

1 Notices

1.1 Limitations of Use:

This software product is classified as For Research Use Only. It is not intended for use in diagnosis or treatment of human or animal diseases.

1.2 Symbols Used In This Guide

Symbol	Meaning
	Attention - Denotes critical information that users need to be aware of.
	Error Icons - Such as the one shown below, are highlighted in red and to continue they must be removed. This is easily accomplished by simply clicking on the icon.

2 In This User Guide

This guide is a manual for using OGT's Interpret NGS analysis software and is designed to be used in conjunction with OGT's range of NGS panel products.

2.1 Overview

This section serves as an introduction to the Interpret NGS Analysis software.

2.2 Before Getting Started

This section outlines requirements prior to the software being used. These requirements are critical to the correct functioning of the software and should be read and understood prior to any data being analysed.

2.3 Detailed Software Features and Workflows

This section deals with the software in detail providing comprehensive step by step instructions in use and administration of the software.

2.4 Technical Support

Please use the contact form in the Help and Support page to request further assistance.

3 Overview

Interpret is a powerful bioinformatic tool designed to allow easy and comprehensive analysis of NGS data generated from OGT's NGS Panel products. The software functionality is limited solely to OGT's NGS panels - other panels CANNOT be loaded. The input is raw data in the form of FASTQ files.

3.1 Resources Required to Operate Interpret

Access to the software is via a web browser, OGT recommend Google Chrome, however the data is processed and stored by OGTs NGS analysis pipeline. OGT provides an installer to manage the process of deploying the software. Please consult the separate installation guide for further information on this.

Users must ensure that browsers have pop-ups enabled .

The computational resources required for processing data with Interpret will depend on the number of samples being processed and the depth of sequencing. As a guideline we would recommend:

Requirement	**Minimum**	**Recommended**
Memory (RAM)	16 GB	24 GB
CPU Cores	8	16
Storage (Disk space)	500 GB	2 TB
Operating System	Windows 7 or newer	
Other	Virtualisation (VT-x) to be enabled in the BIOS	

Table: Minimum and recommended hardware requirements for running Interpret

3.2 Preparing Data For Analysis

- **Demultiplexed FASTQ files – REQUIRED**

For each sample to be processed a pair of corresponding paired end FASTQ files are required. The software expects the FASTQ files to be compressed with gzip .

Software to demultiplex FASTQ files is not part of the functionality provided by Interpret and must be implemented prior to loading of the FASTQ files into Interpret. It is assumed that this will be provided by the sequencing instrument vendor.

- **Target Regions File – REQUIRED**

The Target Regions File is supplied by OGT with each of our NGS Panel products. The Target Regions File defines the regions covered by its associated panel. Interpret will use the regions within the panel file to define range of the analysis.

The Target Regions Files supplied by OGT are in a proprietary format; ONLY files supplied by OGT can be used by Interpret.

- **Protocol File – REQUIRED**

The Protocol File defines settings used by the NGS analysis pipeline. A default protocol is supplied with the panel but additional protocols can be created and stored within the software.

4 Accessing Interpret

Interpret is accessed through a web browser, all up to date browsers should work, however **we recommend using** Google Chrome.

To discover the correct web address of the software please consult the Installation Guide.

4.1 Login to Interpret

The login screen for Interpret will appear as shown below. To login submit the user name and password for your account.

If you do not have an account please contact your administrator to have one created.

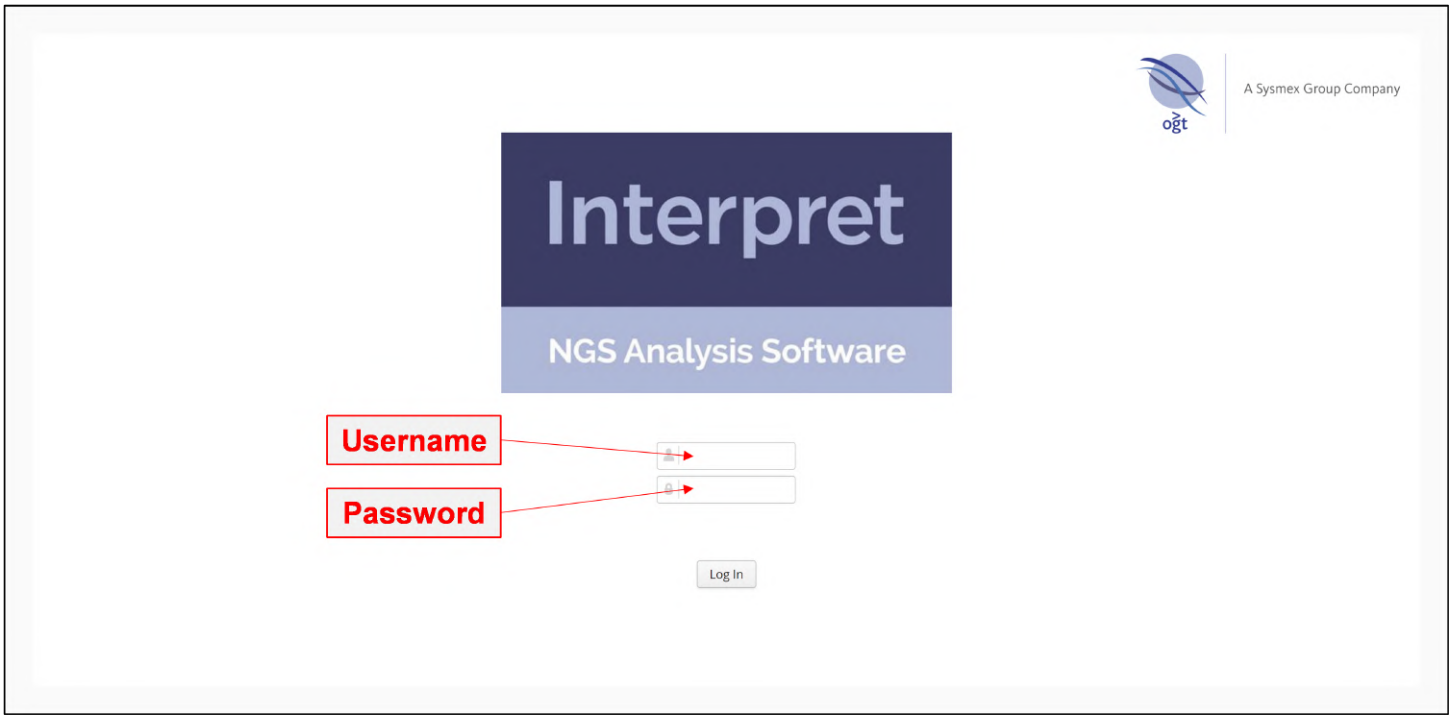


Figure : The login window of Interpret

4.2 The Dashboard

After successful login the default dashboard page will be displayed as shown below:

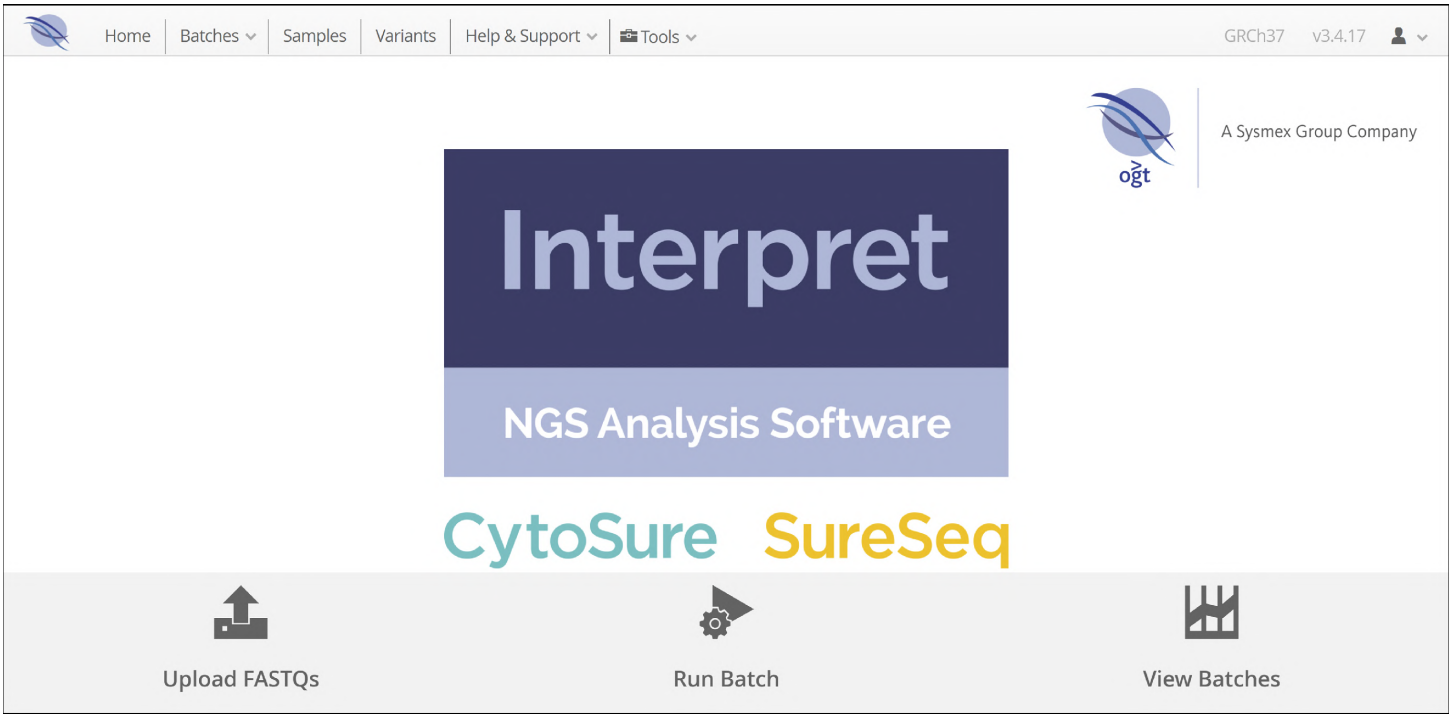


Figure: The dashboard view of Interpret

4.3 Logging Out of Interpret

To logout of the software move the mouse to the user icon on the top right of the dashboard page and in the drop-down select 'Logout'

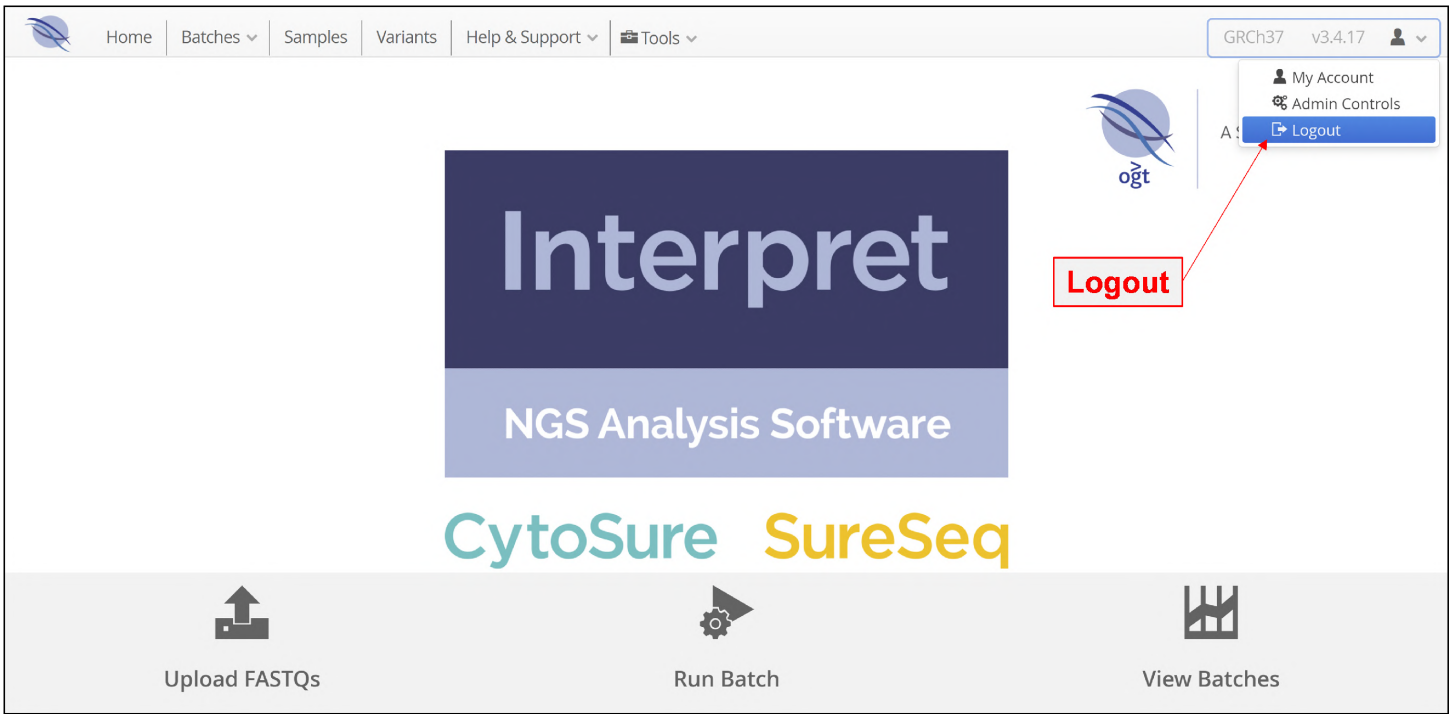


Figure: Accessing the logout option of Interpret

5 The Dashboard View

The dashboard view displayed below comprises 3 sections

- Menu Bar
- Dashboard buttons to provide function shortcuts
- User account options

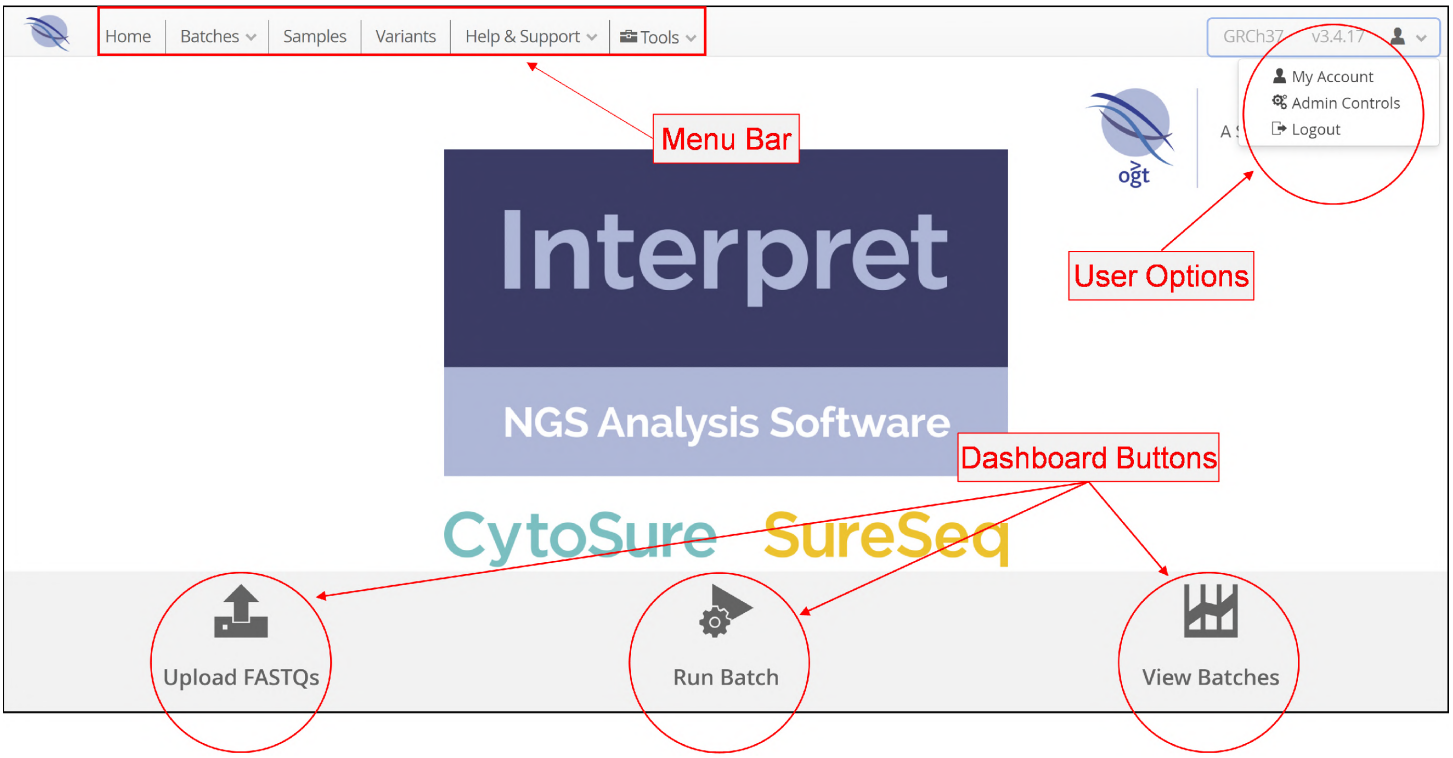


Figure: Annotated view of the dashboard

Menu Bar

The menu bar provides access to the functionality:

- Home - Link back to the Dashboard View
- Batches - Setting up and reviewing analysis batches
- **Samples -** Sample related functions
- **Variants-** Provides a means to view all data from a variant centric view
- Help & Support - A means to provide feedback as well request support
- **Tools -** Access to any additional tools

Figure: The menu bar from the dashboard

Dashboard Buttons

These provide shortcuts to the common actions required by users.

- Upload FASTQs - Select and upload FASTQ files.
- Run Batch - Run an analysis of a batch of loaded sample files
- View Batches - View the results of the batch analyses

Figure: Shortcut icons on the dashboard view

User Options

The User Options drop down menu gives the user access to a range of administration tools. Additionally this section of the dashboard displays the build of the genome being used as well as the version of the software. In this case it is GRCh37 and v3.3.61.

The drop down options are as follows:

- **My Account -** Your account details
- Admin Controls - Additional options described in detail in the Admin Options section of the guide
- **Logout -** Return to the Login page

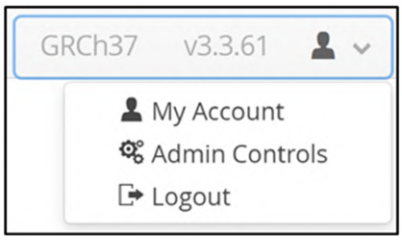


Figure: User account options

6 Loading FASTQ Files

To load FASTQ files, on the dashboard either select 'Upload FASTQs' in the drop down from the 'Batches' menu item.

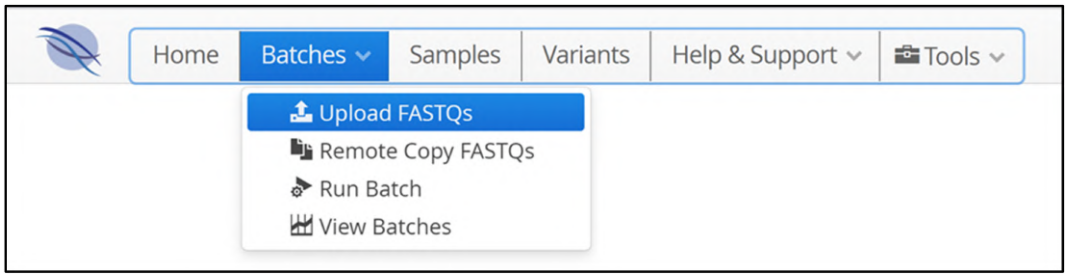


Figure: Accessing FASTQ uploads from the menubar

Or, click on the 'Upload FASTQs' icon on the dashboard page

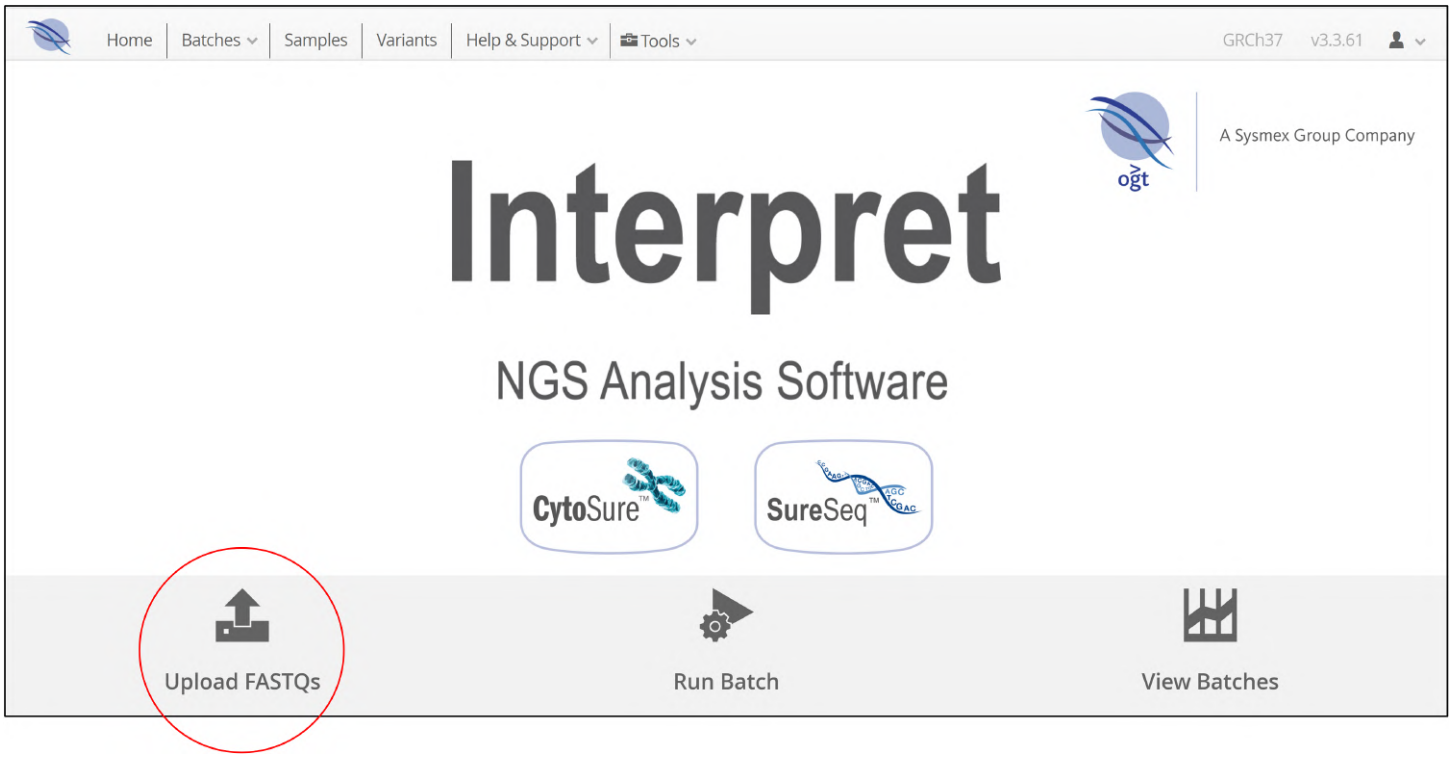


Figure: Accessing FASTQ uploads via the dashboard short-cut

Either choice opens the Upload FASTQs window.

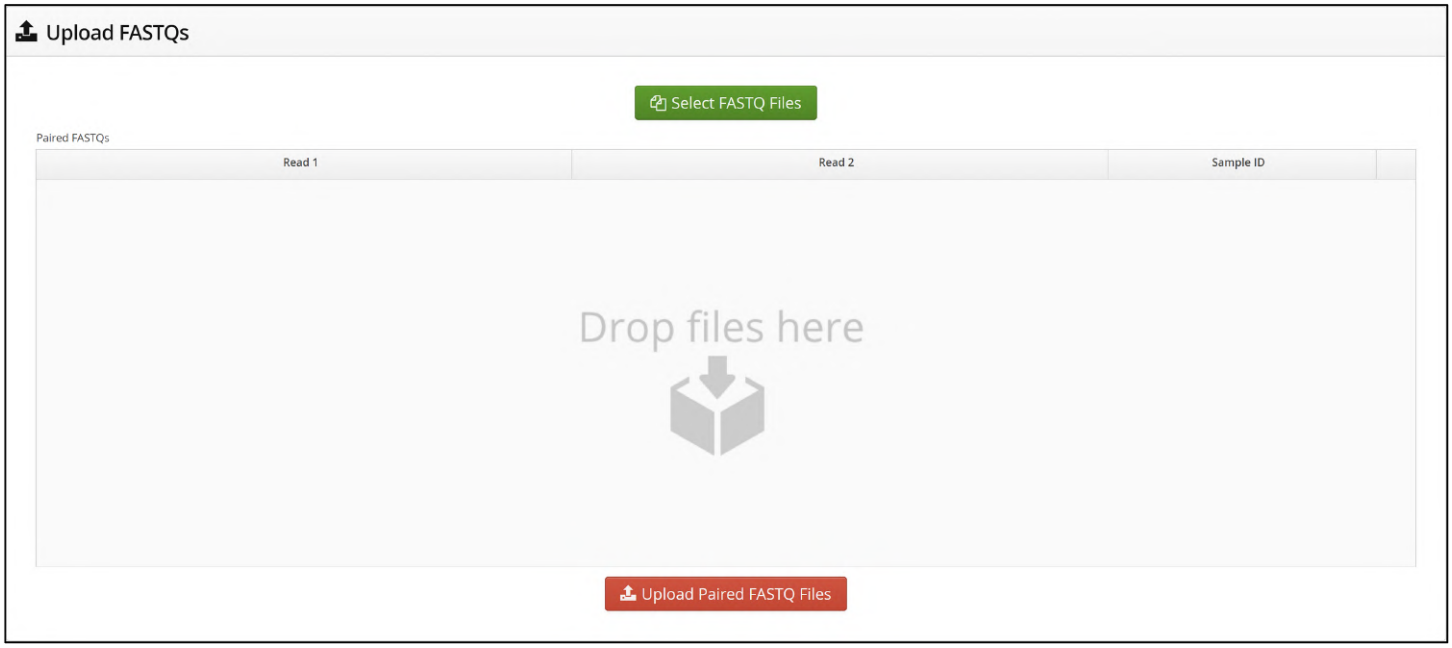


Figure: Upload FASTQs form

Initially the window shows an empty table with the heading "Paired FASTQs" with column headers for Read1, Read2 and Sample ID.

Figure: Initial view of the FASTQ upload table

To load FASTQ files users can either click on the 'Select FASTQ files' button to open a file browser or they can drag and drop files directly.

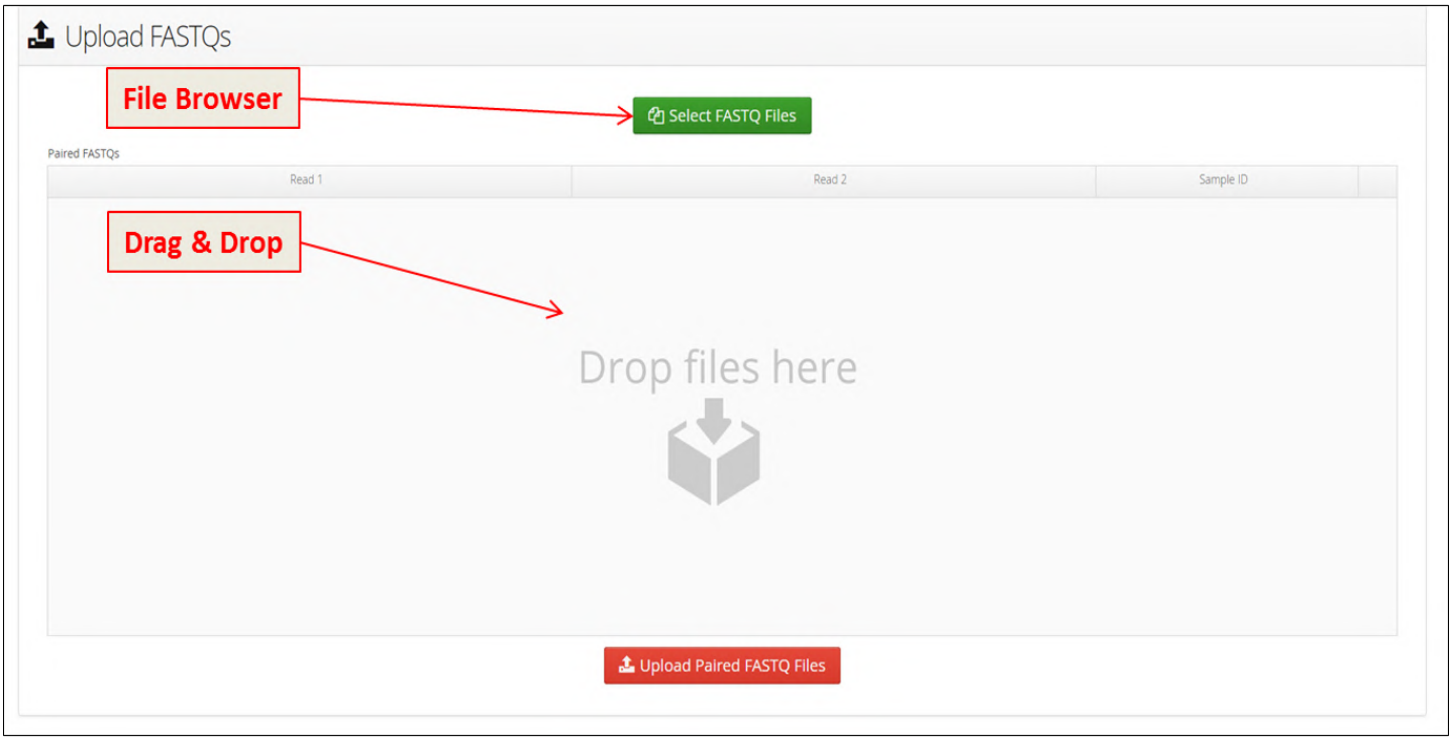


Figure: Methods of uploading files to Interpret

The software requires all FASTQ files to be compressed by gzip. Any file without this file extension will not be loaded.

When FASTQ files are loaded the software will try to automatically pair them into Read 1 and Read 2. This is based on the filenames automatically generated by Illumina sequencers.

The file matching protocol assumes that a file pair shares the same name up to the annotation for whether the file is for Read1 data or Read2 data which for Illumina would be either _R1 or _R2.

For example, with the following pair of FASTQ files:

Read1: Sample-400_R1.fastq.gz

Read2: Sample-400_R2.fastq.gz

The software would be automatically select the Sample ID as the portion of the file name highlighted in magenta.

Sample-400 _R1.fastq.gz

Sample-400 _R2.fastq.gz

As there is a Read1 file and Read2 file with the matching file names the software will automatically pair them. In the first instance the Sample ID is set to Sample-400, though, this can be easily changed.

These can be seen now populating the "Paired FASTQs" table.

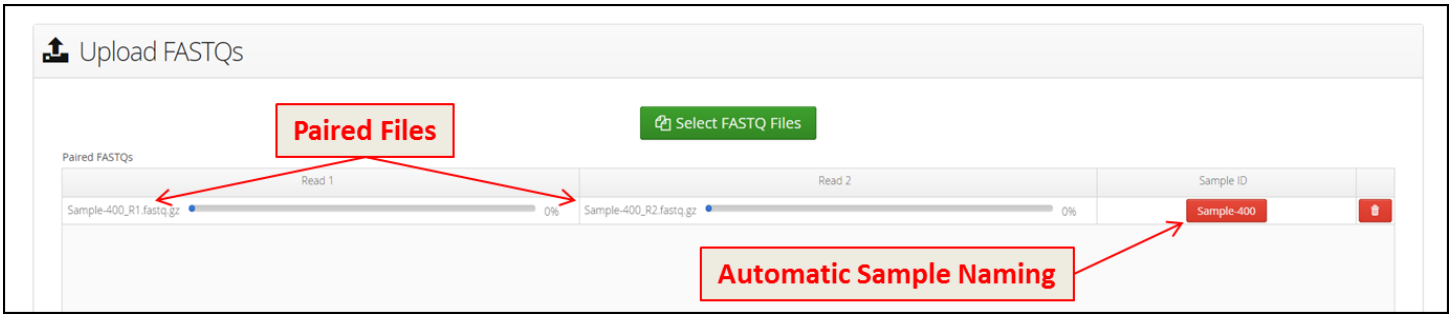


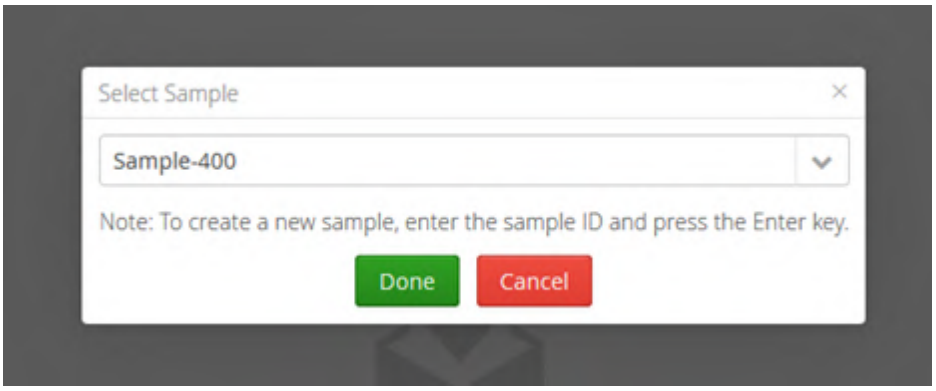
Figure: Automatic pairing of FASTQ files

6.1 Renaming a Sample ID

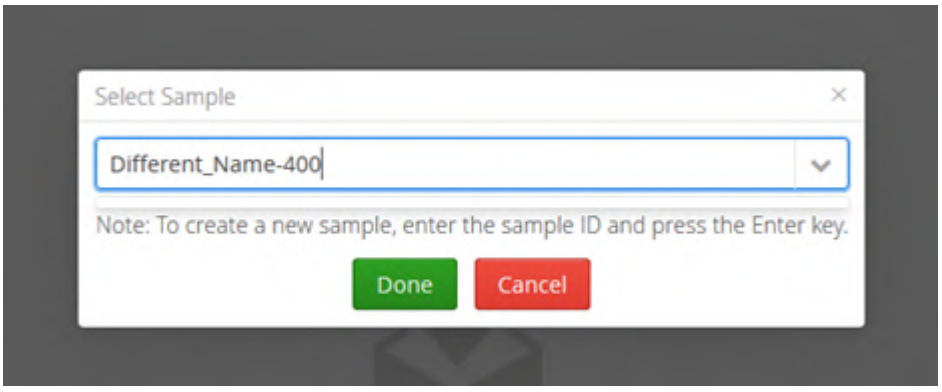
As you can see in the previous screenshot a Sample ID is initially displayed with a red background as this is the ID that has been automatically generated.

If the Sample ID is not correct then it can be easily modified by clicking on the Sample ID that needs to be modified.

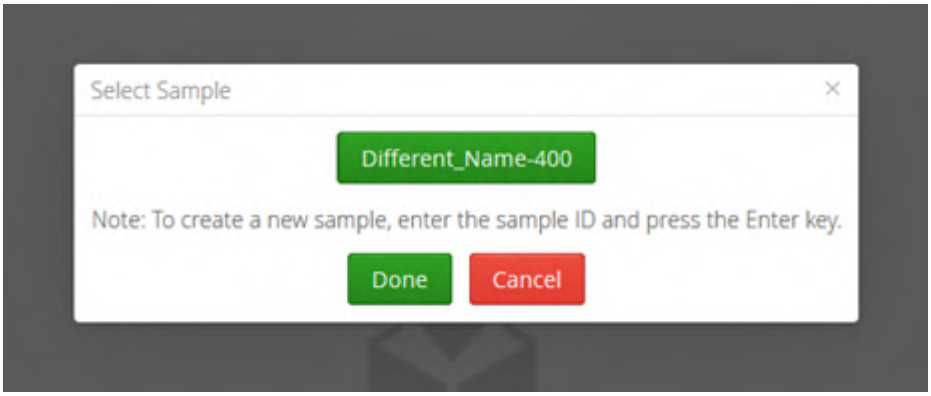
In the popup, the current Sample ID will be shown.



In order to change this, enter the required name in the text box and press enter.



The updated name is then displayed.



Selecting Done updates the display with the new Sample ID

In the example below the background colour has changed to green to represent the Sample ID for that pair of files has been modified by the user.

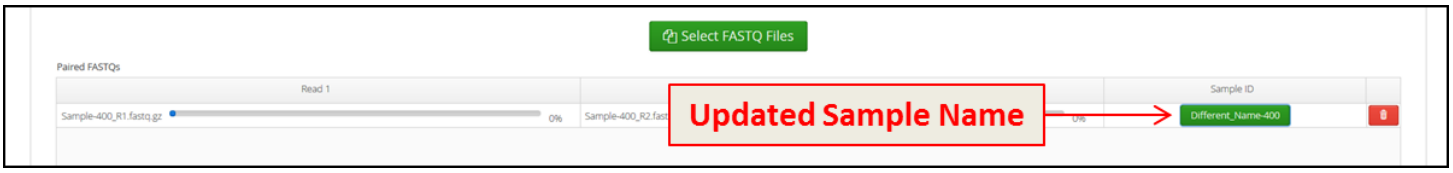


Figure: Table with updated Sample ID

6.2 Pairing 'unpaired' FASTQ Files

Sometimes it may be the case that files have been paired incorrectly.

If this is the case and there are files that are not a pair then clicking on the bin icon  will remove the samples from the Paired FASTQs table and move them to a new table called "Unpaired FASTQs".

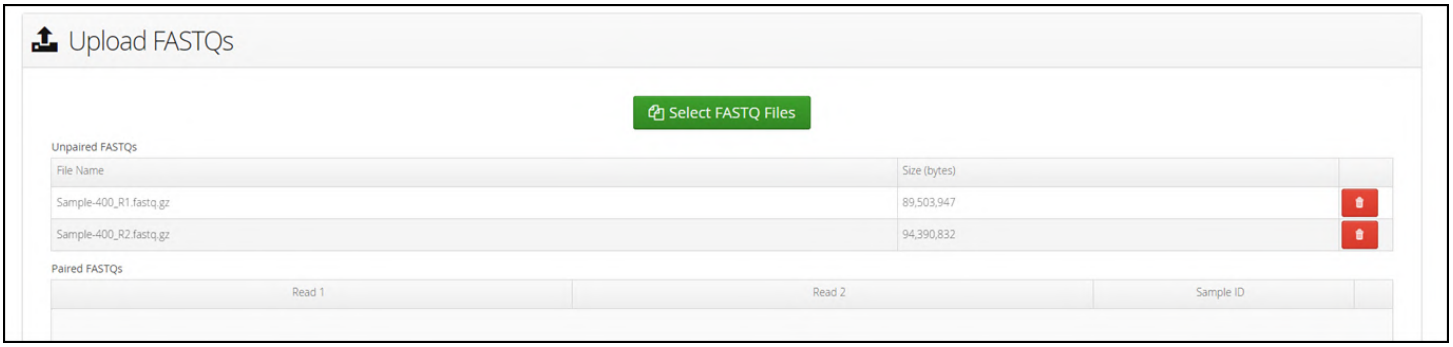


Figure: Table with unmatched FASTQ files

From here they can easily combine as a pair by highlighting one file and then clicking on the second. When this is done, they are considered to be a pair again and moved back to the Paired FASTQs table.

Alternatively clicking on the bin icon  for files in the Unpaired FASTQs table removes them from the Upload FASTQs page.

When files without matching file names are loaded, they are initially displayed in a panel for Unpaired FASTQs.

However, an alternative pair of read files without matching names such as:

This_Fastq_R1.fastq.gz

and

That_Fastq_R2.fastq.gz

would not be automatically paired. This needs to be completed by the user.

Figure: Table with unmatched FASTQ files highlighted

Selecting two of these files by shift clicking automatically denotes them to be a pair and they are moved to the Paired FASTQ panel.

The software will select the Read1 file in the pair to be the first FASTQ file that is selected, so in the example below ONE_Fastq_R1.fastq.gz is selected so it becomes the Read1 file.

However, initial selection the other file ANOTHER_Fastq_R2.fastq.gz would have resulted in the alternative situation

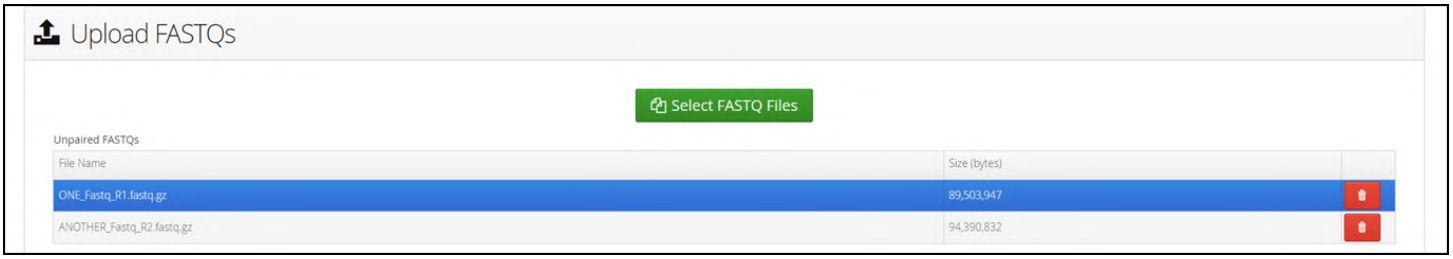


Figure: Selecting an unpaired FASTQ file

The software creates a name for the Sample ID based on the file name for the file denoted as the Read1 file, but this can be changed by following the protocol for changing a Sample ID.

Figure: Table showing the pairing of unmatched FASTQ files and the automatically assigned name

6.3 Proceeding with File Upload

Now that the files have been paired and correct Sample IDs have been assigned, they are ready to be uploaded.

Figure: Manually paired FASTQ files ready for upload

At the bottom of the Paired FASTQs table are 3 parameters that allow tracking of the upload.

These are:

- An estimate of the amount of time remaining.
- The amount of data uploaded as an amount in Mb out of the total as well as a percentage completed statistic.
- The proportion of files uploaded.

Figure: Three different metrics for upload of FASTQ files can be monitored

Selection of the "Upload Paired FASTQ Files" button starts the upload process, though there is an additional check run by the software when there is a Sample ID that has been automatically assigned.

In this case the user must confirm the Sample ID is correct before the upload is initiated.

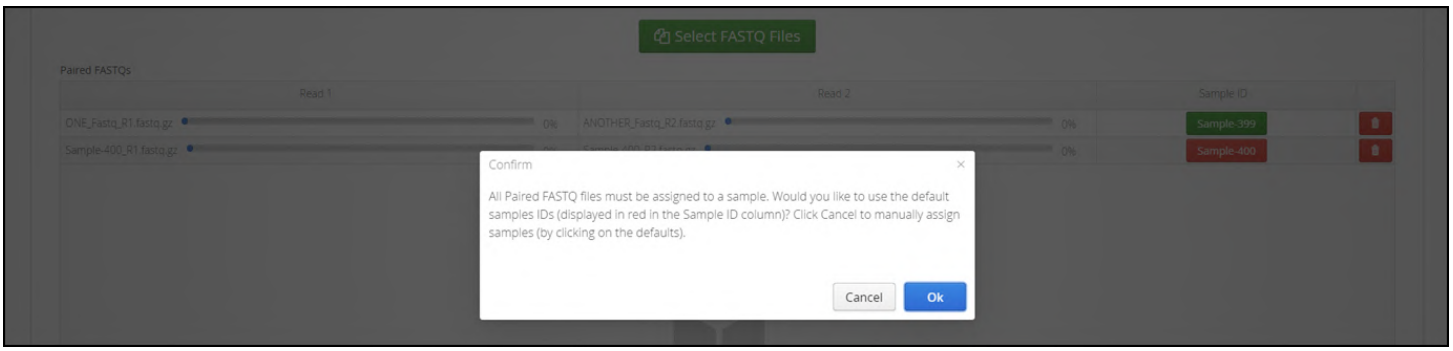


Figure:

Once it has started users are able to track progress of the upload by looking at the blue progress bars for each of the files. In order for the download to progress it is important to keep the browser tab open (or in a separate window) while uploading FASTQs, otherwise the upload will pause.

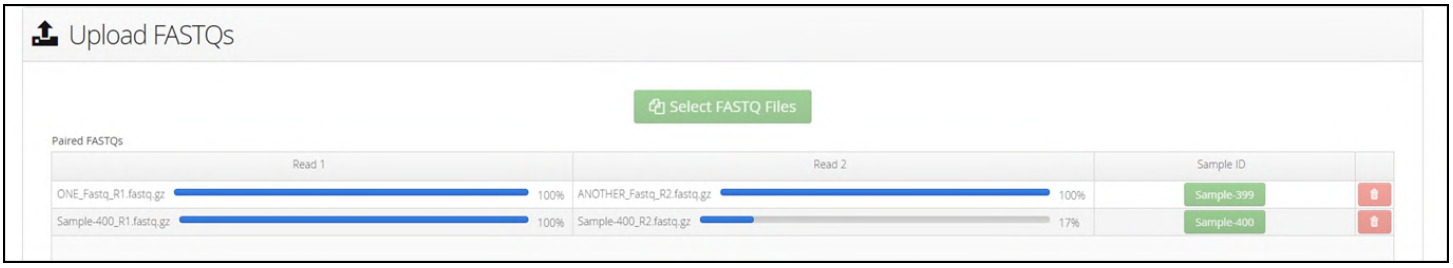


Figure: Progress bars tracking file uploading

Alternatively there are the tracking metrics displayed which can give a more precise estimation of how the upload is progressing.

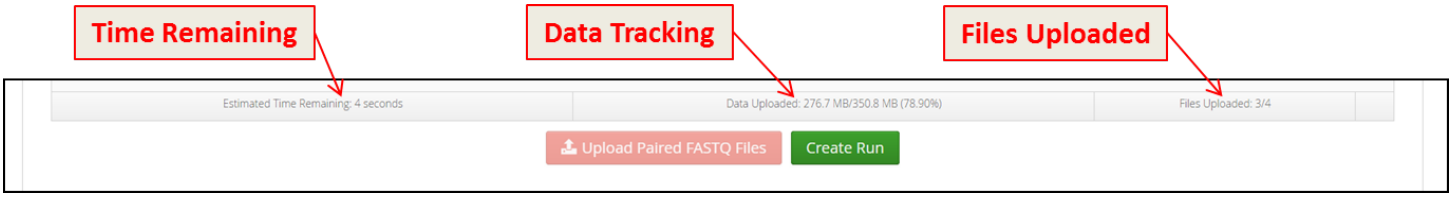


Figure: Metrics for tracking file uploading

Once the upload is complete, the software asks whether the user would like to create a new analysis run or to wait. Setting up an analysis will be covered in the next section.

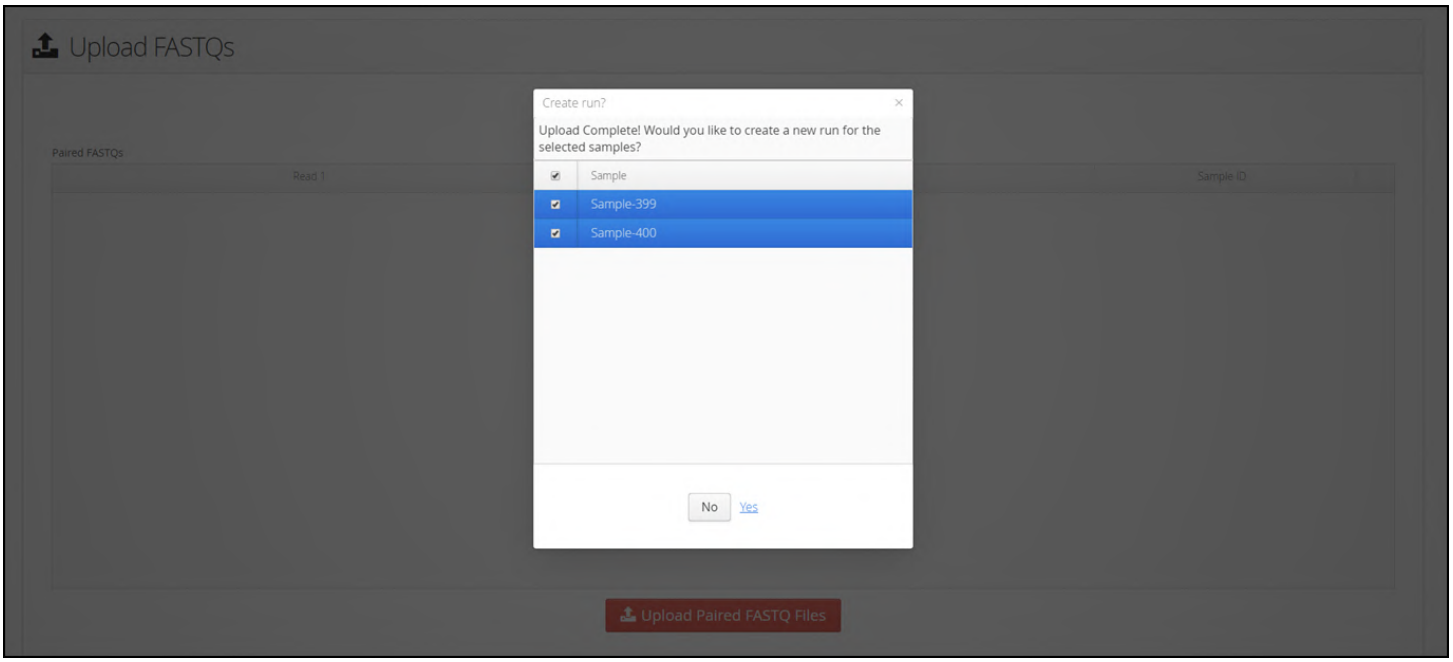


Figure: Popup menu displayed once FASTQ file upload is complete

6.4 Remote Copy of FASTQ files

In addition to previous method of loading FASTQ files it is also possible to run a remote copy of files accessible on the network mounted to Interpret.

This is accessed by Batches drop-down in the menu bar.

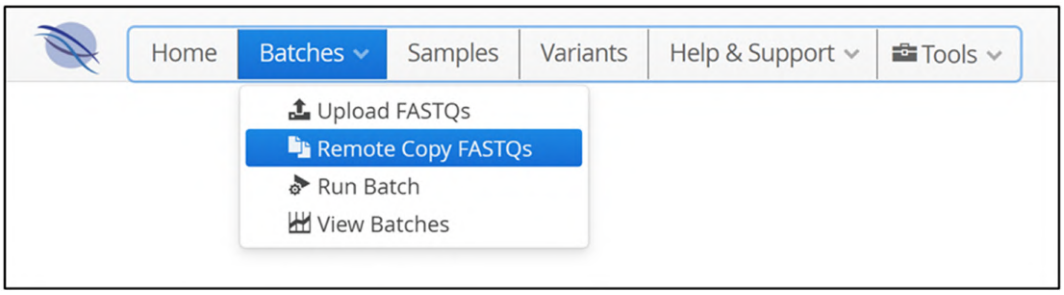


Figure: Accessing the 'Remote Copy FASTQs' function from the menu bar

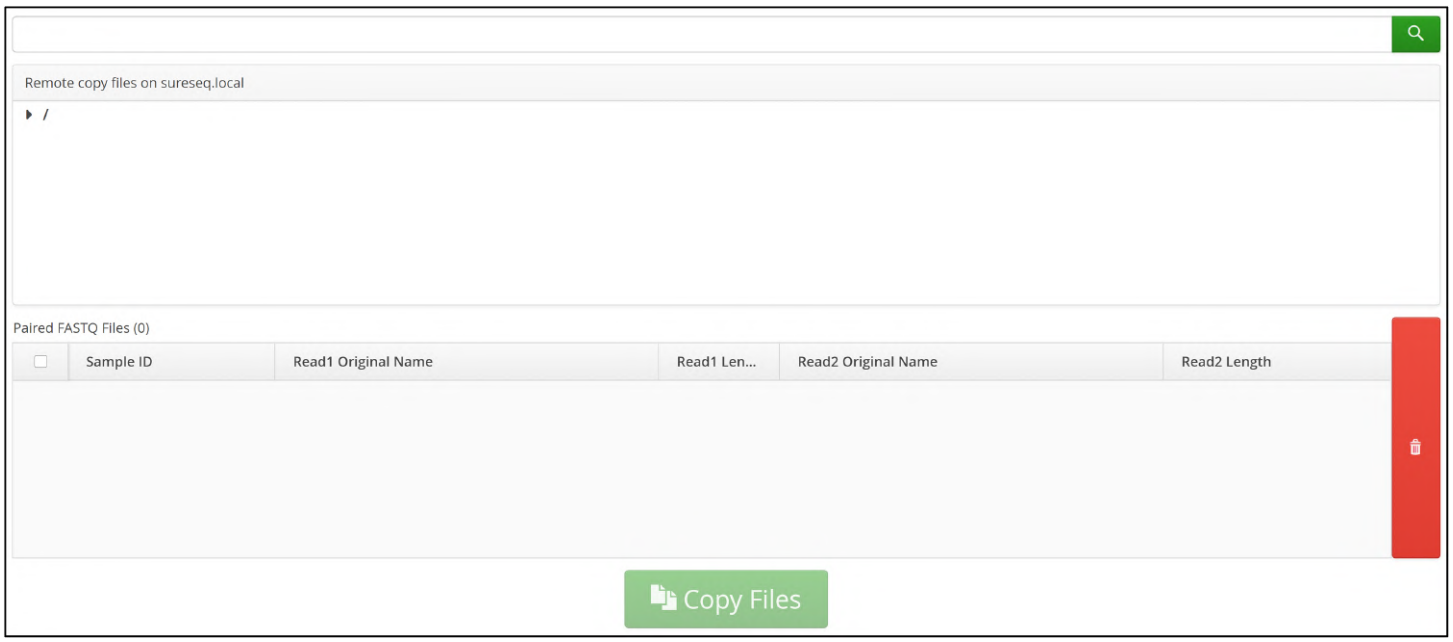


Figure: The remote FASTQ file copy interface

If input FASTQ files are located in a folder which is accessible by the web application, it is possible to navigate the file system and select FASTQ files for upload to the system directly from this folder. Upload using this method has the advantage of avoiding the need to maintain an active web browser during the upload process, as files are copied in the background. Whether such folders are accessible to the web application will depend on many factors, including how the software has been installed. It may be the case that it is not possible to access input data folders, but if the required folders are not accessible, contact OGT for assistance.

To identify FASTQ files, expand the file tree to the required location. FASTQ files will be displayed in green, and can be added to the list by double-clicking. Alternatively, double-click on a folder to add all FASTQ files in that folder. In some browsers, double-clicking on a folder will instead set the "root" location of the file tree to the selected folder, and it may be necessary to double-click a second time in order to add the files to the list.

Files can be removed from the list of Paired FASTQ files by selecting the associated checkbox and clicking the red delete button. Once the required set of FASTQ files have been selected, click the Copy Files button to begin the upload

process. As with the Upload FASTQ Files section, users then have option of creating a batch from the selected FASTQ files.

> Sample IDs
> Please note that the Remote Copy interface does not provide a means to modify the sample ID associated with each FASTQ file - these are instead automatically determined by the name of the FASTQ file based on the standard nomenclature.

7 Viewing Samples

Once the FASTQ files have been uploaded the samples will be available to view in the Samples page

Accessing this is via the Samples button on the dashboard menu bar shown in the figure below.

Figure: Selection of Samples from the Dashboard menu bar

Samples and status are displayed on the left hand side of the window and when a sample is selected further information is displayed on the right hand side.

When a sample is first loaded there is no additional information present.

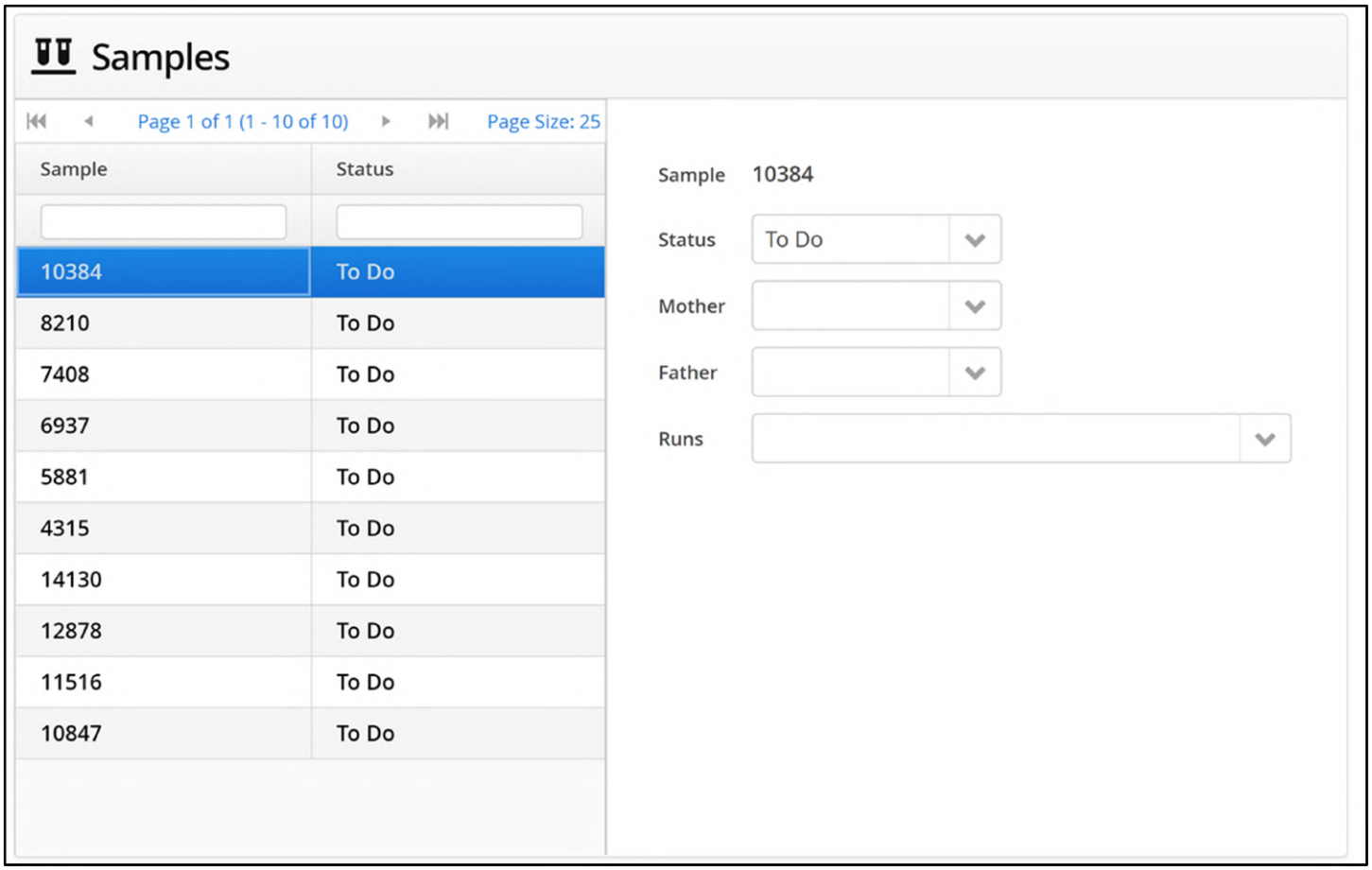


Figure: The initial view of samples in the Sample page

Samples can be searched by entering a part of the sample name in the search box. In the example below all samples containing "10" are displayed.

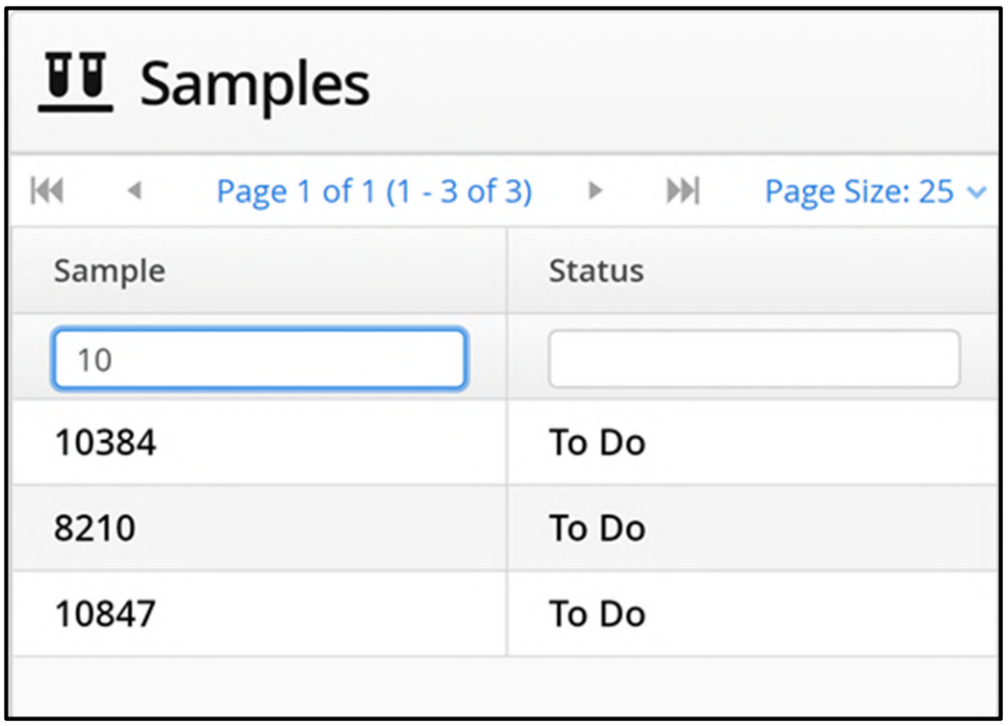


Figure: Searching for samples containing 10

The status of a sample can be updated. When first loaded, the status will be set to "To Do" and will be updated to "Running Pipeline" once processing has begun. Once sample processing is complete, the status will be further updated to "In Review". Users can assess the results if the analysis and manually update the sample status to "Completed" as required.

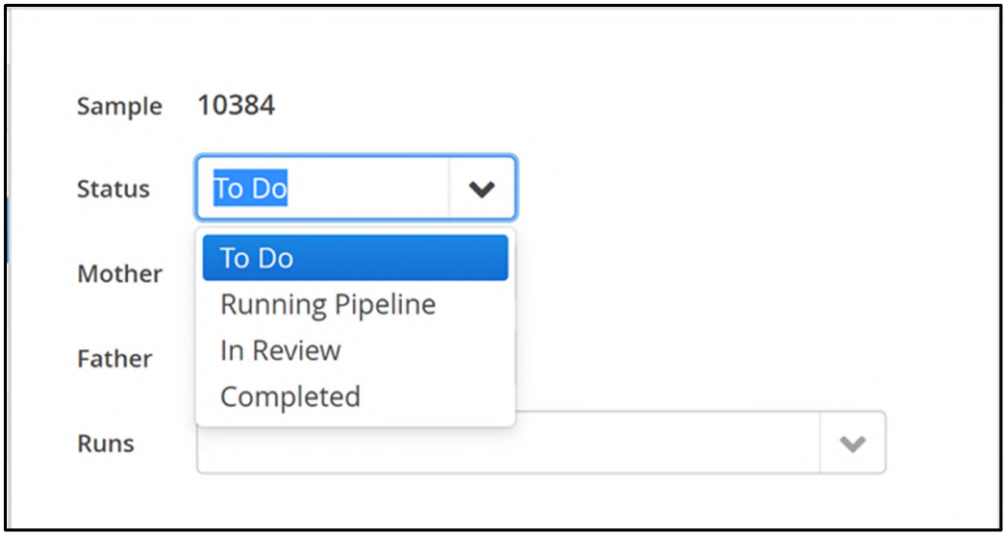


Figure: Modifying the status of a sample in the Sample view

It is also possible to specify the mother and father of a sample if they are also loaded in Interpret.

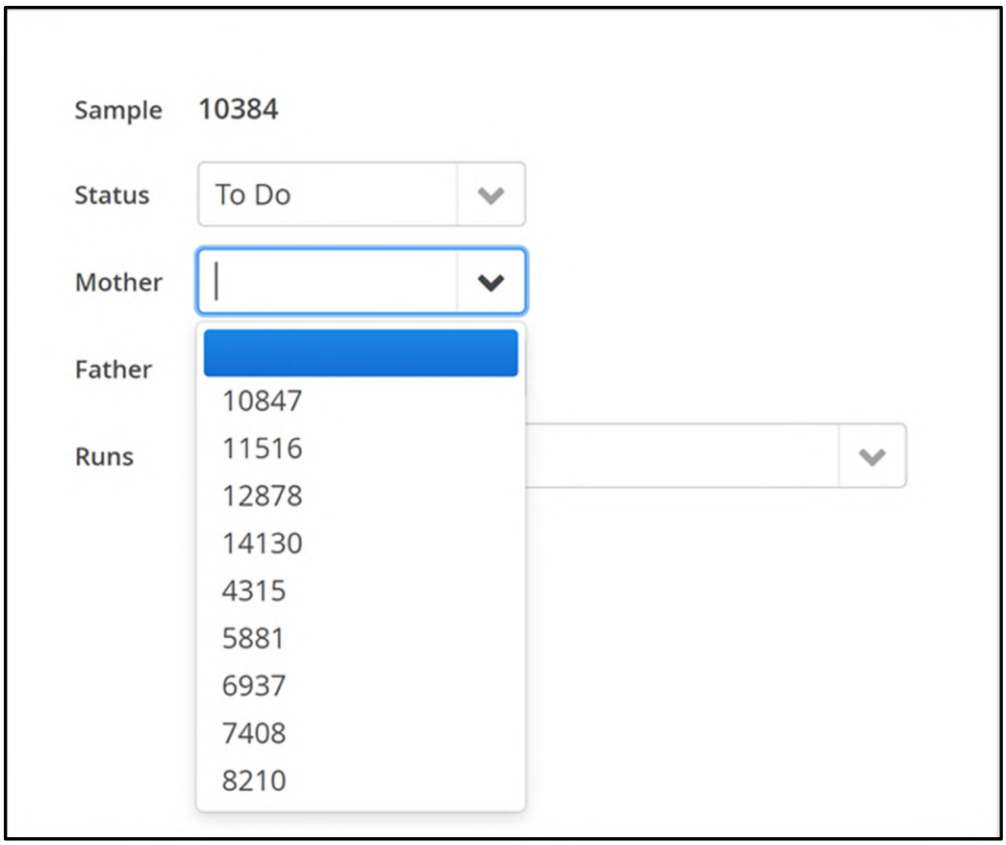


Figure: Specifying the mother of a sample

Initially, before any analysis, the run drop-down list will be empty.

Figure: A sample that has not been processed yet having no run data listed

When a sample has been analysed, each run can be accessed from the drop-down menu. Each run will have a set of data which is displayed by 3 tabs. These are for general run information, QC metrics and results of the analysis.

The General tab displays basic information about the analysis and provides a link to batch view.

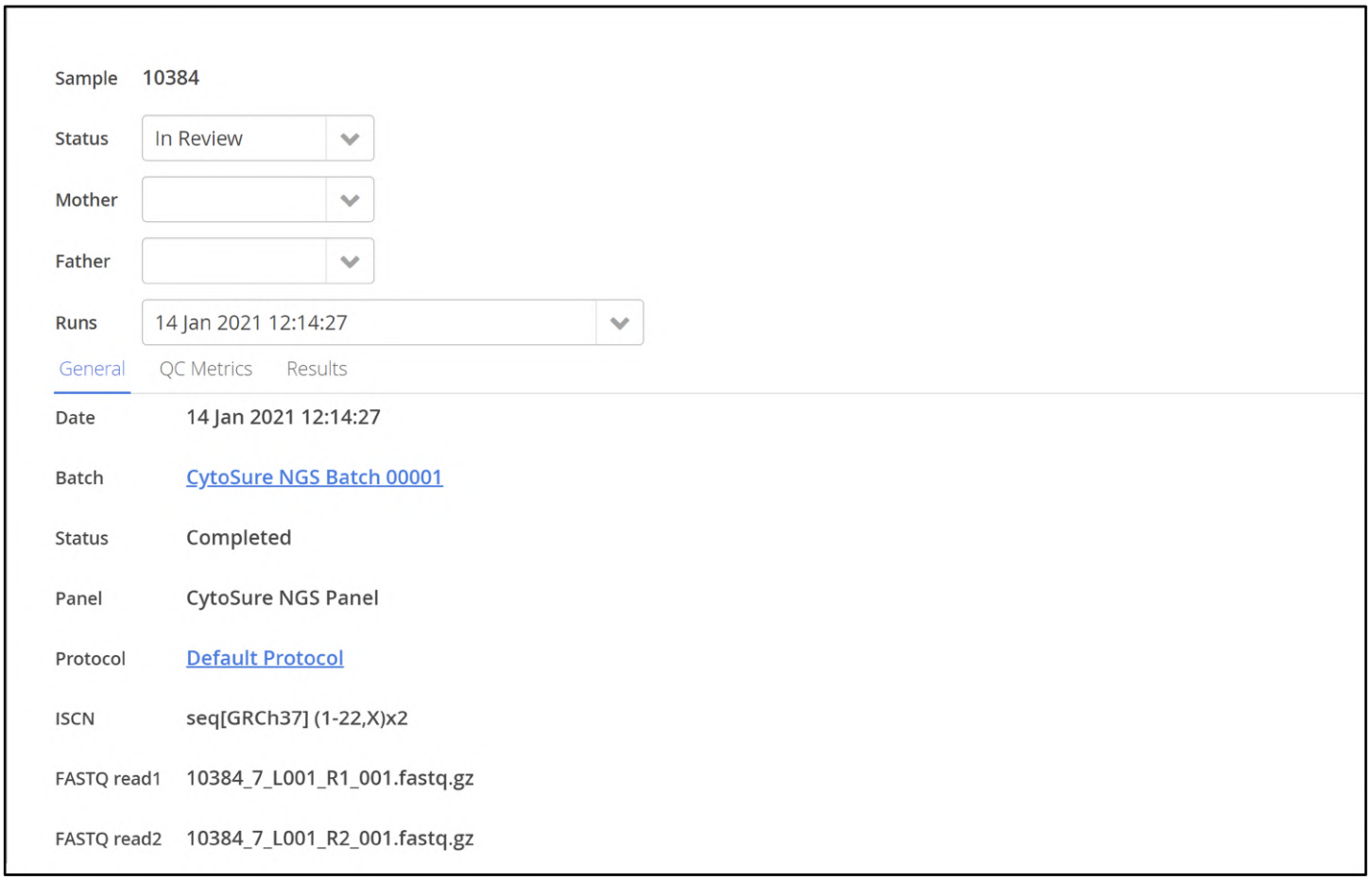


Figure: Viewing the General tab for a sample run

The QC Metrics tab gives an overview of the metrics of the sample. The data will be colour-coded according to the metric set that was defined for the analysis protocol.

There is further information on metric sets in the section of the guide that covers the admin options.

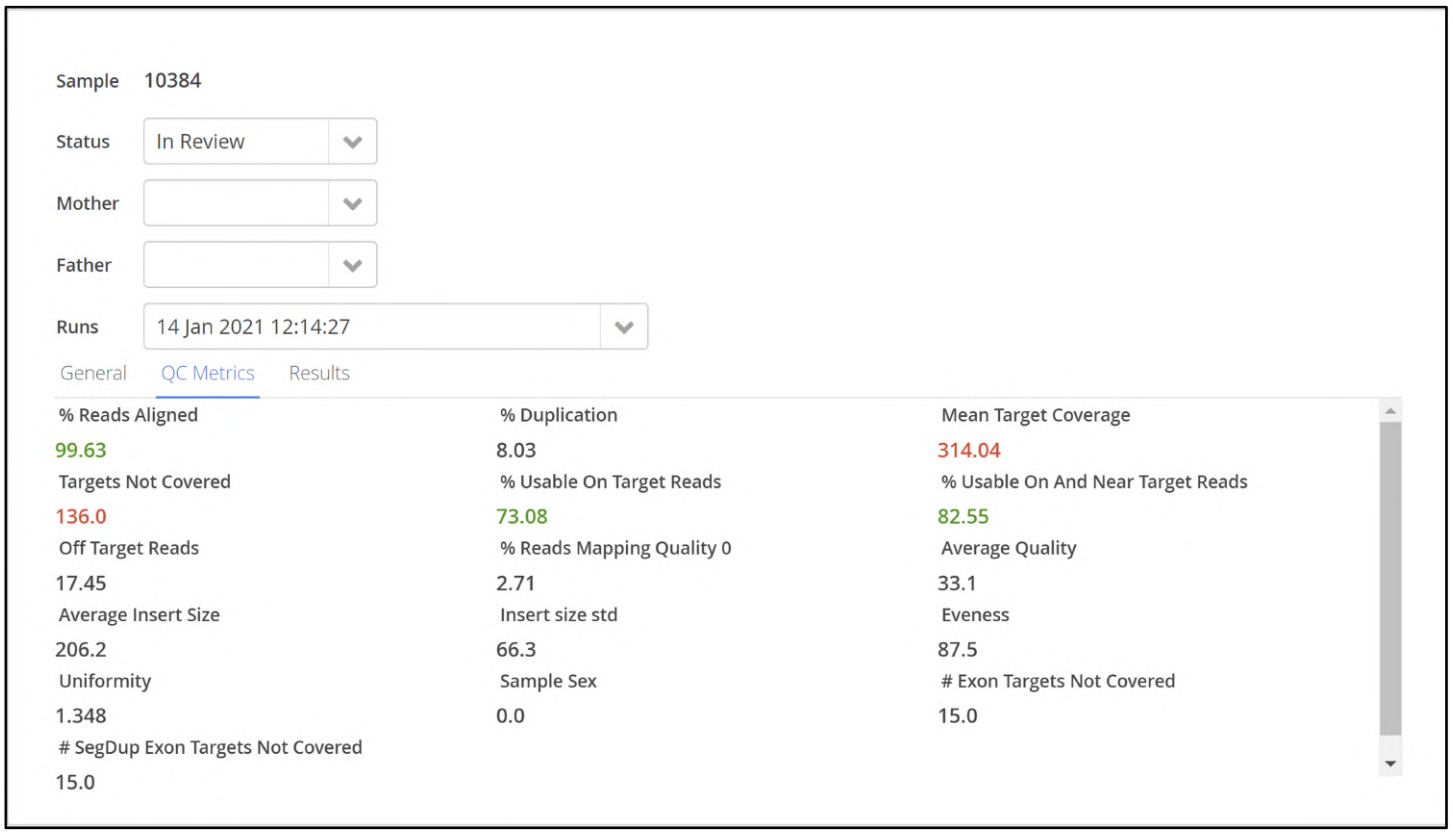


Figure: Viewing the QC Metrics tab for a sample run

Finally, the Results tab provides links (in green) to download files from the analysis as well to view (in blue) the different variants that have been detected.

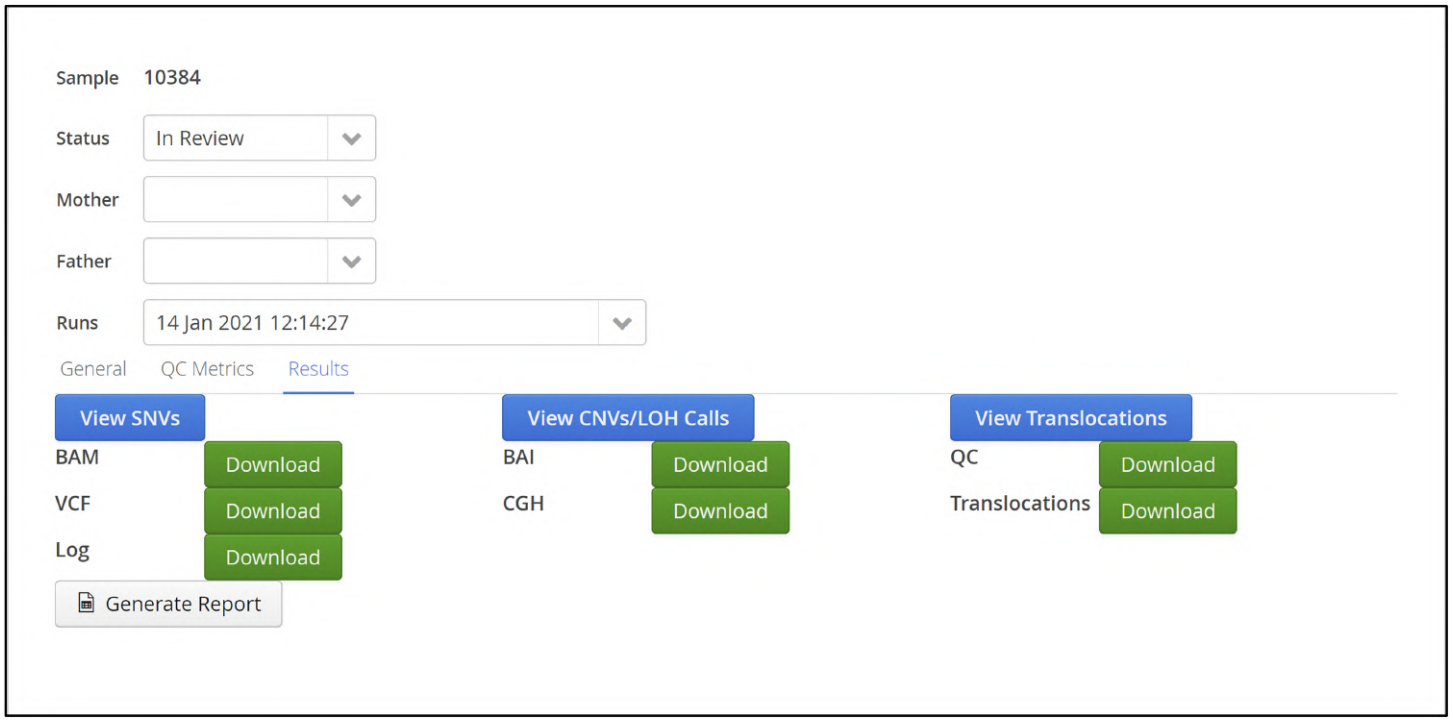


Figure: Viewing the Results tab for a sample run

7.1 Adding User-defined Variables

In order to enable the user to capture and report custom information related to samples processed in Interpret, the admin controls section provides a means to create variables of different data types via Admin Controls > Analysis > Manage Samples > Variables.

Figure: The manage samples page in the Admin Controls

Selecting Add New Variable Category provides a text box to name the new variable and clicking Add adds a new sub-tab to the Variables tab.

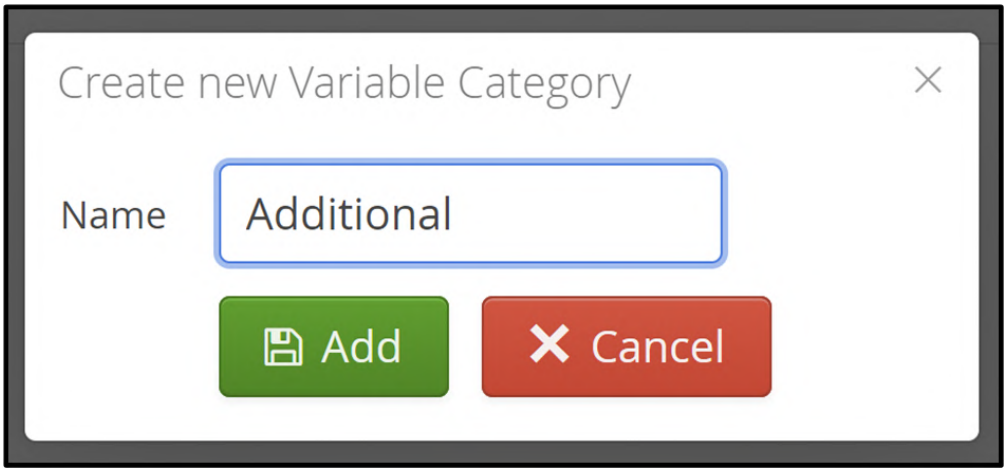


Figure: Creating a new variable category named Additional

With the new category called "Additional" generated, users can create variables associated with the category by selecting Add New Variable .

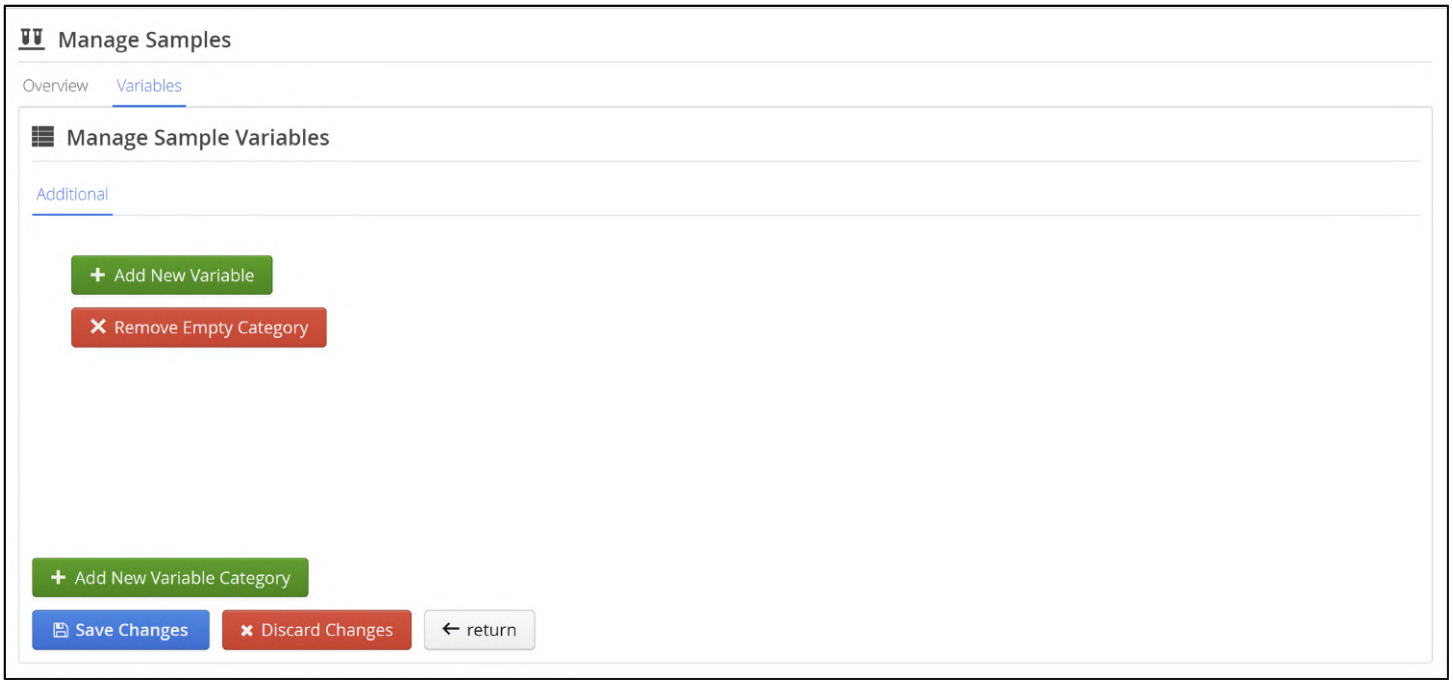


Figure: An empty custom category named "Additional".

To delete an empty category, click the Remove Empty Category button. To create new variables in the category, click the Add New Variable button, assign a Name to the variable, confirm the Category with which it should be associated, and select the appropriate data type from the DataType drop-down box, then click the **Add** button.

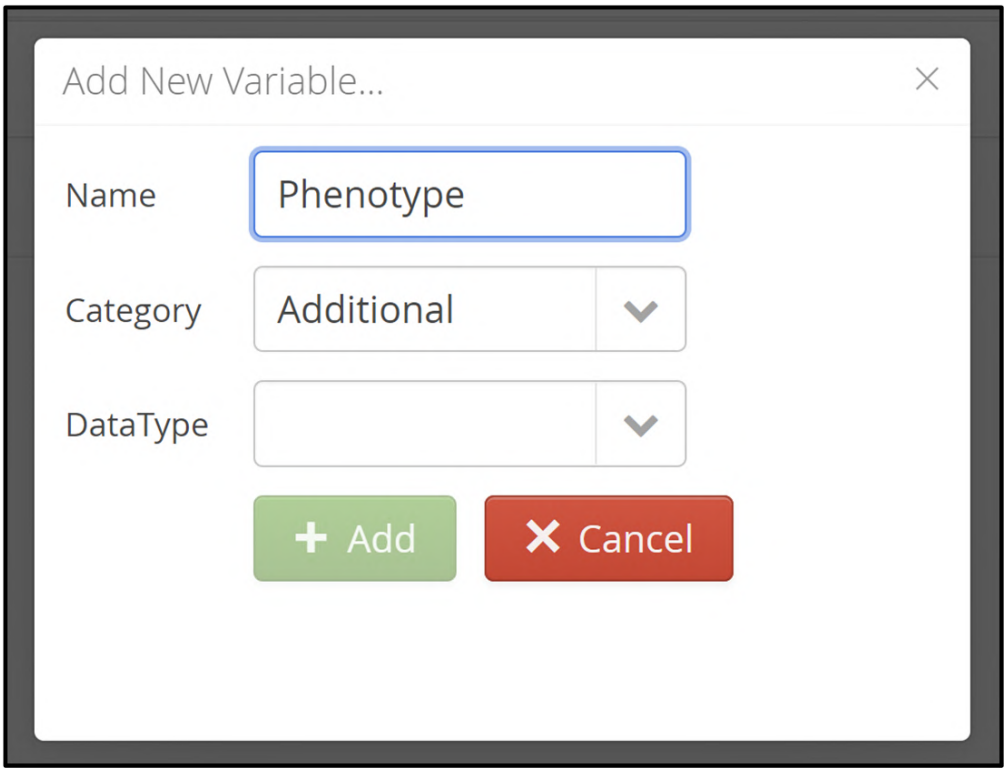


Figure: Creating a new variable named "Phenotype" in a category named "Additional".

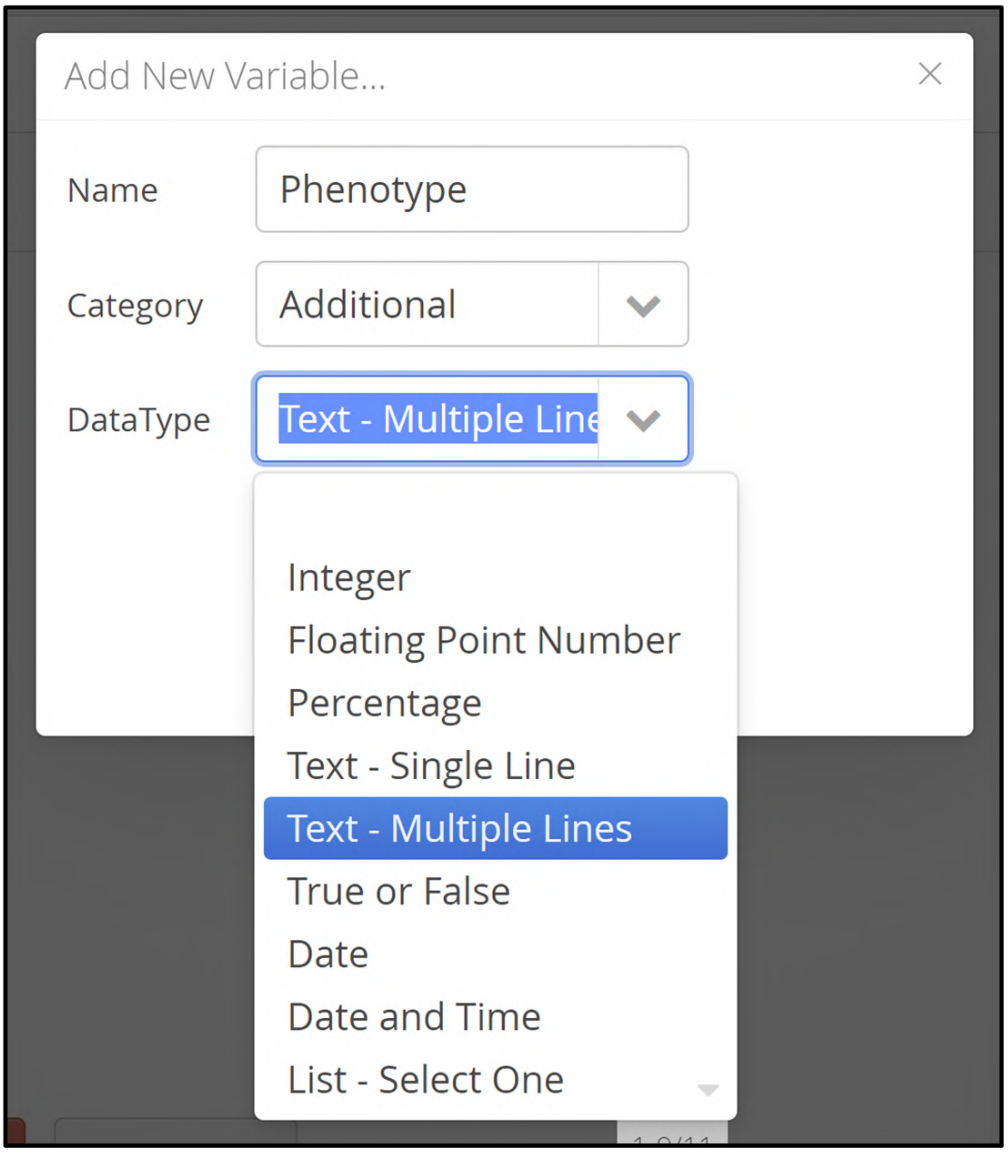


Figure: Selecting the appropriate data type for the new variable - in this case, "Text - Multiple Lines".

Once a variable has been created, it will be listed, along with its data type, in the appropriate category in the Manage Sample Variables section. To delete a variable, click on the cross next to the variable name.

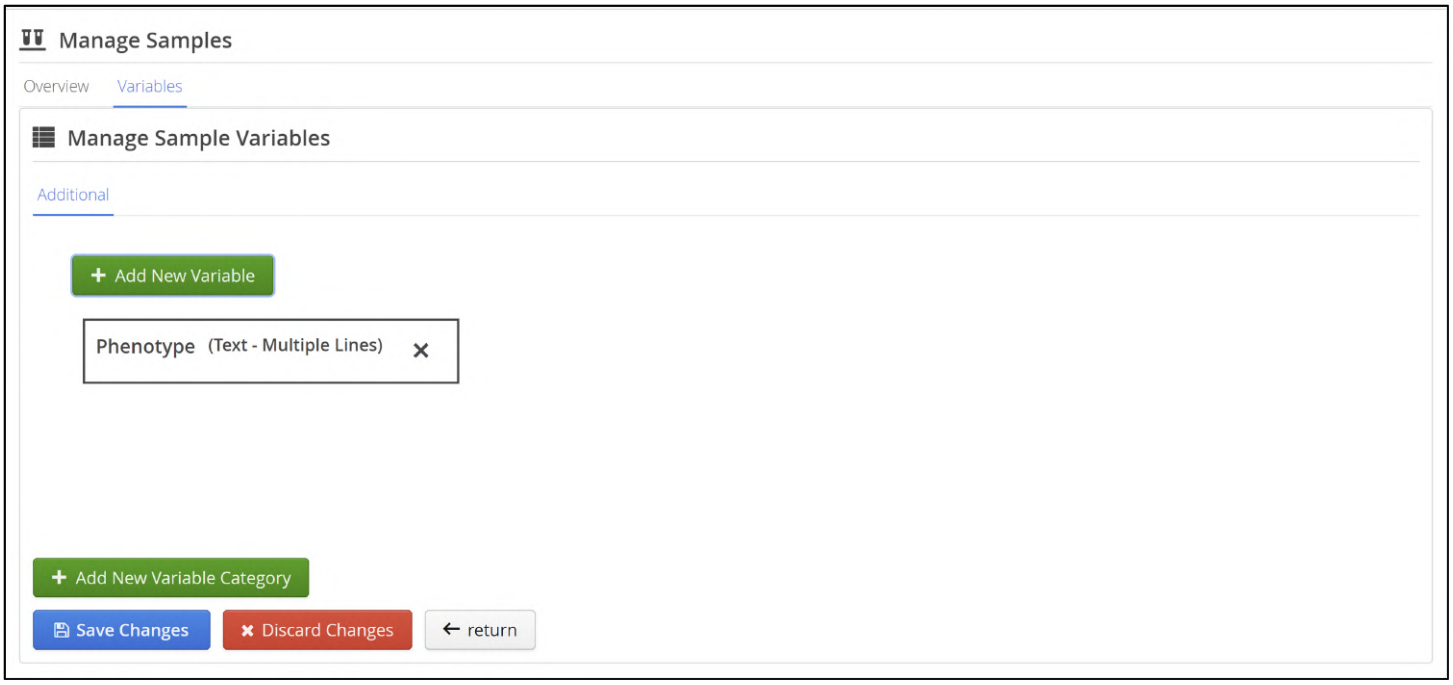


Figure: A custom field named "Phenotype" listed under the "Additional" category.

Having been created in the system, custom fields may be populated for each sample in the Samples view, and will also be displayed in the sample run page whenever the sample has been processed in a batch (accessible by clicking on the sample row in the Completed Samples table in the Batch Overview page).

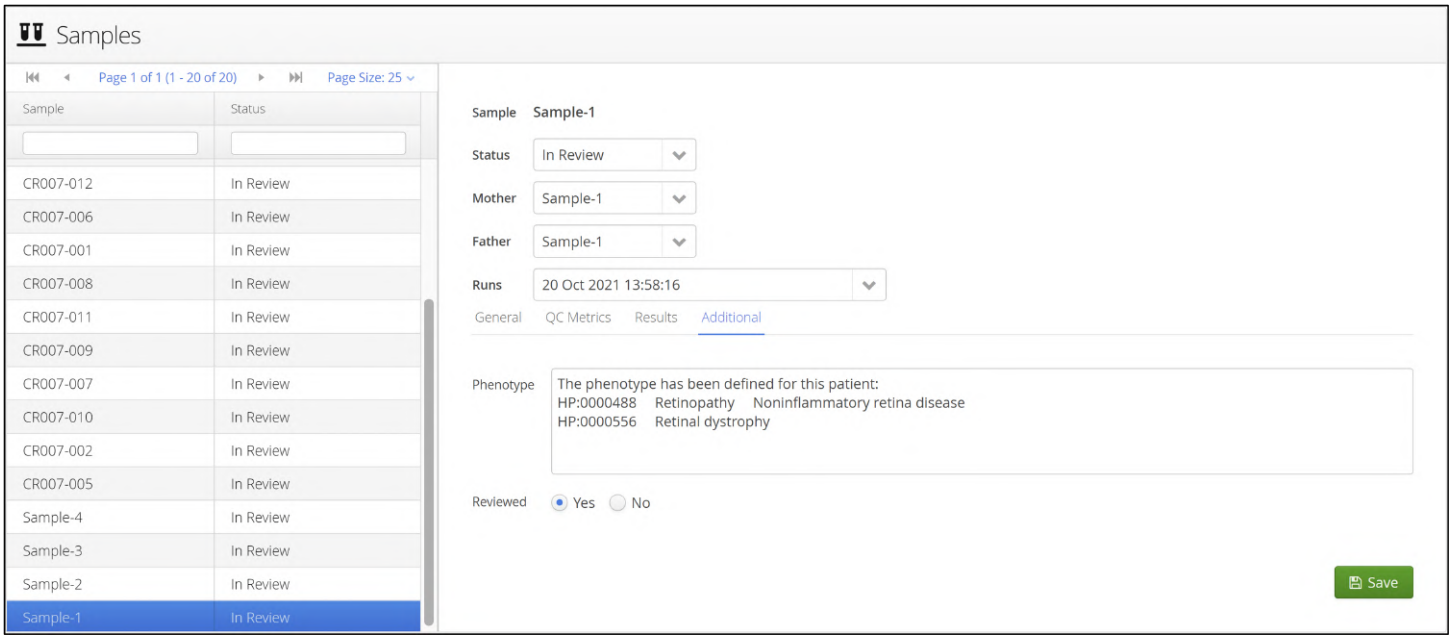


Figure: Editing the content of the "Phenotype" field in the Samples page.

Interpret also provides a framework enabling the development of plug-ins to import sample data in bulk from other sources, such as spreadsheets, text files or a LIMS. If you are interested in importing data in bulk, contact OGT - a suitable plug-in may be available, or it may be possible to develop a plug-in to satisfy your requirements.

8 Running an Analysis

On the dashboard either select "Run Batch" in the drop down from the 'Batches' menu item.

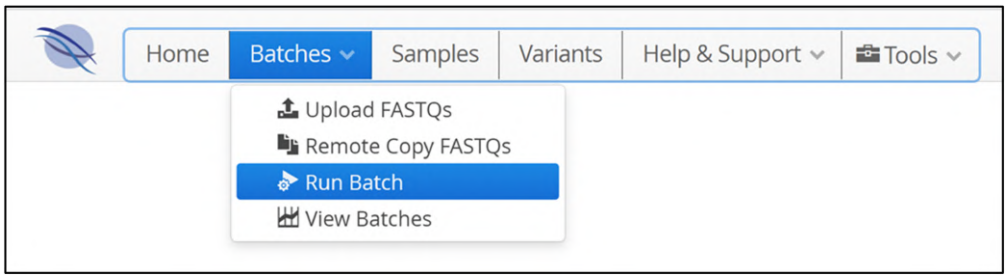


Figure: Selection of Run Batch from the Dashboard menu bar Batches drop down menu

Or, click on the 'Run Batch' icon on the dashboard page

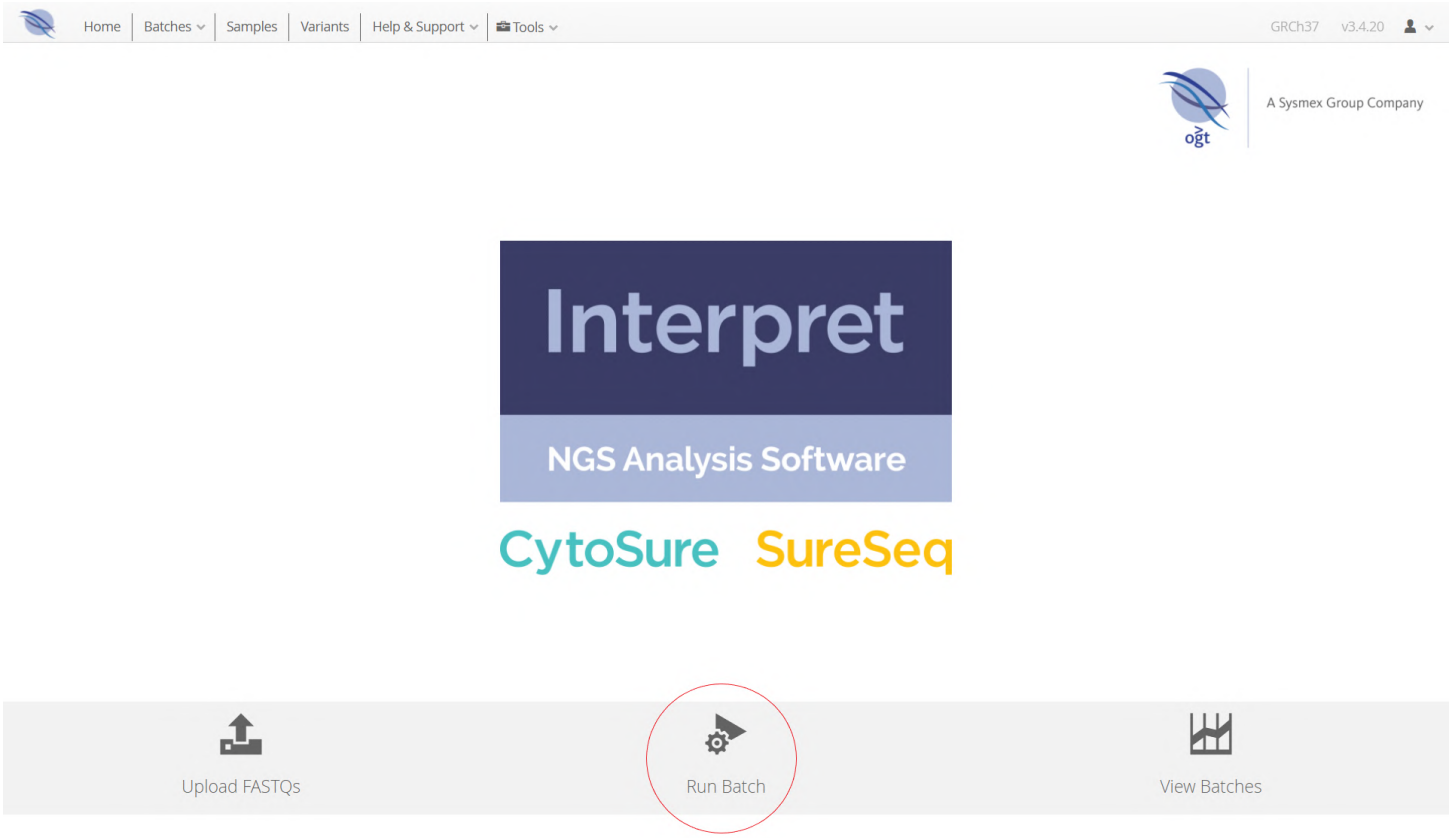


Figure: Selection of Run Batch from the dashboard short-cut buttons

Either choice leads to the initial Run Batch page is as follows:

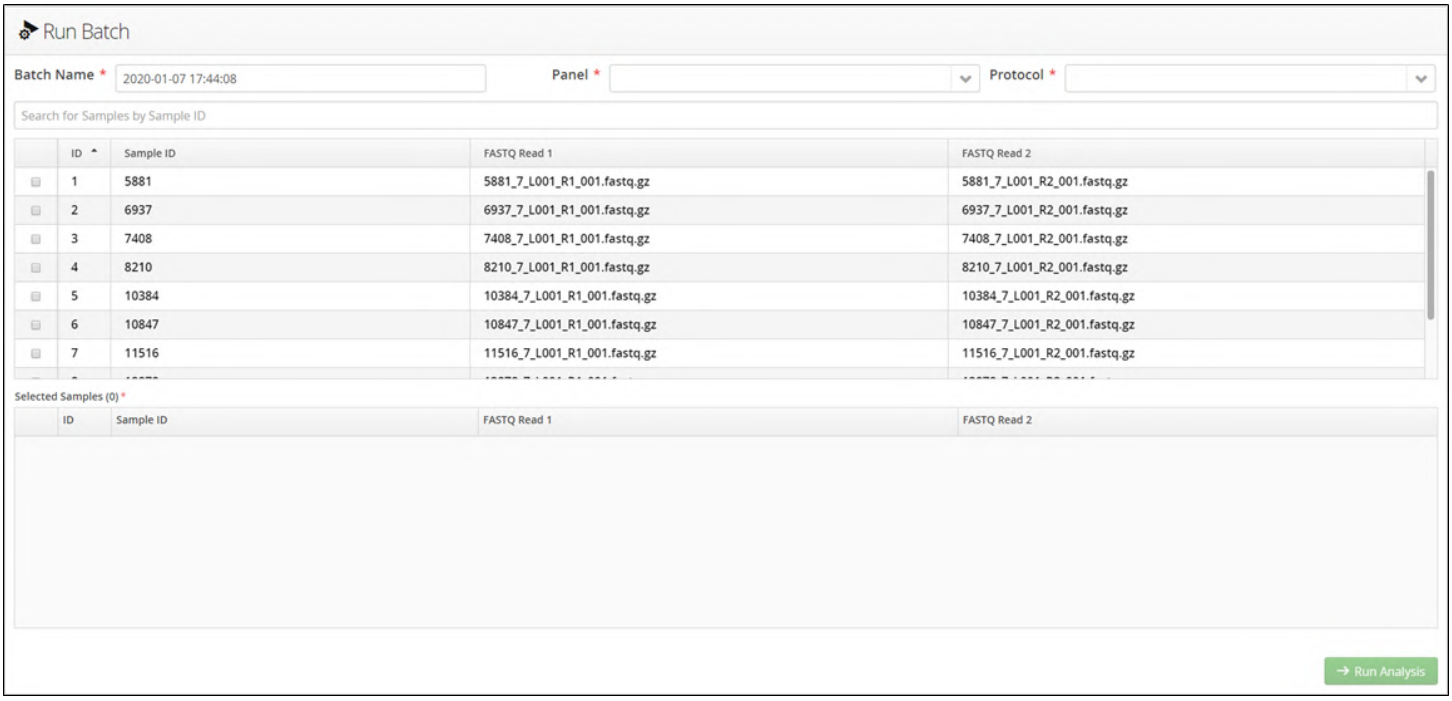


Figure: Initial view of the Run Batch window

Besides showing the list of available samples there are additional text fields and drop-down menus.

In order to run an analysis the user needs to

1. Select samples for the analysis
2. Select the correct panel for the samples
3. Select the analysis protocol

Figure: Input fields for the Run Batch window

Optionally users can specify a name for the batch analysis. A default batch name is provided with the date followed by the time in the format YYYY-MM-DD HH:MM:SS.

In the example below the user has created the batch name CytoSure NGS Batch 00001

Figure: Entering a batch name

The samples have been processed with OGTs CytoSure NGS panel so that is the selection to make from the Panel dropdown menu.

Figure: Selecting the appropriate Panel

The user now specifies the protocol that will be used, in this case the Default Protocol.

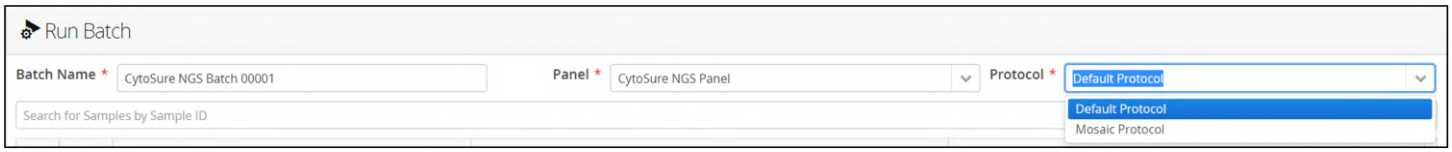


Figure: Selecting the protocol to use for processing the batch

> Panel-Protocol Compatibility
> Only protocols whose Pipeline Type are included in the list of pipeline types supported by the selected Panel will be listed in the Protocol drop-down list. Additionally, if any **Pipeline Capabilities** supported by the protocol are not supported by the selected panel, a warning will be displayed indicating which processes will not be run.

Lastly, the user specifies the samples to be analysed.

There may be a large number of samples loaded into the system, so to enable easier sample selection it is possible to add a search term. In this case the user is looking for all samples containing the number 5. Additionally, search terms are independent of the case used.

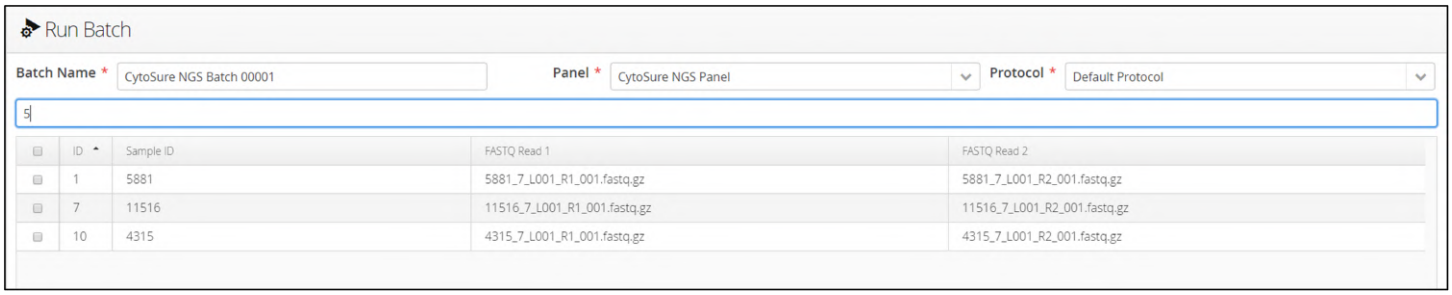


Figure: Filtering loaded samples with a search term

Selecting the checkbox next to a sample moves a loaded sample into the Selected Samples table.

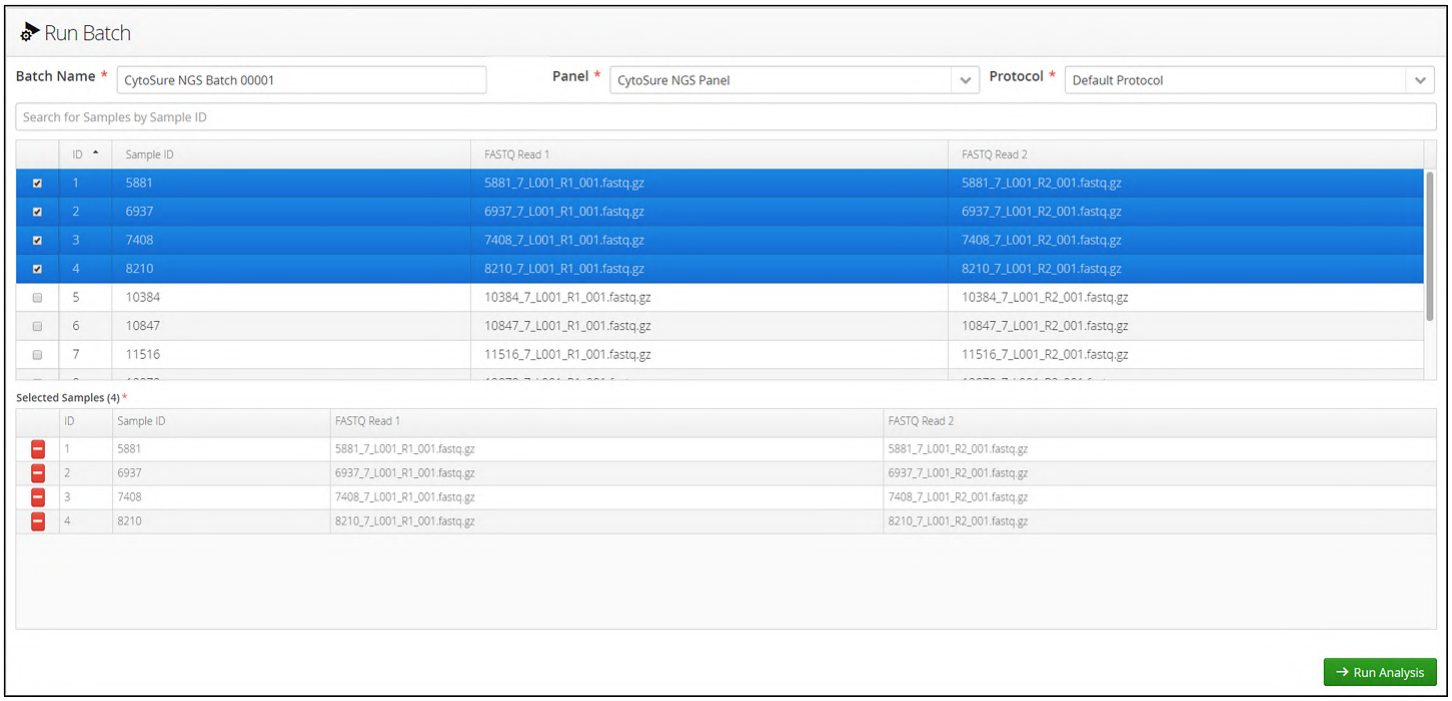


Figure: Adding a sample to an analysis batch

Clicking on the minus icon ⊟ will remove the sample from the Selected Samples tables.

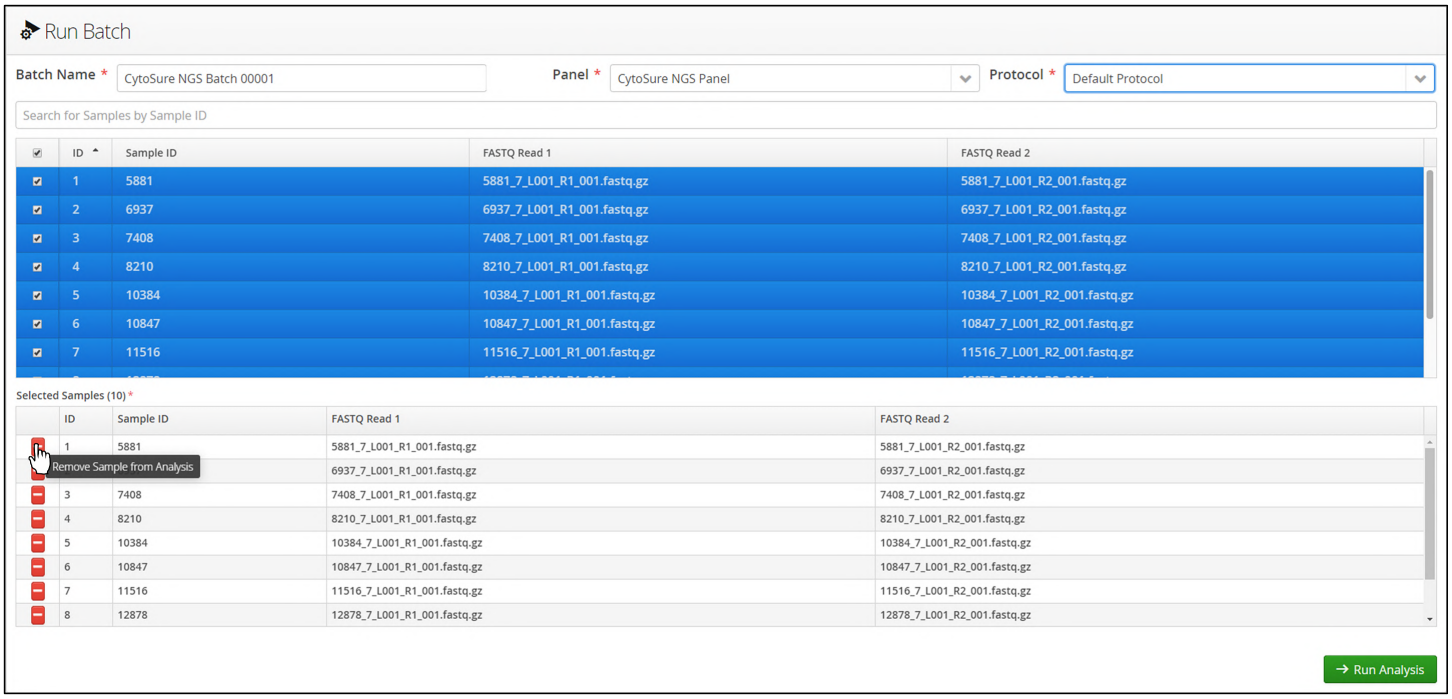


Figure: Removing a sample from an analysis batch

When all selections have been made the run can be started by selecting → Run Analysis

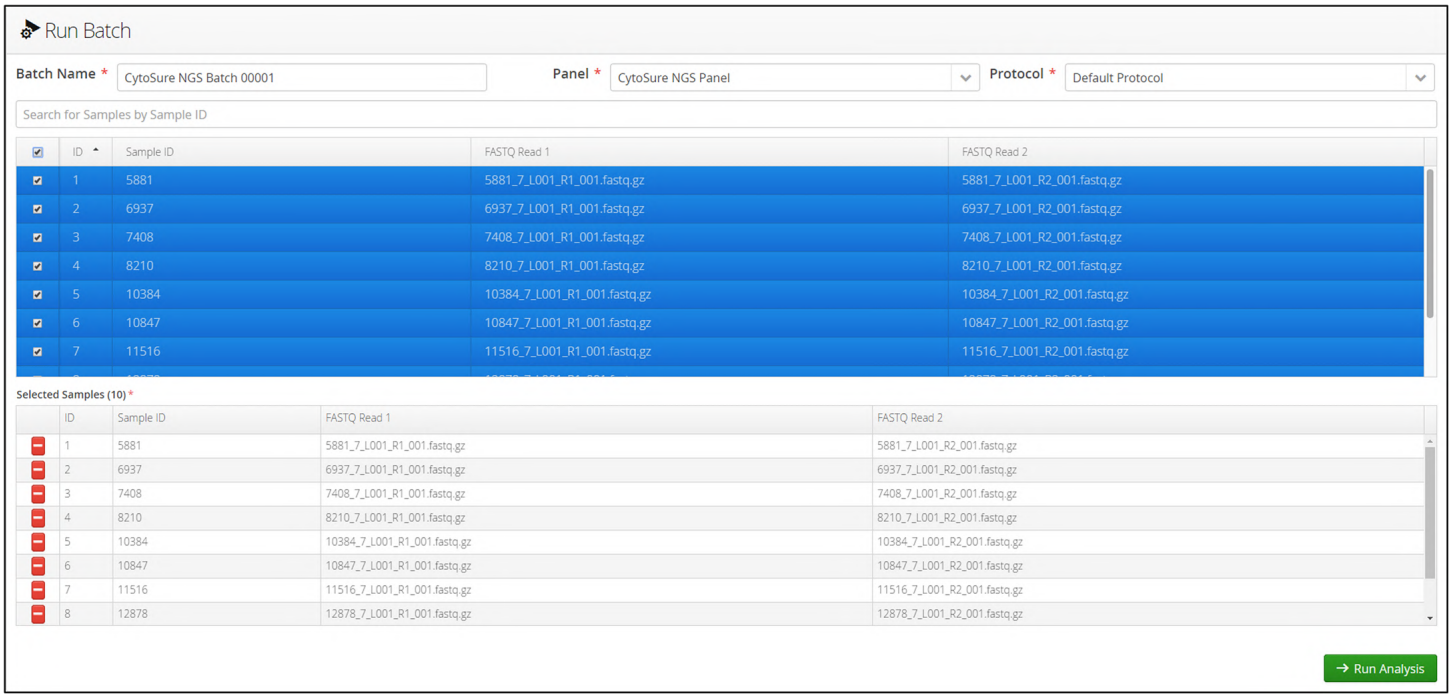


Figure: Starting an analysis

If the selected protocol has "Enable CNV and LOH Calling" set to "Yes", CNVs will be detected by comparison with a set of reference samples which need to be defined in the protocol as either "All Batch Samples" or a specific set of reference samples whose FASTQ files have already been uploaded and designated to the system by the user.

In the latter case, OGT may provide a set of data files that can be used as a reference set for CNV analysis. As more samples are processed users may extend the reference pool by adding any samples they believe are suitable as controls for CNV calling. A user can modify samples designated as reference pool in the protocol in Admin Controls-Manage Samples-Protocols.

If CNV calling is enabled without a reference data set being defined, then, on selecting → Run Analysis, the following error will be displayed.

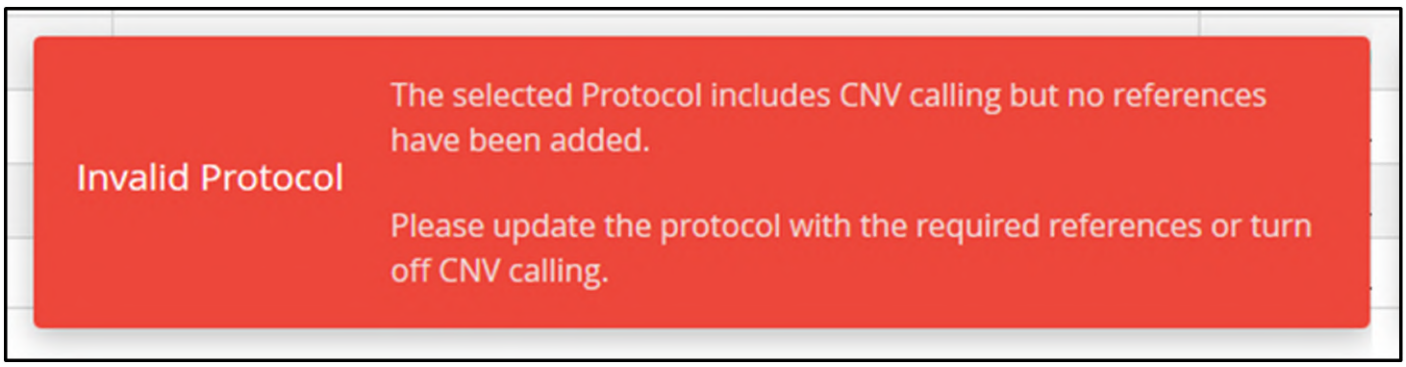


Figure: Error message for using an invalid protocol

Click on the message to remove the warning and select Admin Controls > Analysis > Protocols to set reference samples. More details are in the Protocols section of this User Guide.

Otherwise, a popup presents the chosen files and selected parameters. Following this there is a request for confirmation and upon confirmation the analysis run will be initiated.

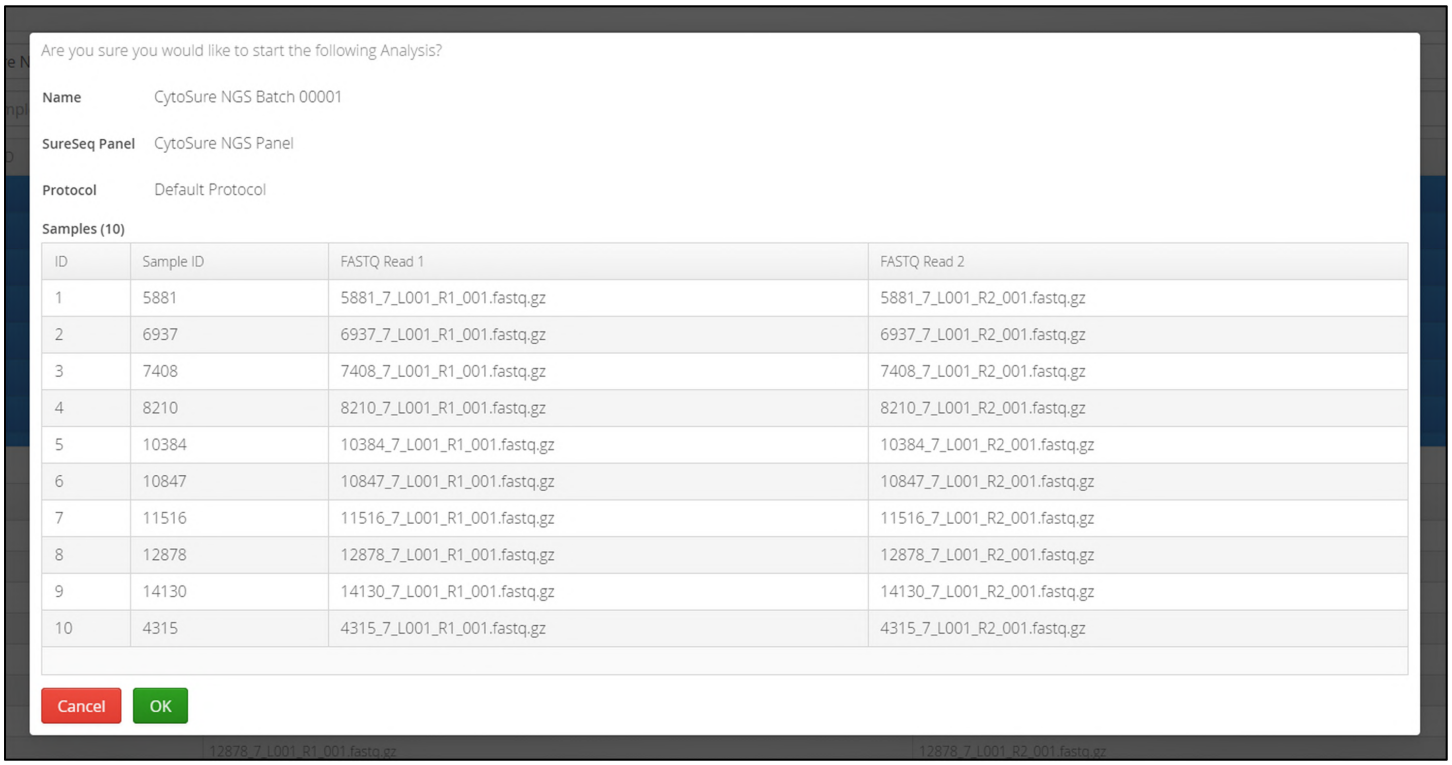


Figure: Window requesting confirmation to run an analysis

Selecting  will start the analysis and the display will change to show information about batch being analysed.

Within this there is an overview window providing an overview of the analysis and a sample window giving information about the status of each sample.

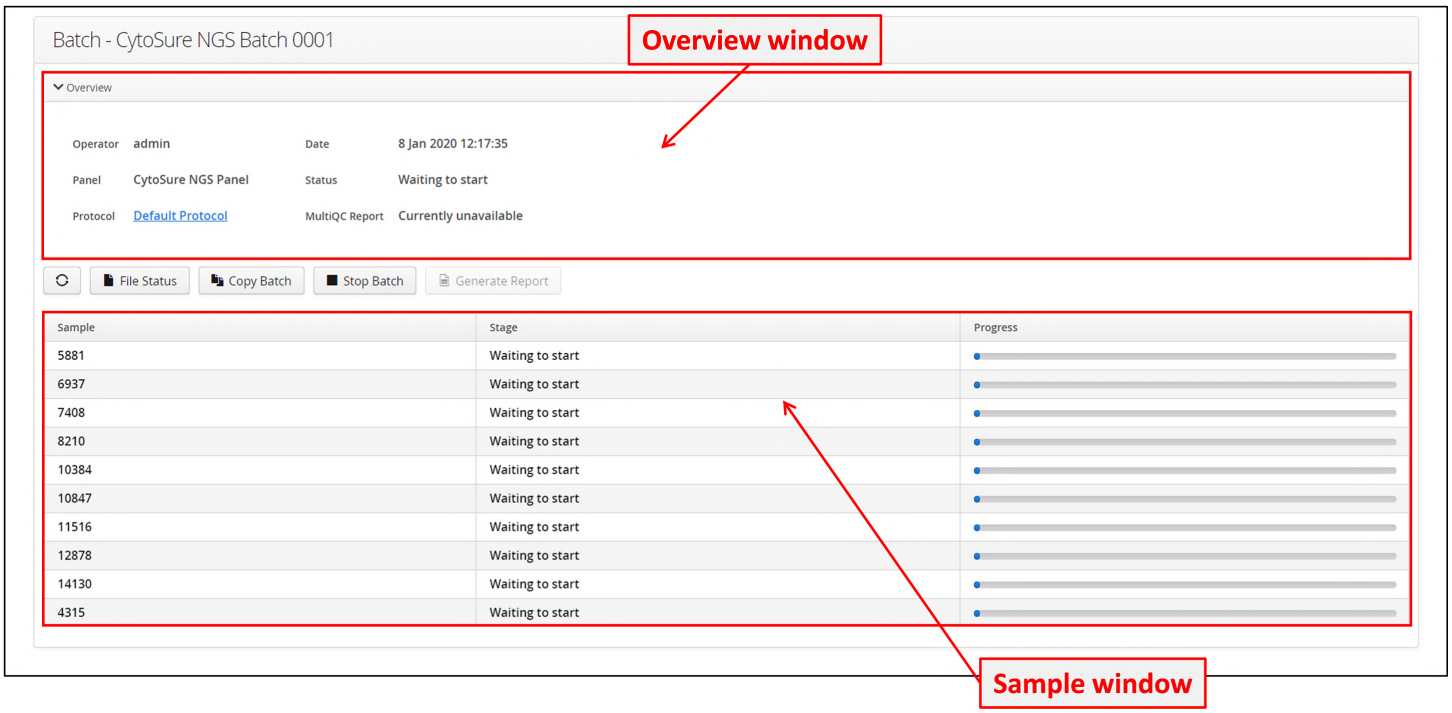


Figure: The batch processing view with the overview window and sample window highlighted

Initially the status of the samples will be listed in the overview window as "Waiting to start"

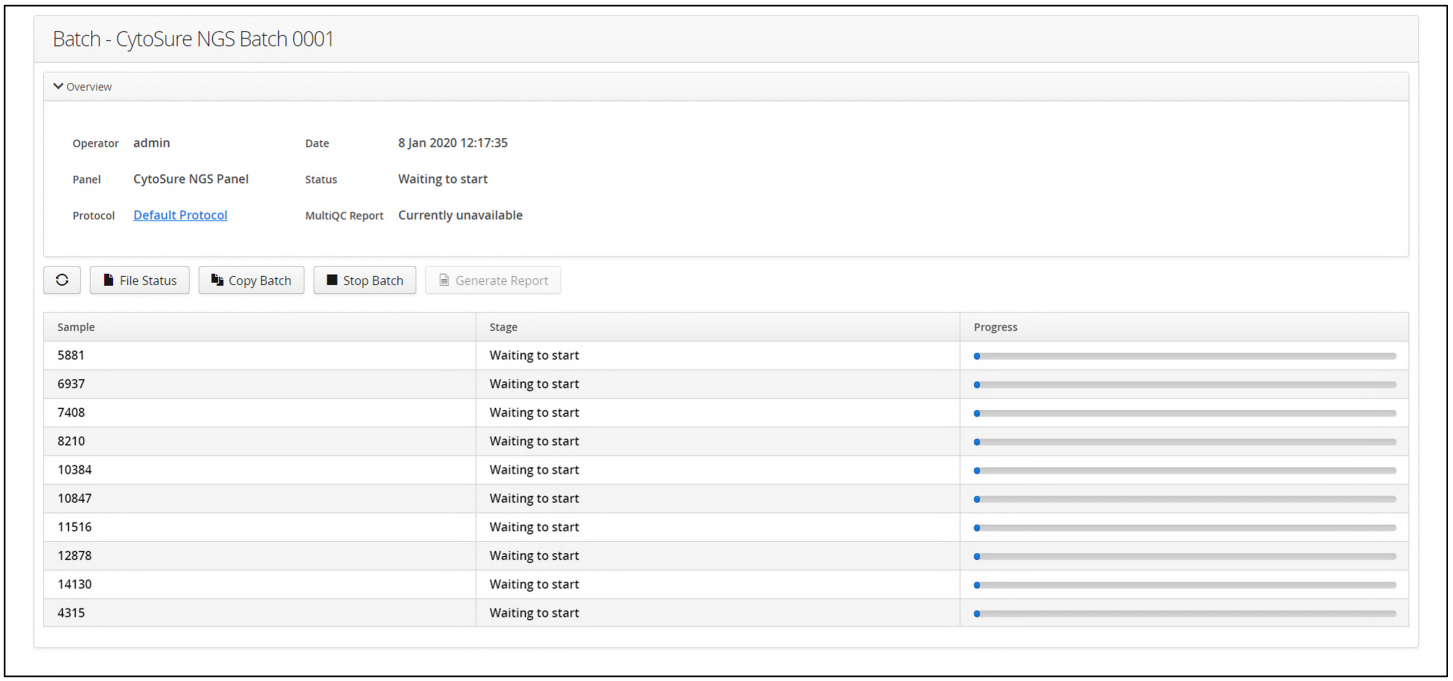


Figure: Initial batch status before analysis starts

Waiting for a reference to be generated

If a reference pool needs to be generated the status shown in the batch overview will report this and provide a means to track progress of the reference pool creation.

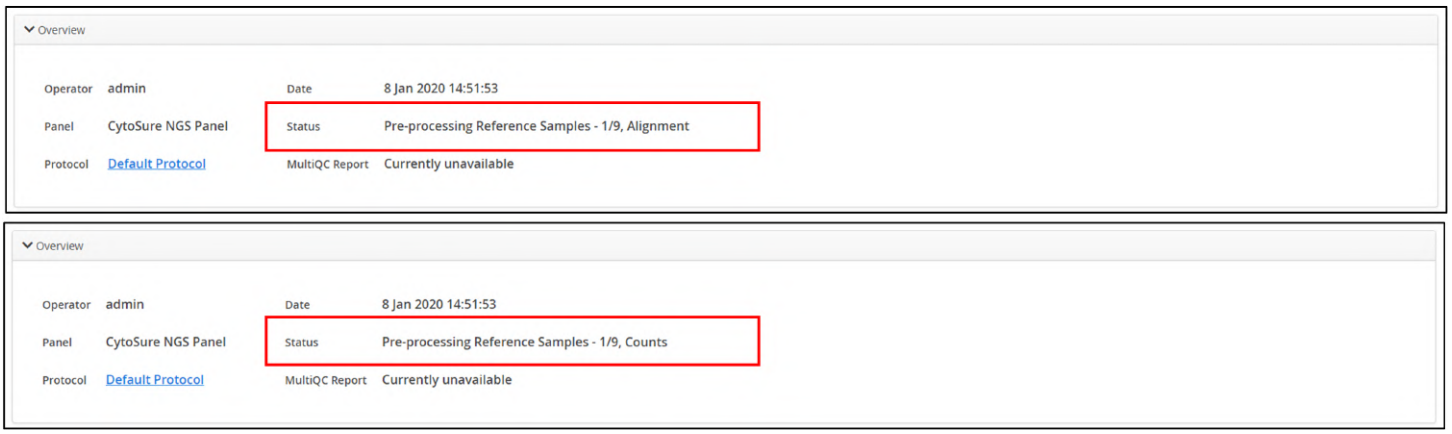


Figure: The analysis status in the overview window, highlighted, showing reporting the status of pre-processing of the reference samples

The status of reference building can also be tracked in the View Batches window which is discussed in the View Batches section of this User Guide.

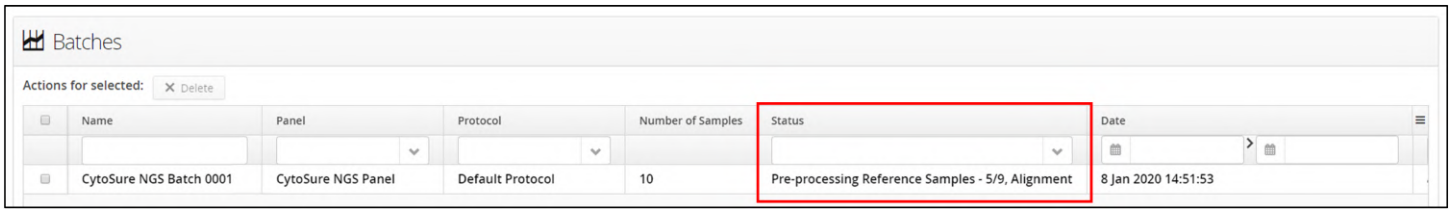


Figure: Reference building status being shown in the View Batches window

If the protocol performs CNV analysis and samples in the analysis are to be used to generate the reference pool against which to make CNV calls then the overview will report the combining of the reference samples.

Once the reference samples have been aligned and counted, they are combined into a pool for the CNV analysis

Figure: The analysis in the over window, highlighted, reporting the combining of the reference samples into a pool.

Samples will be queued until there is capacity available in the pipeline. Once this is available the software will start processing the samples sequentially. The stage of the process is updated and the overall progress can be monitored in the progress bar.

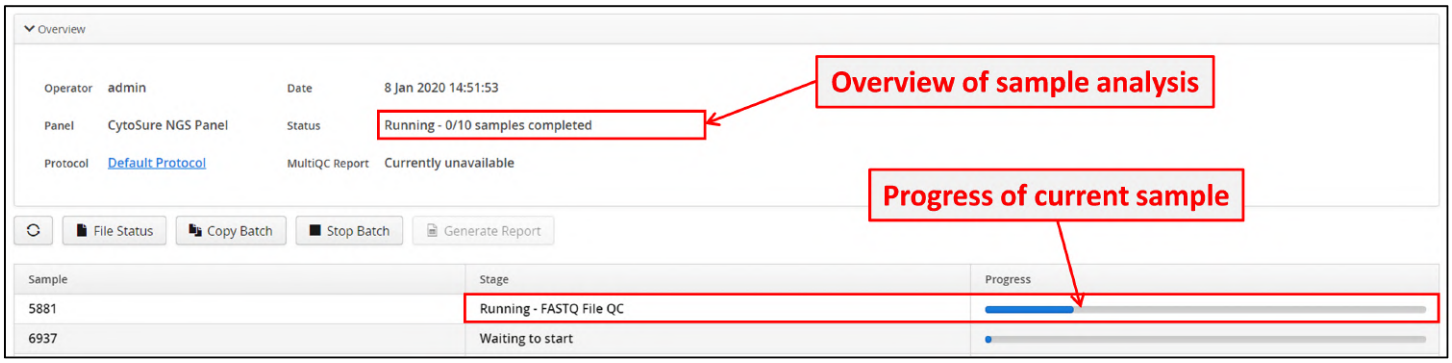


Figure: The batch view showing progress of analysis

Once analysis started the stage of each sample is displayed and can be followed

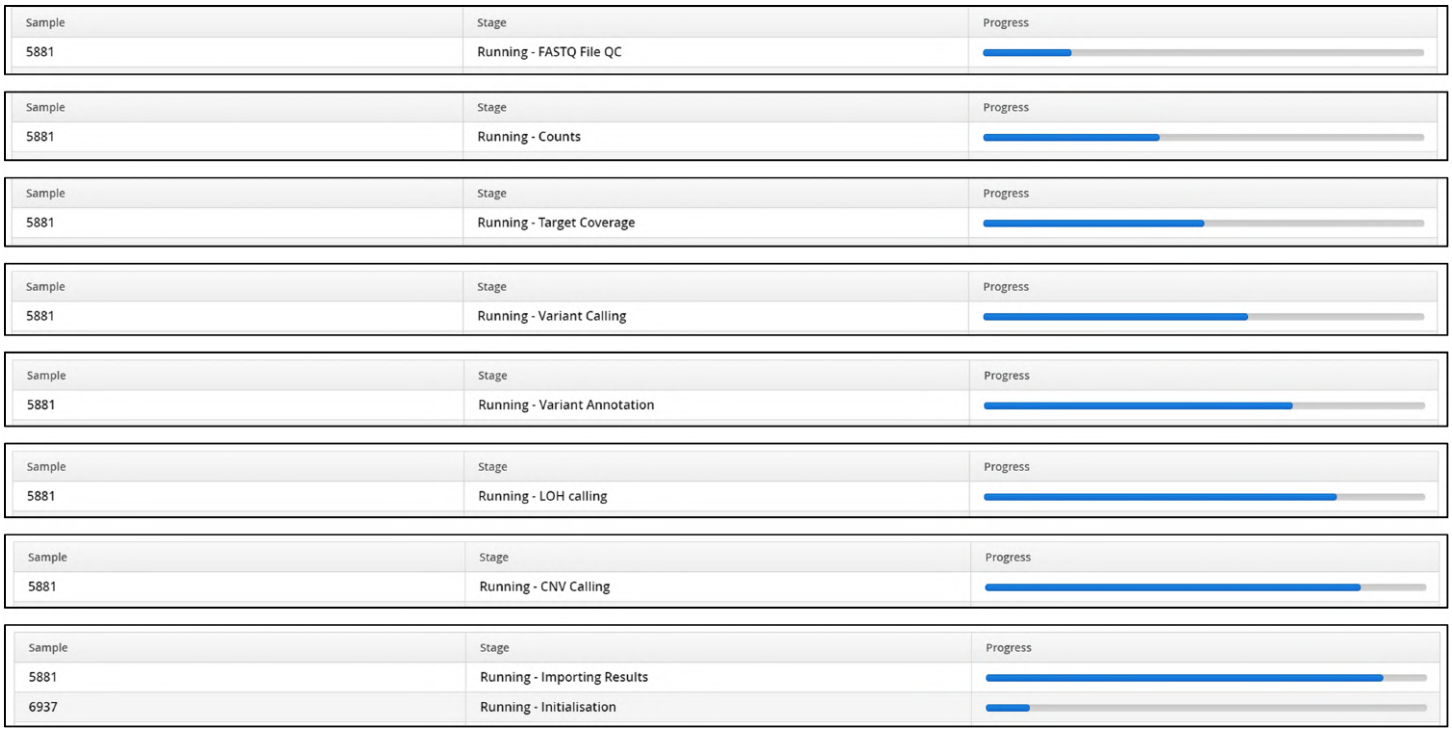


Figure: Tracking progress of a sample processing

Once a sample has been analysed the overview updates the count and a summary of the analysis is displayed in a Completed Samples table.

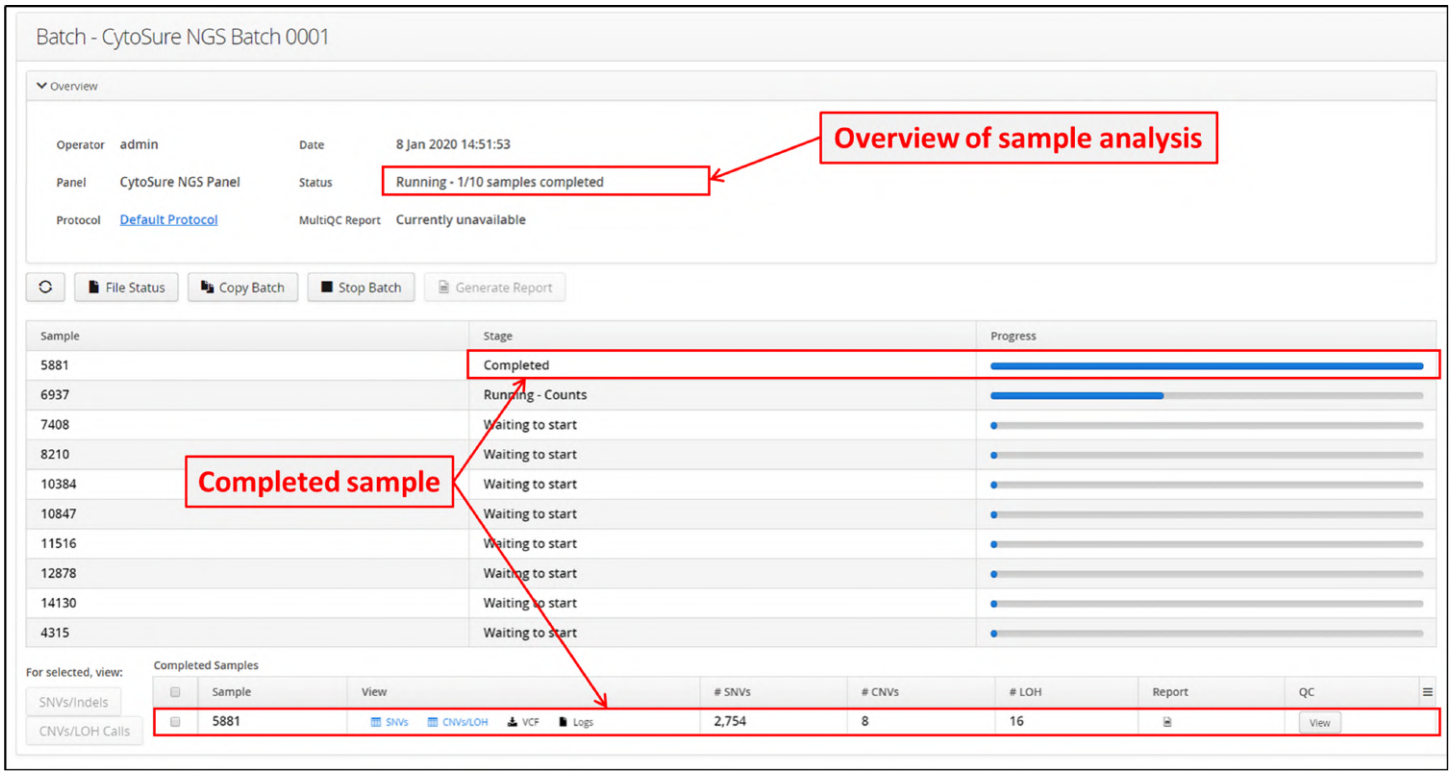


Figure: The first completed sample is displayed below the samples to be processed

When all samples are completed

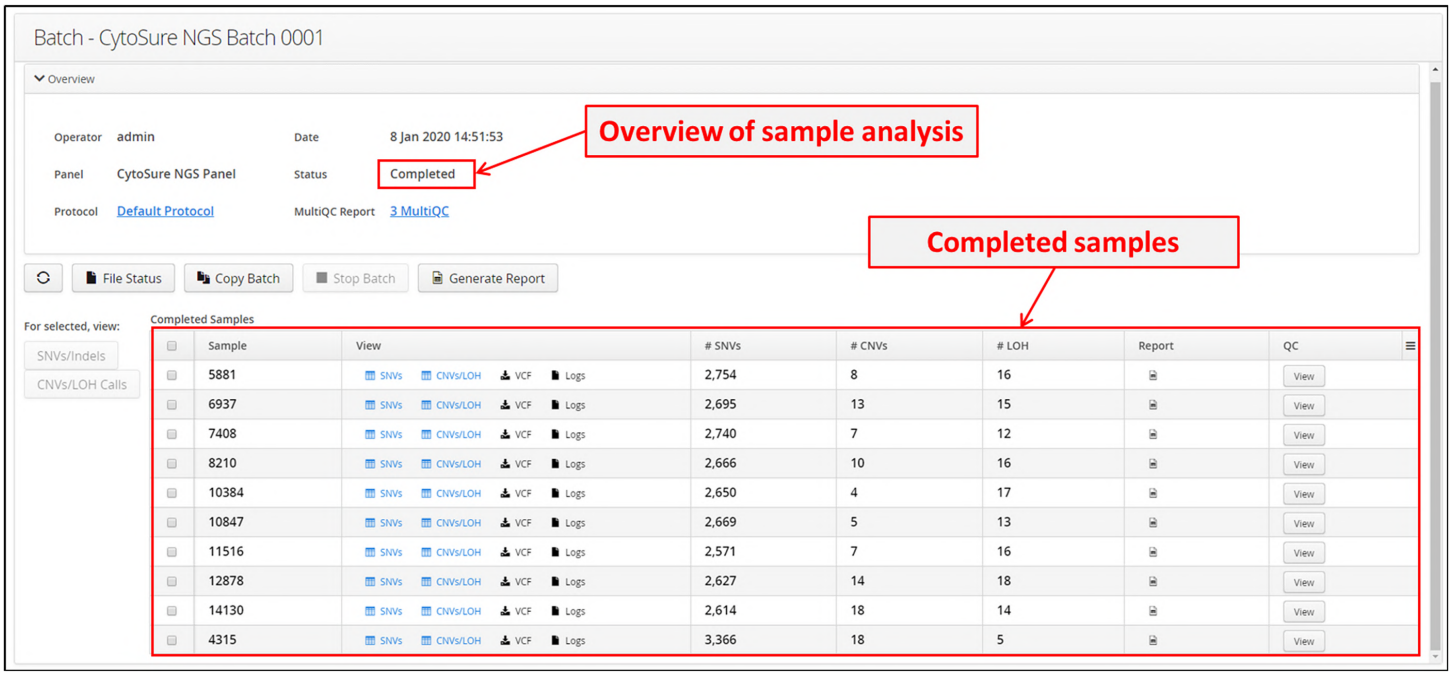


Figure: An analysis with all samples analysed

There is no need to wait until all samples have been processed to view the results for a completed sample.

This will be discussed in the Viewing Analysis Results section of the manual.

9 Viewing Analysis Batches

On the dashboard either select "View Batches" in the drop down from the 'Batches' menu item.

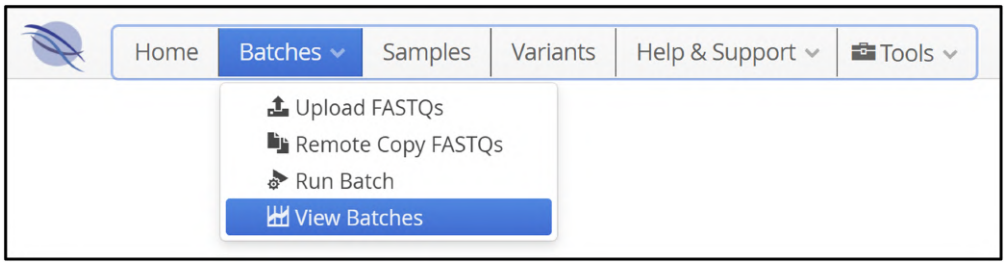


Figure: Selecting View Batches from the menu bar drop down menu

Or, click on the 'View Batches' icon on the dashboard page

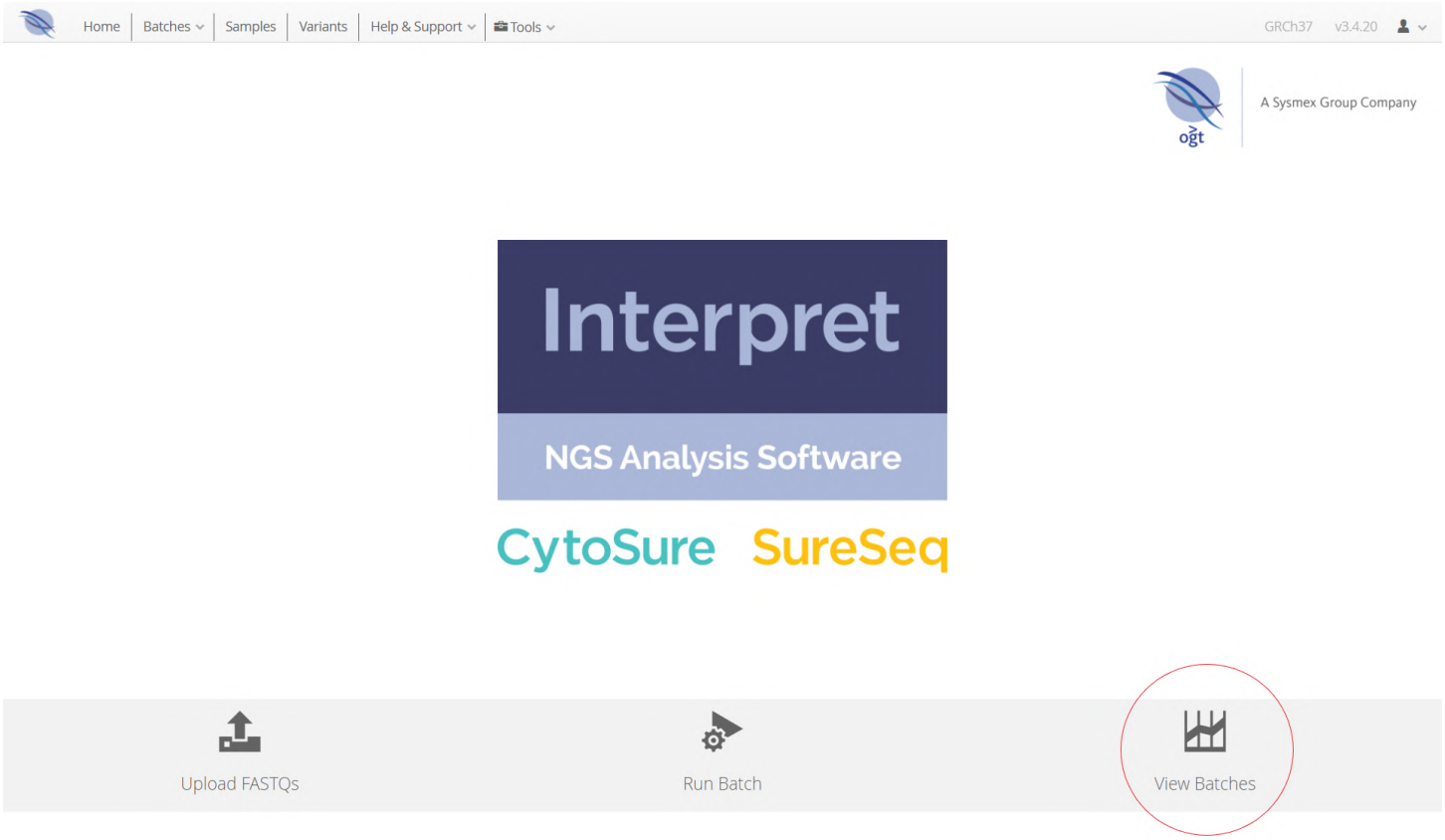


Figure: Selecting View Batches from the dashboard shortcut buttons

The Batches are presented in a table as below.

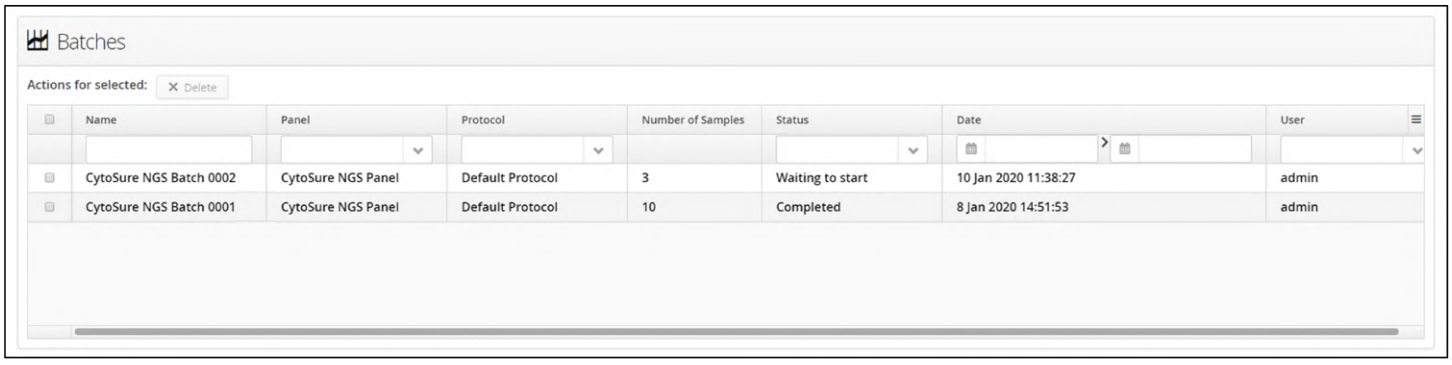


Figure: Initial view of the Batches window

As with other tables in Interpret where there is a column selector icon ≡ a user can add or remove columns from the display

Figure: Column selection options for the Batches window

Column names annotated with a tick are in the current display and changes can easily be made to add or remove columns

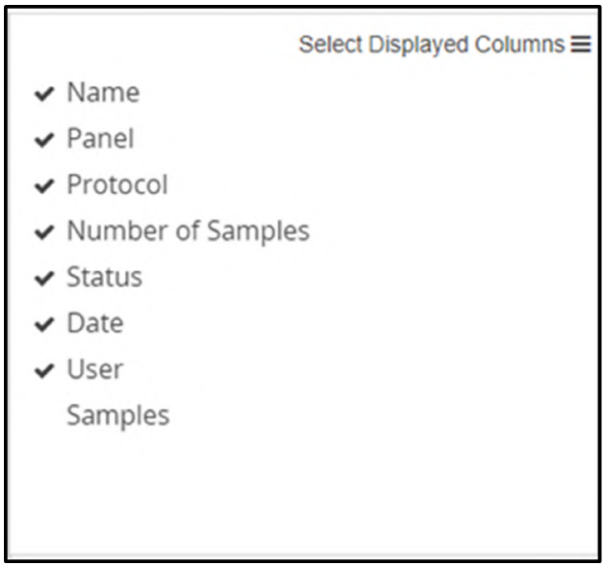


Figure: Selection of columns to display in the Batches window

By default, all batches are presented in the first instance but these can easily be filtered.

Where the column header has a text field, users can type in a search term and all batches with that text contained somewhere in the name, will be retained. The text search is independent of lower- or upper-case letters, "Demo" will return the same samples as "demo".

Alternatively, where there is a drop-down menu selecting one of the values in the menu will lead only to the batches matching the selection being displayed, for example, below only batches that have completed will be displayed.

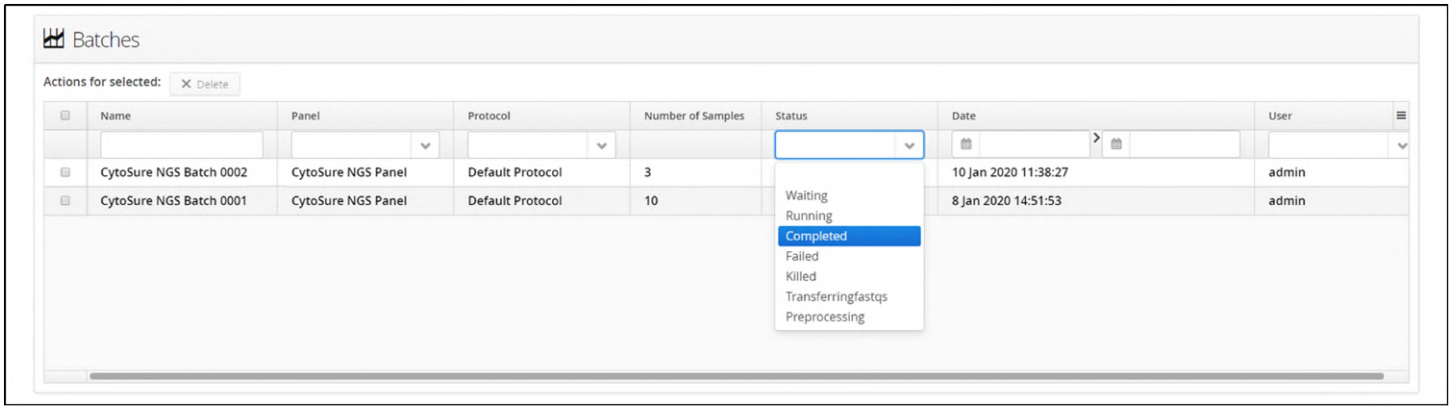


Figure: Filtering batches on status

Lastly, there are date fields, allowing selection of batches run within a set time frame.

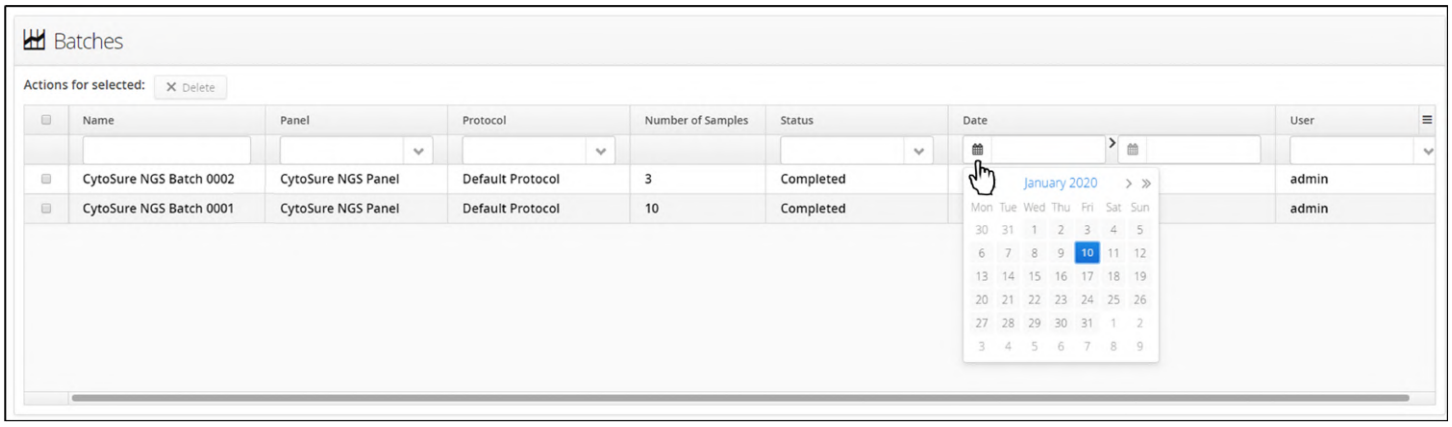


Figure: Filtering batches on date of processing

9.1 Deleting Batches

In the batch view it is possible to delete batches. When first opened there is a greyed out Delete button in the display.

If a batch is selected it is highlighted in blue and Delete button is now active, Clicking the delete button will delete the batch from the software.

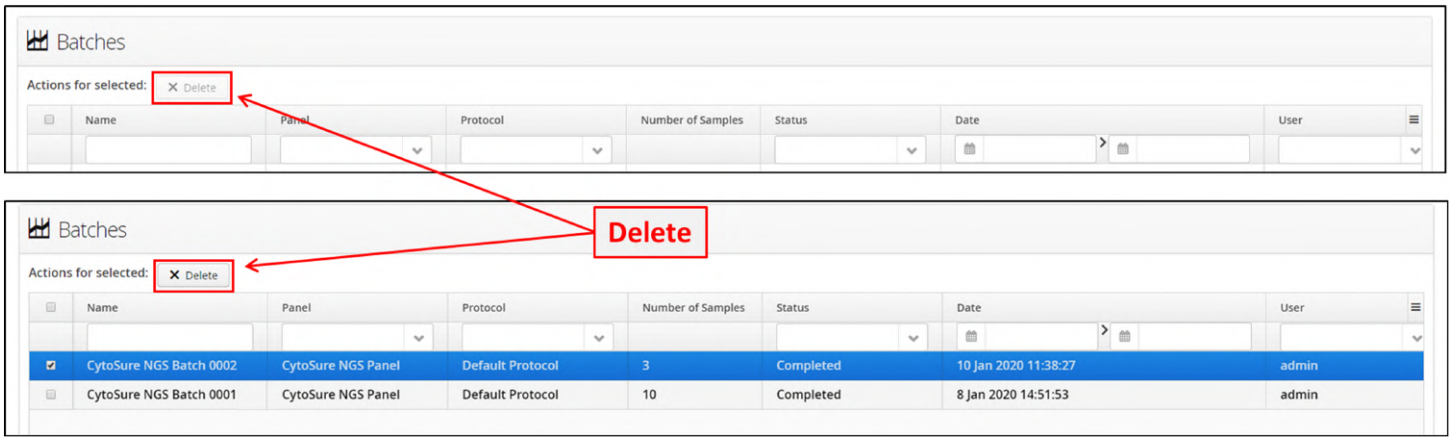


Figure: Selection of a batch to delete highlights the delete button

If the delete button is selected there will be a popup box requesting confirmation of the deletion.

Selecting OK will lead to the batch being deleted.

Once a batch is deleted it CANNOT be recovered.

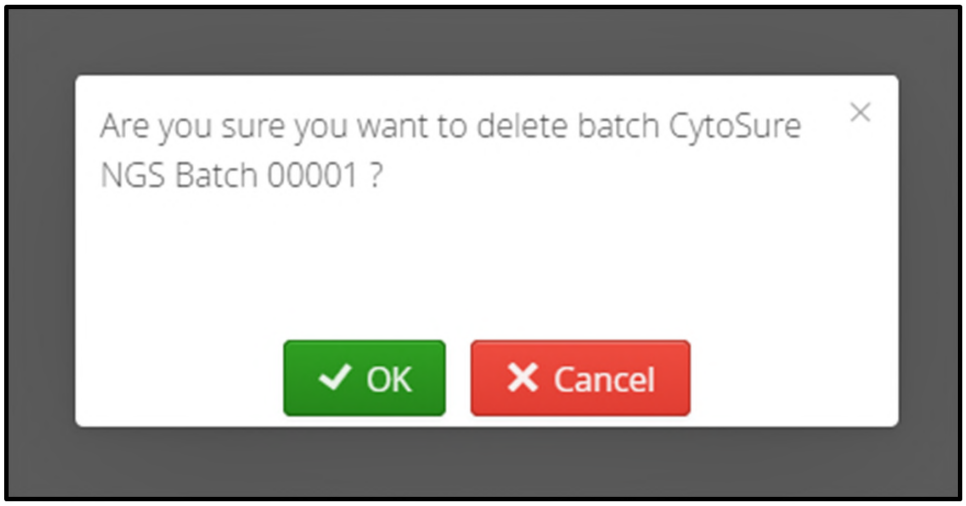


Figure: Popup box requesting confirmation of batch deletion

9.2 Individual Batches

Clicking on a row in the View Batches page will open a new page showing the selected batch in more detail.

There are 3 parts to the information provided,

1. Overview

   The overview provides information about the analysis

   2. Batch Functions

   The batch functions allow users to download files form the analysis, repeat the analysis or generate a report

   3. Sample Details

   In this part there is the headline information form the run about each sample such as the number of SNVs and CNVs called.

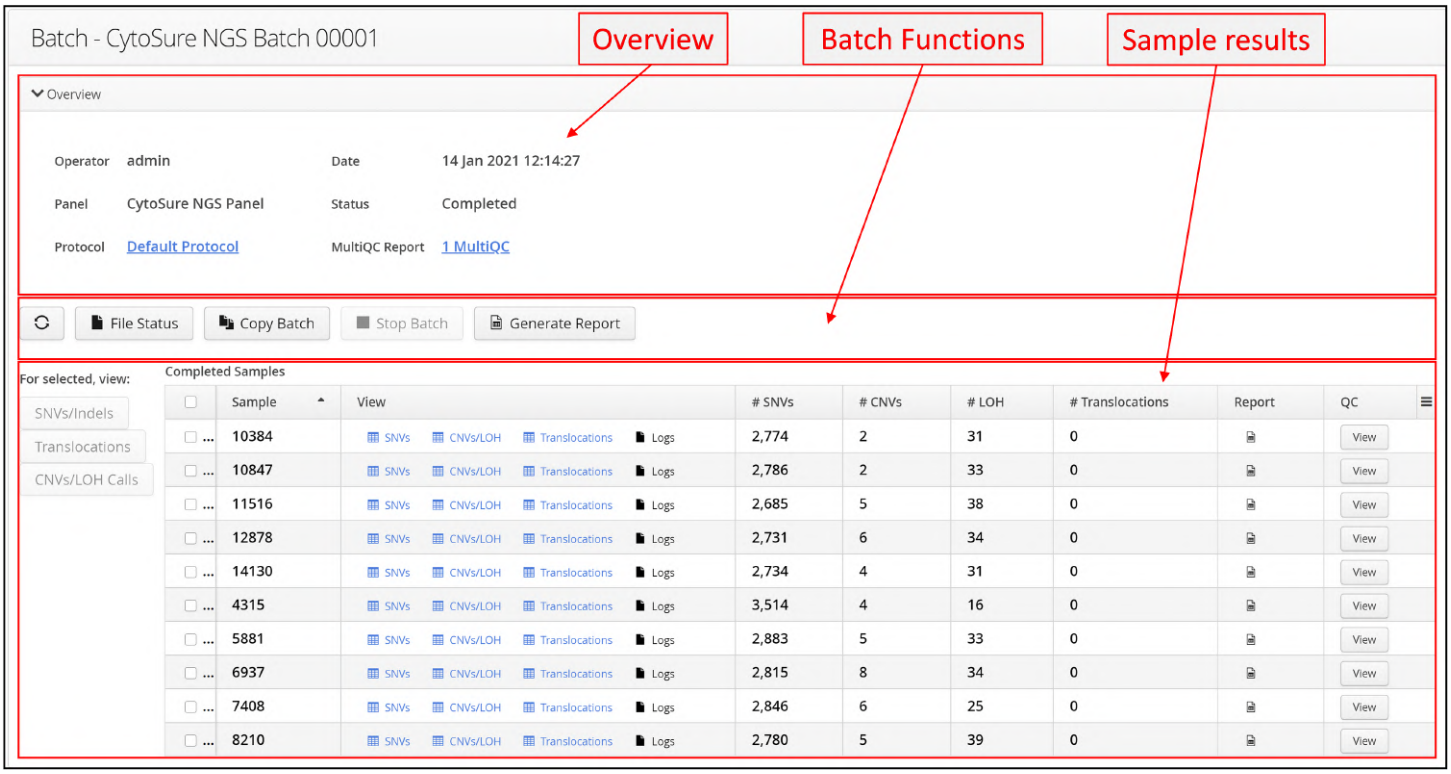


Figure: The sections of the batch analysis window

9.3 Batch QC

Included in the Batch page are two QC reports.

In the batch overview there is a link to a MultiQC report which gives an overview of all the samples that were in the batch.

Additionally, each sample in the completed table has a FastQC report for each read file.

Examples of both of these QC reports are shown below.

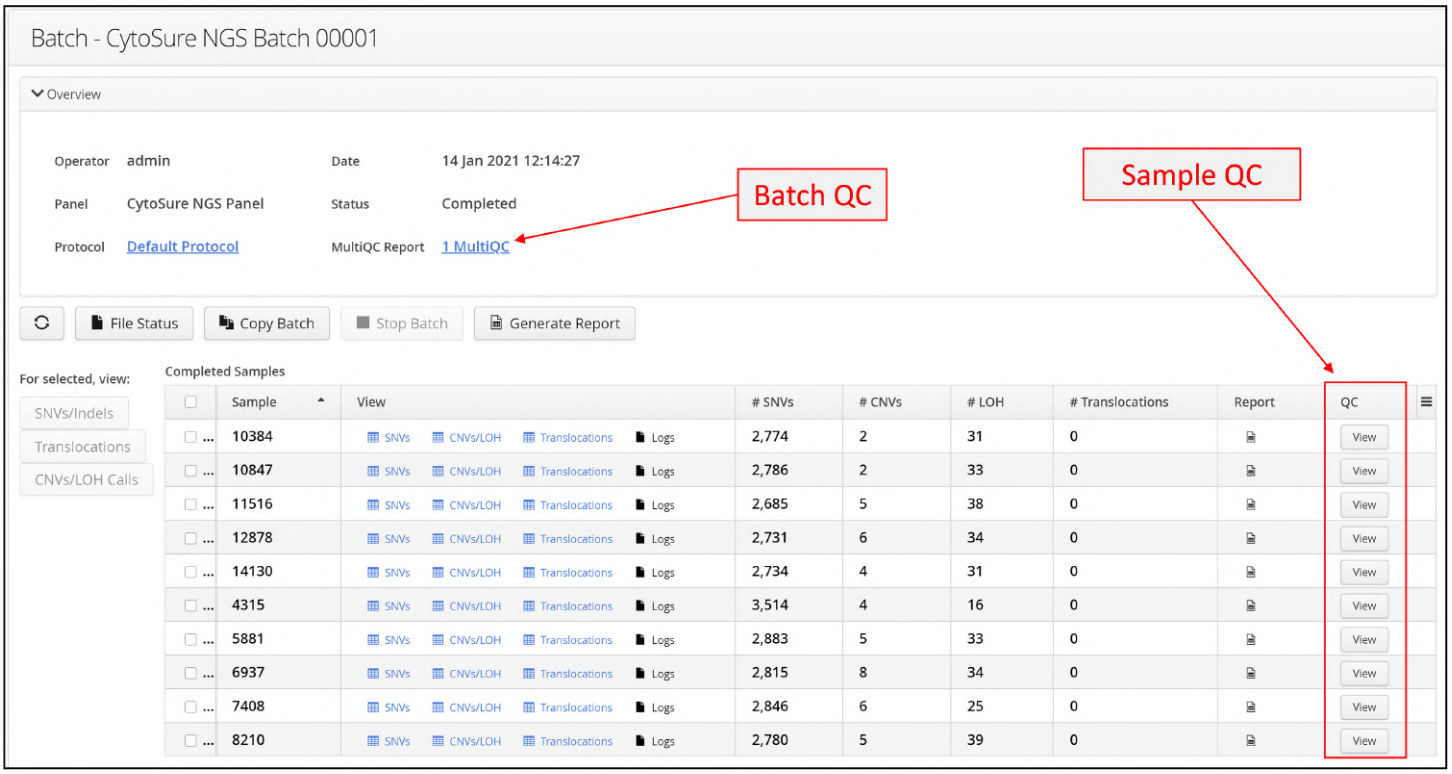


Figure: Links to QC reports for a batch and a sample

9.3.1 QC

9.3.2 Batch QC

MultiQC is a reporting tool for the whole batch of samples. It parses summary statistics from results and log files generated by other bioinformatics tools.

When you launch MultiQC, it recursively searches through any provided file paths for specific files. These files are parsed for relevant information and used to generates a single stand-alone HTML report file. It also saves a directory of data files with all parsed data for further use downstream. To save MultiQC report to user's computer, right click on the page, and choose "Save as...".

Additional information about MultiQC can be found in the next section of this guide.

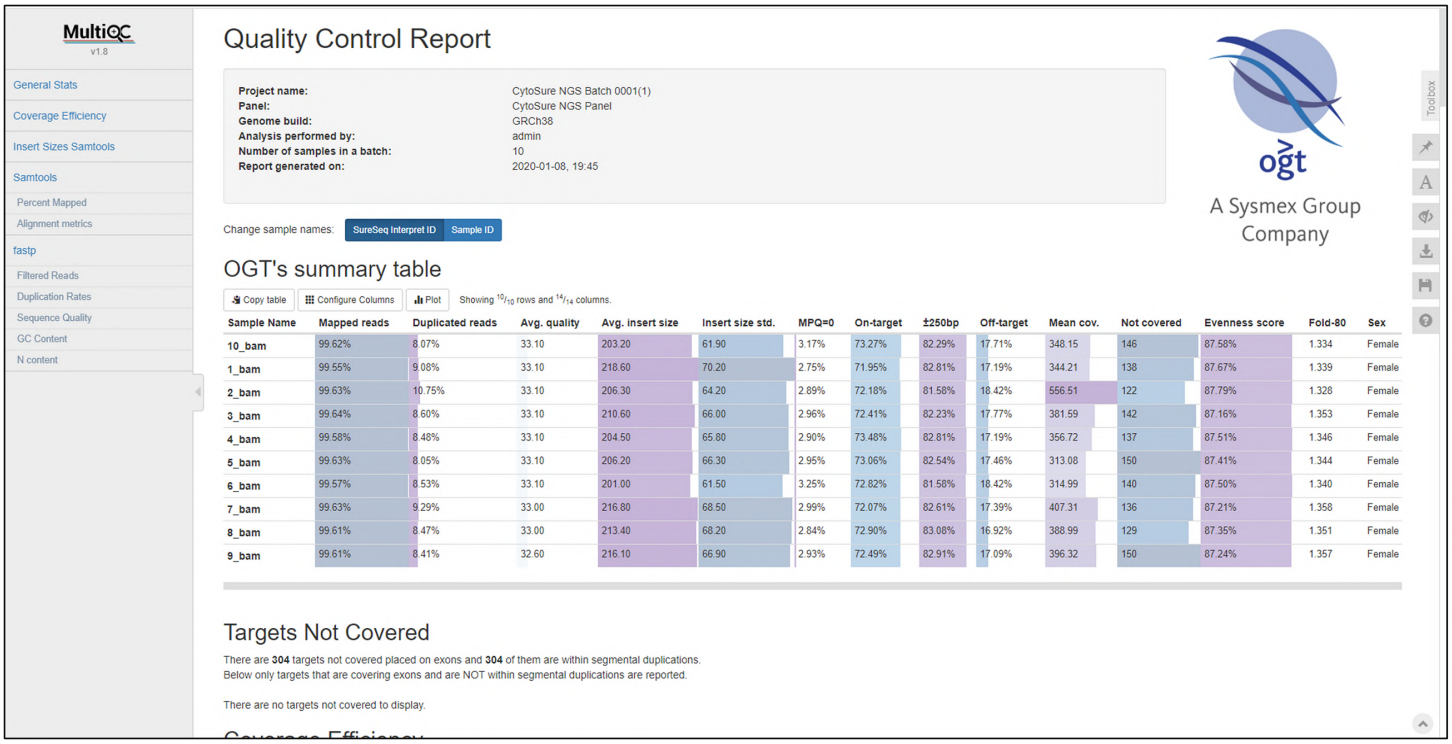
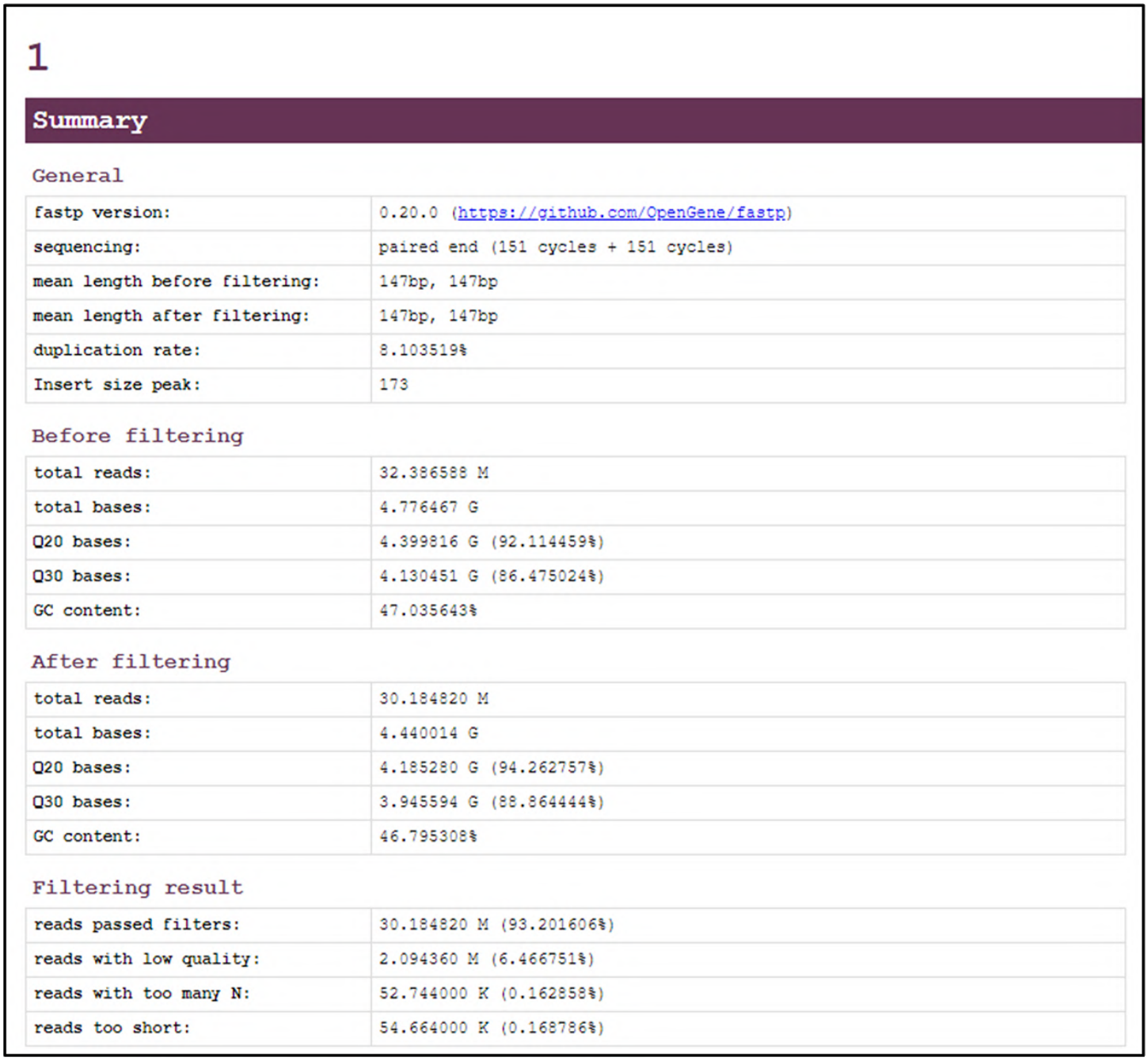


Figure: Example of a MultiQC report

9.3.3 Sample QC

FastP is used for sample QC data generation. Clicking on the 

button on the sample view will open up a new tab in the web browser with the sample QC details.



1

Summary

General

fastp version:	0.20.0 (https://github.com/OpenGene/fastp)
sequencing:	paired end (151 cycles + 151 cycles)
mean length before filtering:	147bp, 147bp
mean length after filtering:	147bp, 147bp
duplication rate:	8.103519%
Insert size peak:	173

Before filtering

total reads:	32.386588 M
total bases:	4.776467 G
Q20 bases:	4.399816 G (92.114459%)
Q30 bases:	4.130451 G (86.475024%)
GC content:	47.035643%

After filtering

total reads:	30.184820 M
total bases:	4.440014 G
Q20 bases:	4.185280 G (94.262757%)
Q30 bases:	3.945594 G (88.864444%)
GC content:	46.795308%

Filtering result

reads passed filters:	30.184820 M (93.201606%)
reads with low quality:	2.094360 M (6.466751%)
reads with too many N:	52.744000 K (0.162858%)
reads too short:	54.664000 K (0.168786%)

Figure: Example of a FastP report

9.3.4 Batch Functions

Below the overview section there are a set of buttons providing a set of option - when the batch has finished processing the Stop Batch button is disabled.

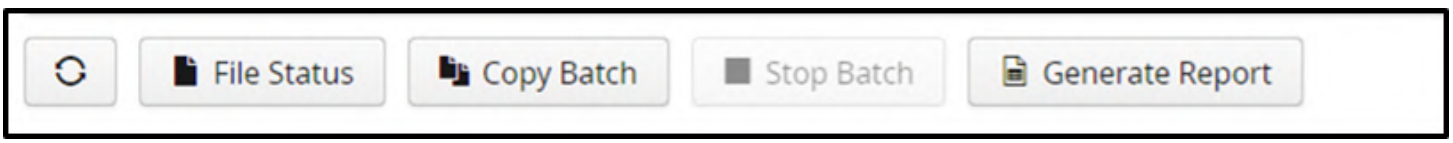


Figure: Batch options

File Status

File status provides shows the files that have been generated for each sample during the analysis.

Files provided are:

1. Alignment files
2. QC files
3. VCF files
4. CGH files for loading into CytoSure Interpret
5. Log files

Where a green tick is displayed, that file is available for download and this can be achieved by clicking on the Download button.

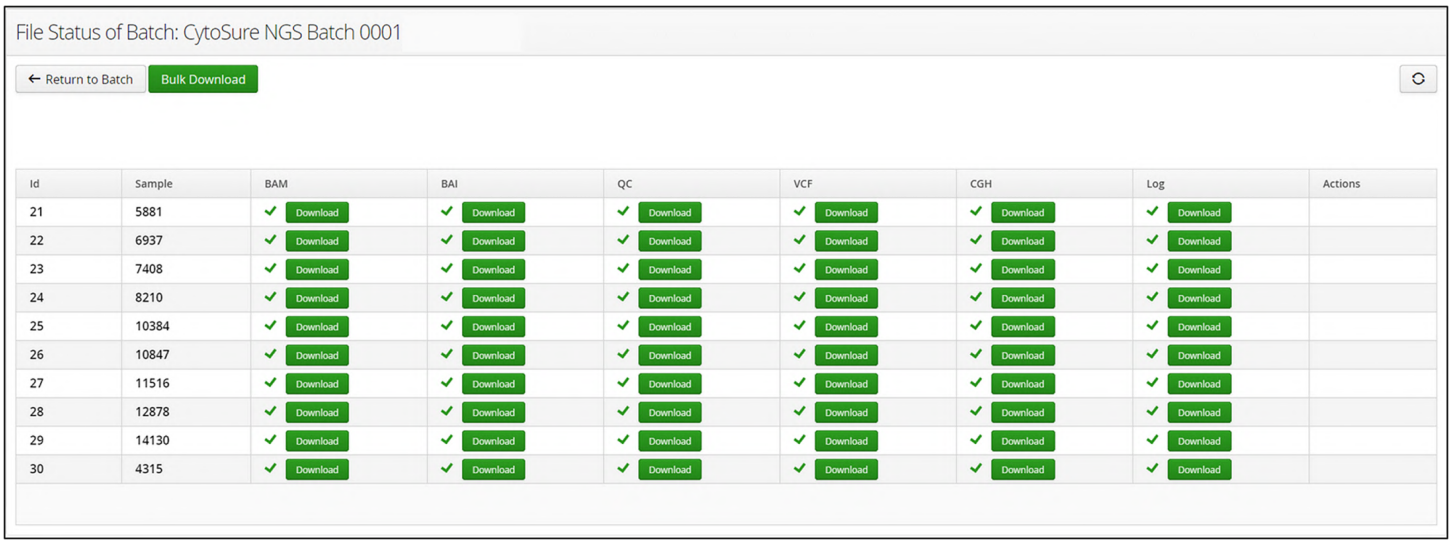


Figure: Status of files generated by the pipeline for each sample

It is possible to download all files, or selected files, simultaneously via the bulk download button Bulk Download

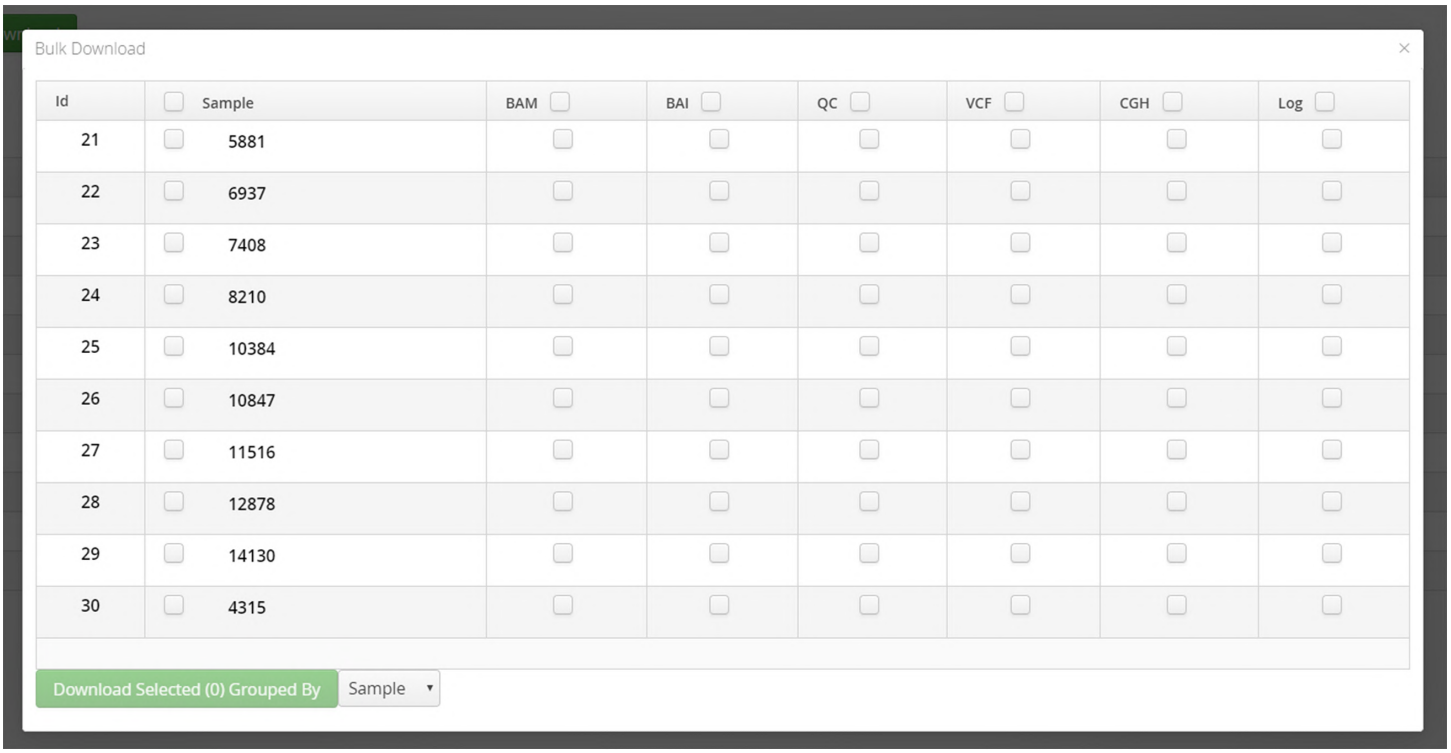


Figure: Bulk download file selector

Specific files can be selected as below

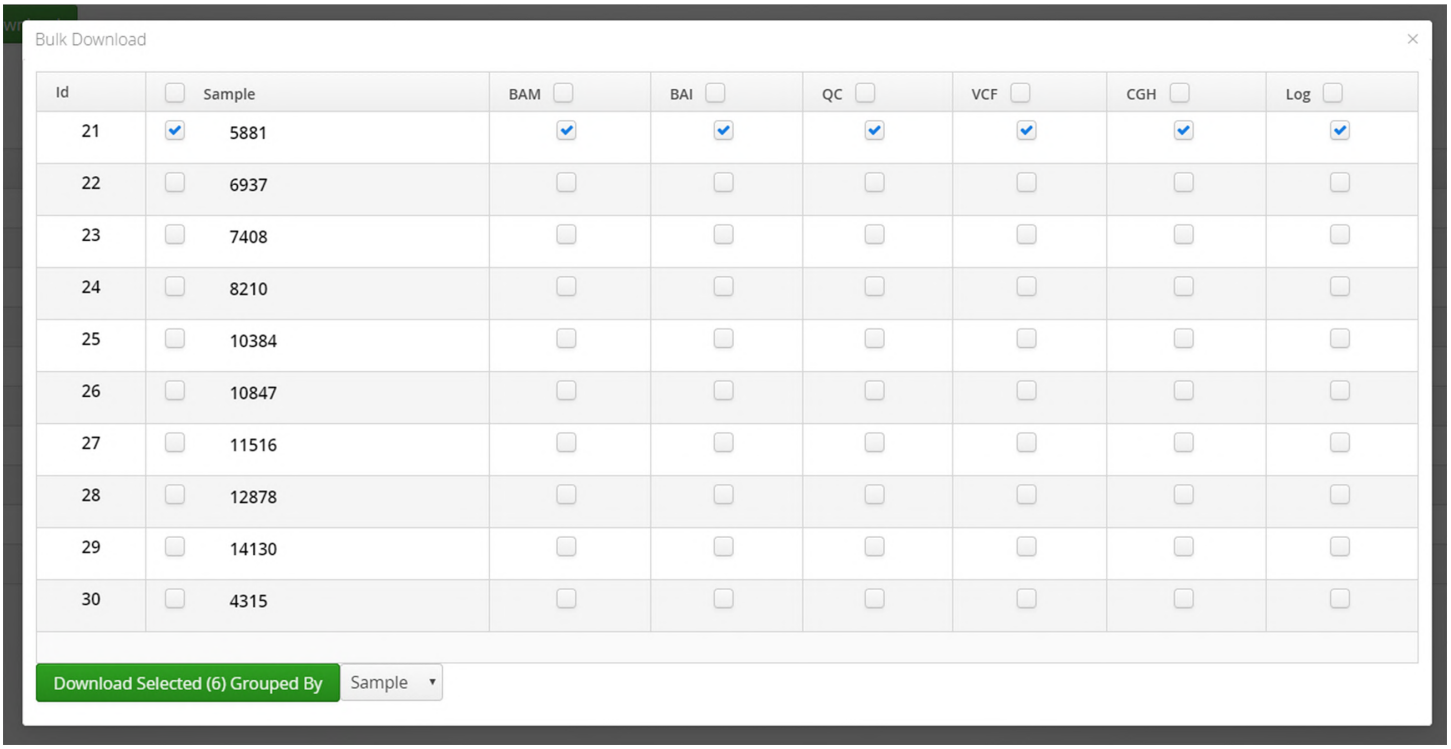


Figure: Bulk download with single sample selected

Alternatively, all files can be selected for download

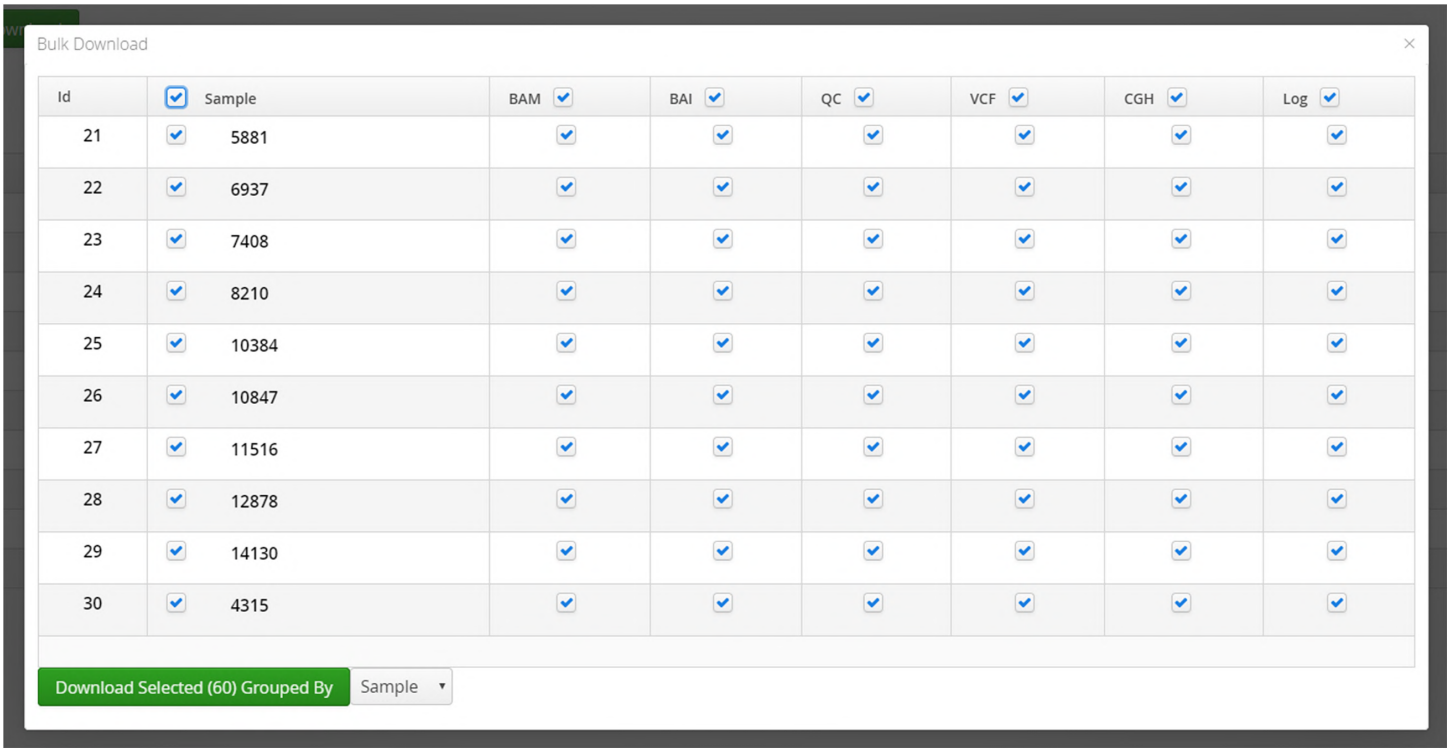


Figure: Bulk download with all files selected

Files downloaded in bulk can be grouped by sample or file type

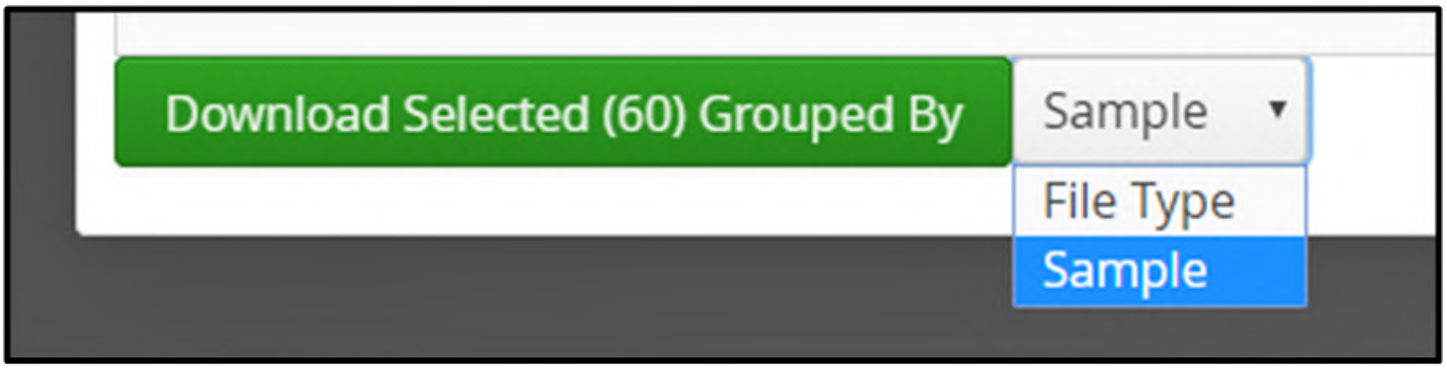


Figure: Bulk download selecting grouping by Sample of downloaded files

Copy Batch

This function allows the user to repeat the batch analysis with the same settings. When selected a Run Batch window opens and if the user selects to Run Analysis the processing will be repeated.

The software will automatically update the Batch Name but otherwise nothing is changed including the time stamp.

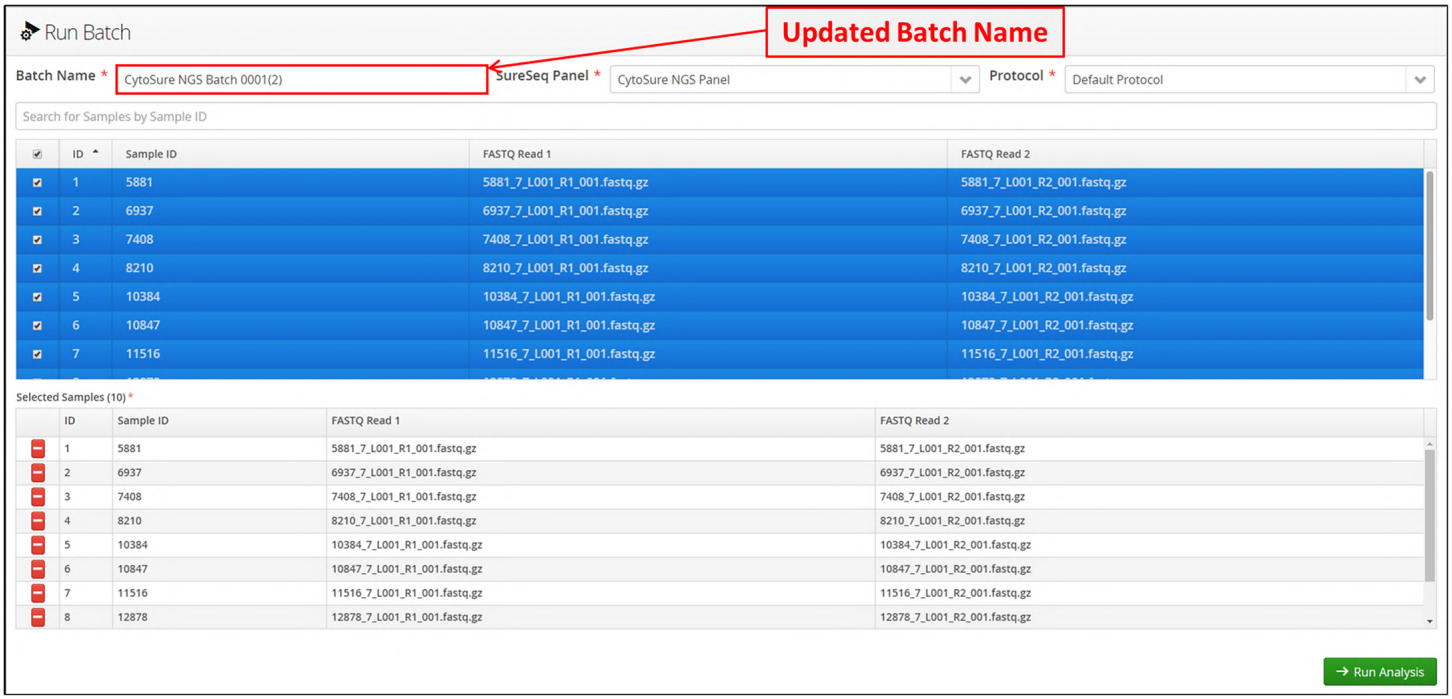


Figure: A batch analysis being repeated using the Copy Batch option showing the updated batch name

Report Generation

Report Generation shows a drop down in which the user can select the report to be generated

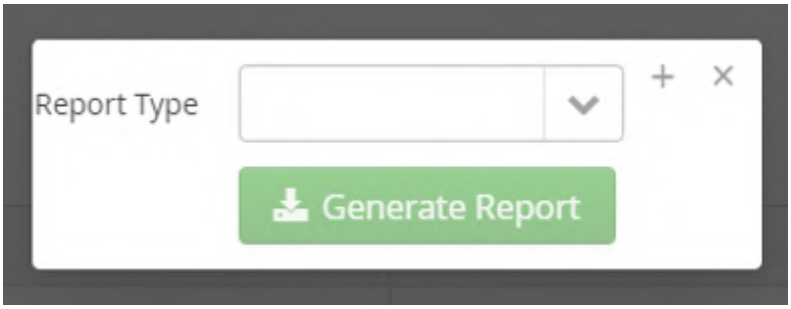


Figure: Initial view of the report options

Currently, the only template loaded is the Batch Report

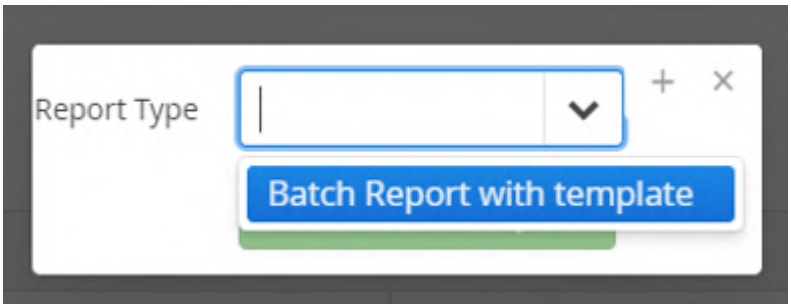


Figure: Selecting a report type

for the QC of the run.

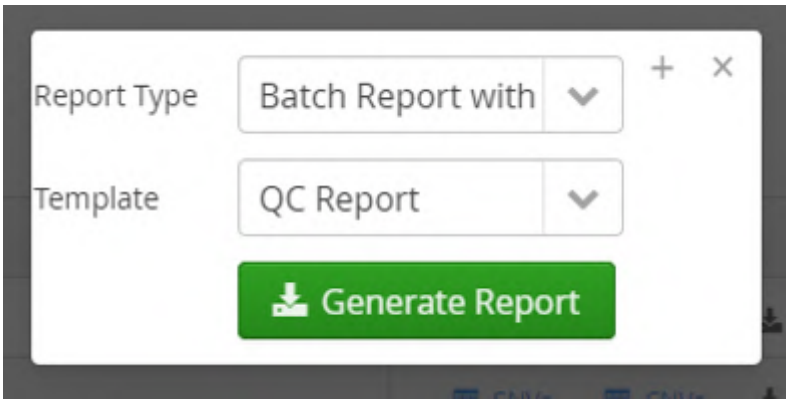


Figure: Selection of a template for the report

When the report is generated the output is a table with a set of metrics for each sample in the batch.

Sample	Percent Reads Aligned	Percent Duplication	Mean Target Coverage	Targets Not Covered	Aligned Reads GC	Aligned Reads Per Base Quality	Usable On Target Reads	Usable On Target Bases
5881	99.4	43.3	536	0	40	37.4	53.9219	35.5563
6937	99.4	43.1	627	0	40	37.4	55.4457	36.6627
7408	99.1	50.2	440	0	40	37.3	39.6108	25.9758
8210	99.1	50.7	418	0	41	37.3	40.9108	26.9532

Figure: Example output of the QC report for a batch

Selecting a completed sample or samples allows viewing of the variant information and this is described in Viewing Analysis Results section.

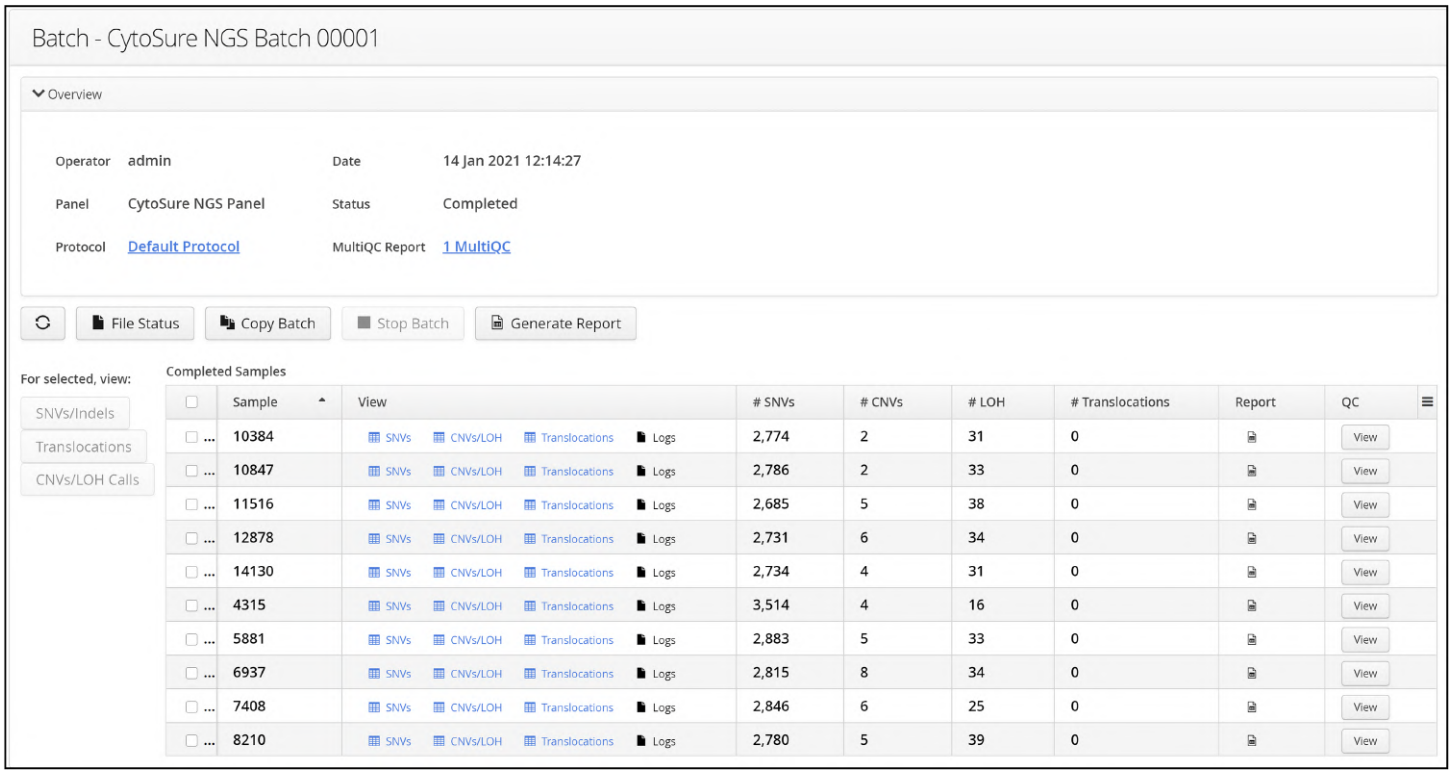


Figure: Selecting a batch to view

10 Viewing Analysis QC

When a batch of samples is processed, besides individual sample metrics that were discussed in the previous section, there is a batch QC report generated. This uses MultiQC and fastp to collate a set of metrics for each sample and merge into a set of graphs and tables.

The report can be accessed from in the batch overview displayed once a batch has completed analysis.

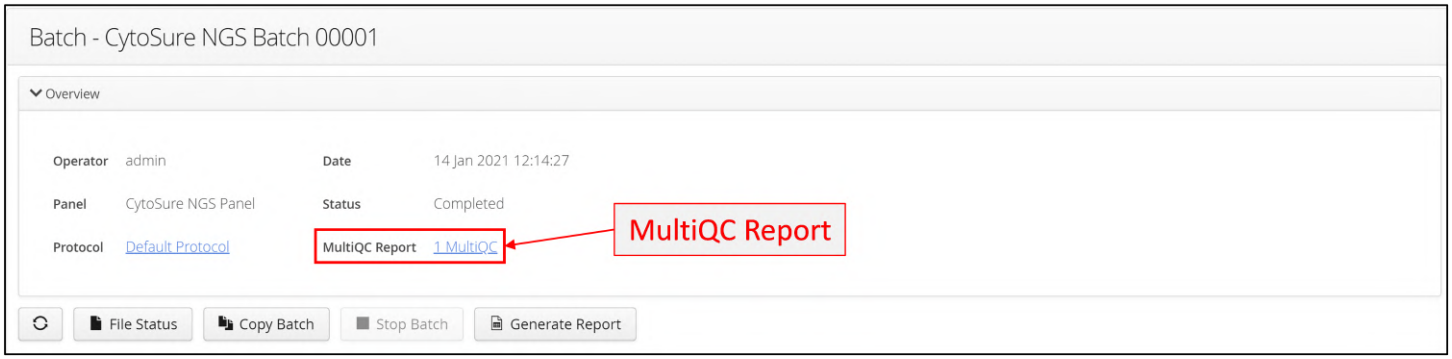


Figure: Accessing the Batch QC report

When the user clicks on the MultiQC Report link a new tab opens up in the browser displaying the QC report. The view is divided into 3 parts - the quality control report for the batch, which comprises the bulk of the display, and 2 tabs that come into the view from the left and the right of the page. These tabs can be viewed and hidden by clicking on their respective buttons which are highlighted in the following figure:

Figure: Image showing the buttons for showing/hiding the tabs in the Quality Control report.

first of these is the .... The second tab provides the MultiQC toolbox for

1. The Quality Control report for the batch
2. The report short cut tab
3. The tool box tab

At the head is the quality control report; this provides general information about the analysis such as the date of the analysis and which user performed it.

Figure: Example batch overview details

OGT's Summary Table

Each sample has a row in the table with some key metrics.

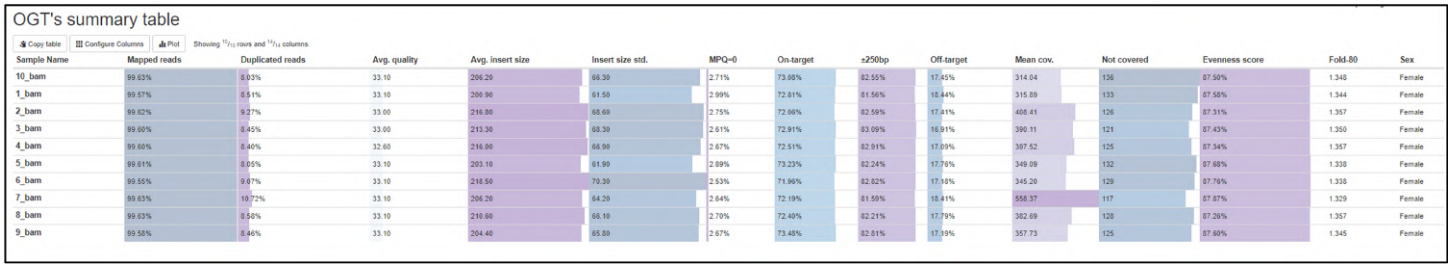


Figure: Example sample QC summary table

The column names from the summary table are listed in the table below with some additional detail as to their meaning.

Column Name	Description
Mapped reads	The percentage of reads that mapped to the reference genome
Duplicate reads	The percentage of reads that were duplicated
Avg. quality	The average read quality reported by Samtools stats
Avg. insert size	The average insert size reported by Samtools stats

Column Name	Description
Insert size std	The standard deviation of the insert size reported by Samtools stats
MPQ = 0	The percentage of reads that were mapped that have a mapping quality of 0
On-target	The percentage of reads that map on target that are not duplicate reads
± 250bp	The percentage of reads that overlap target regions extended by 250bp
Off-target	The percentage of reads that are neither on target nor within the specified flanking region
Mean cov.	The mean target coverage
Not covered	The number of targets with a coverage of less than 1
Evenness score	The fraction of the whole sequencing output that is correctly distributed
Fold-80	The fold of additional sequencing that would be required to ensure that 80% of targeted bases achieve the mean target coverage.
Sex	The chromosomal sex of the sample predicted from the distribution of reads that map to the sex chromosomes

Table: Column names and their description from the QC summary table

Targets Not Covered

Any targets not covered are detailed, providing that they are not within a segmental duplication.

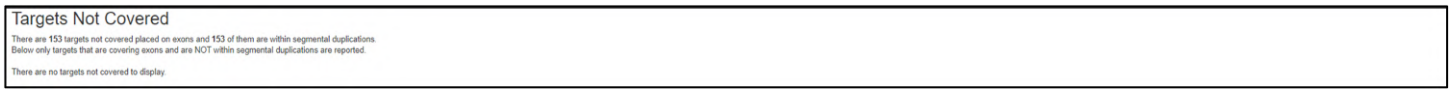


Figure: Example targets not covered summary

Coverage Efficiency

The efficiency of coverage as a measure of depth are displayed.

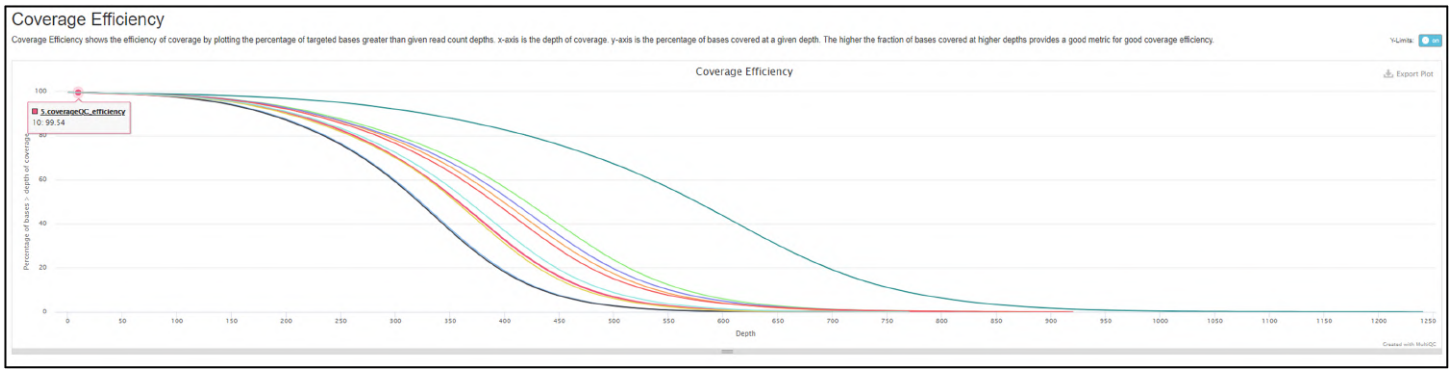


Figure: Example QC report summary

Insert Sizes Samtools

The distribution of insert sizes for each sample is displayed.

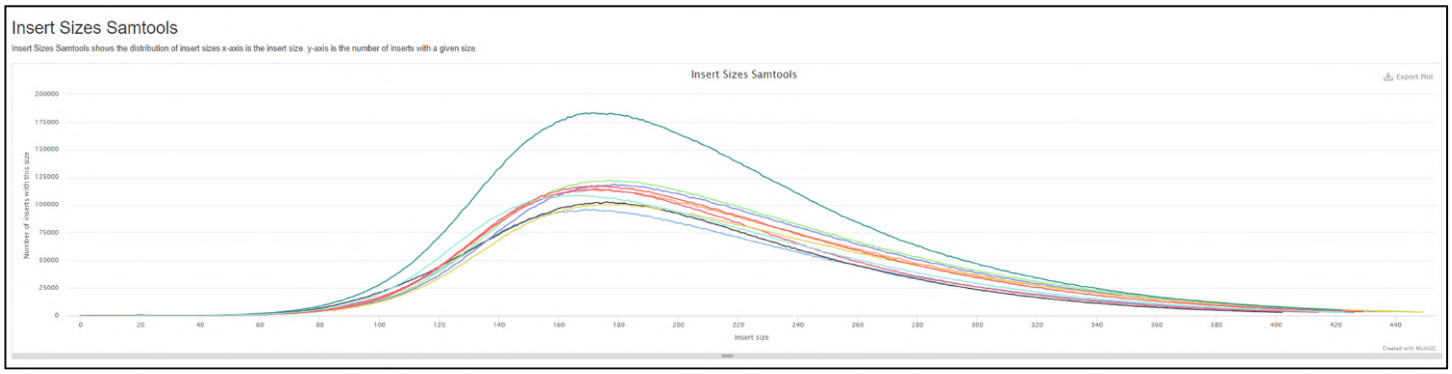


Figure: Example QC report summary

Percent Mapped

The percentage of base calls at each position for which an `N` was called.

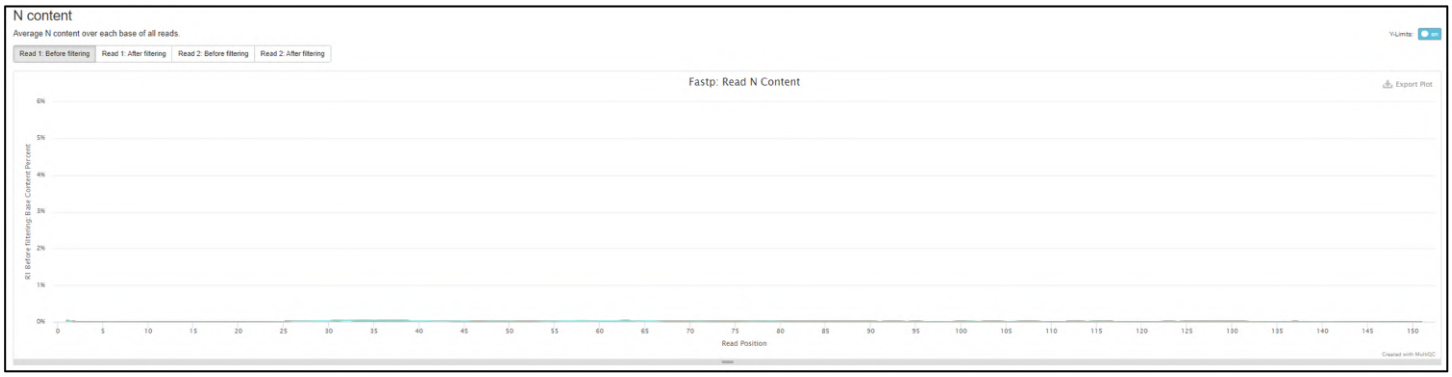


Figure: Example QC report summary

Alignment Metrics

The alignment metrics for all the samples in the batch are plotted.

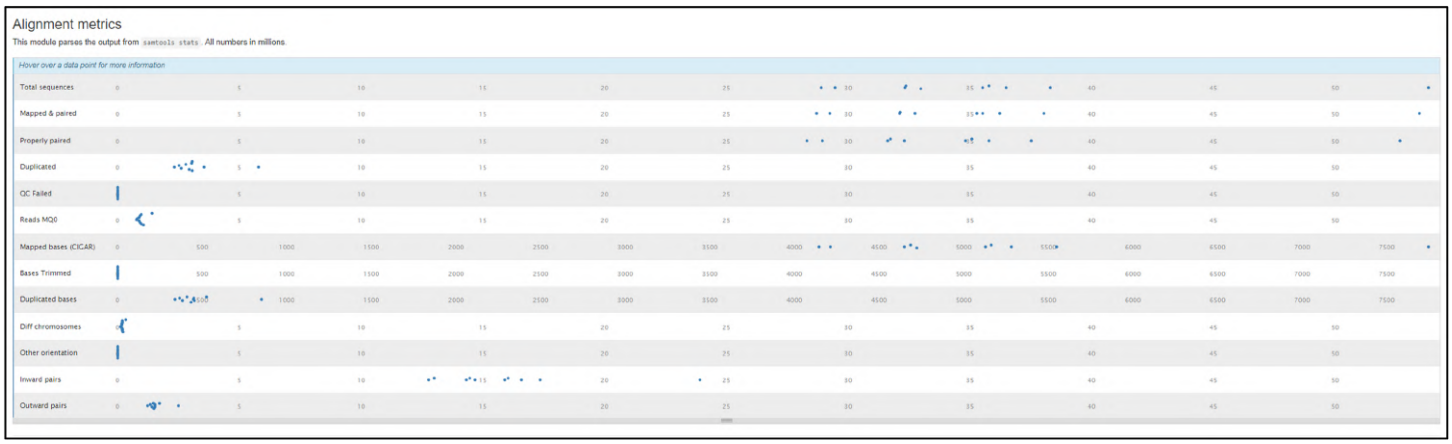


Figure: Example of the alignment metrics

Filtered Reads

The filtered reads graph shows the number or percentage of reads that have been removed by the filter.

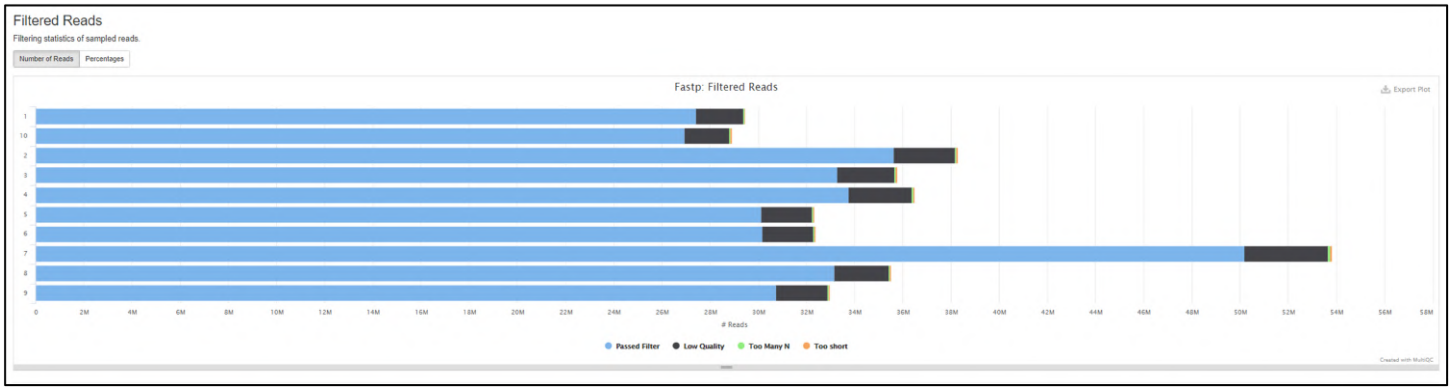


Figure: Example QC report summary

Duplication Rates

The relative level of duplication found for each sample as a percentage.

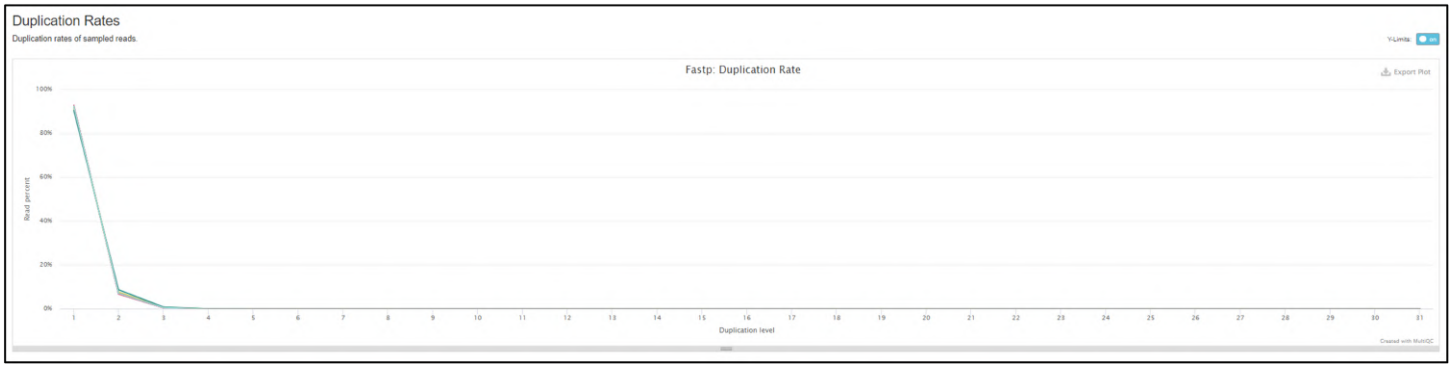


Figure: Example QC report summary

Sequence Quality

The mean sequence quality or Phred score of each base in the insert for each sample.

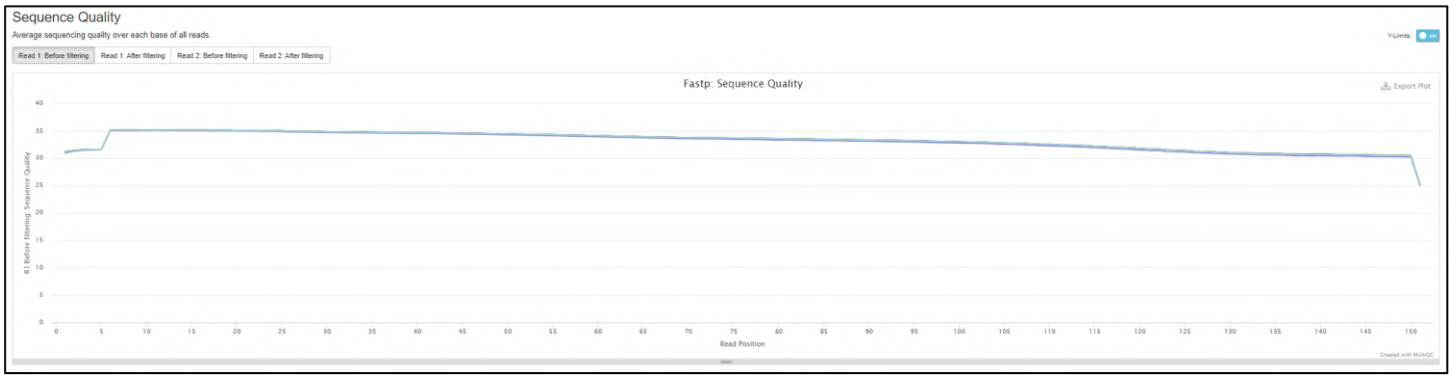


Figure: Example QC report summary

GC Content

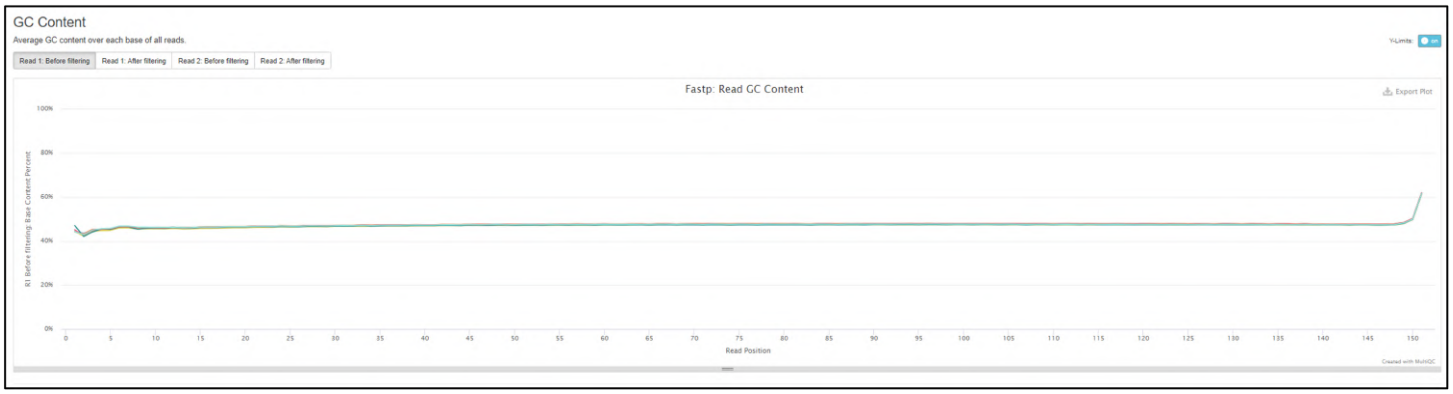


Figure: Example QC report summary

N Content

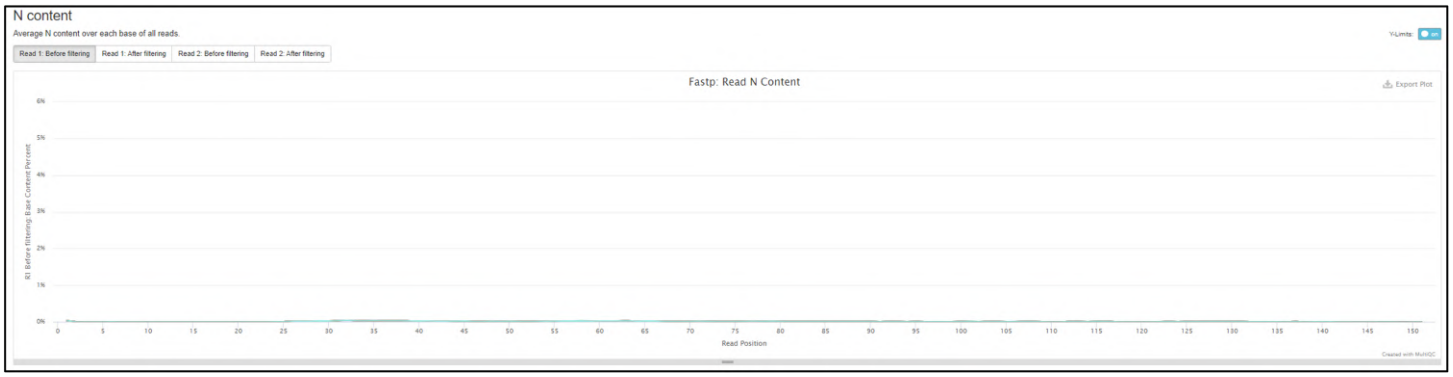


Figure: Example QC report summary

10.1 Sample QC

Sample QC information can also be access via the Sample page. For a particular run selecting the QC metrics tab will provide the relevant information. Colours are defined by the metric set used in the analysis protocol.

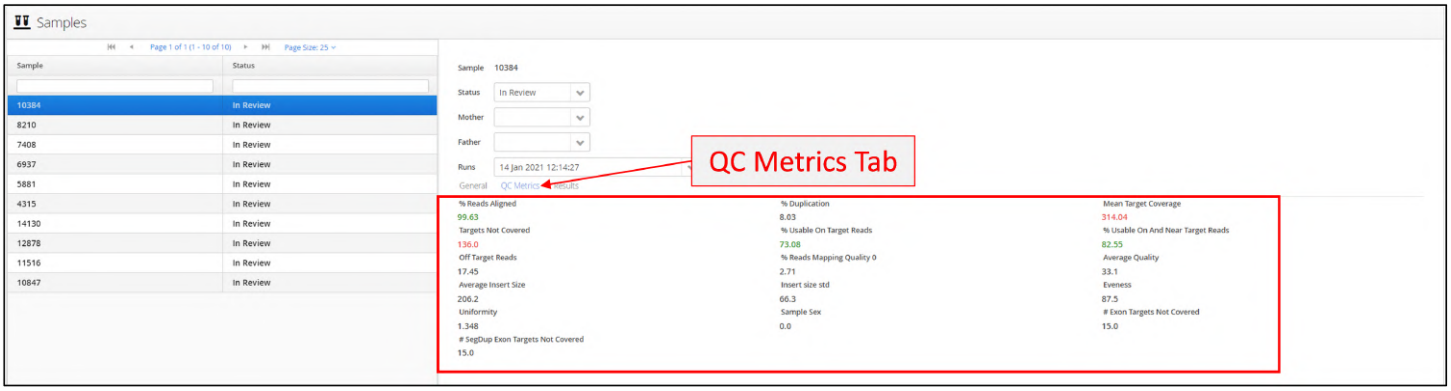


Figure: Viewing the metrics for an individual sample

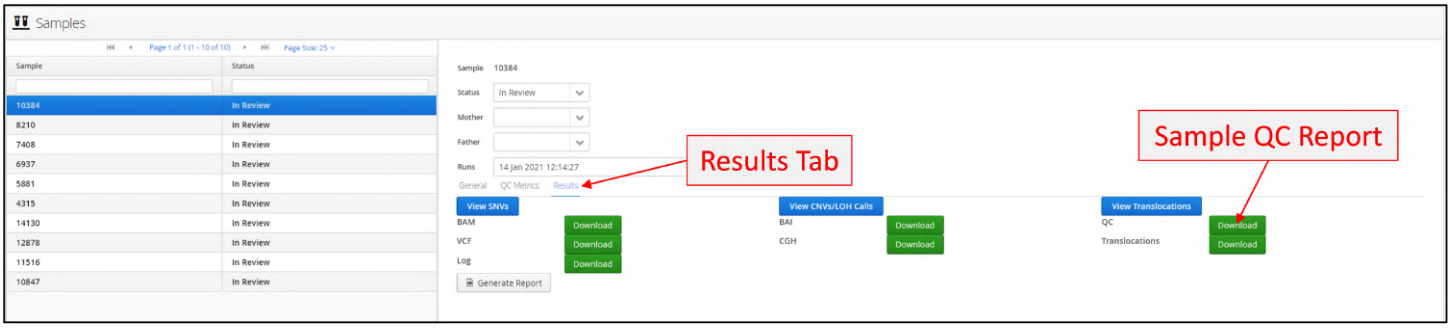


Figure: Accessing the Sample QC report from the sample results tab

10

Summary

General

fastp version:	0.20.0 (https://github.com/OpenGene/fastp)
sequencing:	paired end (151 cycles + 151 cycles)
mean length before filtering:	146bp, 146bp
mean length after filtering:	145bp, 145bp
duplication rate:	7.220084%
Insert size peak:	171

Before filtering

total reads:	28.905316 M
total bases:	4.227335 G
Q20 bases:	3.894346 G (92.122952%)
Q30 bases:	3.658105 G (86.534532%)
GC content:	47.412412%

After filtering

total reads:	26.944780 M
total bases:	3.928181 G
Q20 bases:	3.703276 G (94.274565%)
Q30 bases:	3.492954 G (88.920399%)
GC content:	47.151145%

Filtering result

reads passed filters:	26.944780 M (93.217386%)
reads with low quality:	1.862392 M (6.443078%)
reads with too many N:	46.288000 K (0.160137%)
reads too short:	51.856000 K (0.179400%)

Figure: Start of a FastP report for an individual sample

11 Viewing Analysis Results By Sample

Viewing a Sample

Access to the results from running the pipeline are described in the previous section "View Analysis Batches".

Within each batch are the samples processed in that batch comprising analysed variants and QC metrics.

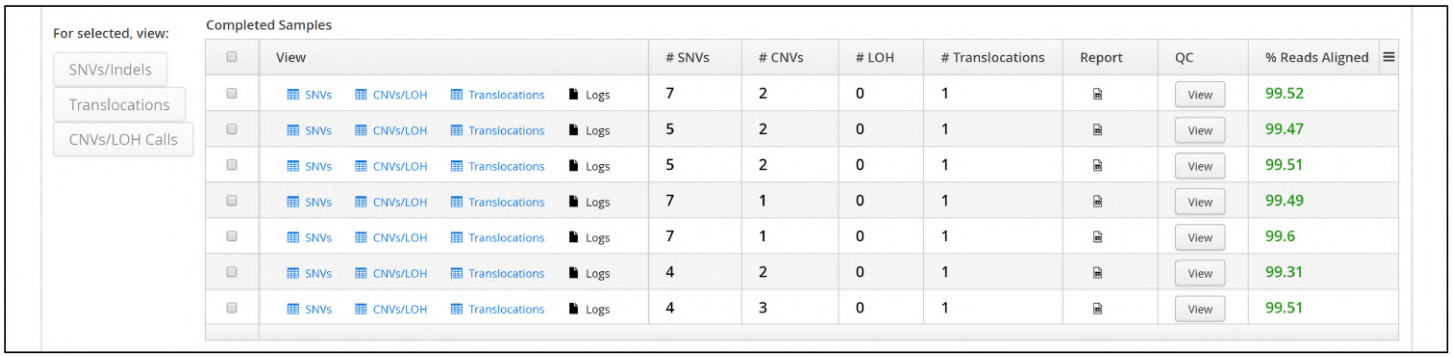


Figure: View of a set of processed samples in the batch view

As with other tables in Interpret, where there is a column selection icon  users can use it to configure which columns are being displayed.

Figure: Column selector button for configuring columns to view in display

The column options for this view are shown in the Figure.

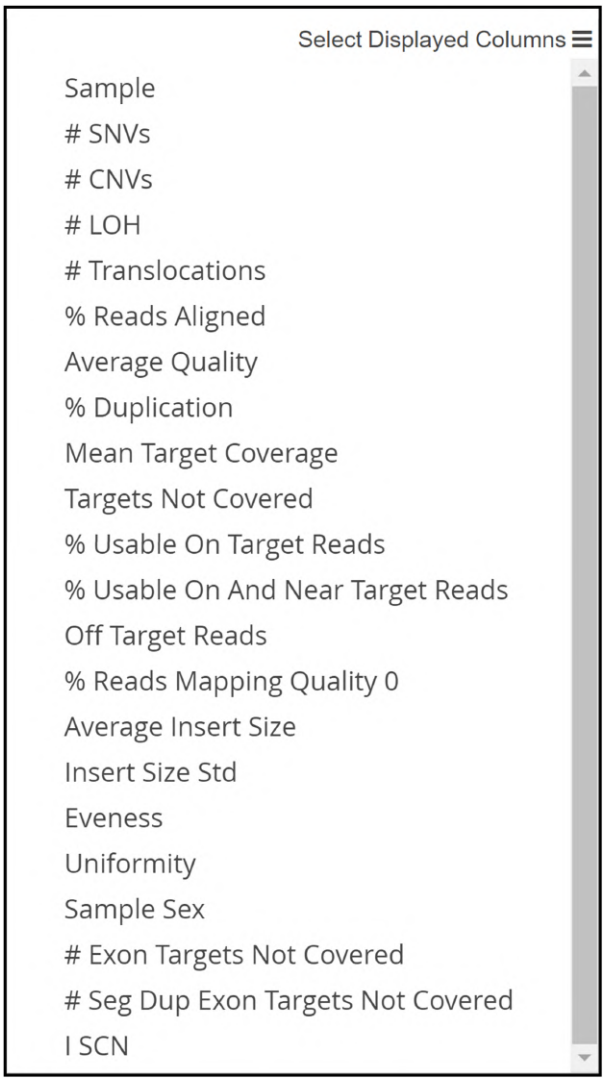


Figure: Columns available for selection to display

There is one completed sample per row and for each sample there is a range of information available to view.

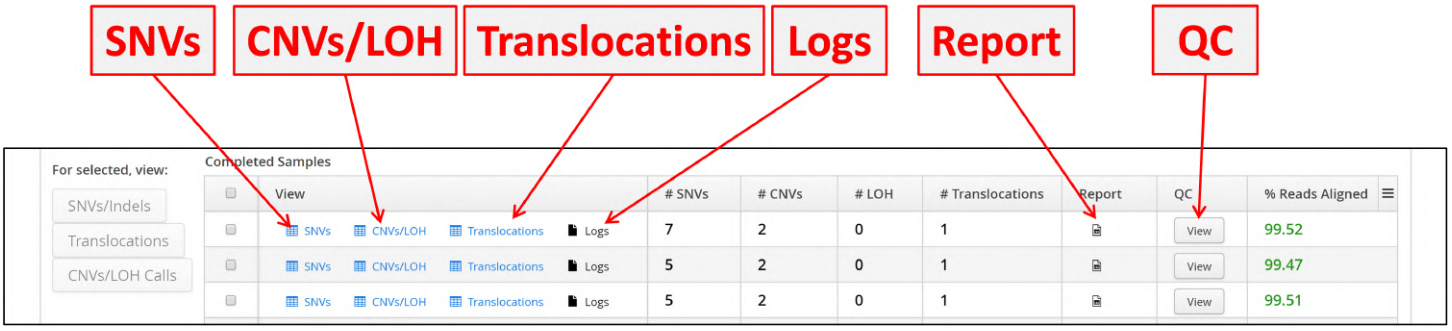


Figure: Information available for each sample

Variants for a sample can be viewed by selecting the SNVs or CNVs/LOH links present in each row.

Multiple samples can be viewed simultaneously by selecting the check boxes of the required samples which will then activate the SNVs and CNVs buttons on the left hand of the view.

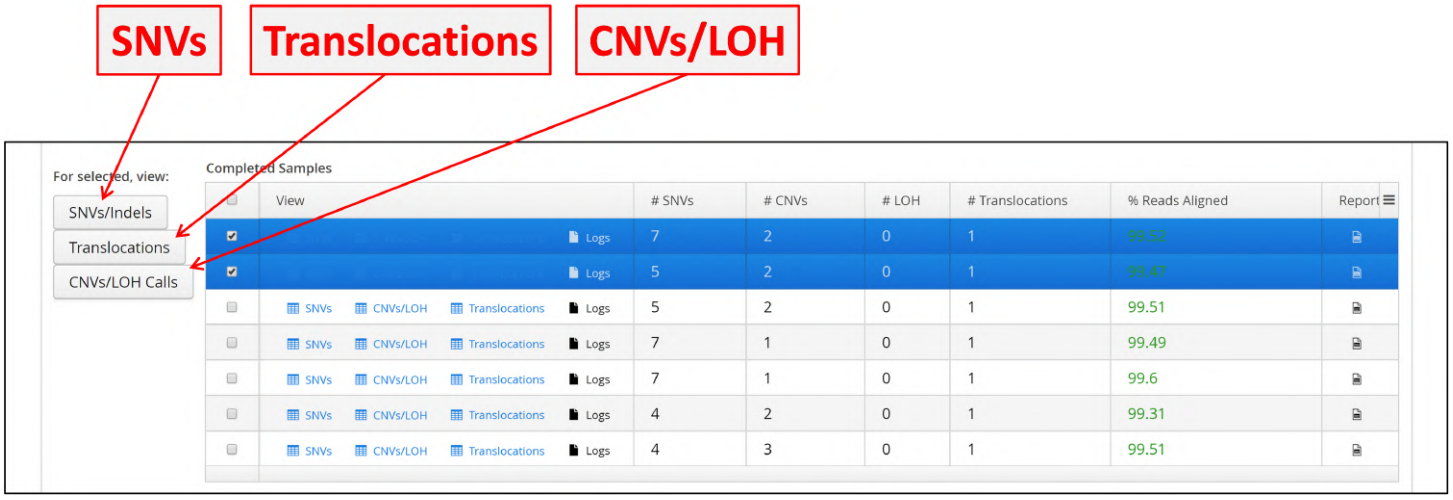


Figure: Selecting multiple samples to view simultaneously

Once selected, the variants will be displayed on a Variants page.

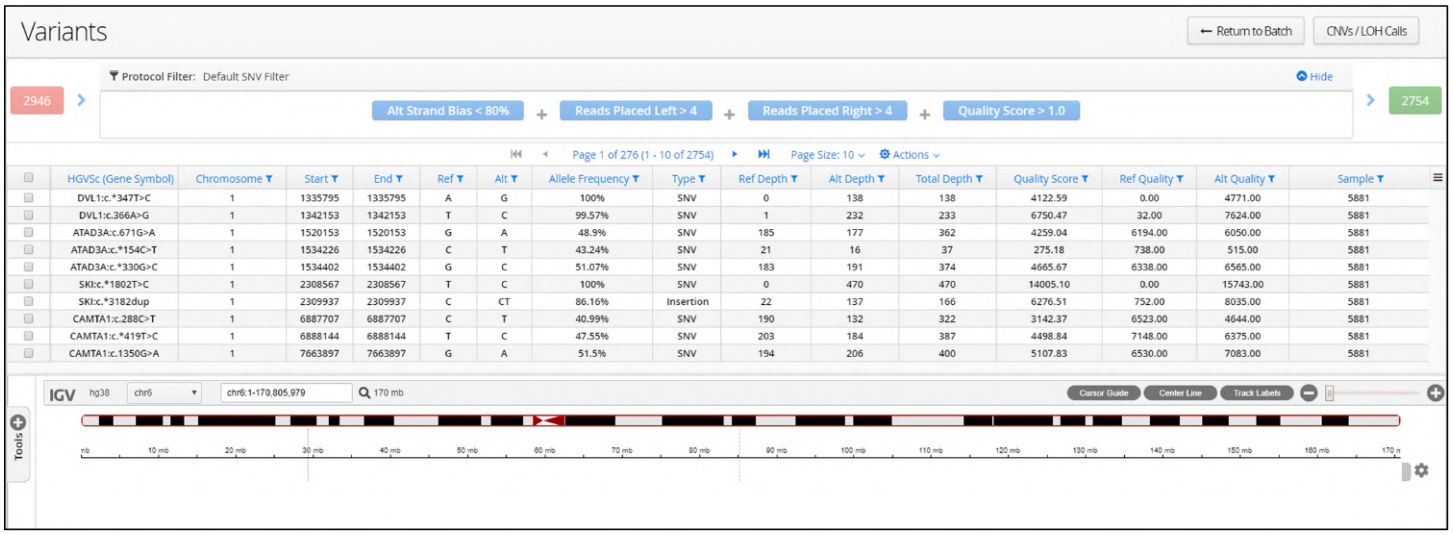


Figure: The initial SNV/Indel display page

At the top of the page are buttons that allow the user to toggle between the CNV and SNV views.

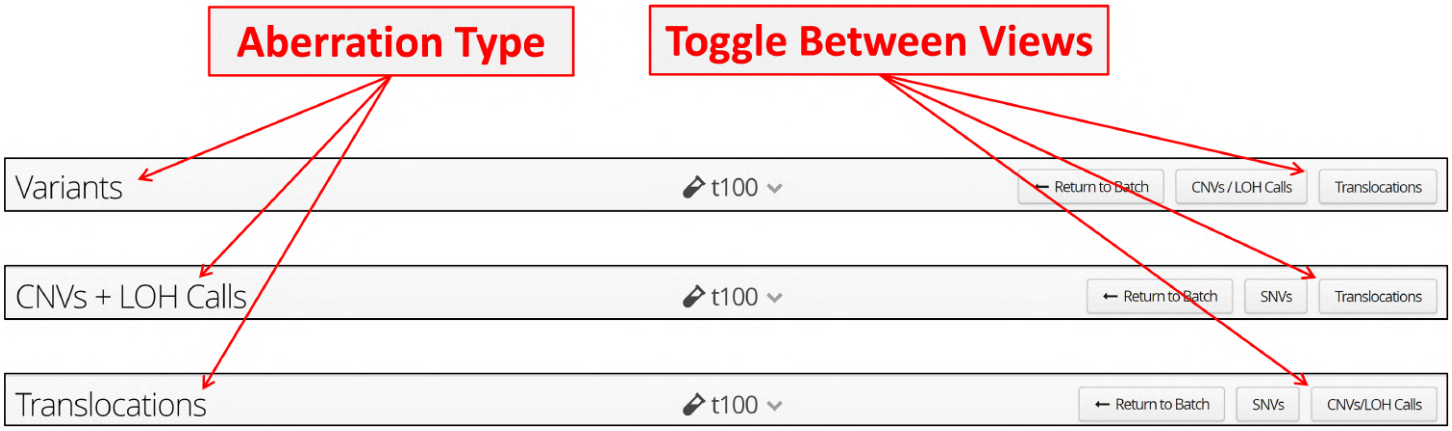


Figure: Toggling between SNV, CNV and Translocation reports

The Variants viewer is divided into three parts:

1. The Protocol Filter, initially this will be showing the Protocol Filter used in the analysis. The filter is modifiable in the Admin Controls (Admin Controls > Analysis > Protocols)
2. The Variant Table, showing the variants, one on each row. In the header there is a drop down "Actions" menu options; these are discussed below.
3. The Integrated Genome Viewer (IGV) that has been embedded in the software. Further details on using IGV are below.

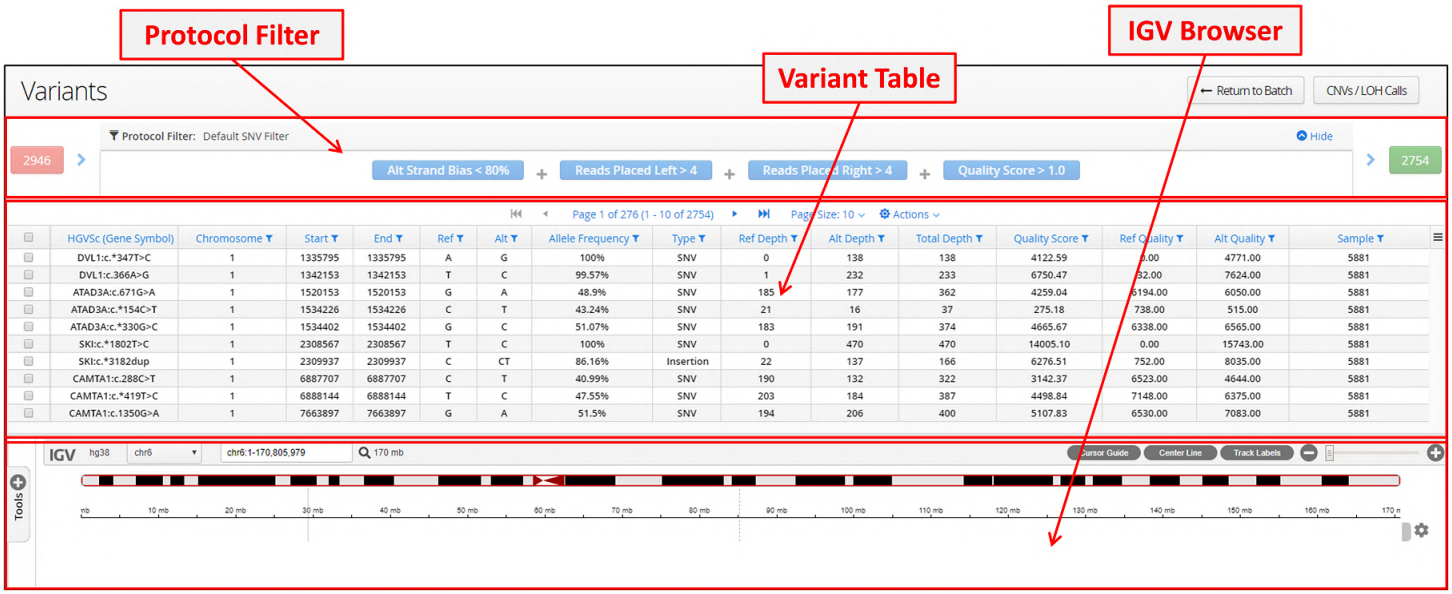


Figure: The sections of a sample results view page

11.1 Viewing SNV and Indel Events

The variant table has a column selector icon  allowing user to configure which columns are displayed.

There are different columns available depending on whether you are viewing the SNV variants page or the CNV/LOH variant page. The options for SNVs are displayed below.

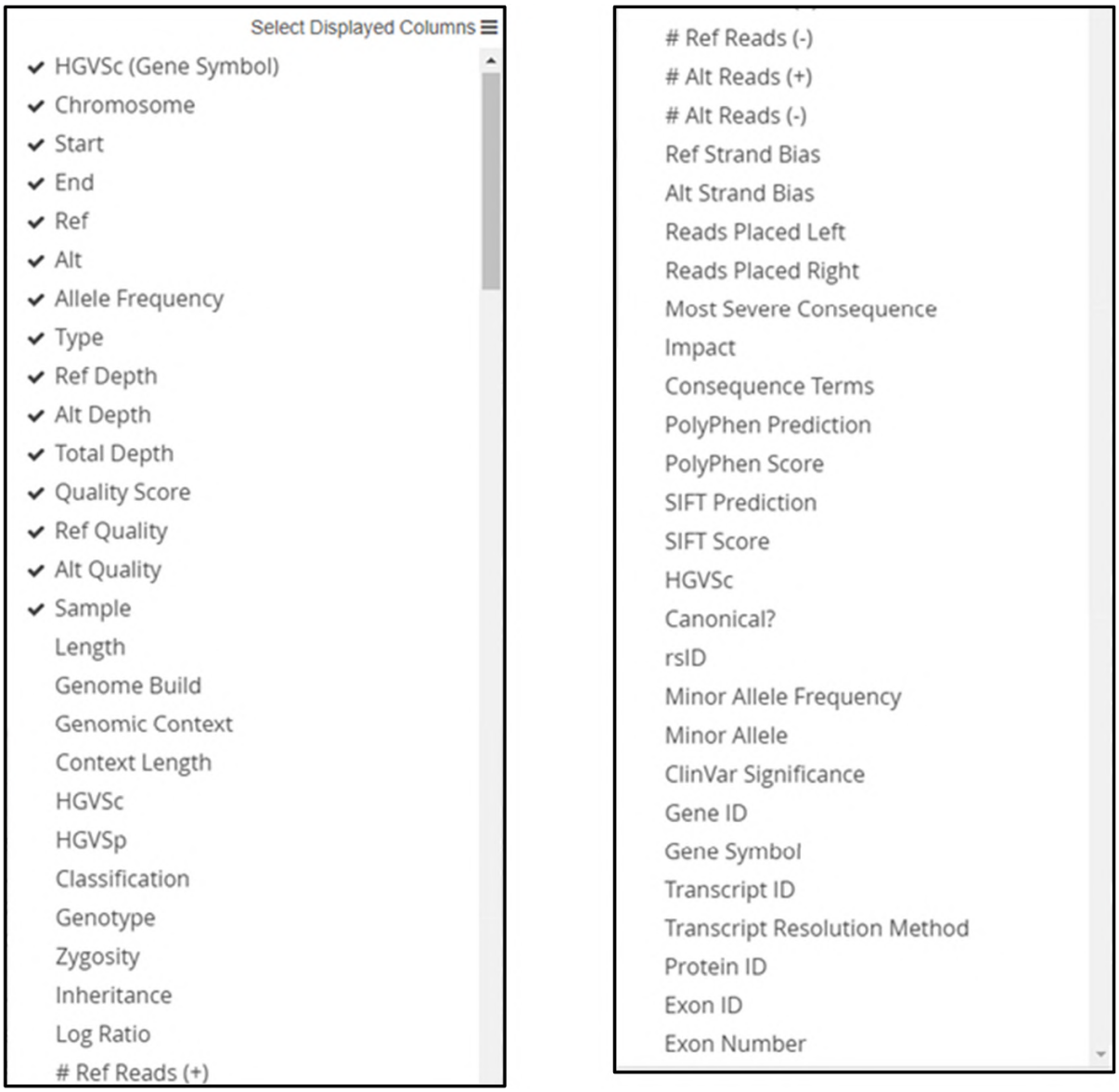


Figure: Columns available to display in the SNVs variant page

Selection of a variant will load the alignment file in IGV allowing review of the alignment.

A range of variants can be displayed and examples of each of these are:

1. SNV

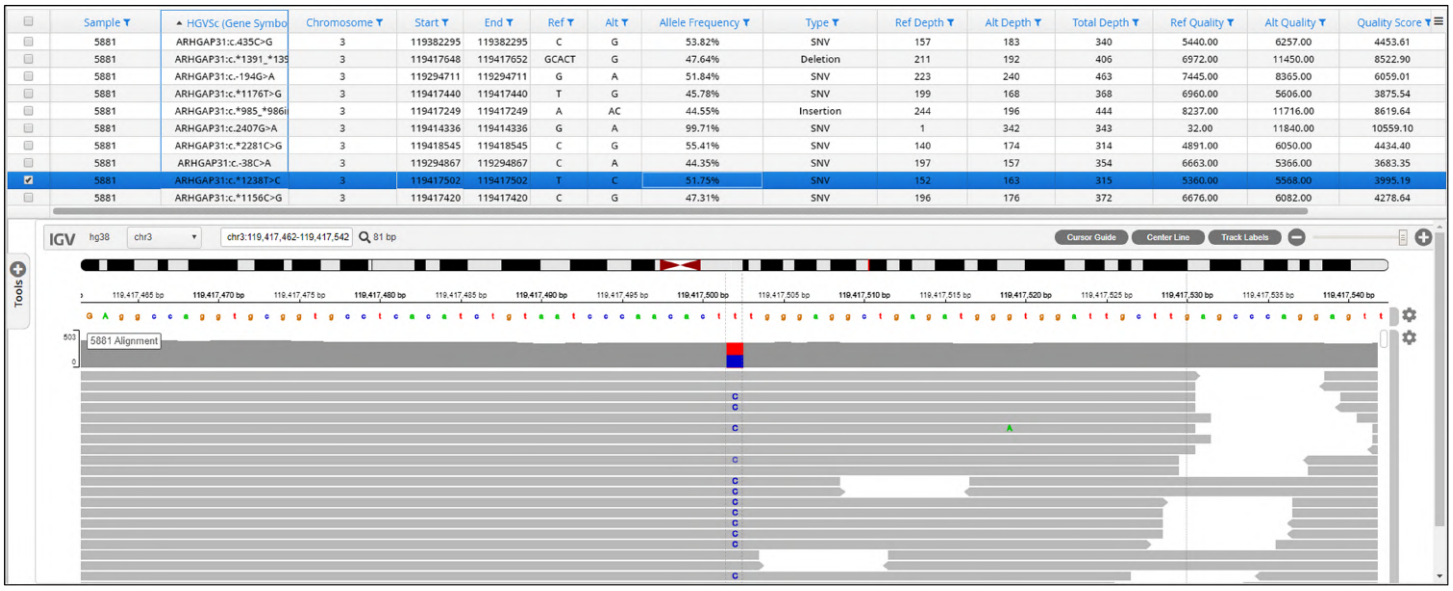


Figure: Example of a SNV being displayed in the IGV browser

2. Deletion

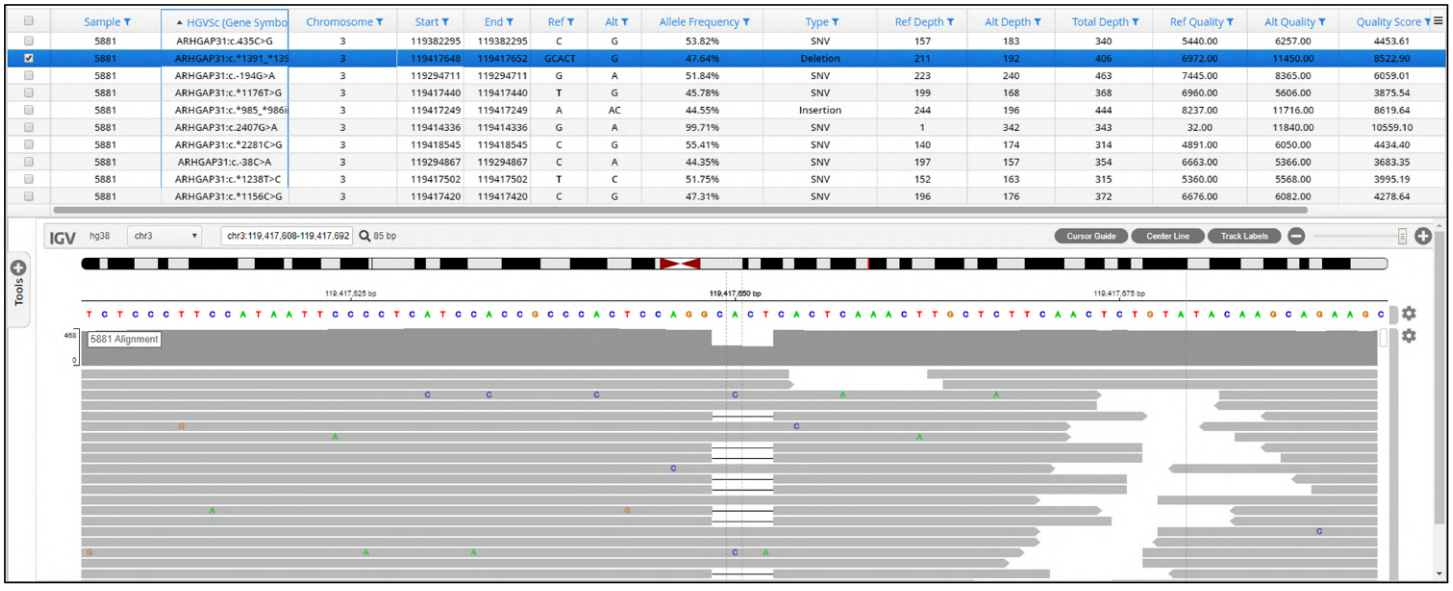


Figure: Example of a deletion being displayed in the IGV browser

3. Insertion

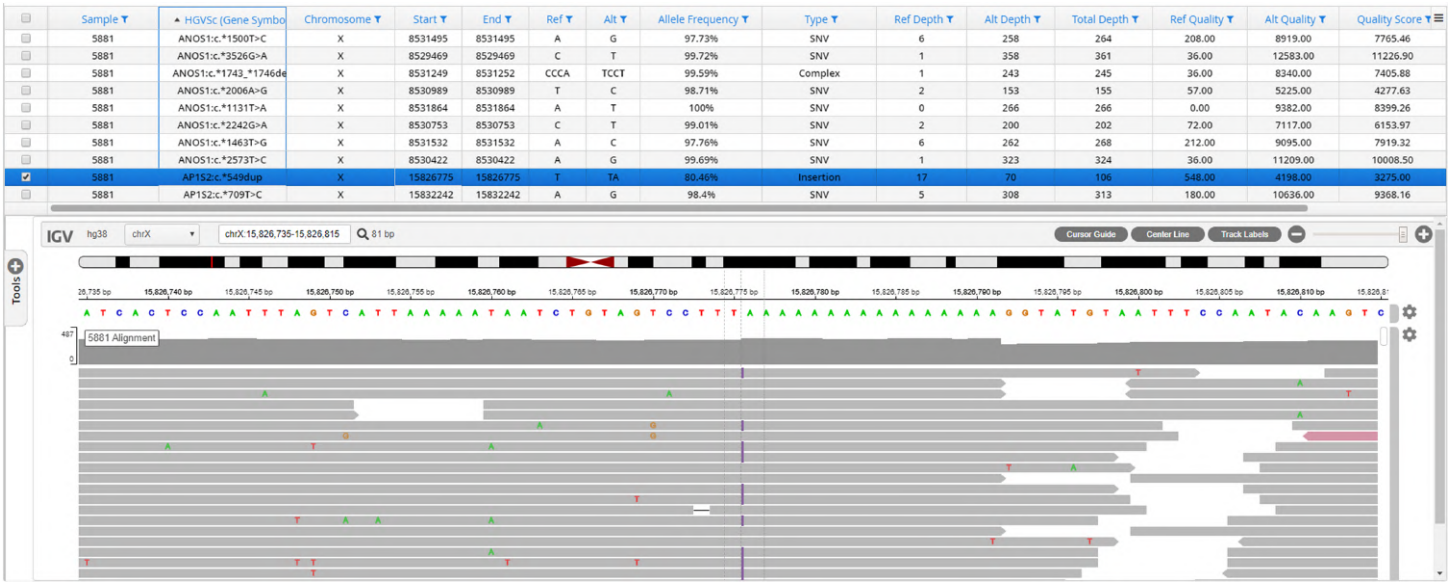


Figure: Example of an insertion being displayed in the IGV browser

4. Complex

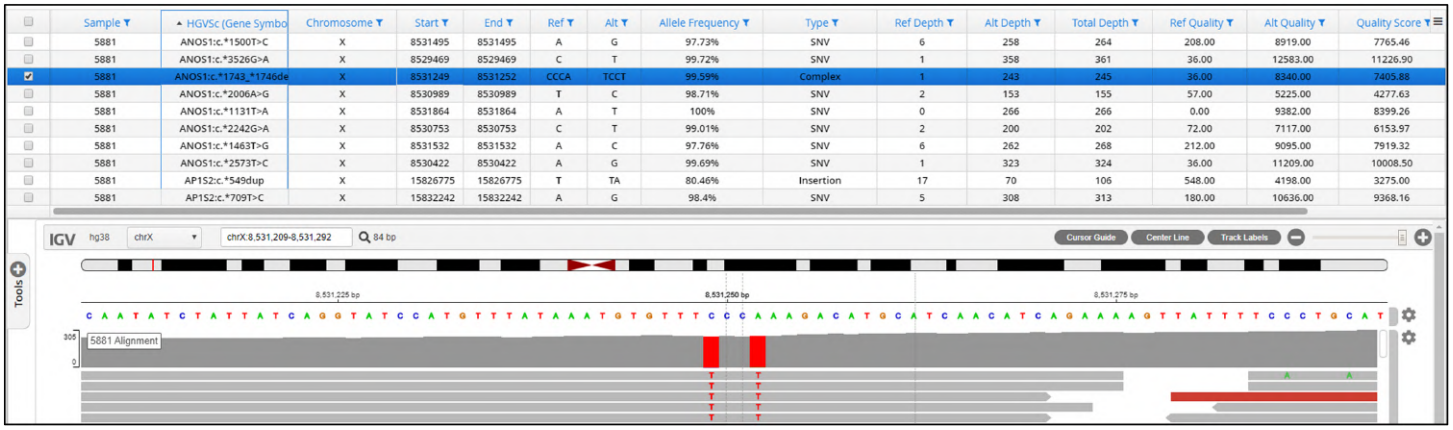


Figure: Example of a complex event being displayed in the IGV browser

5. Multi Nucleotide Polymorphism (MNP)

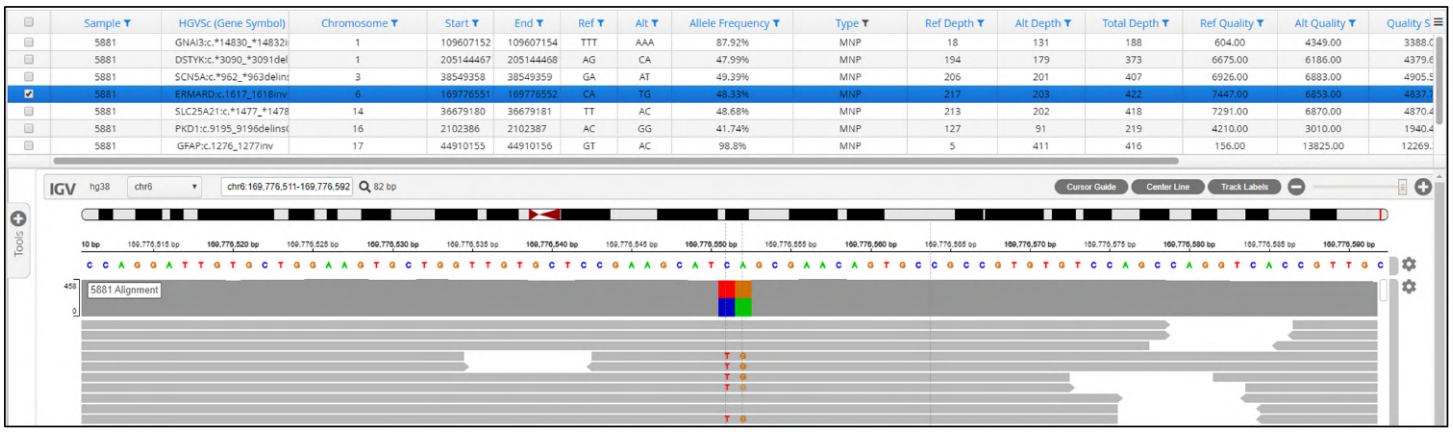


Figure: Example of a MNP variant being displayed in the IGV browser

6. Partial Tanden Duplication (PTD)

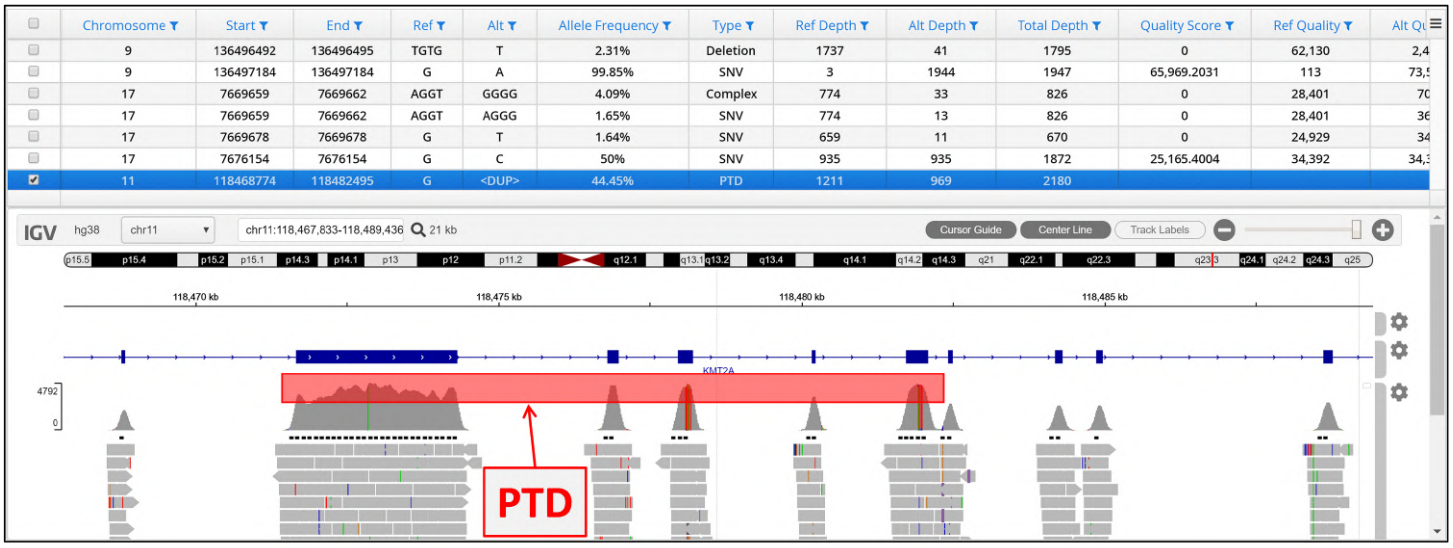


Figure: Example of a PTD being displayed in the IGV browser; the duplication event is highlighted by the transparent red box.

11.2 SNV Options

Right clicking on a row will generate a popup menu with a range of options.

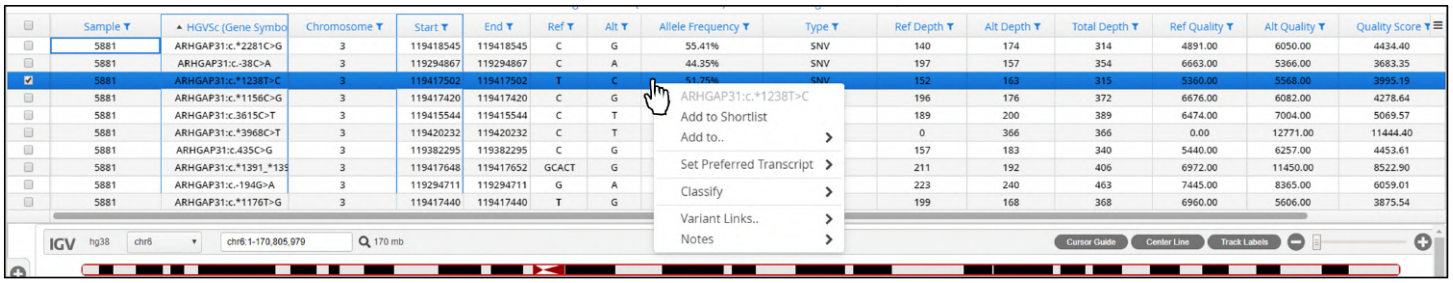


Figure: Options available for each SNV or Indel variant

Add to a Shortlist

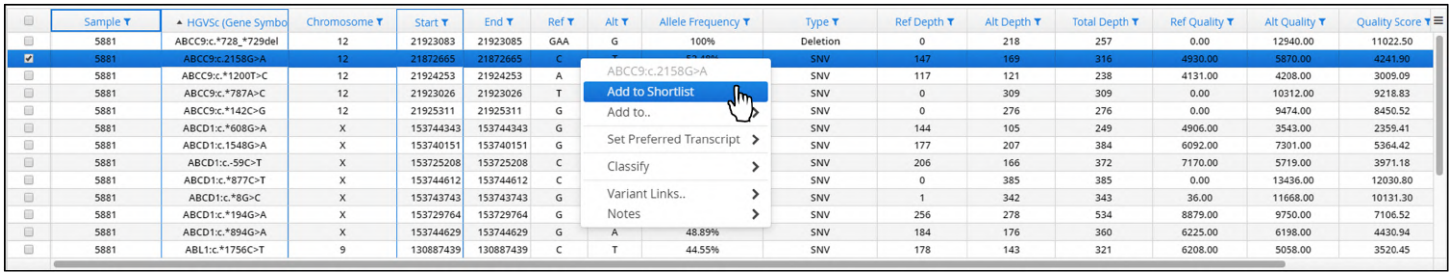
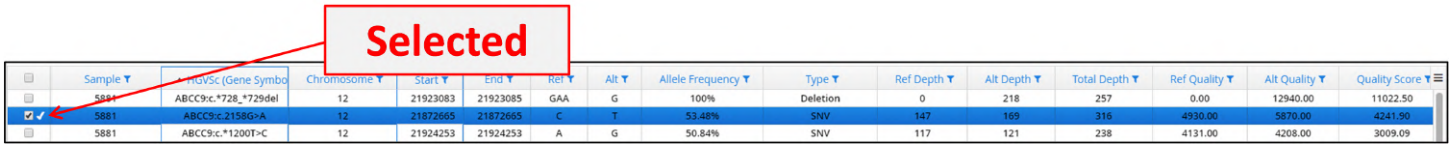


Figure: Adding a variant to shortlist

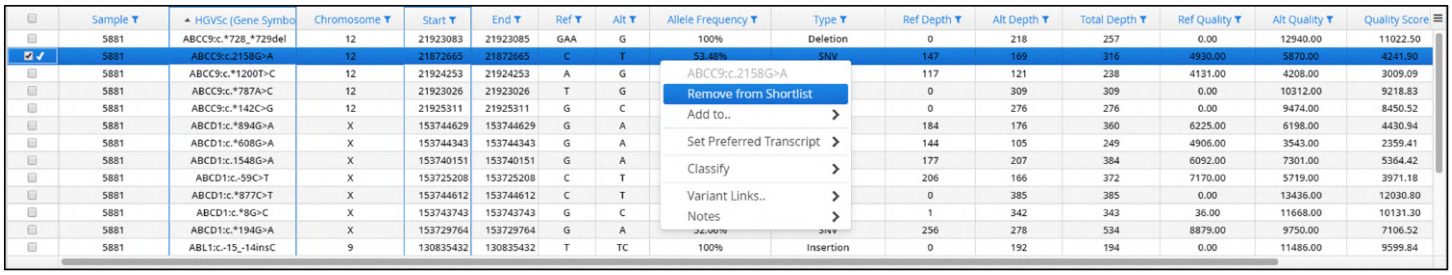
Once a variant has been added to the shortlist it will be annotated with a tick.



Sample	HGVSc (Gene Symbo	Chromosome	Start	End	Ref	Alt	Allele Frequency	Type	Ref Depth	Alt Depth	Total Depth	Ref Quality	Alt Quality	Quality Score
5881	ABCC9:c.*728_*729del	12	21923083	21923085	GAA	G	100%	Deletion	0	218	257	0.00	12940.00	11022.50
5881	ABCC9:c.2158G>A	12	21872665	21872665	C	T	53.48%	SNV	147	169	316	4930.00	5870.00	4241.90
5881	ABCC9:c.*1200T>C	12	21924253	21924253	A	G	50.84%	SNV	117	121	238	4131.00	4208.00	3009.09

Figure: Annotation of a selected variant

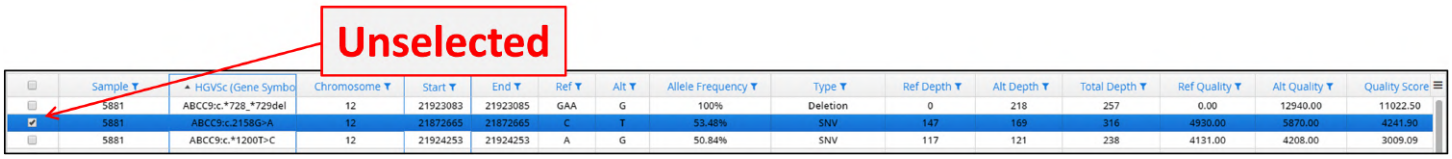
A variant can be removed from the shortlist using the Remove from Shortlist command.



Sample	HGVSc (Gene Symbo	Chromosome	Start	End	Ref	Alt	Allele Frequency	Type	Ref Depth	Alt Depth	Total Depth	Ref Quality	Alt Quality	Quality Score
5881	ABCC9:c.*728_*729del	12	21923083	21923085	GAA	G	100%	Deletion	0	218	257	0.00	12940.00	11022.50
5881	ABCC9:c.2158G>A	12	21872665	21872665	C	T	53.48%	SNV	147	169	316	4930.00	5870.00	4241.90
5881	ABCC9:c.*1200T>C	12	21924253	21924253	A	G			117	121	238	4131.00	4208.00	3009.09
5881	ABCC9:c.*787A>C	12	21923026	21923026	T	G			0	309	309	0.00	10312.00	9218.83
5881	ABCC9:c.*142C>G	12	21925311	21925311	G	C			0	276	276	0.00	9474.00	8450.52
5881	ABCD1:c.*894G>A	X	153744629	153744629	G	A			184	176	360	6225.00	6198.00	4430.94
5881	ABCD1:c.*608G>A	X	153744343	153744343	G	A			144	105	249	4906.00	3543.00	2359.41
5881	ABCD1:c.1548G>A	X	153740151	153740151	G	A			177	207	384	6092.00	7301.00	5364.42
5881	ABCD1:c.-59C>T	X	153725208	153725208	C	T			206	166	372	7170.00	5719.00	3971.18
5881	ABCD1:c.*877C>T	X	153744612	153744612	C	T			0	385	385	0.00	13436.00	12030.80
5881	ABCD1:c.*8G>C	X	153743743	153743743	G	C			1	342	343	36.00	11668.00	10131.30
5881	ABCD1:c.*194G>A	X	153729764	153729764	G	A			256	278	534	8879.00	9750.00	7106.52
5881	ABL1:c.-15_-14insC	9	130835432	130835432	T	TC	100%	Insertion	0	192	194	0.00	11486.00	9599.84

Figure: Removing a variant from a shortlist

Subsequently the tick annotation will be removed.

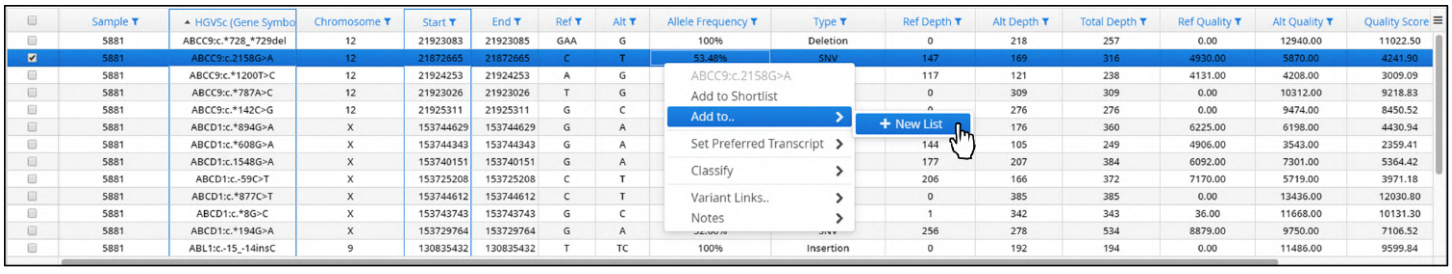


Sample	HGVSc (Gene Symbo	Chromosome	Start	End	Ref	Alt	Allele Frequency	Type	Ref Depth	Alt Depth	Total Depth	Ref Quality	Alt Quality	Quality Score
5881	ABCC9:c.*728_*729del	12	21923083	21923085	GAA	G	100%	Deletion	0	218	257	0.00	12940.00	11022.50
5881	ABCC9:c.2158G>A	12	21872665	21872665	C	T	53.48%	SNV	147	169	316	4930.00	5870.00	4241.90
5881	ABCC9:c.*1200T>C	12	21924253	21924253	A	G	50.84%	SNV	117	121	238	4131.00	4208.00	3009.09

Figure: Annotation denoting shortlisting has been removed

Add to New List

Variants can be added to lists that can be used in software; for example a list of variants can be used in a filter as a means to specifically search for a data set.



Sample	HGVSc (Gene Symbo	Chromosome	Start	End	Ref	Alt	Allele Frequency	Type	Ref Depth	Alt Depth	Total Depth	Ref Quality	Alt Quality	Quality Score
5881	ABCC9:c.*728_*729del	12	21923083	21923085	GAA	G	100%	Deletion	0	218	257	0.00	12940.00	11022.50
5881	ABCC9:c.2158G>A	12	21872665	21872665	C	T	53.48%	SNV	147	169	316	4930.00	5870.00	4241.90
5881	ABCC9:c.*1200T>C	12	21924253	21924253	A	G			117	121	238	4131.00	4208.00	3009.09
5881	ABCC9:c.*787A>C	12	21923026	21923026	T	G			0	309	309	0.00	10312.00	9218.83
5881	ABCC9:c.*142C>G	12	21925311	21925311	G	C			0	276	276	0.00	9474.00	8450.52
5881	ABCD1:c.*894G>A	X	153744629	153744629	G	A			184	176	360	6225.00	6198.00	4430.94
5881	ABCD1:c.*608G>A	X	153744343	153744343	G	A			144	105	249	4906.00	3543.00	2359.41
5881	ABCD1:c.1548G>A	X	153740151	153740151	G	A			177	207	384	6092.00	7301.00	5364.42
5881	ABCD1:c.-59C>T	X	153725208	153725208	C	T			206	166	372	7170.00	5719.00	3971.18
5881	ABCD1:c.*877C>T	X	153744612	153744612	C	T			0	385	385	0.00	13436.00	12030.80
5881	ABCD1:c.*8G>C	X	153743743	153743743	G	C			1	342	343	36.00	11668.00	10131.30
5881	ABCD1:c.*194G>A	X	153729764	153729764	G	A			256	278	534	8879.00	9750.00	7106.52
5881	ABL1:c.-15_-14insC	9	130835432	130835432	T	TC	100%	Insertion	0	192	194	0.00	11486.00	9599.84

Figure: Adding a variant to New List

Initially users will be prompted to create a new variant list by setting the name of the list.

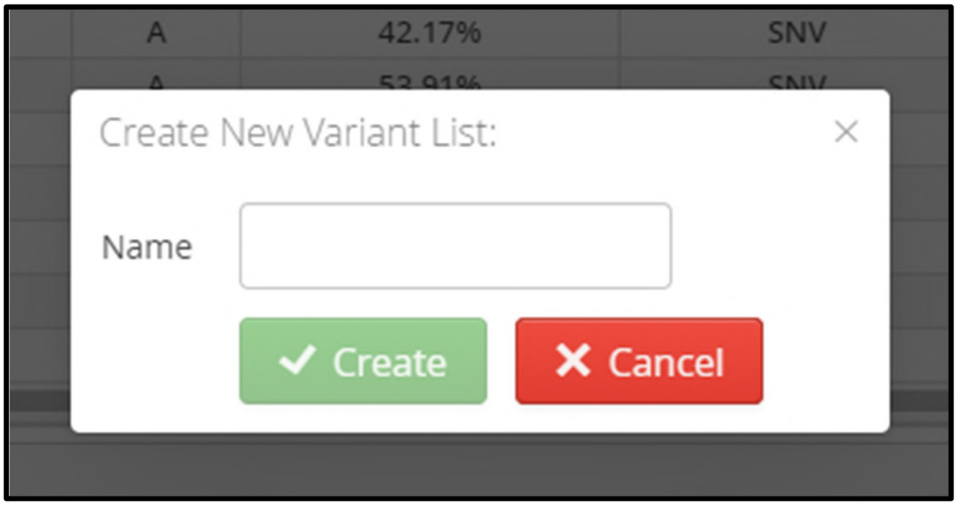


Figure: Creating a new variant list

In the example below a list called New List has been created.

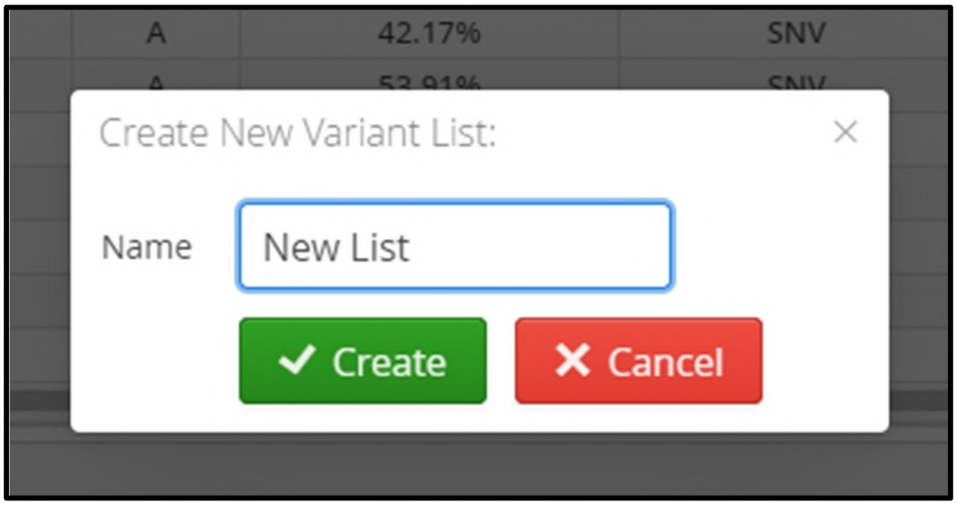


Figure: Setting the name of a new variant list

The list New List is now available and variants can be added to it.

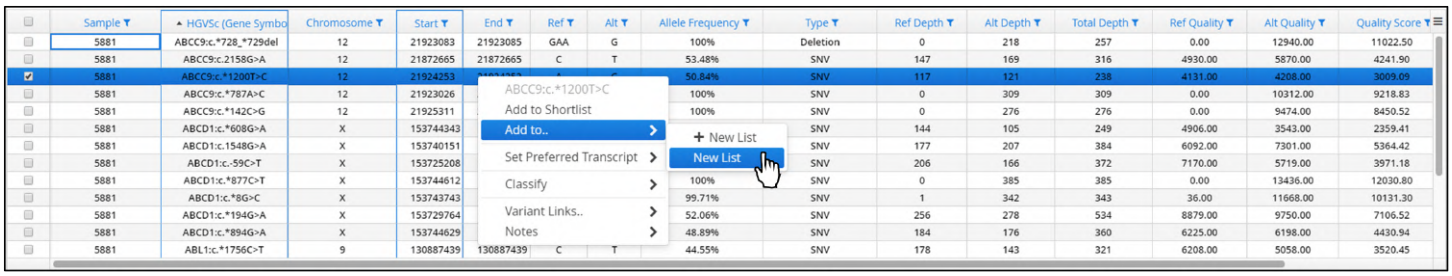


Figure: Adding a variant to the newly created list

Select Transcript

By default the a gene will have the largest canonical transcript set as the preferred transcript.

To change this, users can use the Set Preferred Transcript option to select an alternative transcript from the list available in the database.

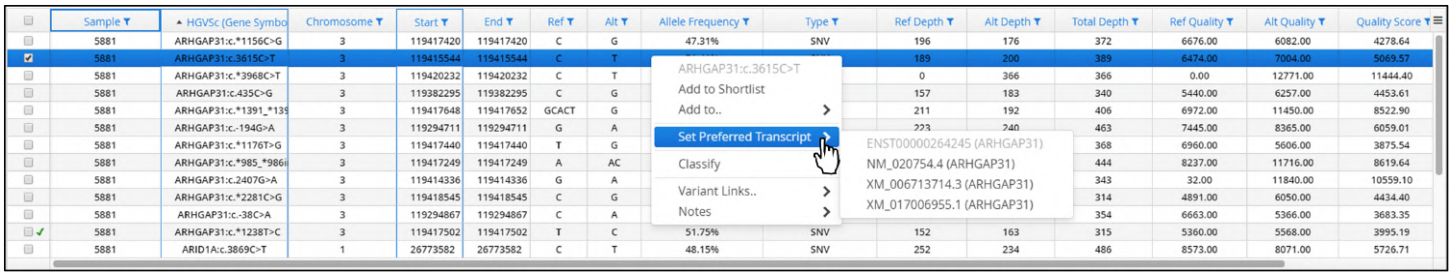


Figure: Setting a preferred transcript for a gene

Variant Classification

Users are able to classify variants in two ways; firstly, a variant can be directly assigned one of the defined classifications.

Additional classifications can be added via the Admin Controls (Admin Controls > Analysis > Classifications).

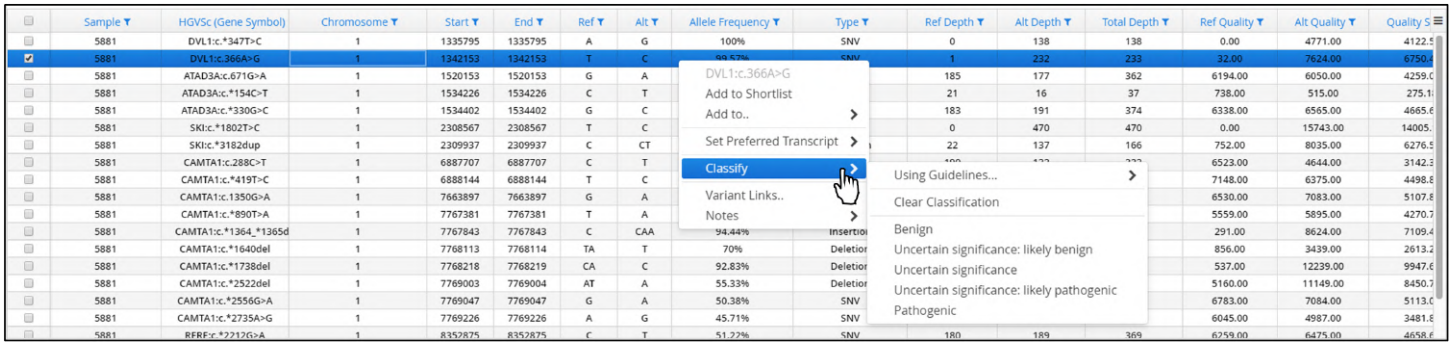


Figure: Variant classification options

A variant classification is selected form the list that is included by default. These are:

- Benign
- Uncertain significance, likely benign
- Uncertain significance
- Uncertain significance, likely pathogenic
- Pathogenic

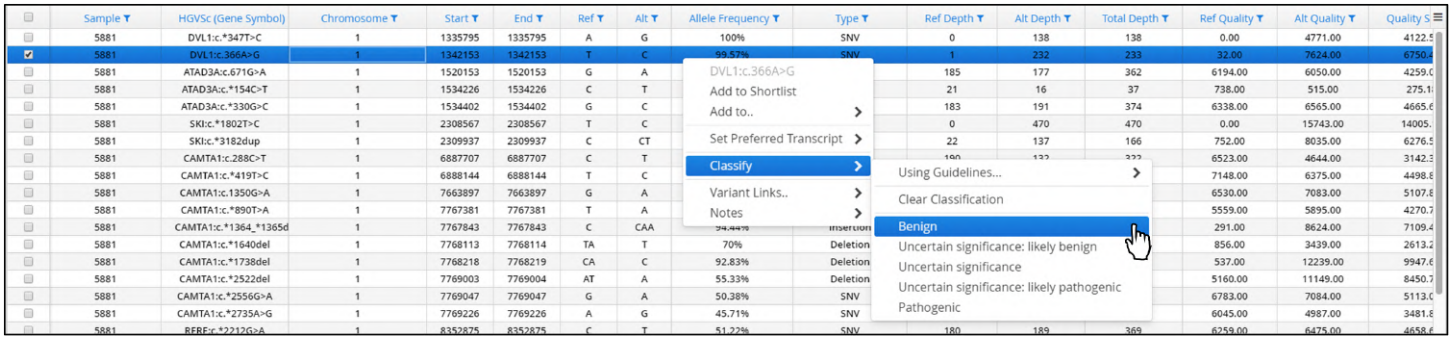
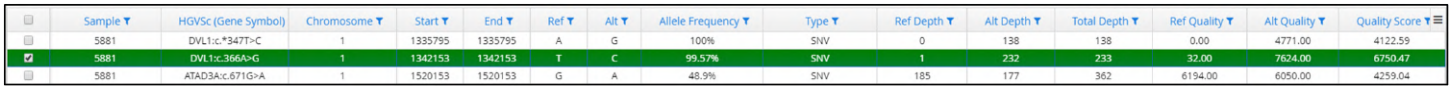


Figure: Annotating a variant as benign

Once this classification has been made the variant will be annotated with the corresponding colour classification. This colour can be changed in the Admin Controls section of the software (Admin Controls > Analysis > Classifications)

This update will be applied to the variant annotation. As a result where the same variant appears in other samples it will have the same colour coding in the table.



Sample	HGVSc (Gene Symbol)	Chromosome	Start	End	Ref	Alt	Allele Frequency	Type	Ref Depth	Alt Depth	Total Depth	Ref Quality	Alt Quality	Quality Score
5881	DVL1:c.*347T>C	1	1335795	1335795	A	G	100%	SNV	0	138	138	0.00	4771.00	4122.59
5881	DVL1:c.366A>G	1	1342153	1342153	T	C	99.57%	SNV	1	232	233	32.00	7624.00	6750.47
5881	ATAD3A:c.671G>A	1	1520153	1520153	G	A	48.9%	SNV	185	177	362	6194.00	6050.00	4259.04

Figure: Update of the annotation to show a variant as benign

A variant classification can be removed using the clear classification method

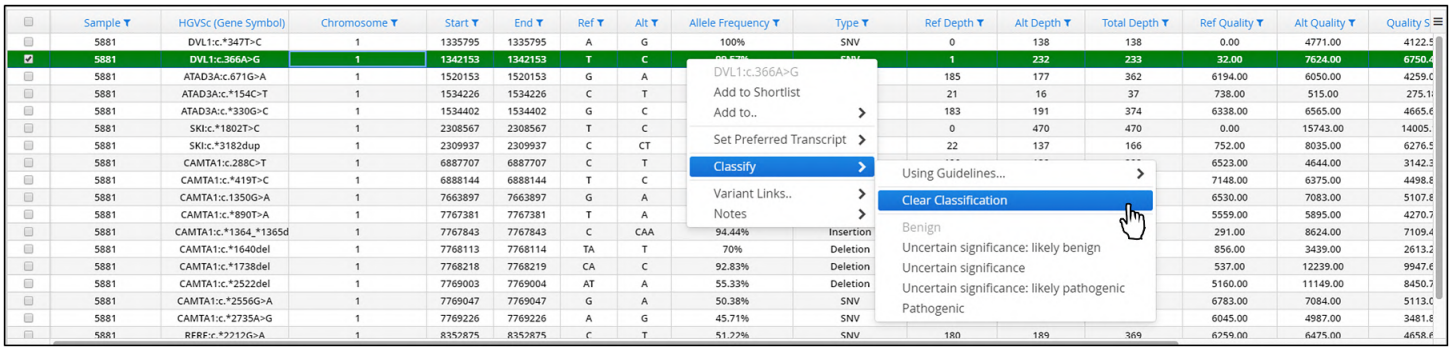
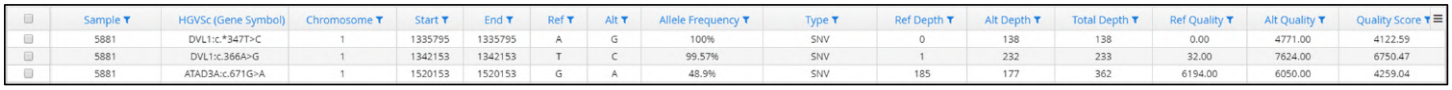


Figure: Removing a variant classification annotation

The classification will be removed for the variant in the table and all other samples with the same variant will be similarly updated.



Sample	HGVSc (Gene Symbol)	Chromosome	Start	End	Ref	Alt	Allele Frequency	Type	Ref Depth	Alt Depth	Total Depth	Ref Quality	Alt Quality	Quality Score
5881	DVL1:c.*347T>C	1	1335795	1335795	A	G	100%	SNV	0	138	138	0.00	4771.00	4122.59
5881	DVL1:c.366A>G	1	1342153	1342153	T	C	99.57%	SNV	1	232	233	32.00	7624.00	6750.47
5881	ATAD3A:c.671G>A	1	1520153	1520153	G	A	48.9%	SNV	185	177	362	6194.00	6050.00	4259.04

Figure: Update of the annotation to show variant without any classification

Using the American College of Medical Genetics and Genomics (ACMG) Guidelines

An alternative means to derive a classification for a variant is via guidelines described by the ACMG. These guidelines are included with Interpret.

To follow the ACMG guidelines the user provides answers to a specific set of questions. Each answer will navigate the user through the conditions of the guidelines until a classification of the variant can be made.

Selecting to use the guidelines option leads to a new window opening.

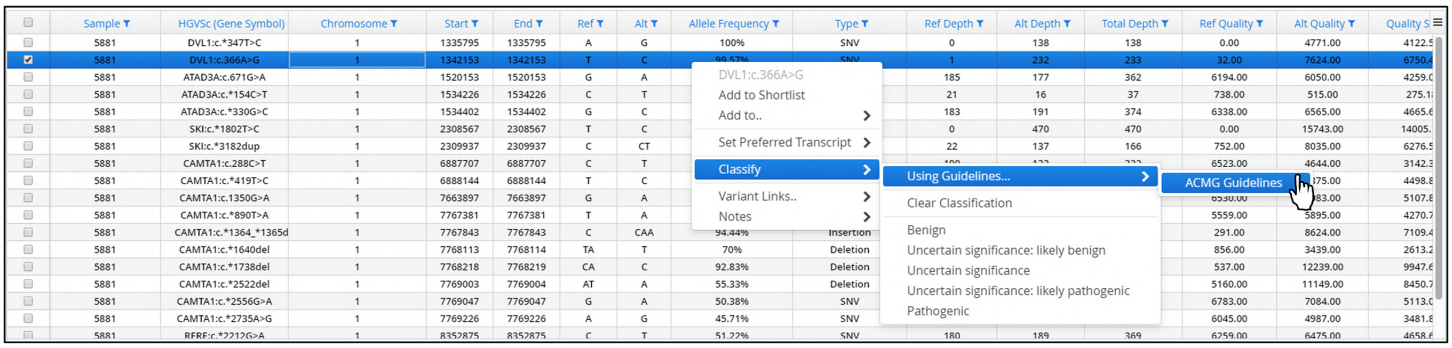


Figure: Selection of the inbuilt ACMG classification guidelines

The initial ACMG window, shown below, consists of a progress bar that will report how close to a classification

- Progress Bar - Showing the progress of the classification.
- Questions - These are the questions to be answered.
- Toggle - Allowing the display to be a graph view or a table view.

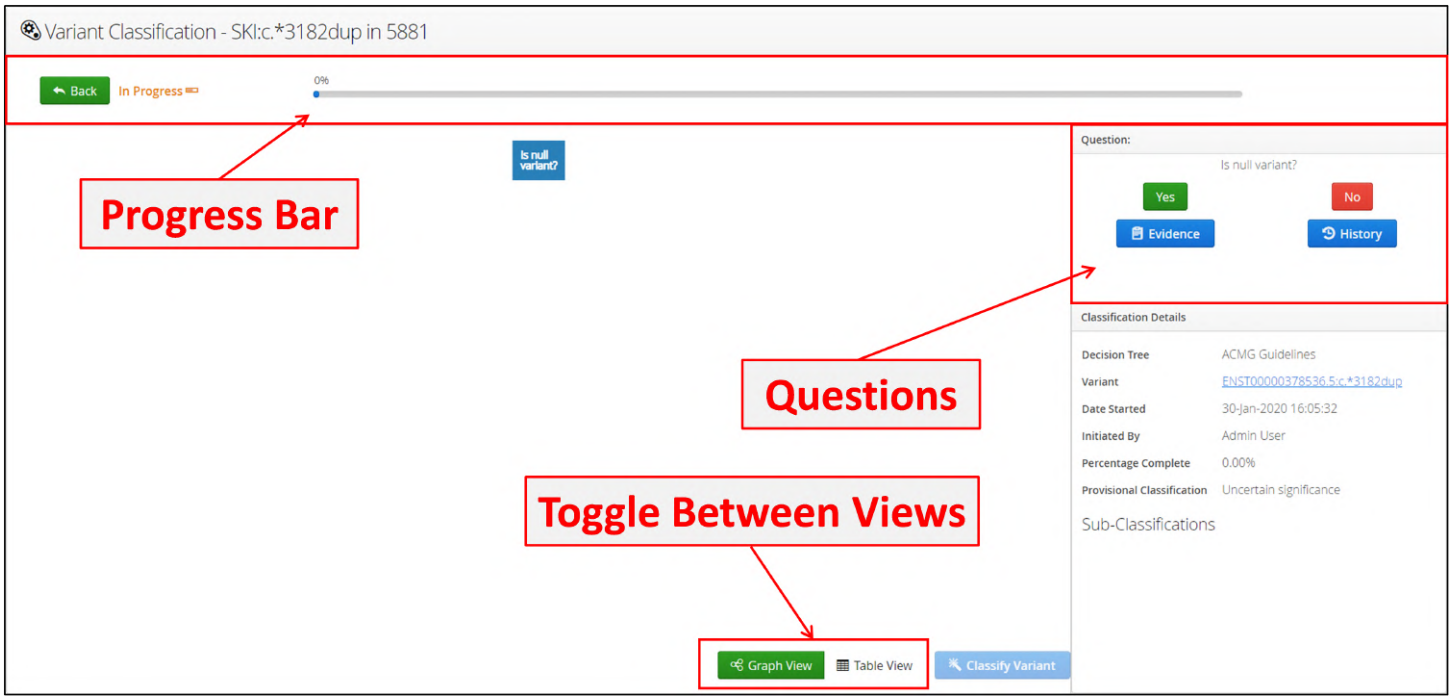


Figure: The initial ACMG classification window

As the user answers questions the progress bar will update.

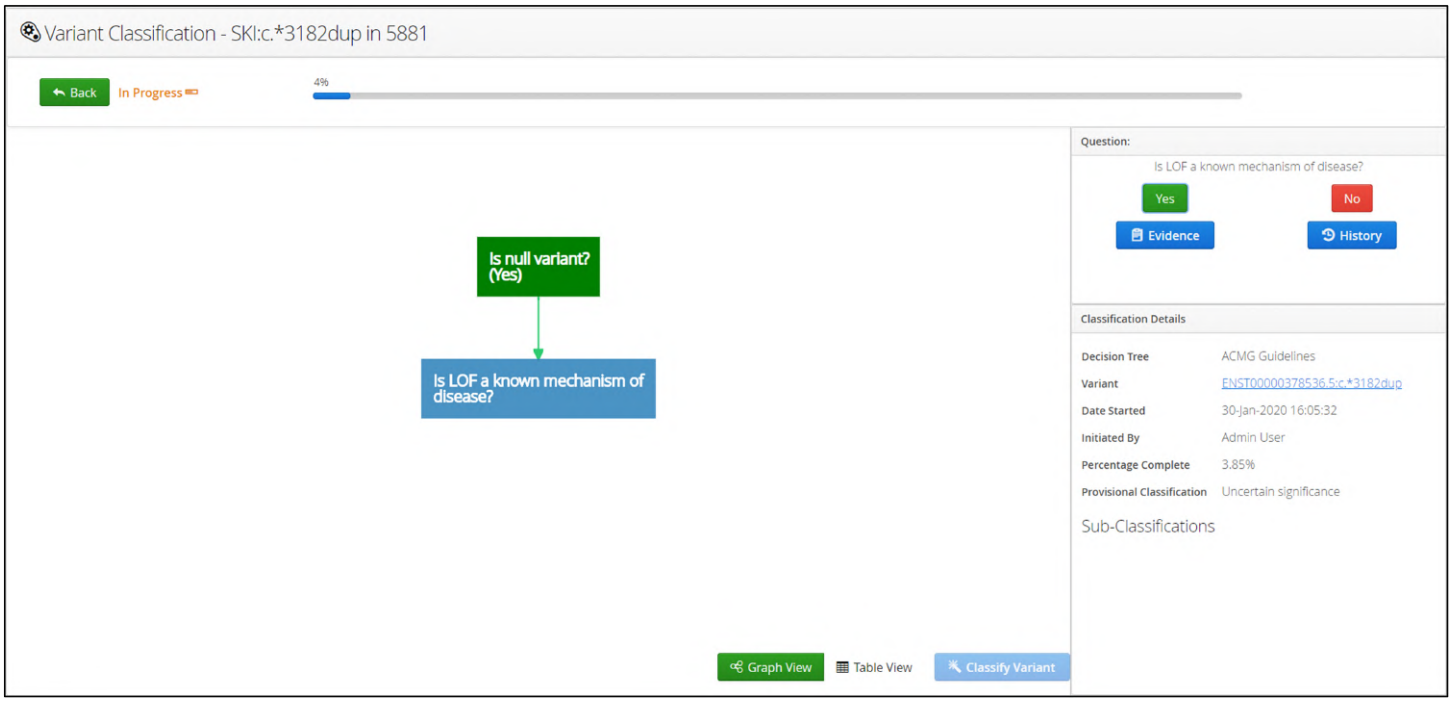


Figure: A classification at the start of the guidelines, with 4% progress

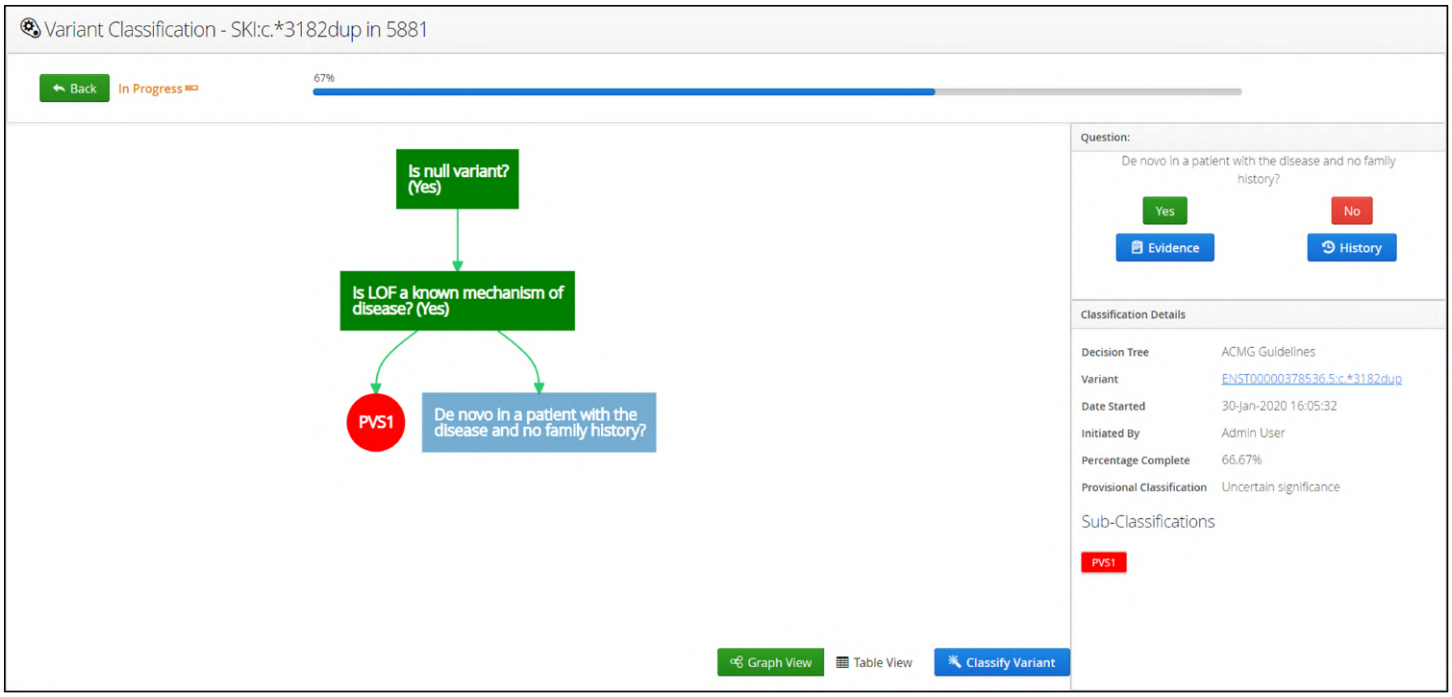


Figure: A classification with 67% progress

When sufficient questions have been answered to allow a classification the progress bar will update to show 100% and say Ready to be classified.

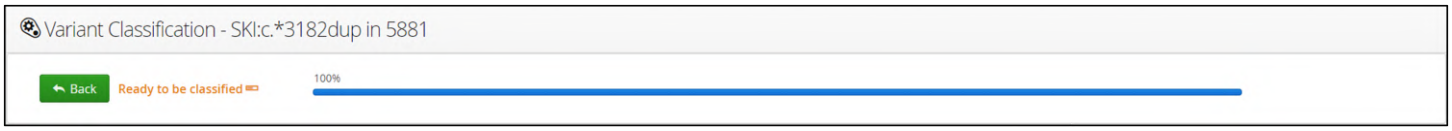


Figure: A completed ACMG classification

A window will appear showing the classification and give the user the option of making the classification or cancelling.

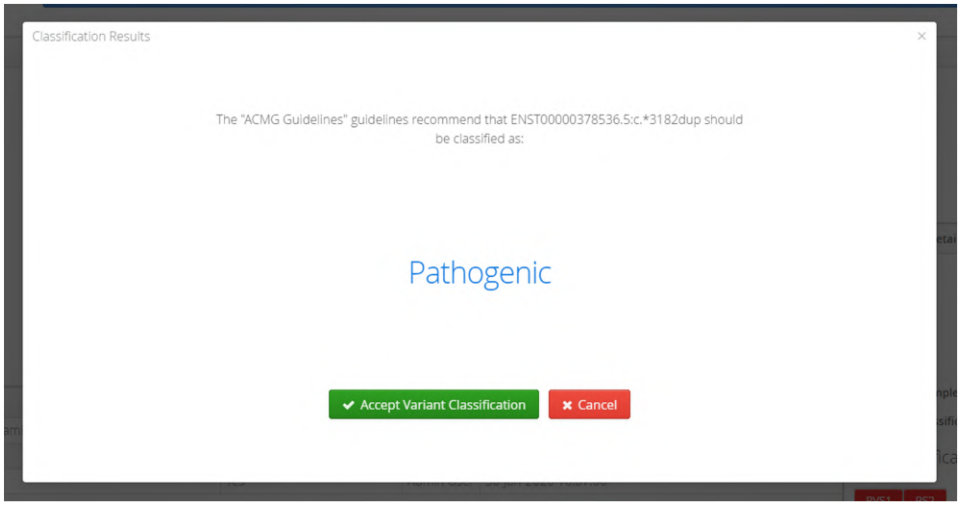
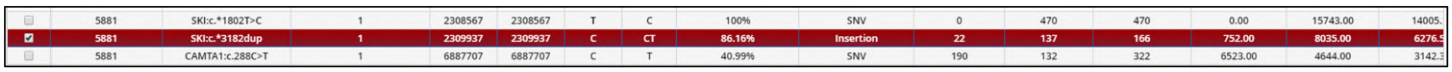


Figure: A completed classification showing a Pathogenic all has been made

Selecting the classification will update the variant's annotation accordingly.



	5881	SKI:c.*1802T>C	1	2308567	2308567	T	C	100%	SNV	0	470	470	0.00	15743.00	14005.
☑	5881	SKI:c.*3182dup	1	2309937	2309937	C	CT	86.16%	Insertion	22	137	166	752.00	8035.00	6276.
	5881	CAMTA1:c.288C>T	1	6887707	6887707	C	T	40.99%	SNV	190	132	322	6523.00	4644.00	3142.

Figure: Updated annotation for the variant to show its status as pathogenic

It is possible to review the choices made in the guidelines; using the table view, users can see which questions were asked and how they were answered by whom and when.

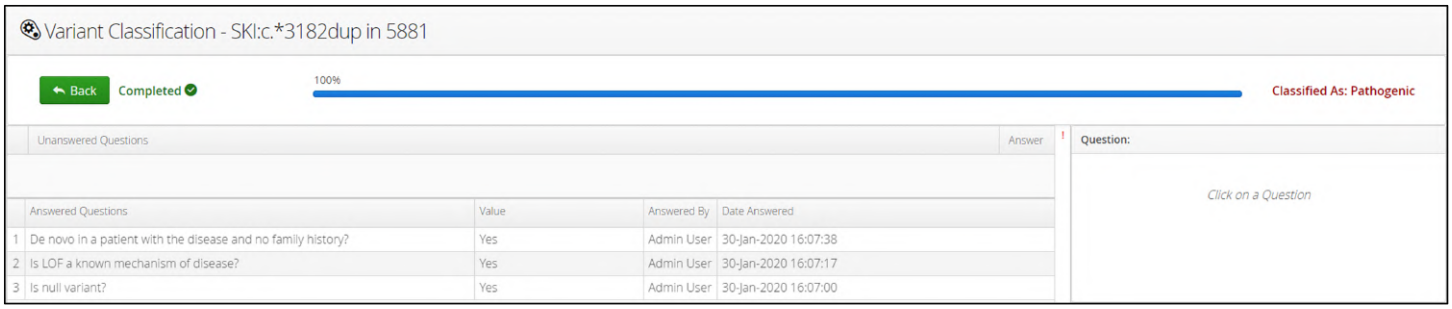


Figure: Table view of a completed classification

Selecting a row from the table view allows a result to be modified if that is required.

Alternatively, evidence can be added to support the answer to the question

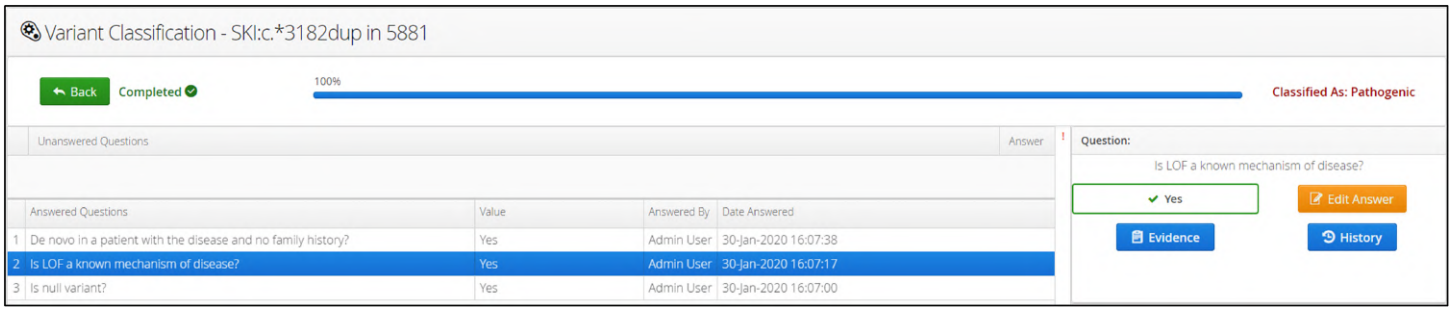


Figure: Reviewing an answer in the table view

Variant Links

The software allows users to link out to external sources of documentation. Currently included are:

- EnsEMBL
- ClinView
- ExAC

Additional resources can be added in the Admin Controls (Admin Controls > Analysis > Manage Links).

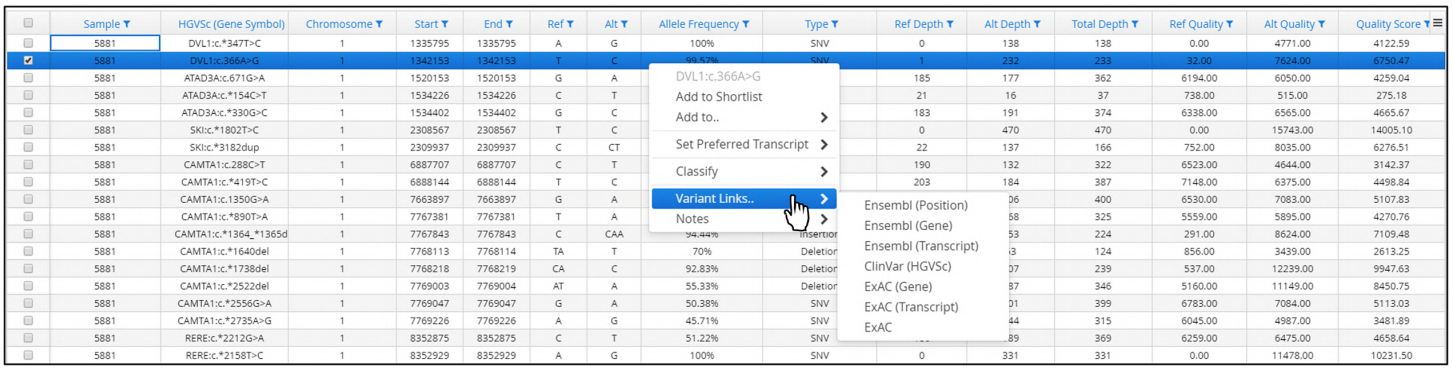


Figure: Variant inks available in the software

If a source is selected Interpret will show the information in a separate tab in the web browser.

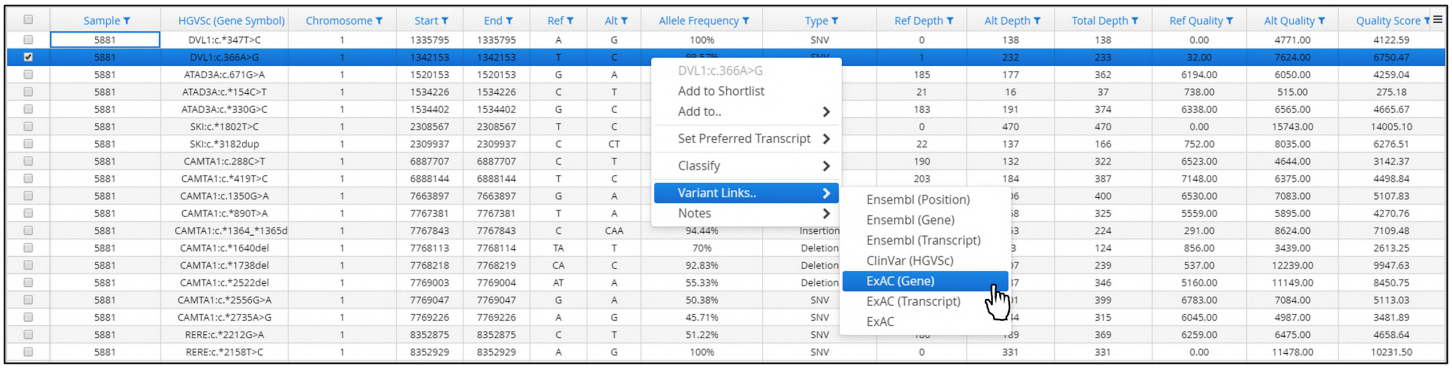


Figure: Selection of ExAC as an external resource for the gene selected

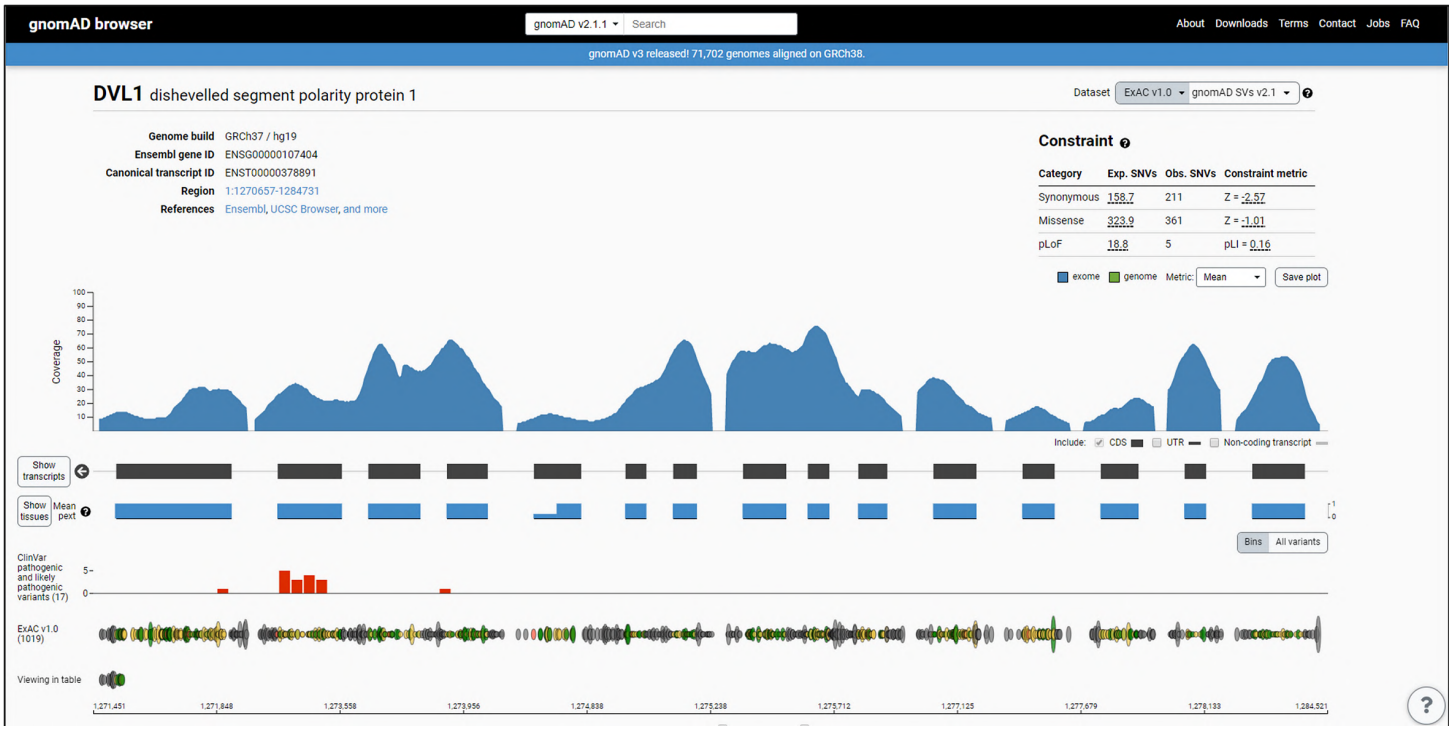


Figure: Example of the software linking out to an external data source, in this case the GnomAD for the gene containing the variant in Interpret

Add Notes

Interpret allows users to add notes for a variant and to also edit notes on the system. This is accessed through the Notes menu item.

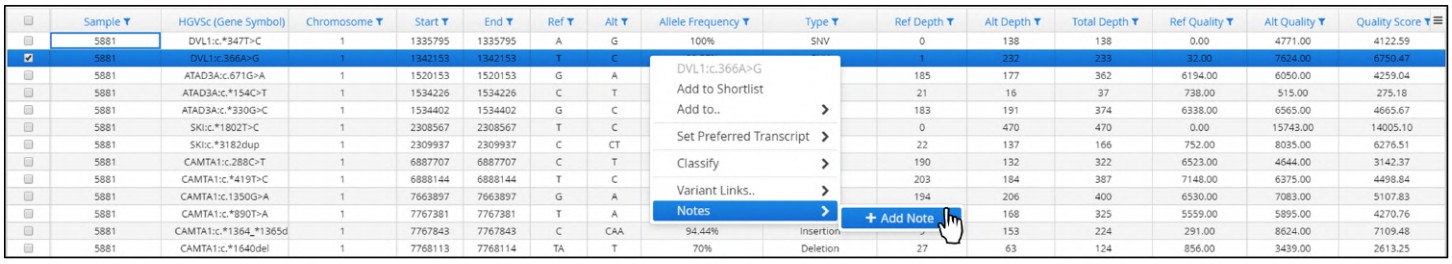


Figure: Adding a note to a variant

Selecting the Add Note will generate a popup window

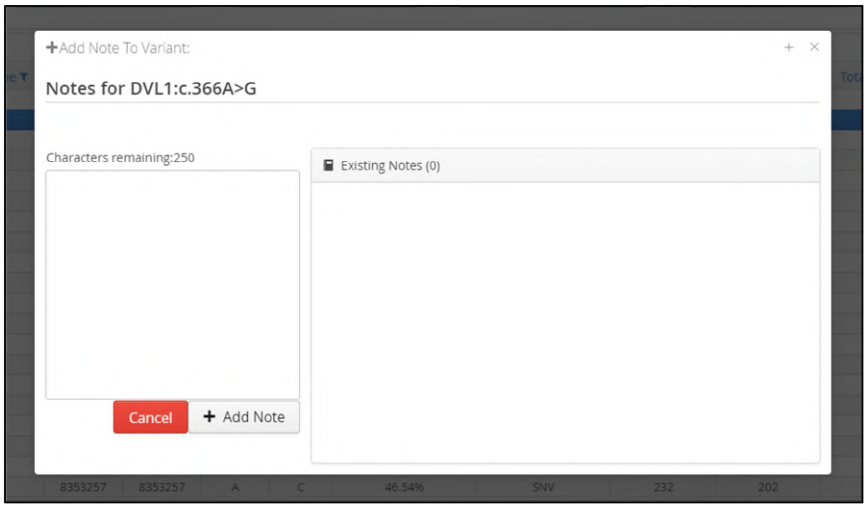


Figure: The note template window for the selected variant

Users can add the required text, up to 250 characters, in the text box

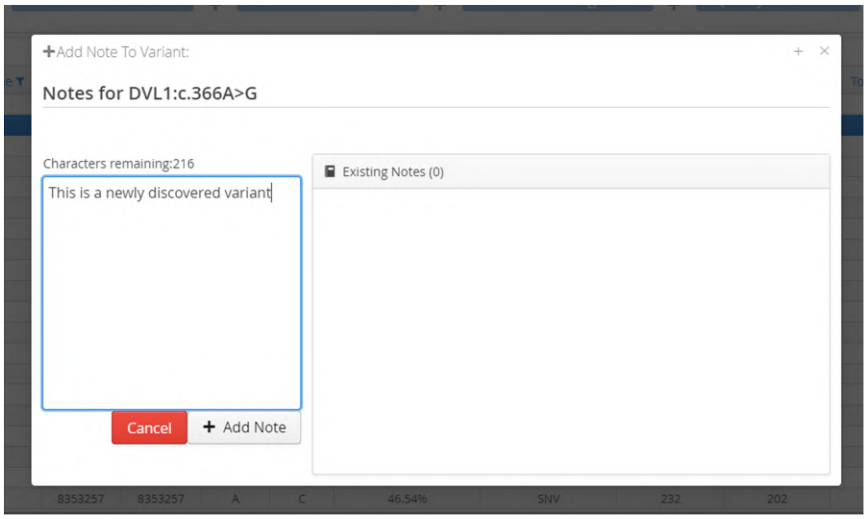


Figure: Addition of a note to a variant

Selecting Add Now will append the note to the variant.

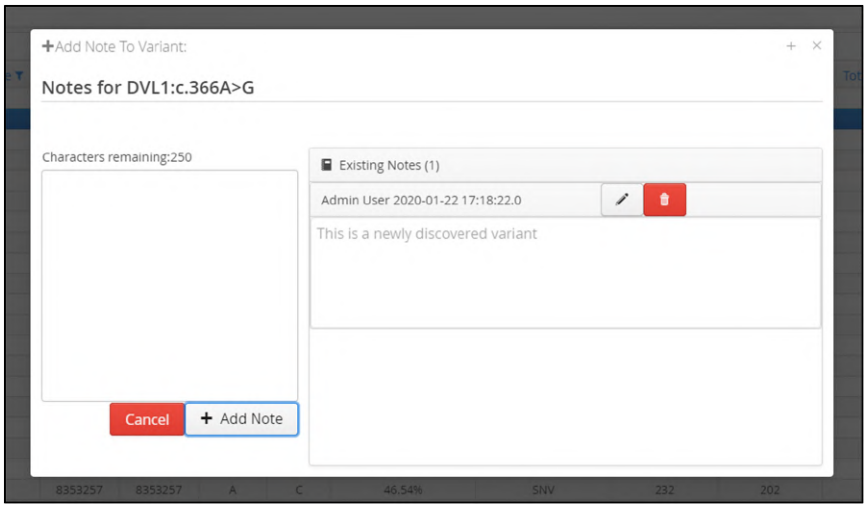


Figure: An example of a note on the system

The additional text will now be displayed

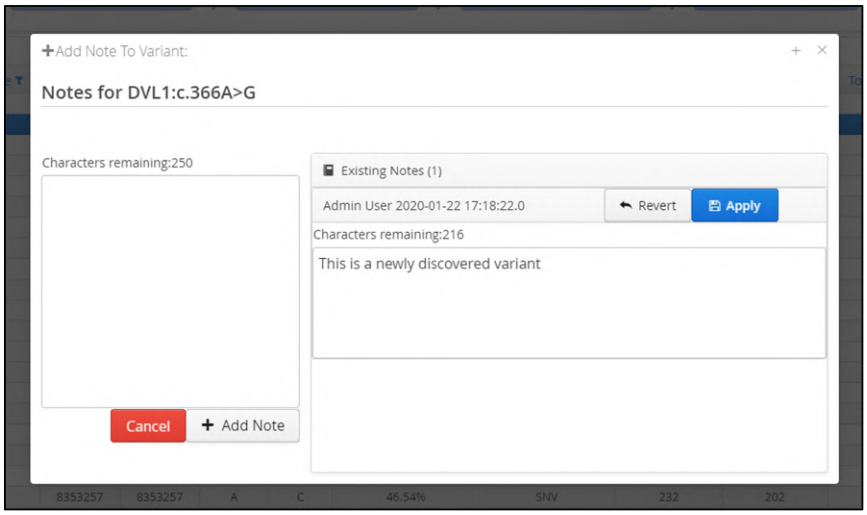


Figure: Appending text to an existing notation

Notes can be modified by clicking on the pen icon. This makes the text box editable

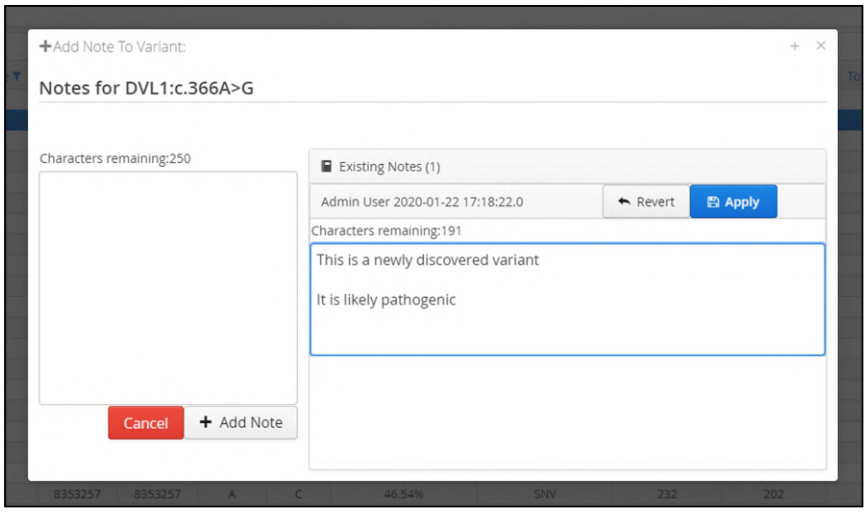


Figure: Adding an update to a note

Once any update has been made, selecting Apply will incorporate the changes.

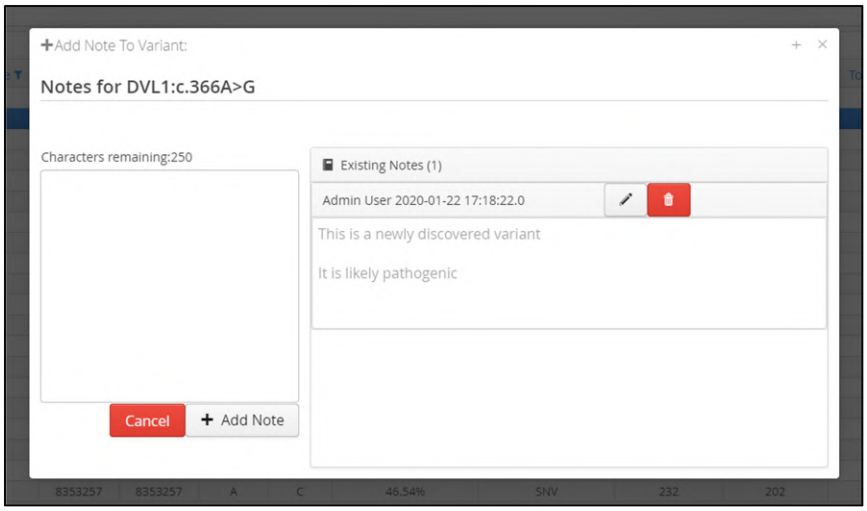


Figure: A note showing the updated annotation

Similarly, a note can be deleted through the red bin icon.

Users are asked to confirm the delete request after which the note will be removed.

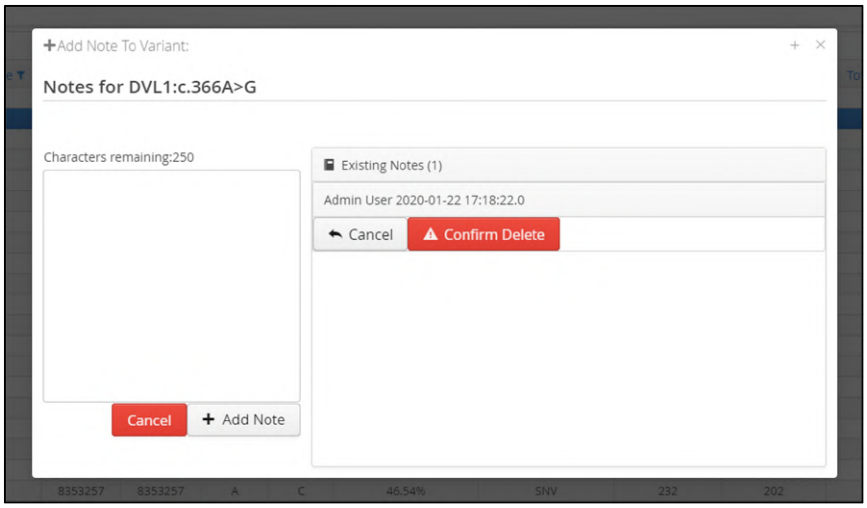


Figure: Deleting a variant note

Where there is a note for a variant the note can be viewed through the Notes options seen when right clicking on the variant.

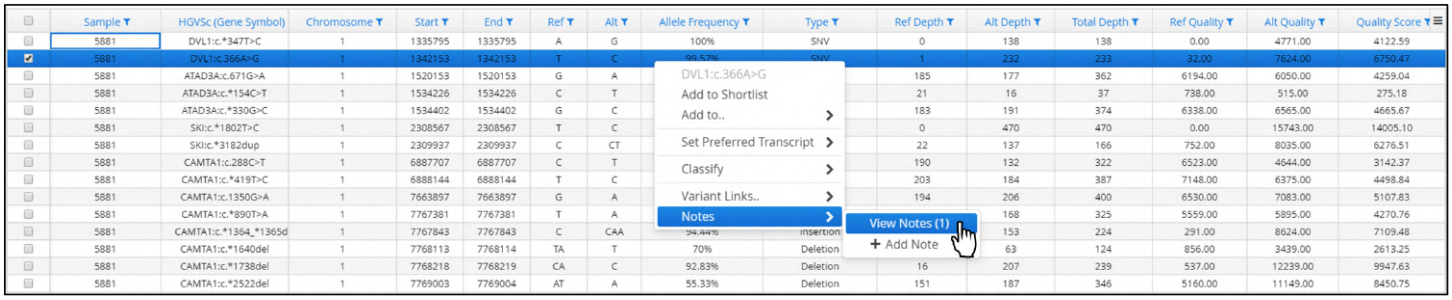


Figure: Selecting a Note to view

The note is displayed on the screen.

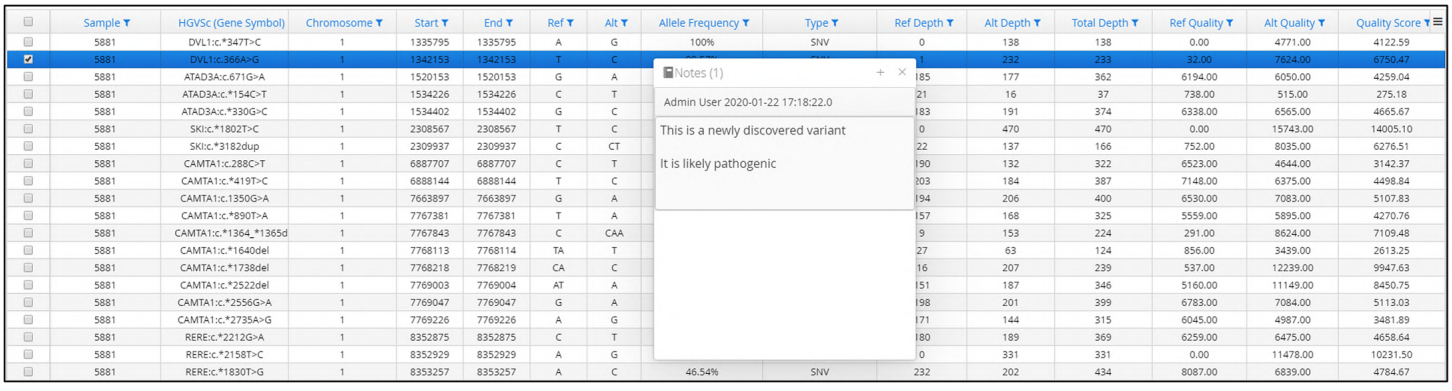


Figure: Viewing of a note that has been added to a variant

11.3 Viewing CNV and LOH Events

The variant table has a column selector icon allowing user to configure which columns are displayed. The figure below shows the columns available for display.

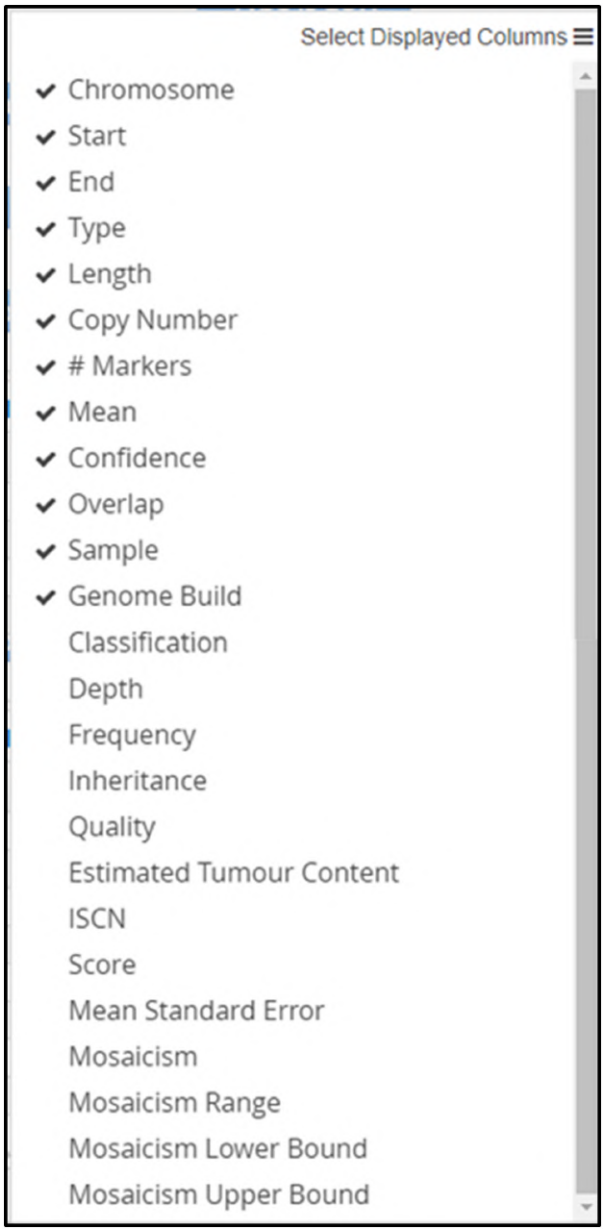


Figure: Columns available to select for display in the CNV/LOH variants page

Selecting a variant will show it in IGV, a track for both CNV and LOH will displayed.

Sometimes there will only be a CNV call as in the example below.

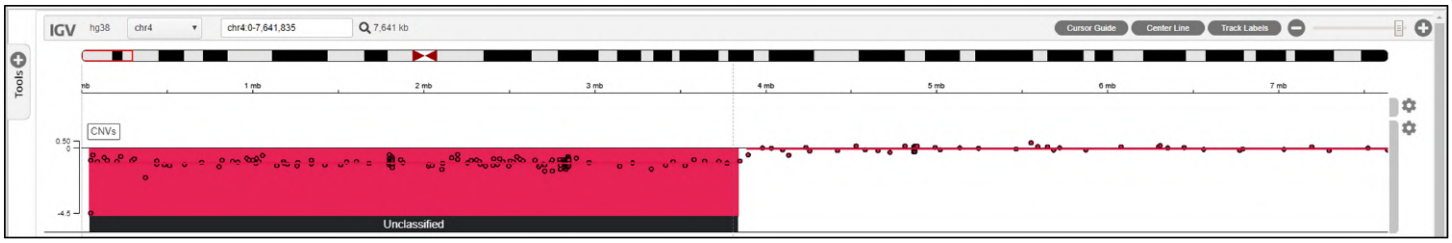
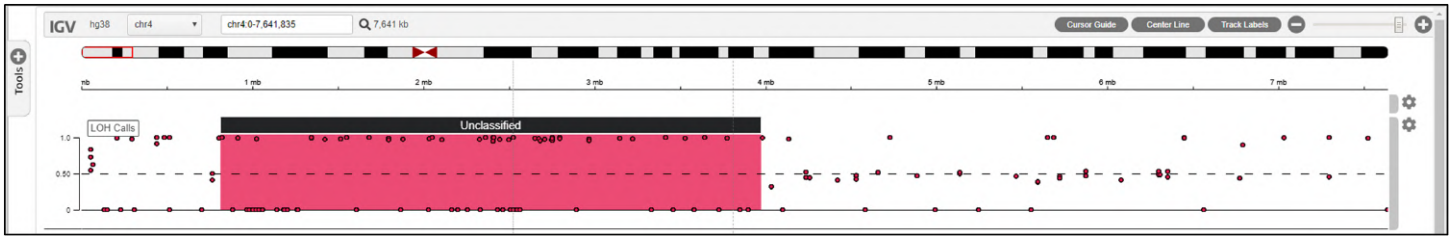


Figure: Example of a CNV call only

Sometimes there will only be a LOH call



Figur e: Example of a LOH call only

Sometimes there will be CNV and LOH calls

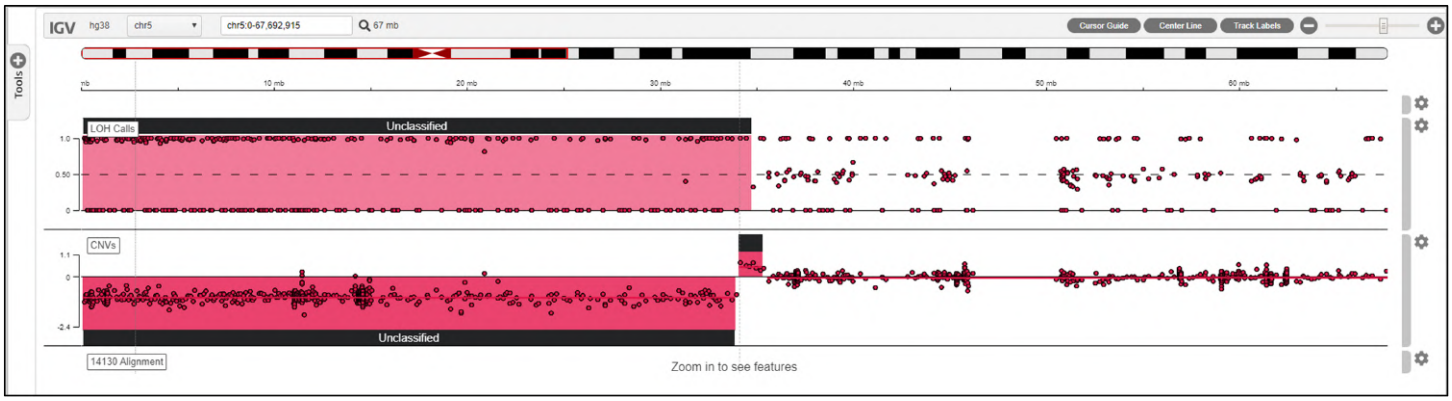


Figure: Example of a sample with a CNV call and a LOH call in the same genomic location

CNV and LOH Options

As with the page displaying SNV and Indel calls there are options available for each variant called by the CNV/LOH pipeline,

Right clicking on a variant will provide a menu of the possible options.

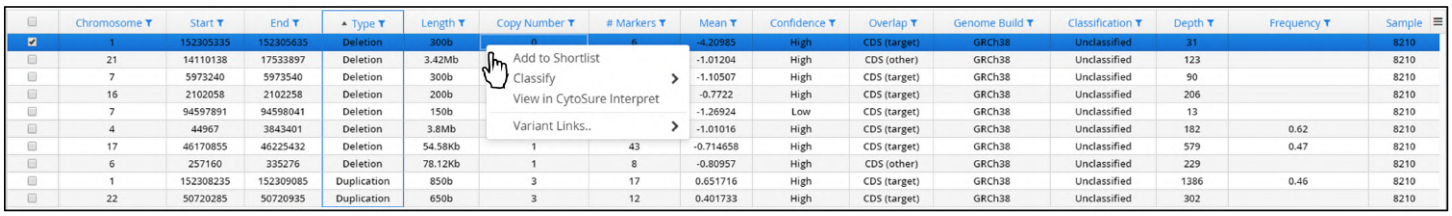
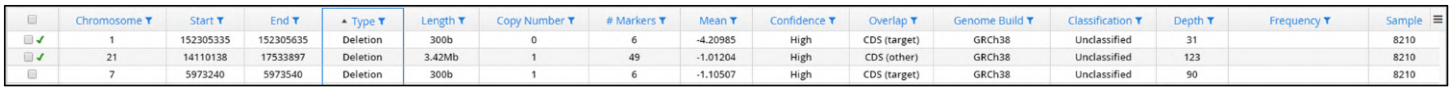


Figure: Available options for a selected CNV or LOH call

Adding to a shortlist

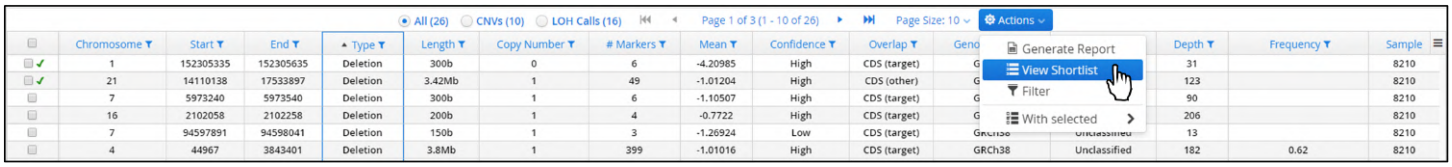
Variants added to a shortlist are annotated with a tick



Chromosome	Start	End	Type	Length	Copy Number	# Markers	Mean	Confidence	Overlap	Genome Build	Classification	Depth	Frequency	Sample
1	152305335	152305635	Deletion	300b	0	6	-4.20985	High	CDS (target)	GRCh38	Unclassified	31		8210
21	14110138	17533897	Deletion	3.42Mb	1	49	-1.01204	High	CDS (other)	GRCh38	Unclassified	123		8210
7	5973240	5973540	Deletion	300b	1	6	-1.10507	High	CDS (target)	GRCh38	Unclassified	90		8210

Figure: Variants added to the shortlist displayed

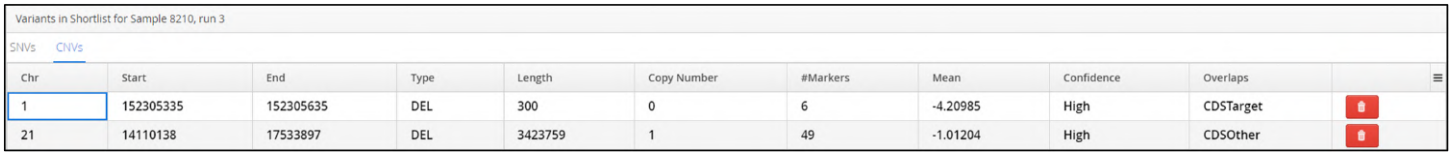
Shortlisted variants can be viewed.



Chromosome	Start	End	Type	Length	Copy Number	# Markers	Mean	Confidence	Overlap	Genome Build	Classification	Depth	Frequency	Sample
1	152305335	152305635	Deletion	300b	0	6	-4.20985	High	CDS (target)	GRCh38	Unclassified	31		8210
21	14110138	17533897	Deletion	3.42Mb	1	49	-1.01204	High	CDS (other)	GRCh38	Unclassified	123		8210
7	5973240	5973540	Deletion	300b	1	6	-1.10507	High	CDS (target)	GRCh38	Unclassified	90		8210
16	2102058	2102258	Deletion	200b	1	4	-0.7722	High	CDS (target)	GRCh38	Unclassified	206		8210
7	94597891	94598041	Deletion	150b	1	3	-1.26924	Low	CDS (target)	GRCh38	Unclassified	13		8210
4	44967	3843401	Deletion	3.8Mb	1	399	-1.01016	High	CDS (target)	GRCh38	Unclassified	182	0.62	8210

Figure: Accessing the shortlist of selected variants

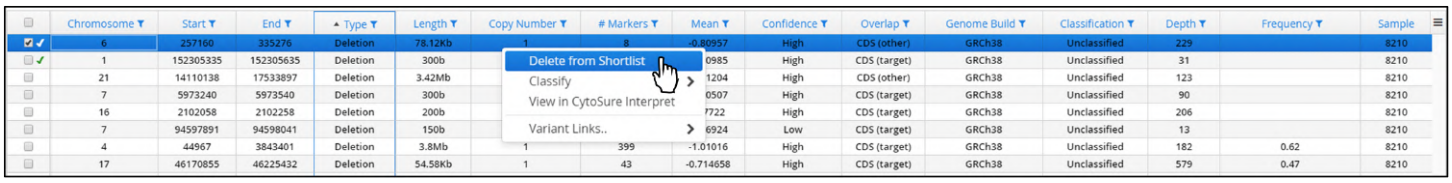
The shortlist opens in a separate view. A variant can be removed from the shortlist be clicking on the red bin icon in the shortlist view.



Chr	Start	End	Type	Length	Copy Number	#Markers	Mean	Confidence	Overlaps
1	152305335	152305635	DEL	300	0	6	-4.20985	High	CDSTarget
21	14110138	17533897	DEL	3423759	1	49	-1.01204	High	CDSOther

Figure: Viewing the shortlist of CNV or LOH variants

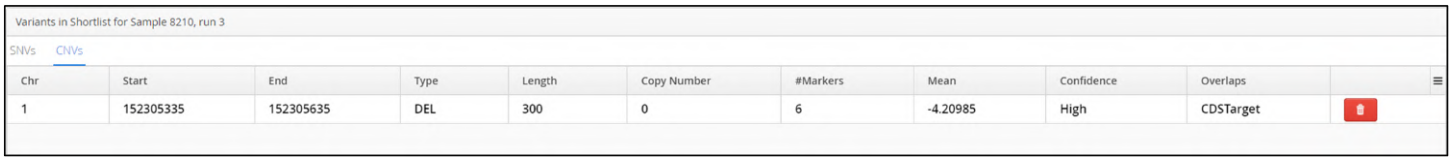
Alternatively, a variant can be removed from the shortlist using the CNV options menu.



Chromosome	Start	End	Type	Length	Copy Number	# Markers	Mean	Confidence	Overlap	Genome Build	Classification	Depth	Frequency	Sample
6	257160	335276	Deletion	78.12Kb	1	8	-0.80957	High	CDS (other)	GRCh38	Unclassified	229		8210
1	152305335	152305635	Deletion	300b			0985	High	CDS (target)	GRCh38	Unclassified	31		8210
21	14110138	17533897	Deletion	3.42Mb			1204	High	CDS (other)	GRCh38	Unclassified	123		8210
7	5973240	5973540	Deletion	300b			0507	High	CDS (target)	GRCh38	Unclassified	90		8210
16	2102058	2102258	Deletion	200b			7722	High	CDS (target)	GRCh38	Unclassified	206		8210
7	94597891	94598041	Deletion	150b			6924	Low	CDS (target)	GRCh38	Unclassified	13		8210
4	44967	3843401	Deletion	3.8Mb	1	399	-1.01016	High	CDS (target)	GRCh38	Unclassified	182	0.62	8210
17	46170855	46225432	Deletion	54.58Kb	1	43	-0.714658	High	CDS (target)	GRCh38	Unclassified	579	0.47	8210

Figure: Deleting a variant from the shortlist

The shortlist will be updated to reflect the removal of a variant.



Chr	Start	End	Type	Length	Copy Number	#Markers	Mean	Confidence	Overlaps
1	152305335	152305635	DEL	300	0	6	-4.20985	High	CDSTarget

Figure: The shortlist showing that the variant has been removed

Variant Classification

A variant can be classified from the list that is included by default. These are:

- Benign
- Uncertain significance, likely benign
- Uncertain significance
- Uncertain significance, likely pathogenic
- Pathogenic

Additional classifications can be added in the Admin Controls section of the software (Admin Controls > Analysis > Classifications)

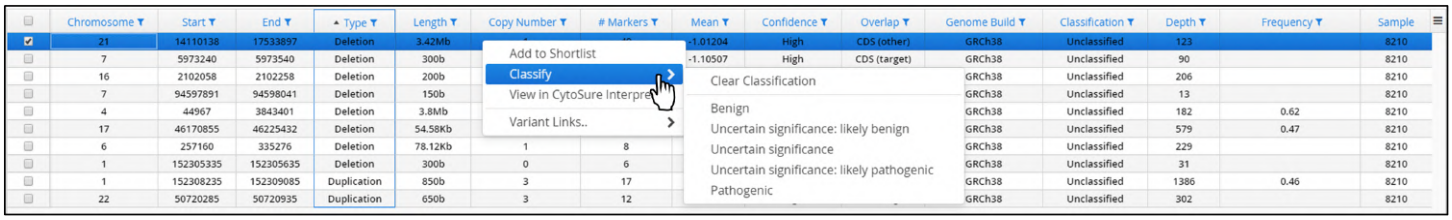


Figure: Default classifications

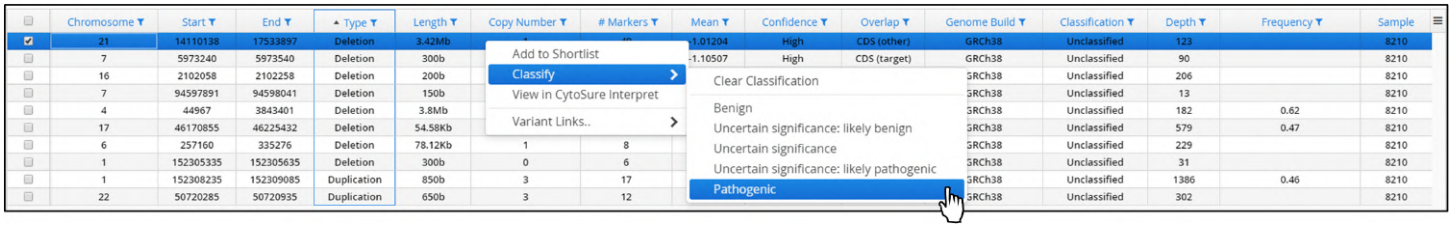


Figure: Classifying a CNV deletion as pathogenic

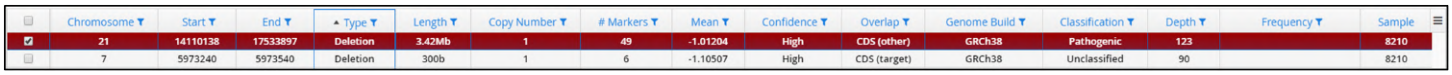


Figure: Updating of the variant to show the new classification

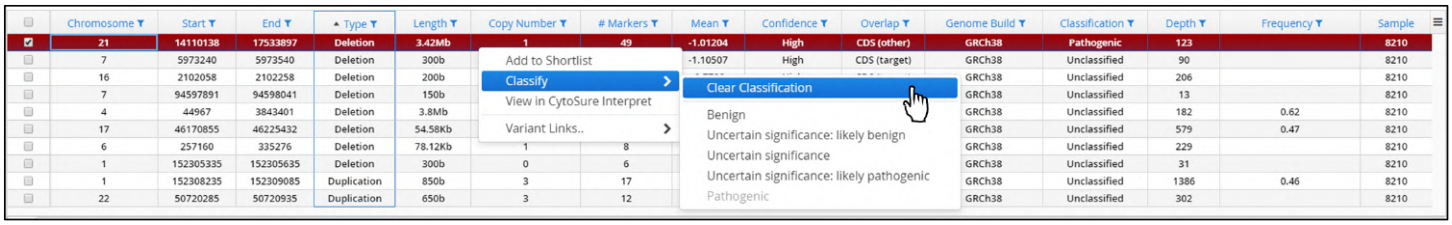


Figure: Removing a variant classification

View Classification History

User can review the classification of a variant by selecting that option in the menu.

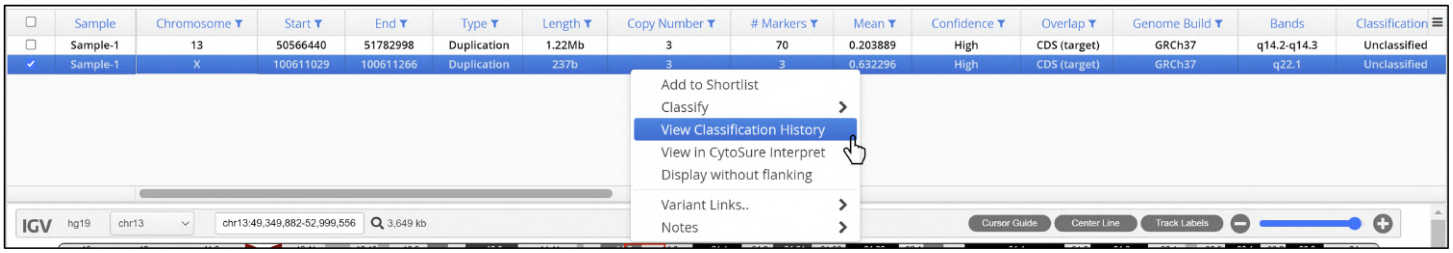


Figure: Selecting view classification history option

When chosen a table appears displaying how a variant has been classified, who made the classification and when any changes were made

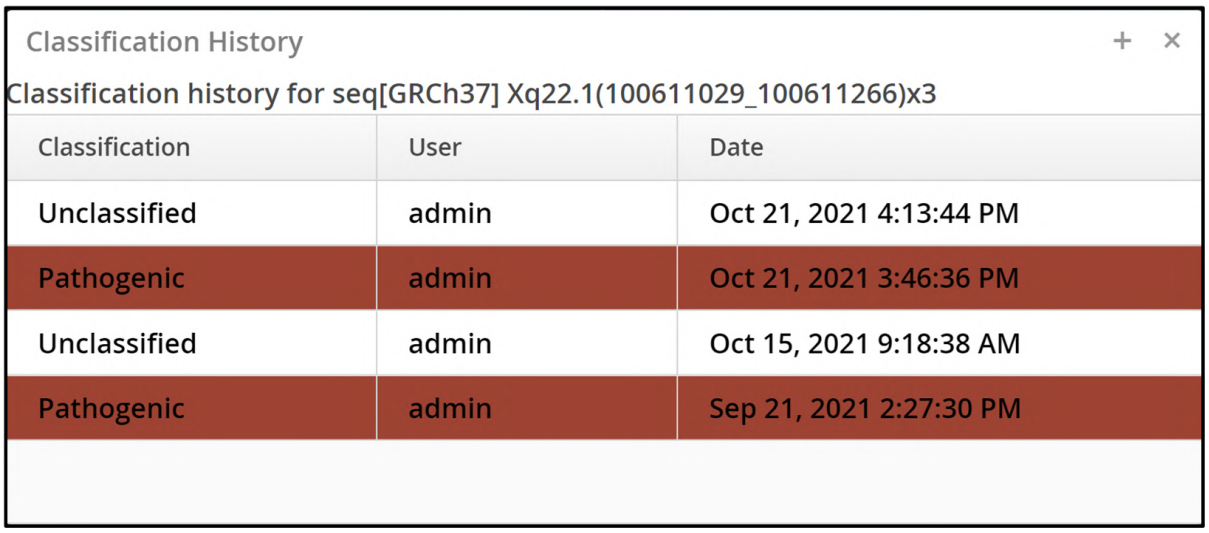


Figure: An example of a variant's classification history

View in CytoSure TM

CytoSure Interpret is OGTs class leading microarray software analysis platform. For existing microarray customers, CNV and LOH events can be loaded into CytoSure Interpret.

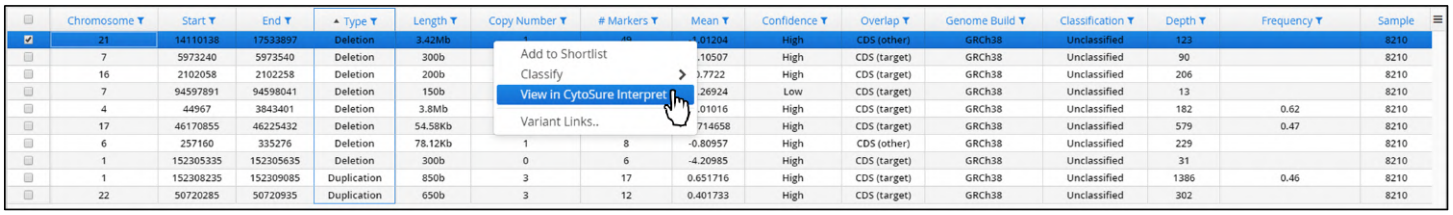


Figure: Selecting to view a CNV deletion in CytoSure Interpret microarray software

It is necessary to have CytoSure Interpret open prior to selecting this option.

If it is not yet running Interpret will issue a prompt to the user.

Figure: Prompt from Interpret if trying to load data in CytoSure Interpret when it is not running

Variant Links

The software allows users to link out to external sources of documentation. Currently included are:

- EnsEMBL

Additional resources can be added in the Admin Controls (Admin Controls > Analysis > Manage Links).

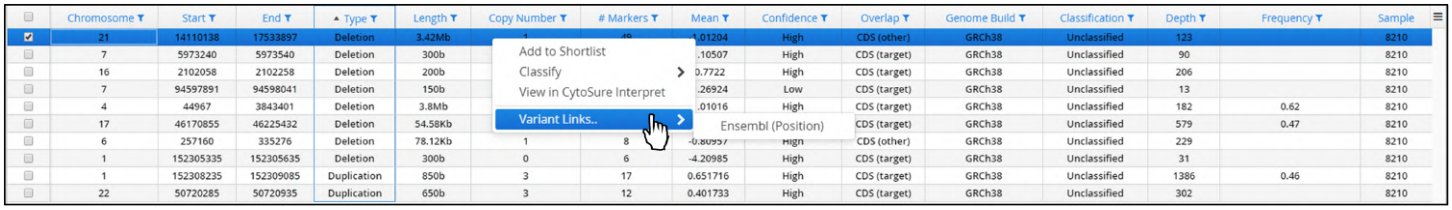


Figure: Accessing variant links

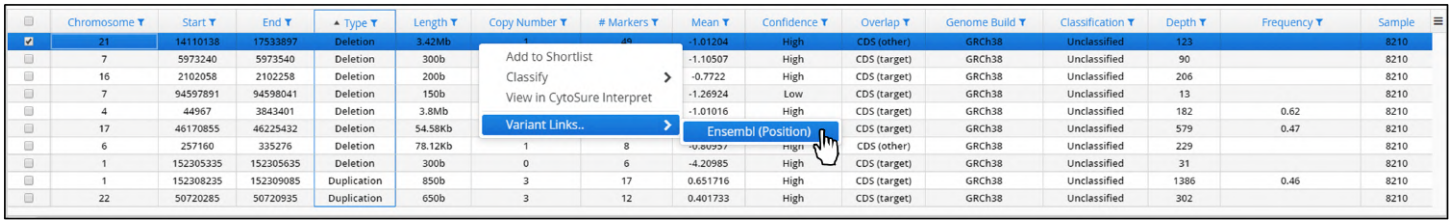


Figure: Accessing EnsEMBL as an external source for data annotation

11.4 Adding Notes to CNVs

Users can add notes to CNVs

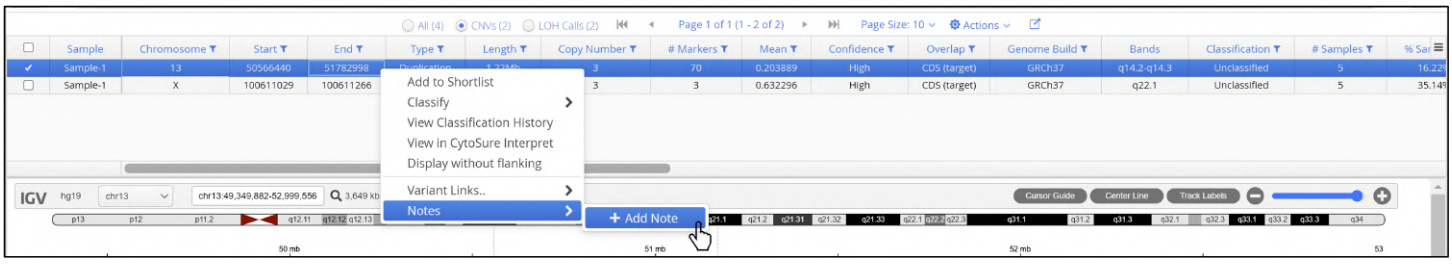


Figure: Selecting the option to add notes to a CNV

When chosen a text editor is displayed as well any pre-existing notes. At the top is an option to choose a file and to then upload it. Below is the text box where details can be entered.

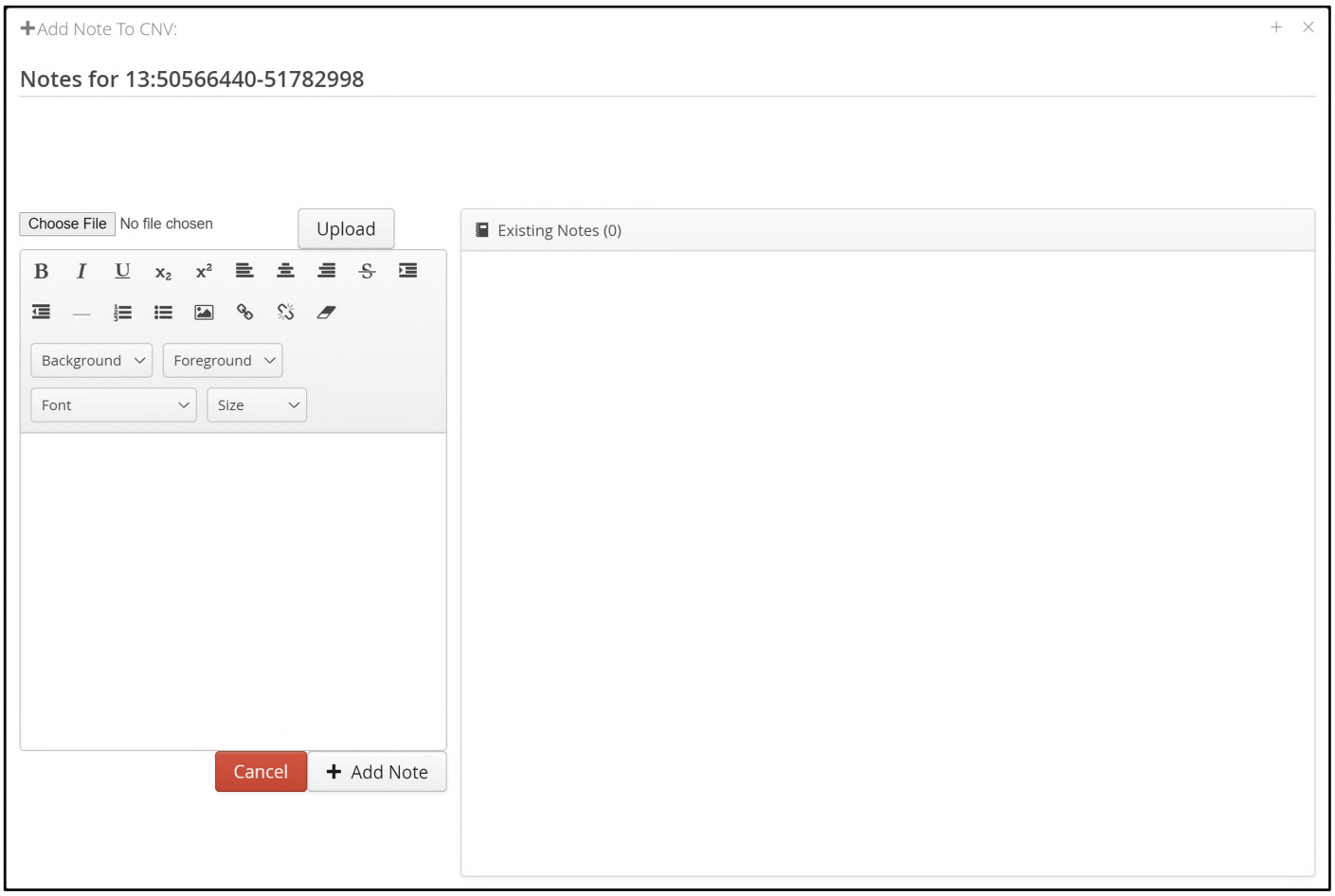


Figure: A blank template for creating a note

In the example below a file has been uploaded and text entered.

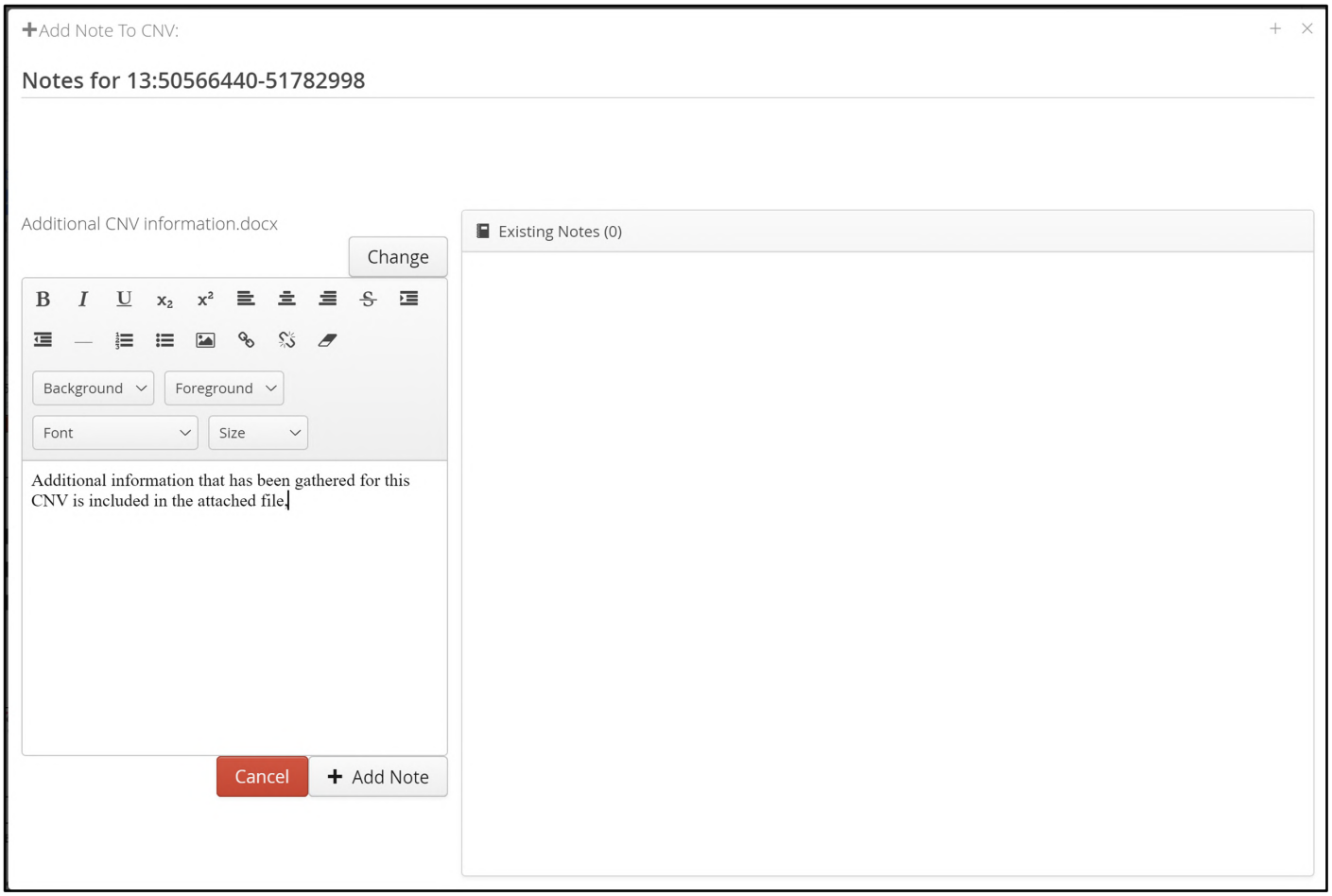


Figure: Addition of text and a file to a note

Selecting + Add Note completes creation of the note and it is added to the existing notes section. Any file that has been uploaded with the note is shown and can be downloaded if required.

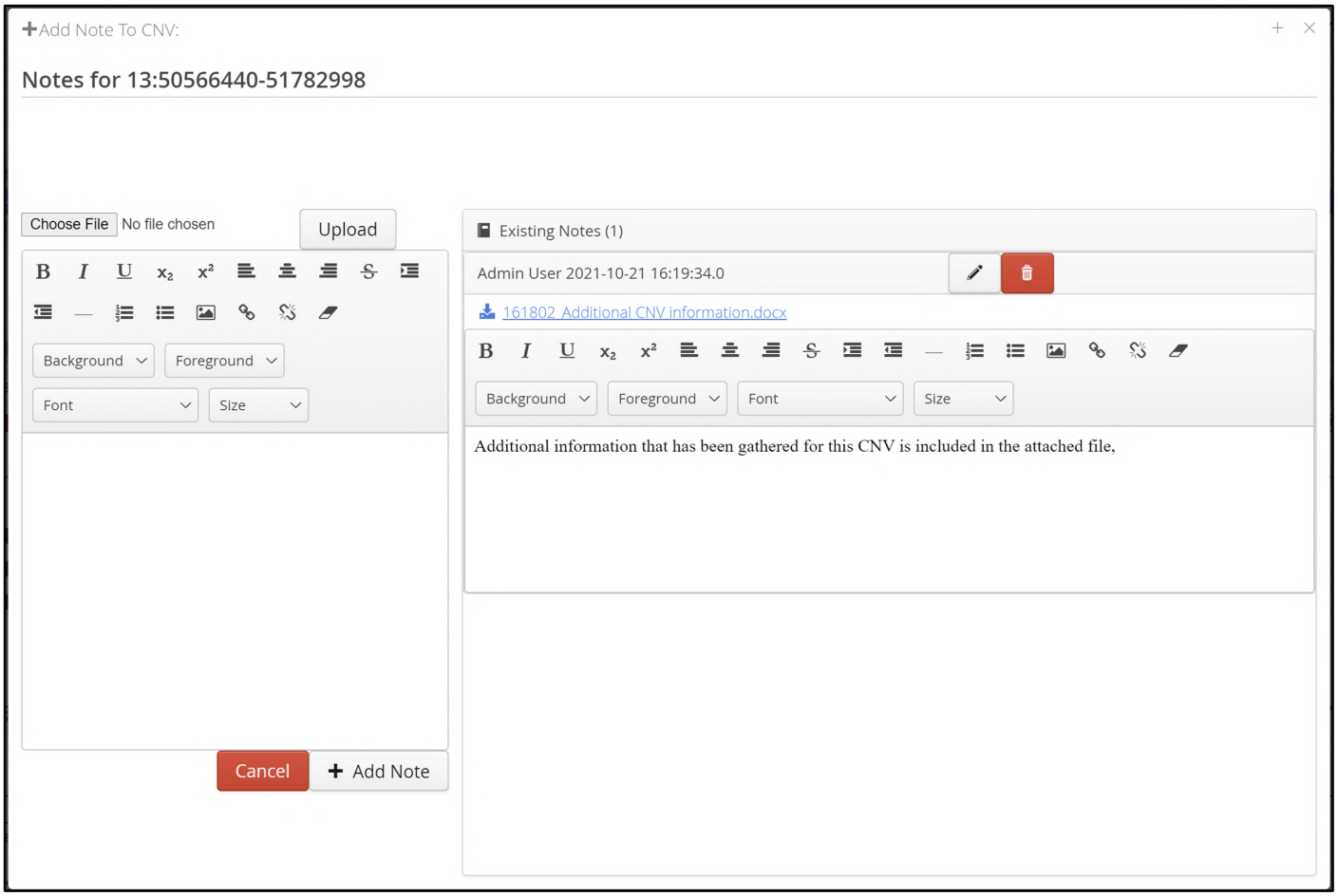


Figure: A new note is shown.

Now when the user accesses the menu for the variant there is an additional option providing the display of any notes.

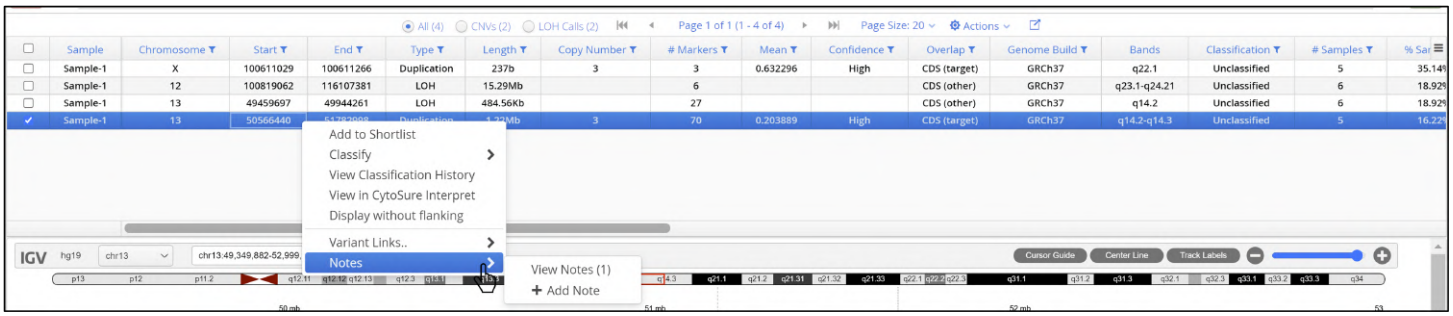


Figure: Existing notes are available to view

If selected, the note(s) are shown in a separate box.

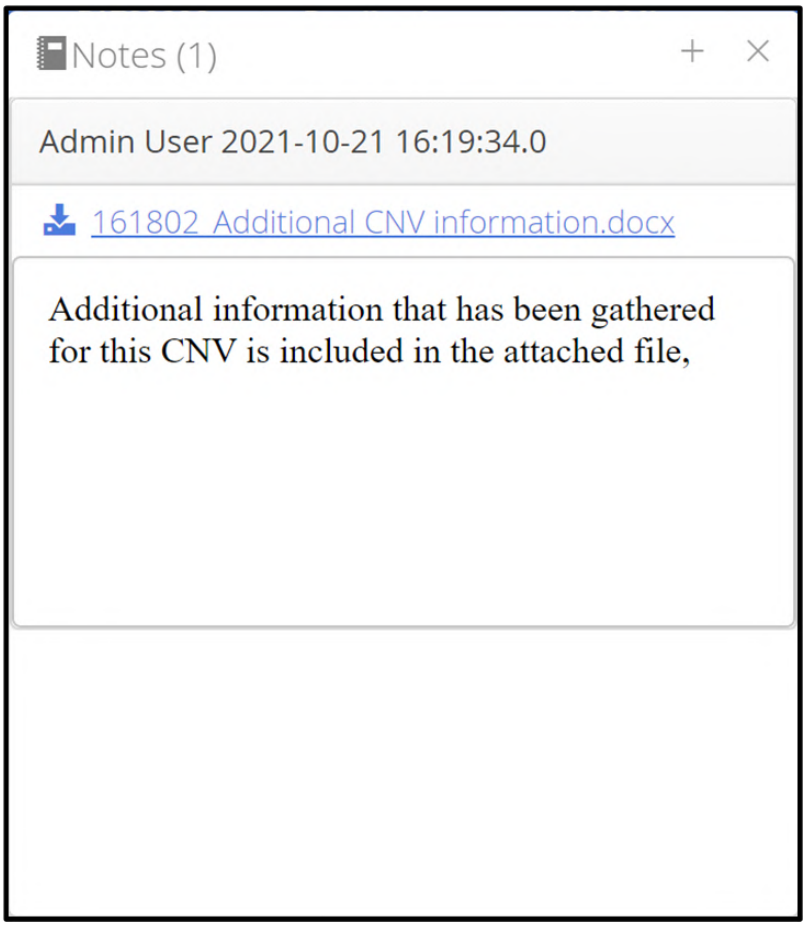


Figure: Viewing an existing note

Additional notes can be added as shown below.

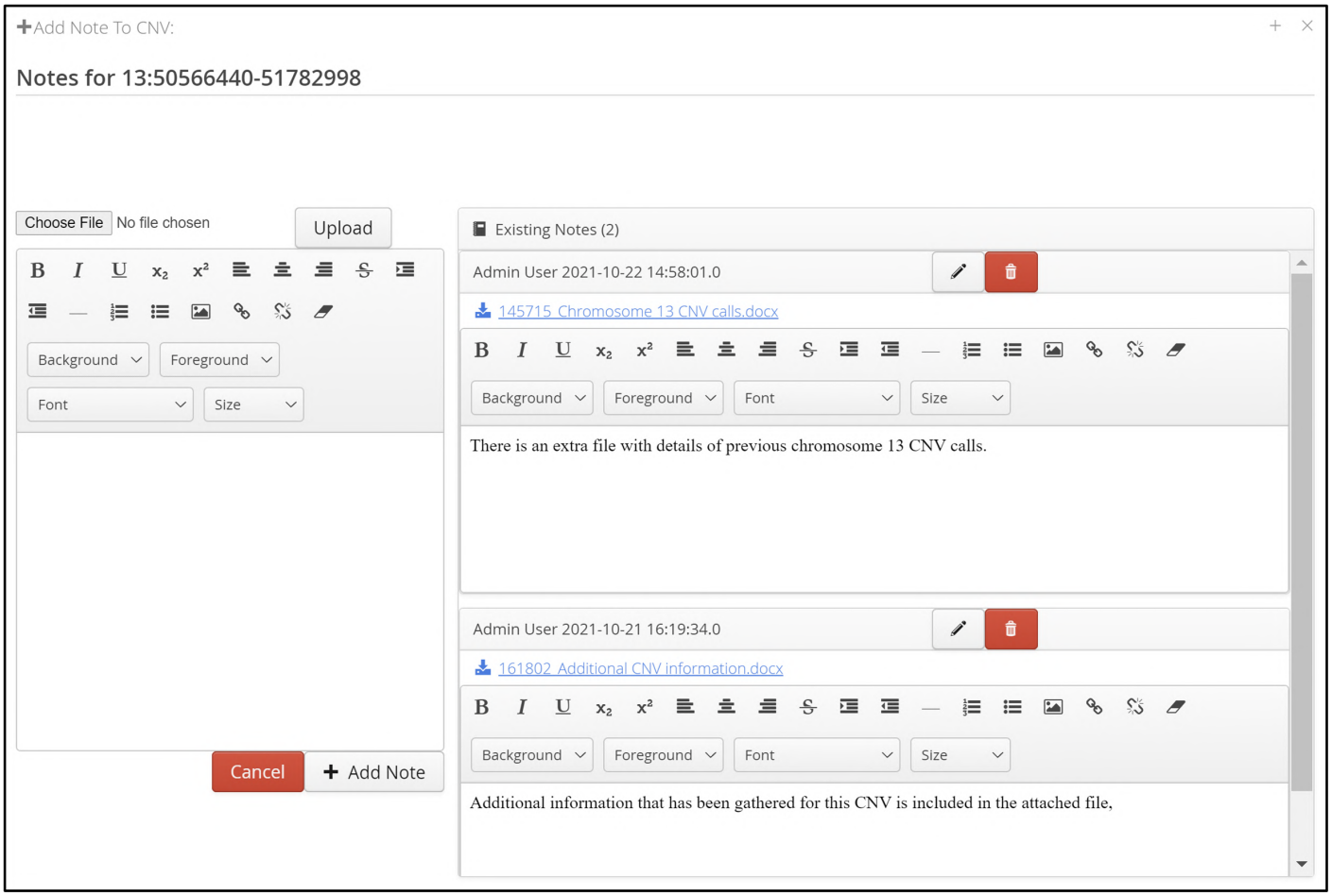


Figure: A blank note template with two existing notes

11.5 Manual Creation of CNVs

It may be that the user believes, based on the visual representation of the CNV data, that the software has missed a CNV call and would like to manually generate it. For example, a user may believe that the region highlighted in the screenshot below represents a CNV, but it has not been automatically detected by the software.

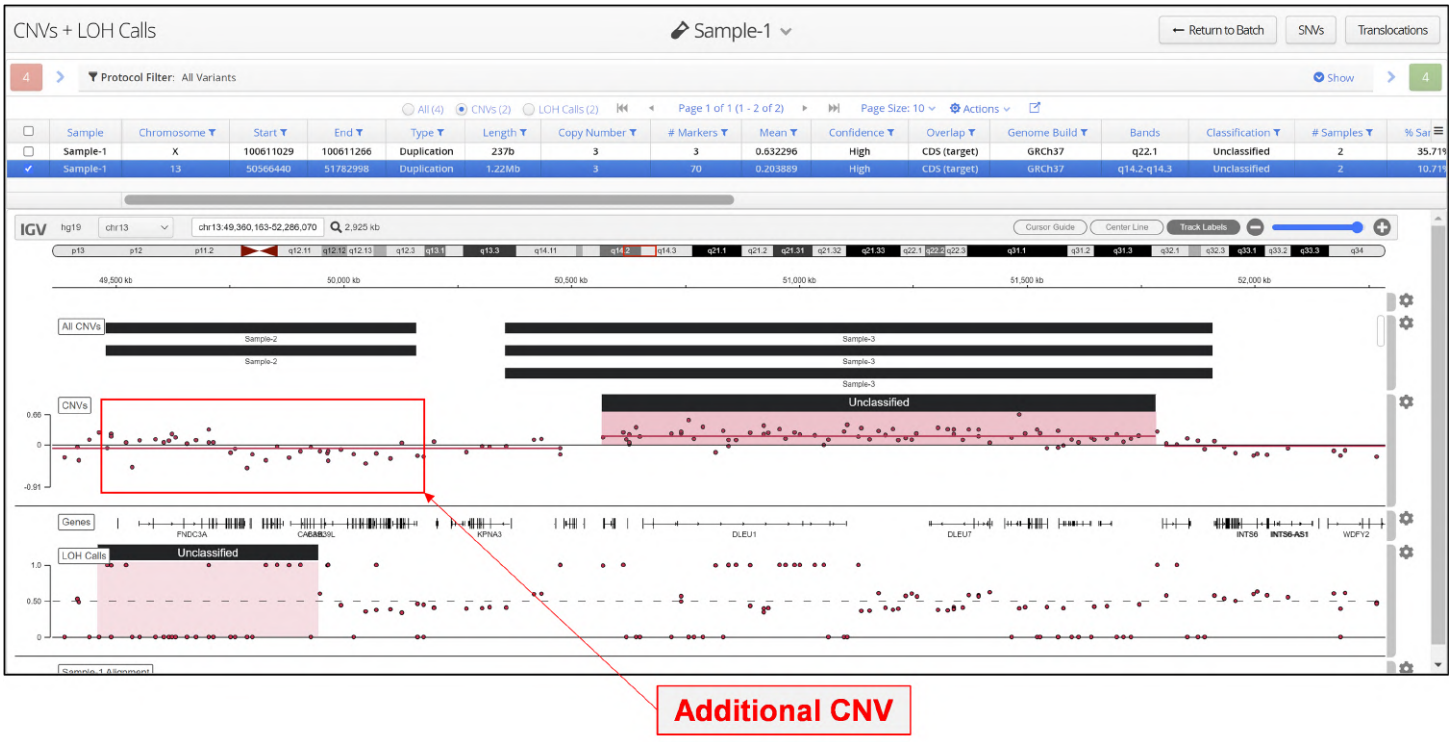


Figure: A region, not called by the software as a CNV, that the user wants to manually define as a CNV

In order to manually create the CNV call, the user defines the CNV region by using the mouse to select the region in the chromosome ruler track. Alternatively, the coordinates can be provided in the text box in the menu bar of IGV.

Figure: Using the ruler region in IGV to select the region to display

The IGV window resets to the size of the required region

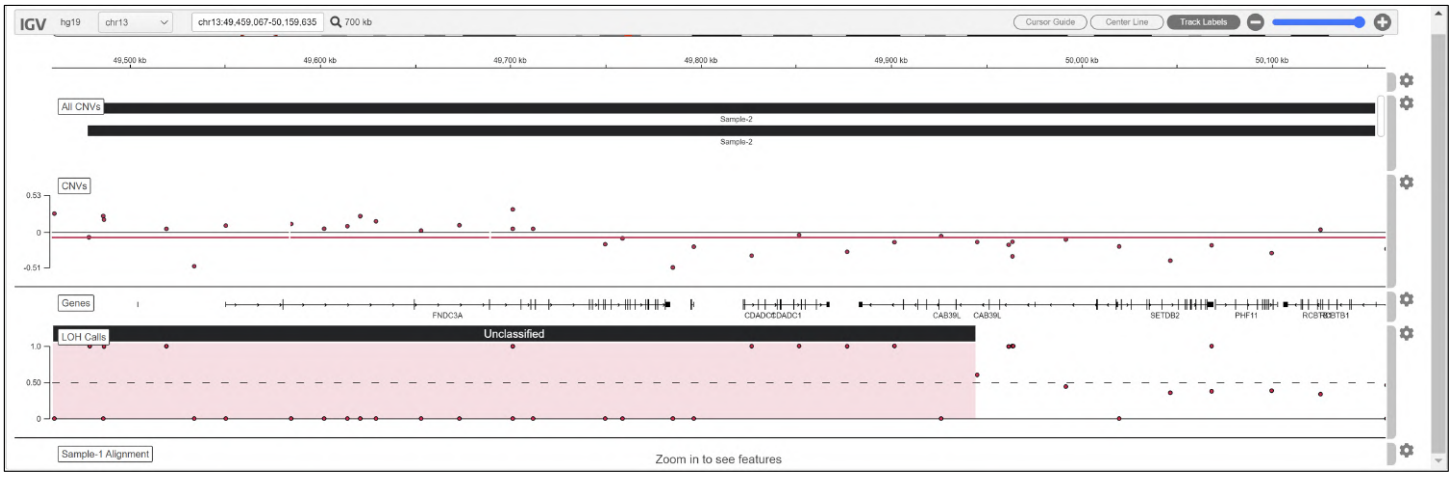


Figure: IGV set to the boundaries of the region to be defined as CNV

Select the Add CNV option in the Actions menu.

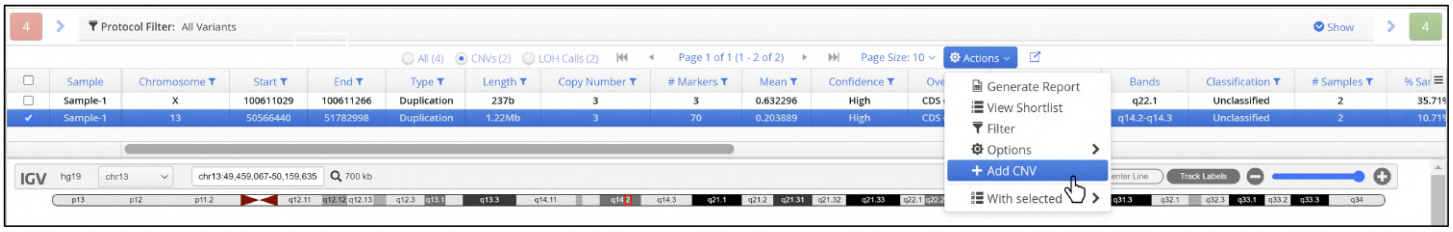


Figure: Selection of the Add CNV option from the Actions menu in the variant table header

Users have the option to define the entire region in IGV as the CNV or the software will snap to the nearest probe at each end.

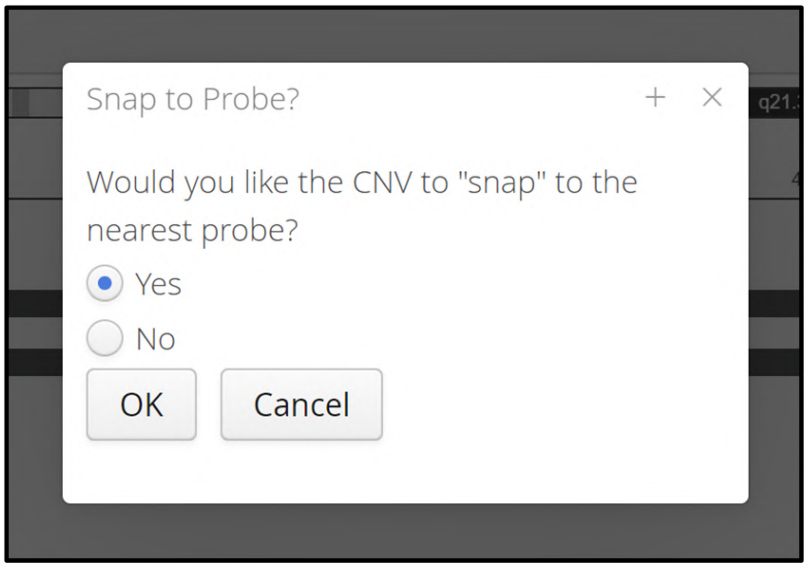


Figure: The option to snap to the nearest probe or to keep the region as that displayed in IGV

Once a CNV has been created manually it is **NOT** automatically displayed in the software. The creation of the CNV can be confirmed by the number of CNVs detected that feed into the protocol filter; in the figure below the number has incremented by one to five.

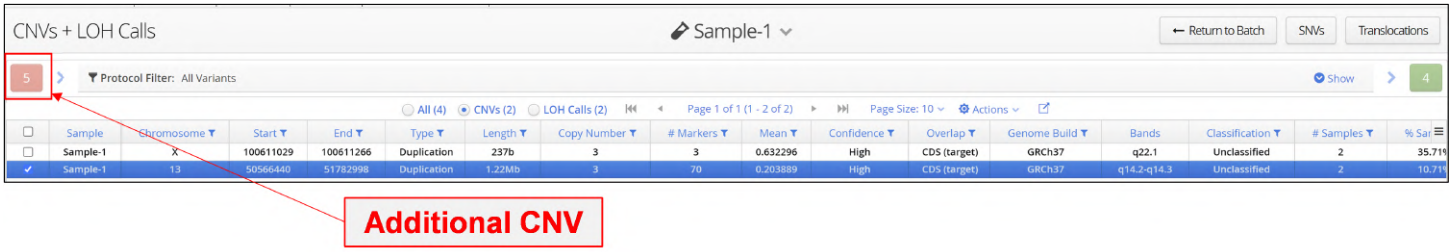


Figure: The number of CNVs has been incremented by one.

However, in order to display it in the variant table additional steps need to be taken. In the original analysis the protocol and filters did not select the manually defined region as being a CNV and so in order to include it the default protocol CNV filter need to be updated. This can be done in the Admin Controls > Analysis > Protocols section.

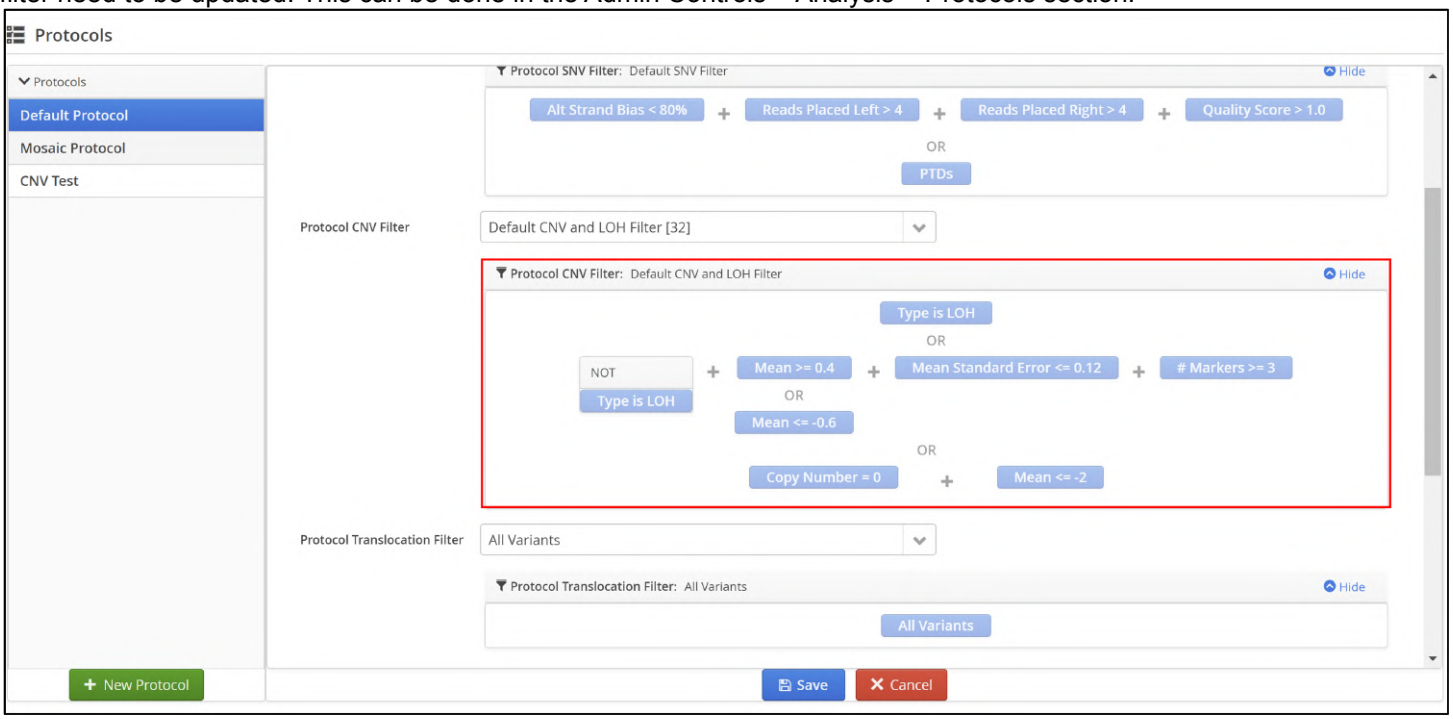


Figure: The default CNV filter in the default analysis protocol

Users need to create a filter that allows manually created CNVs to be included and this can be added to the default CNV filter. This is shown in the figure below.

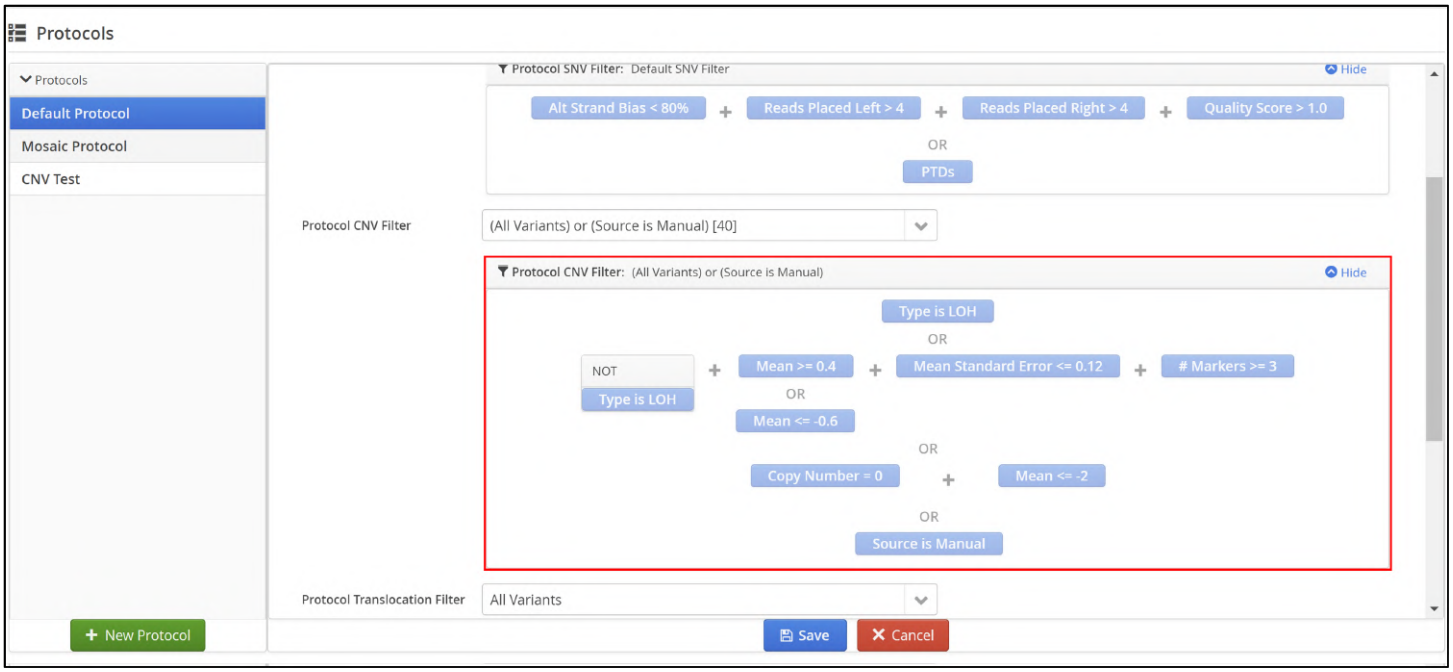


Figure: The CNV filter in the Default Protocol has been edited to include OR Source is Manual

Repeating the analysis with the updated filter will result in any manual CNVs being added to the sample variant list.

11.6 Merging CNV calls

There are occasions when CNVs are called with small regions in between that the user would like to combine into a single larger CNV.

In order to do this, adjust the scaling in the IGV window such that both CNVs are visible, then right clicking between them in the track will generate a popup menu with the option to Merge Displayed provided.

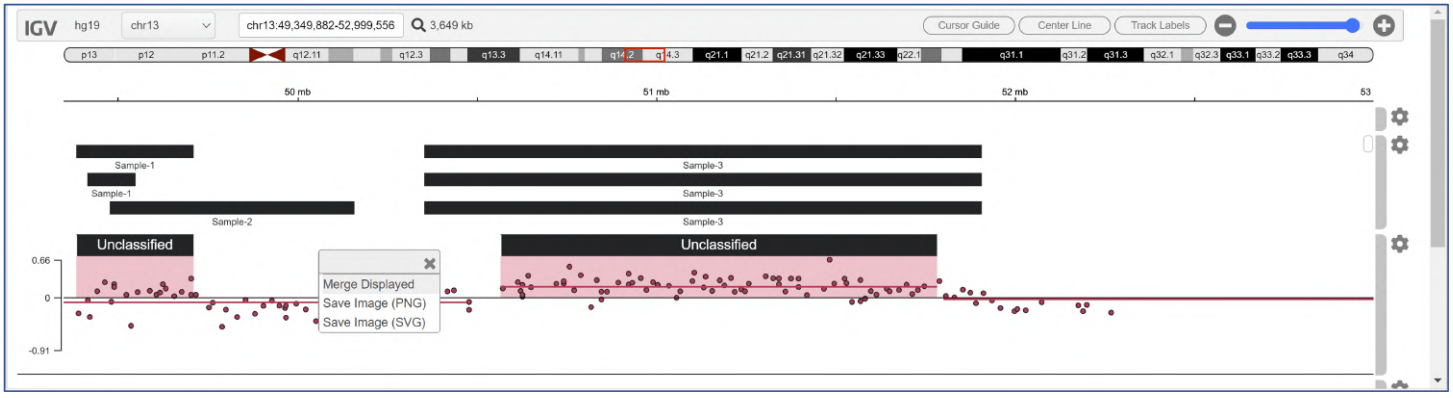


Figure: Selecting the option to merge displayed CNVs

If selected the software will request confirmation of the merge option.

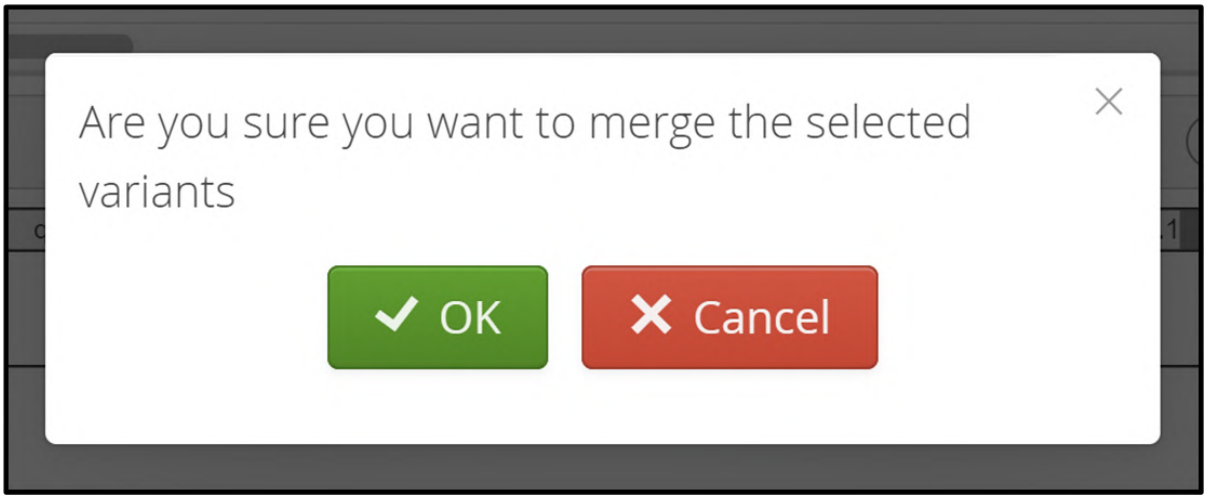


Figure: Confirmation of merger option

Following confirmation of the merge option the variant table will be updated. There will be a single row containing the new merged CNV that spans the two previously separate calls. Additionally, the variant counts above the table will be updated.

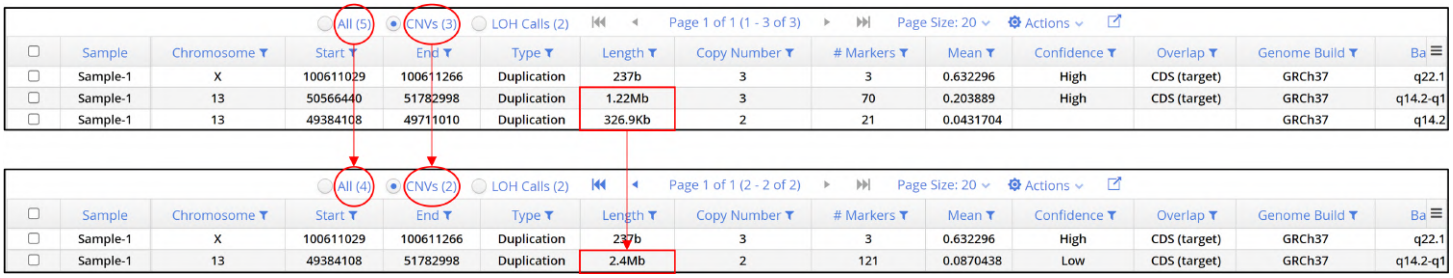


Figure: An updated variant table showing the merging of 2 CNVs to a single row in the table as well as the decrease in the number of All variants and CNV variants.

Likewise, in the IGV window the two calls are now combined.

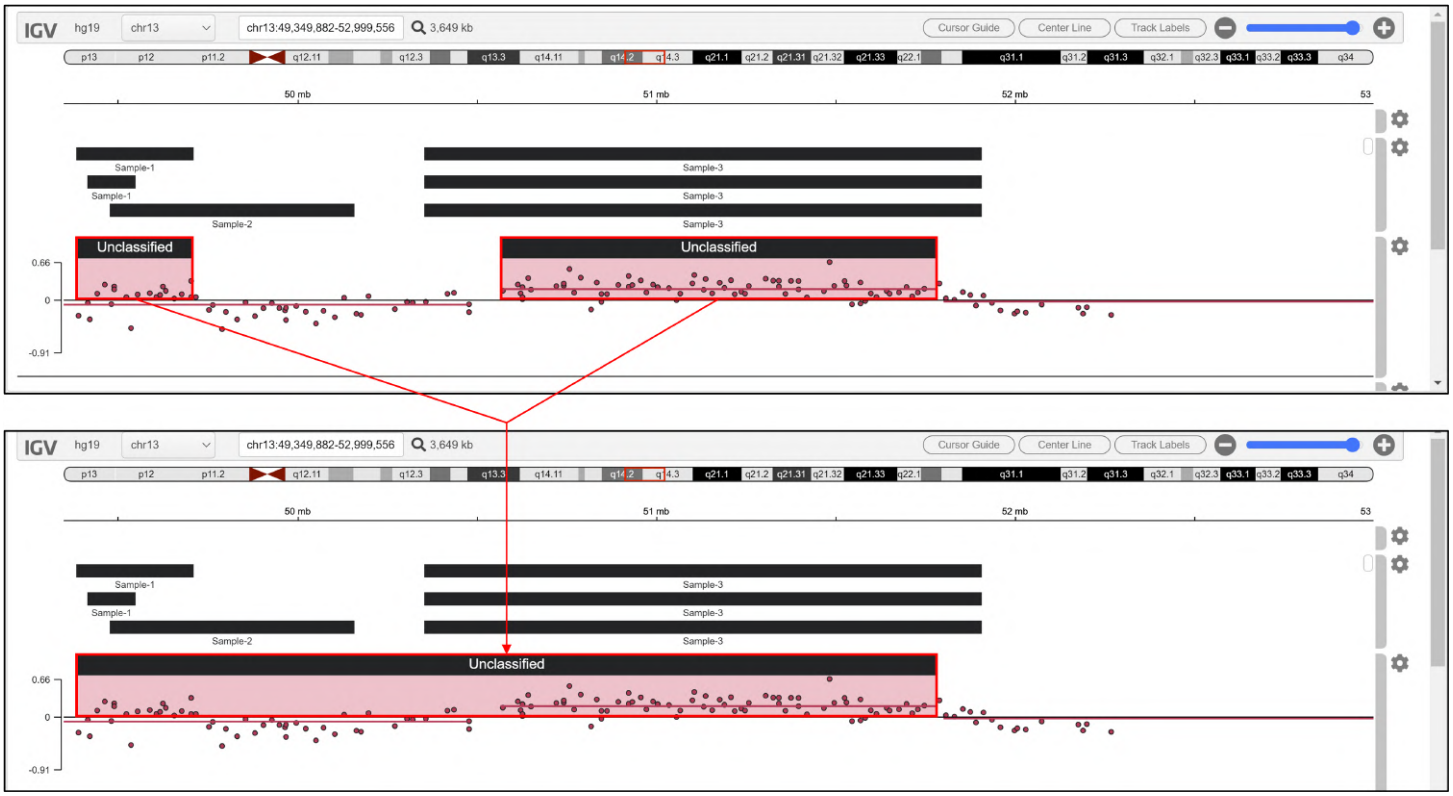


Figure: Following the merger of 2 CNVs a single CNV is now displayed.

11.7 Separating Merged CNV calls

Having been created, users are able to dissolve a merged CNV. Right clicking on the CNV row in the variant table will display the standard popup menu but now with an additional option of Dissolve which will split the merged CNV back into the original separate calls.

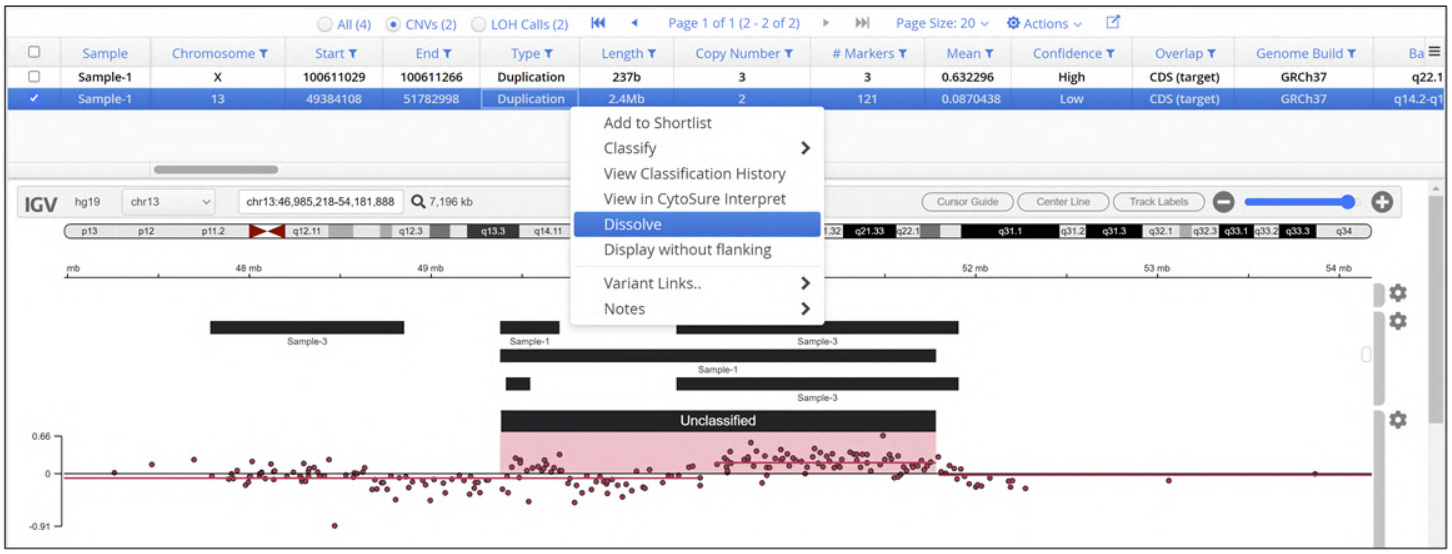


Figure: Selecting the dissolve option in CNV variant table.

11.8 Aneuploidy Plots

Interpret is able to provide aneuploidy plots in order for the user to assess whether there is a difference in chromosome number in a set of patients.

This functionality is accessed through the Tools sub-menu in the software dashboard as shown below.

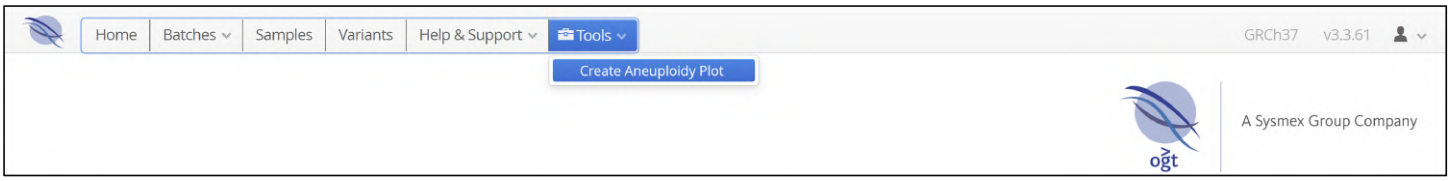


Figure: The create aneuploidy plot option

The aneuploidy plot option can only be used when the user is viewing CNV data. If this is not the case then the following error message will be displayed. As with all error messages in the software it can be removed by clicking on it.

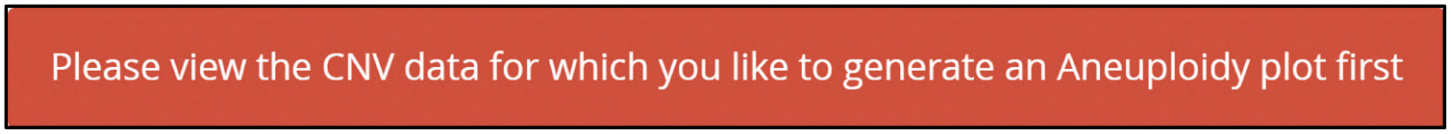


Figure: Error message from aneuploidy plot option

The figure below illustrates the correct view from which to launch an aneuploidy plot.

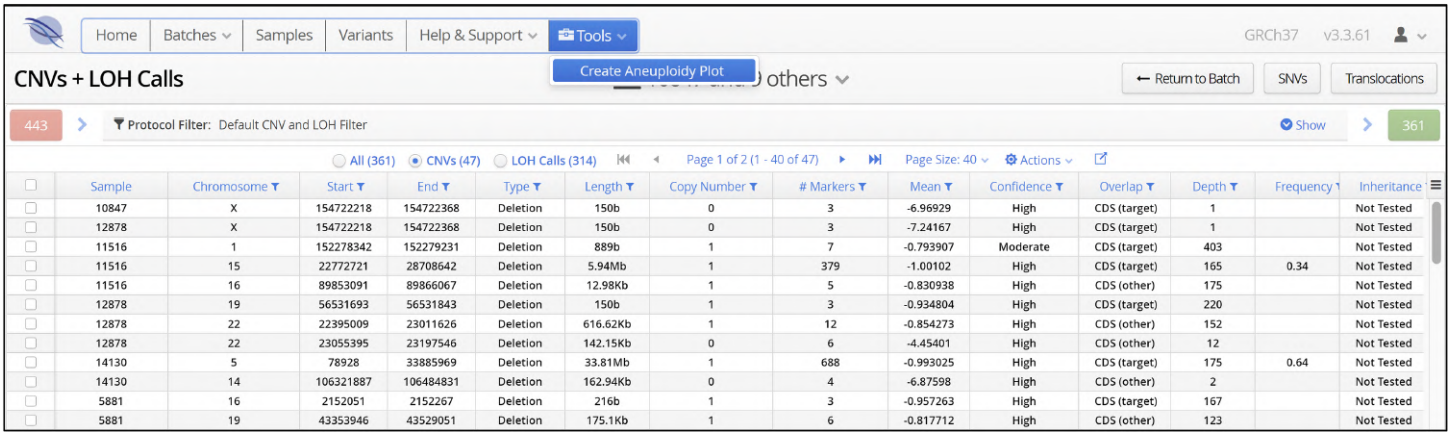


Figure: Selection of aneuploidy plot from menu bar Tools

The user will then be asked to provide the region list that needs to evaluated for plotting. Information on creating a region list is documented in this manual in the Administration Controls > Analysis > Region lists.

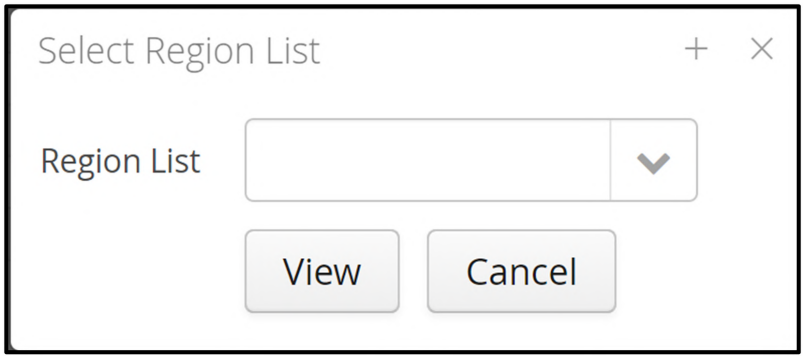


Figure: Select a region list pop-up

The available regions lists will be displayed in a drop-down menu; in this example there is a single region list created in the software.

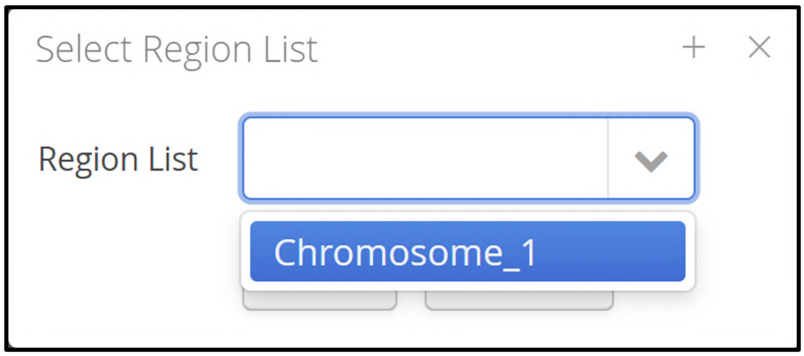


Figure: Selection of a region from the drop-down list

Following selection of the required region list the user clicks on the View button.

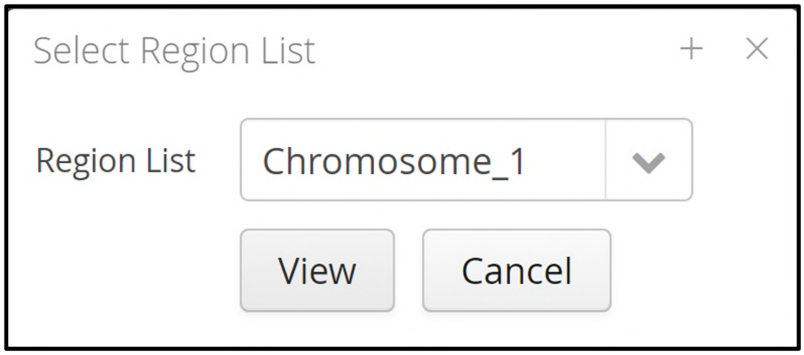


Figure: Selection of the region Chromosome_1

An example aneuploidy plot is displayed below.

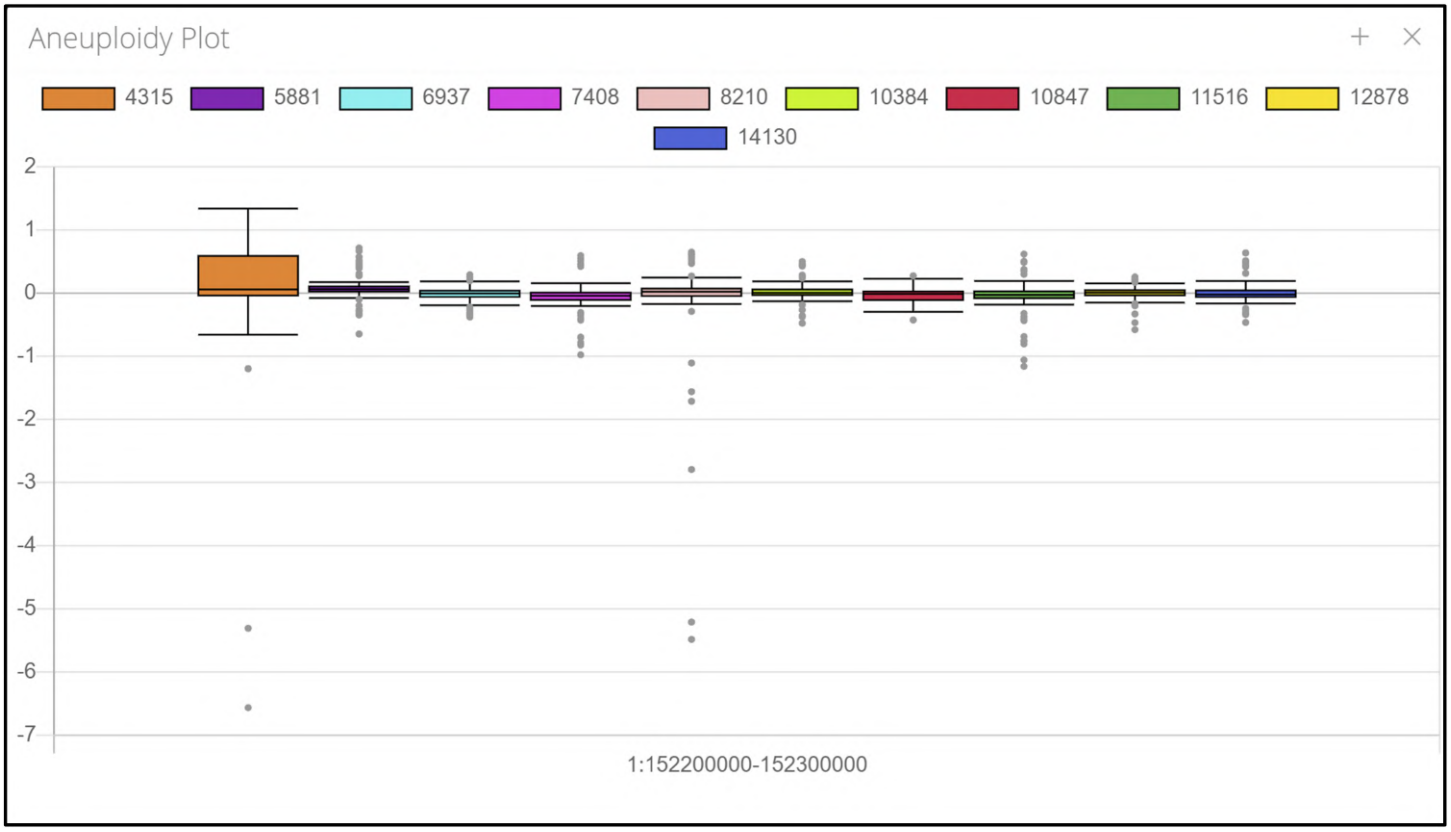


Figure: An example of an aneuploidy plot for the chosen region across the samples in the batch

11.9 Viewing Translocation Events

The variant table has a column selector icon allowing user to configure which columns are displayed. The figure below shows the columns available for display.

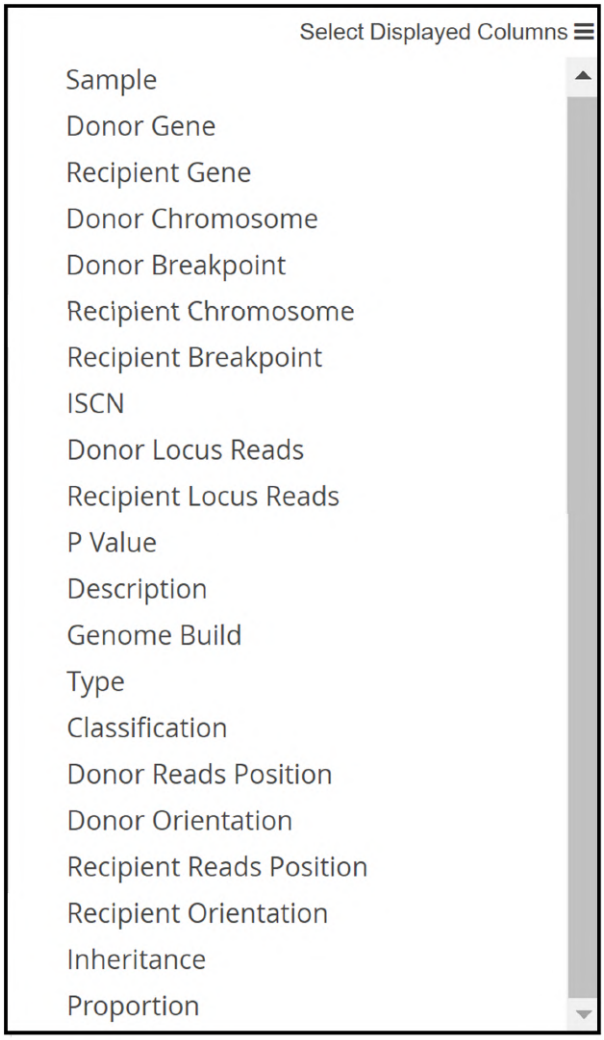


Figure: Columns available to select for display in the translocations variant page

Figure: Example of a translocation

Translocation Options

As with the page displaying SNV and Indel calls there are options available for each translocation variant called by the software,

Right clicking on a variant will provide a menu of the possible options.

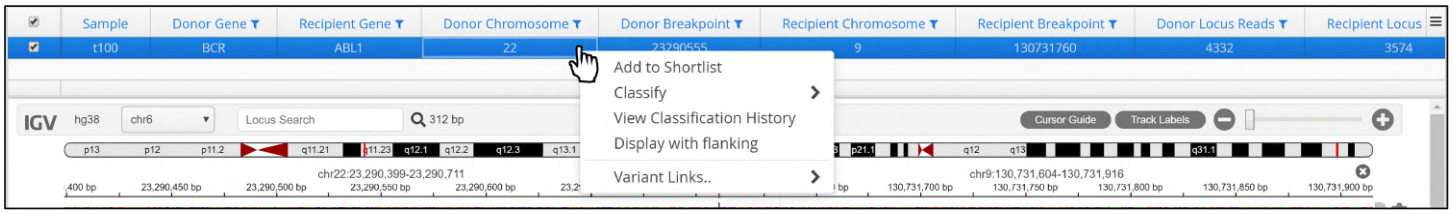


Figure: Translocation options

Adding to Shortlist

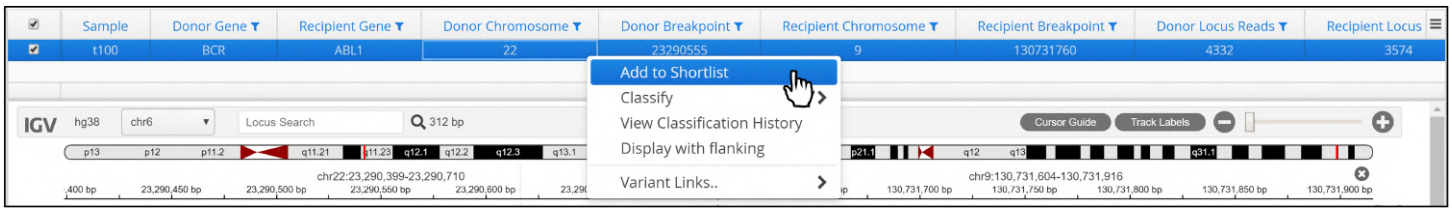


Figure: Adding a translocation to the shortlist

Once a variant has been added to a shortlist the available option is updated to now allow that variant to be deleted from the shortlist.

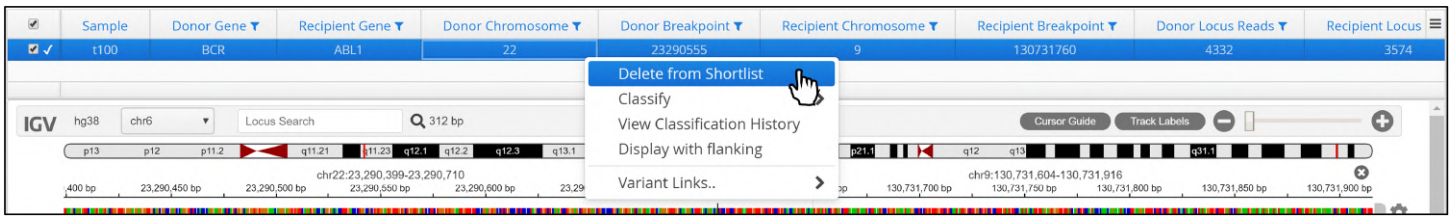


Figure: Selecting to delete a variant from the shortlist

Variant Classification

A variant can be classified from the list that is included by default. These are:

- Benign
- Uncertain significance, likely benign
- Uncertain significance
- Uncertain significance, likely pathogenic
- Pathogenic

Additional classifications can be added in the Admin Controls section of the software (Admin Controls > Analysis > Classifications)

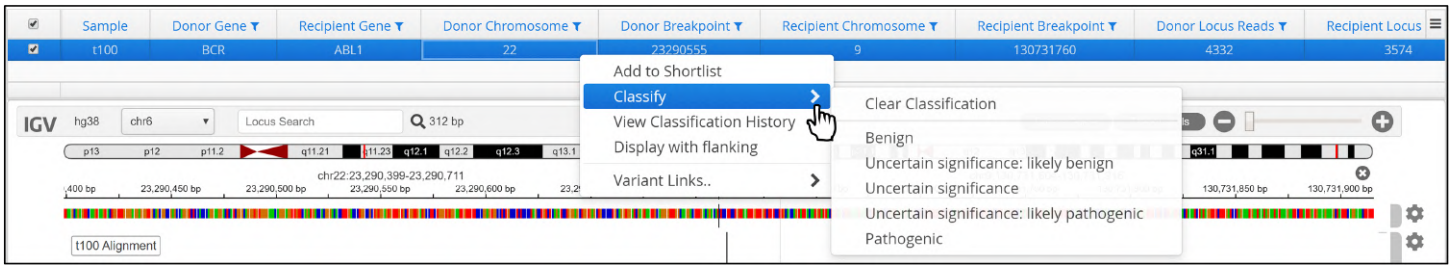


Figure: Classify the translocation

A variant classification may change over time and it is possible to track the changes and view the classification history.

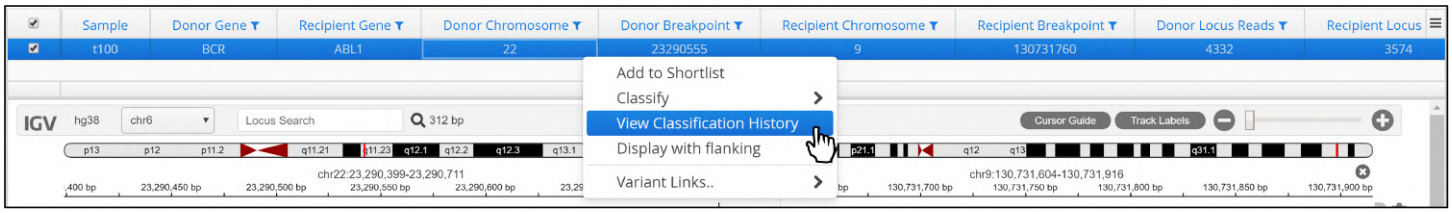


Figure: Viewing a variant's classification history

Initially, the classification will be blank.

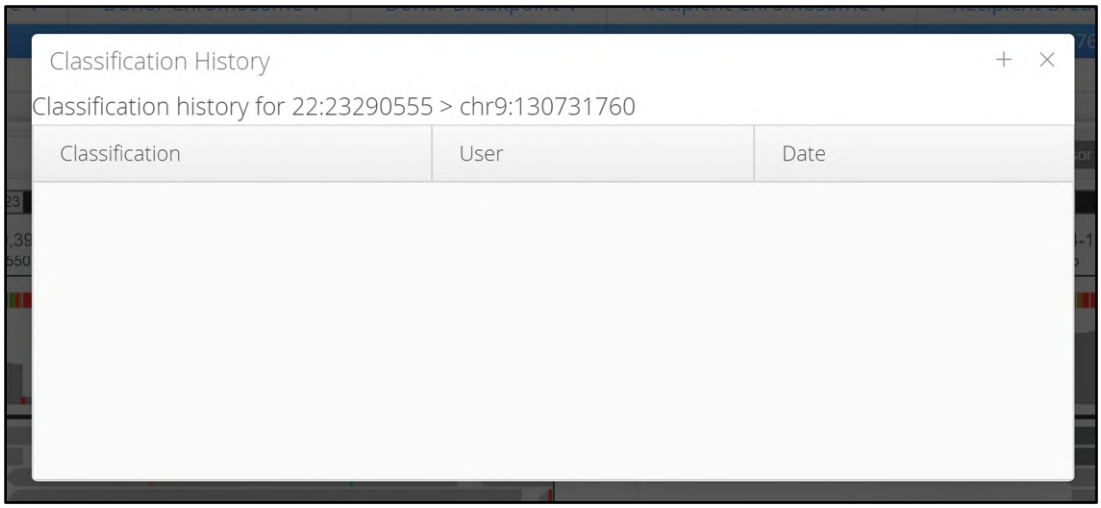


Figure: A variant with no classification history

When a classification is made the history table will show the classification type, who made it and when it was made.

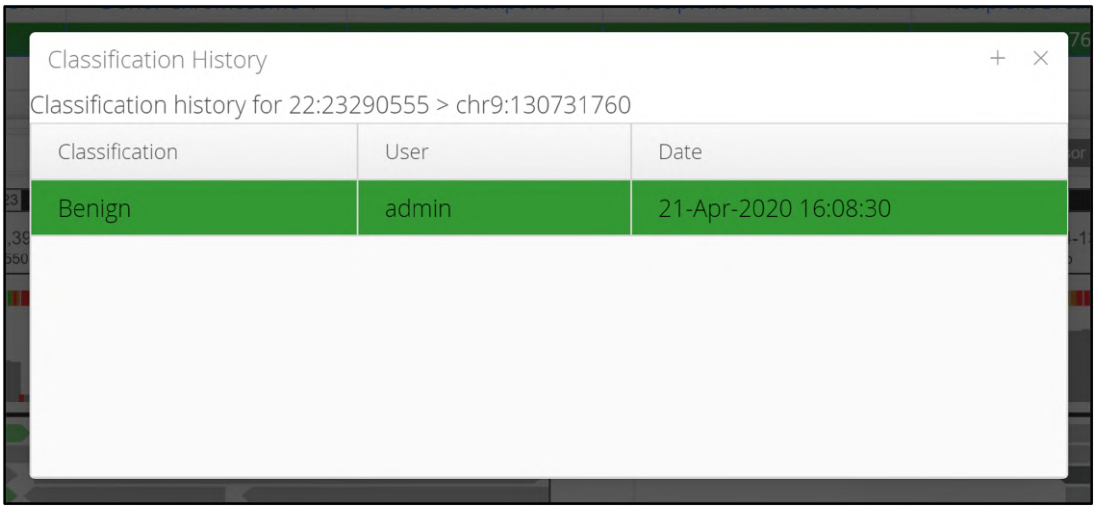


Figure: Example of a benign classification

Any updates to the classification will be recorded with previous designations retained.

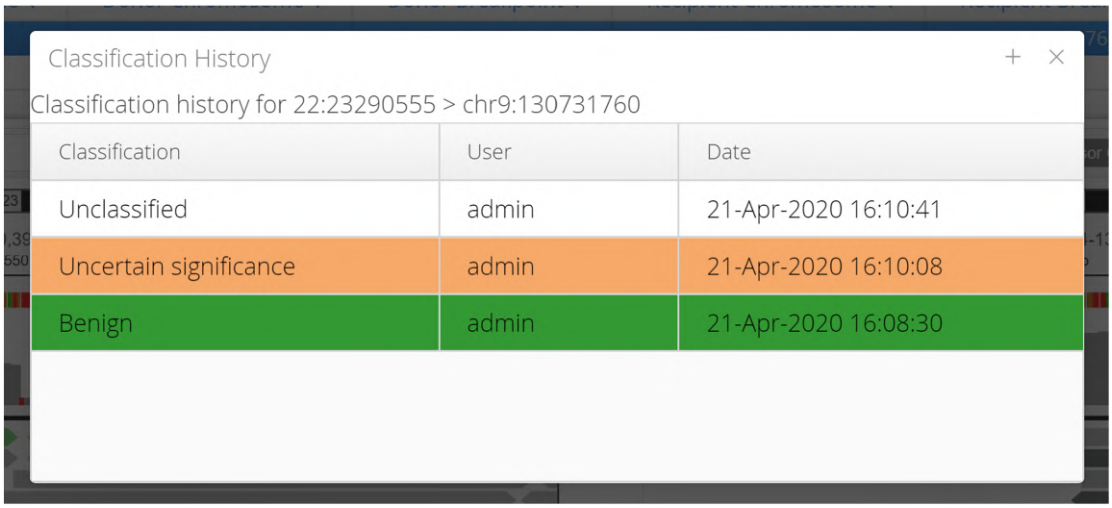


Figure: Example of a tracking a translocation classification change

Display with Flanking

Users can select to view translocations with flanking sequence

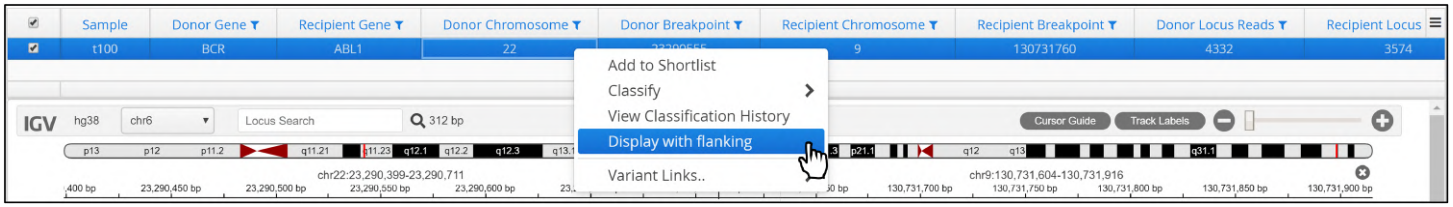


Figure: Selecting to show a translocation with flanking sequence

Variant Links

Links to external data sources are available; these are managed in Admin Controls > Analysis > Manage Links

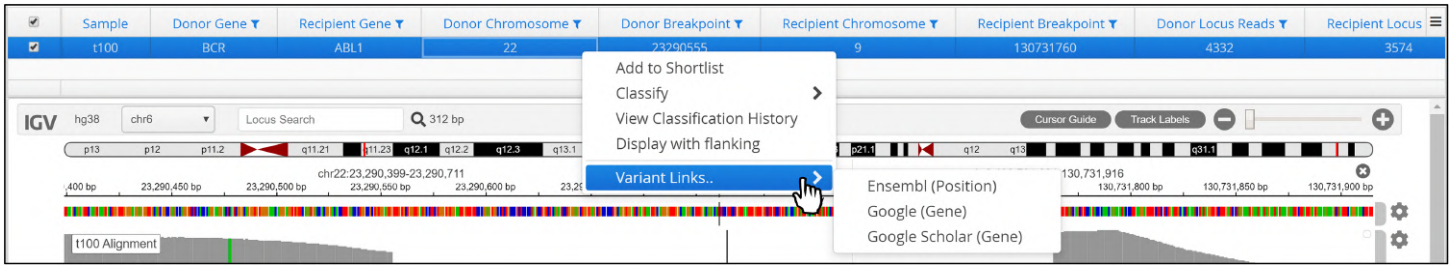


Figure: Linking out to external data sources

Variant Table Options

Column Sorting

Rows in the variant table can be sorted using the column header. In the example below the results have been sorted by decreasing and increasing allele frequency.

Currently, data can only be sorted by one column.

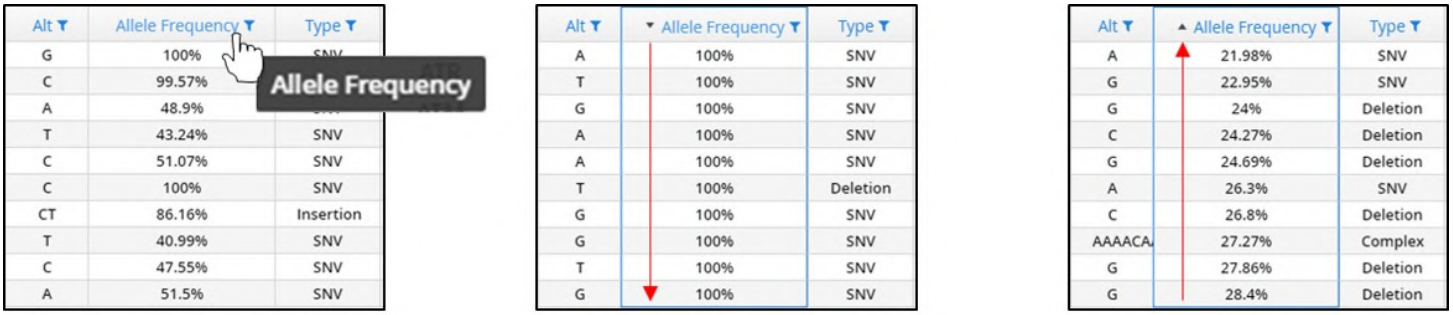


Figure: Sorting by Allele Frequency

Dynamic Filtering

As shown previously the variants page displays the Protocol Filter, the number of variants detected by the pipeline and presented to the filter is depicted in a red box and the number remaining in a green box.

In the image below you can see that there are 2946 variants (in the red box) detected by the pipeline that are to be filtered based on the settings in the protocol. Subsequently there are 2754 remaining (as shown in the green box).

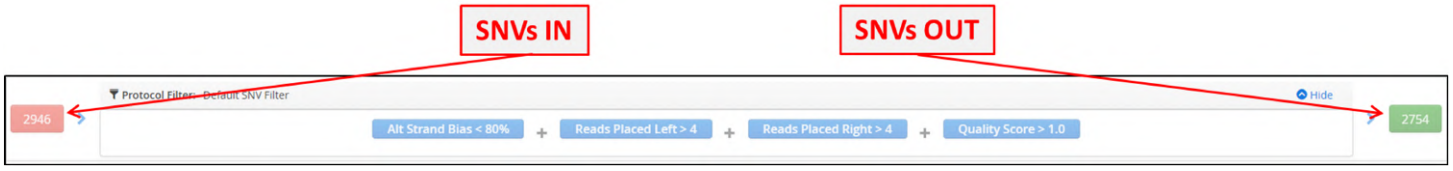


Figure: The filter used by the protocol in the analysis of the sample displayed in the Variants page

However, the user is able to implement additional filtering dynamically. Any column header with the funnel icon can be used as a filter. For example, a user may want to filter on Gene Symbol

Figure: Selection of the funnel icon for the Gene Symbol column

In this case they want to see only variants found when the total depth is greater than 200.

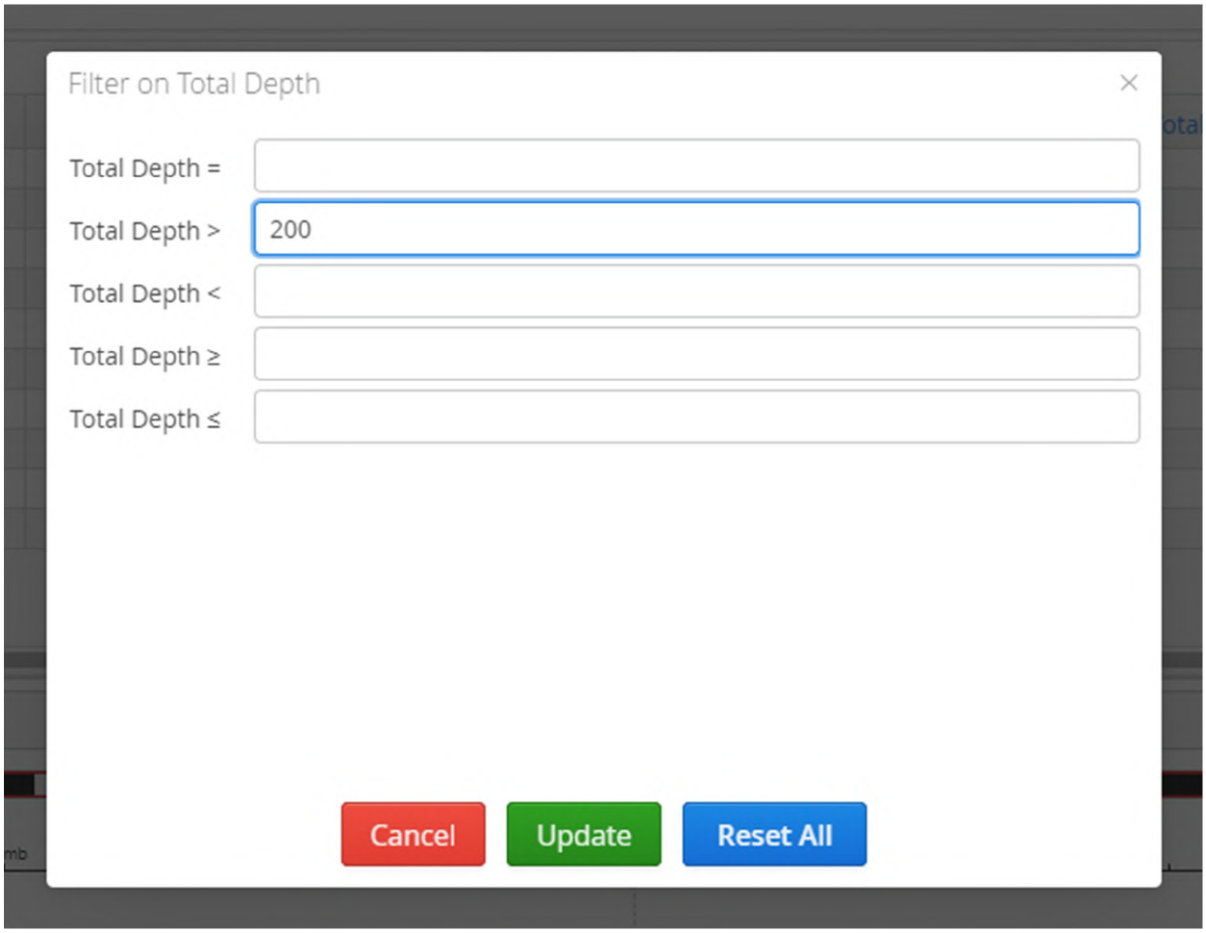


Figure: Dynamic filtering of the variants using the Total Depth column

After updating the Variants view now shows a Dynamic Filter window and within it is the "Total Depth > 200" filter.

From the 2754 variants generated by the protocol using the default filter, it can be seen that a further 265 have been removed filtered with 2489 remaining.

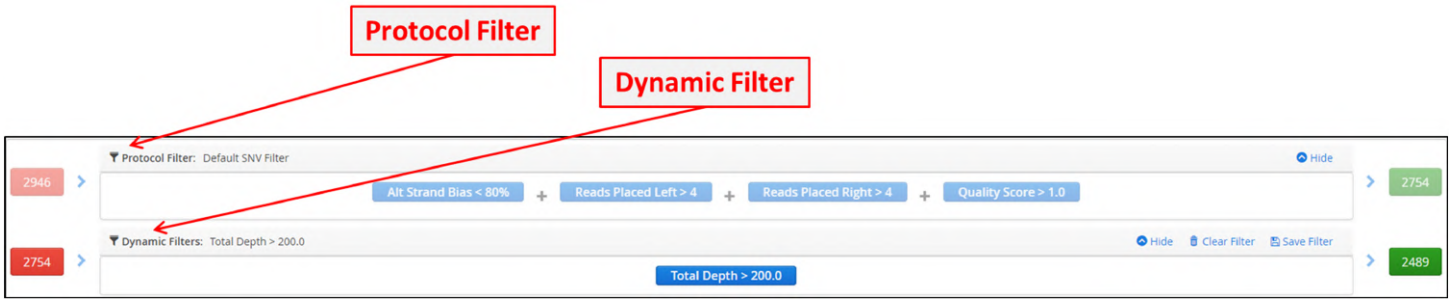


Figure: Example of Variants filtered by the protocol filter and a dynamic filter

Dynamic filters can be chained together so additional filters can be added for instance an Allele Frequency greater than 80%

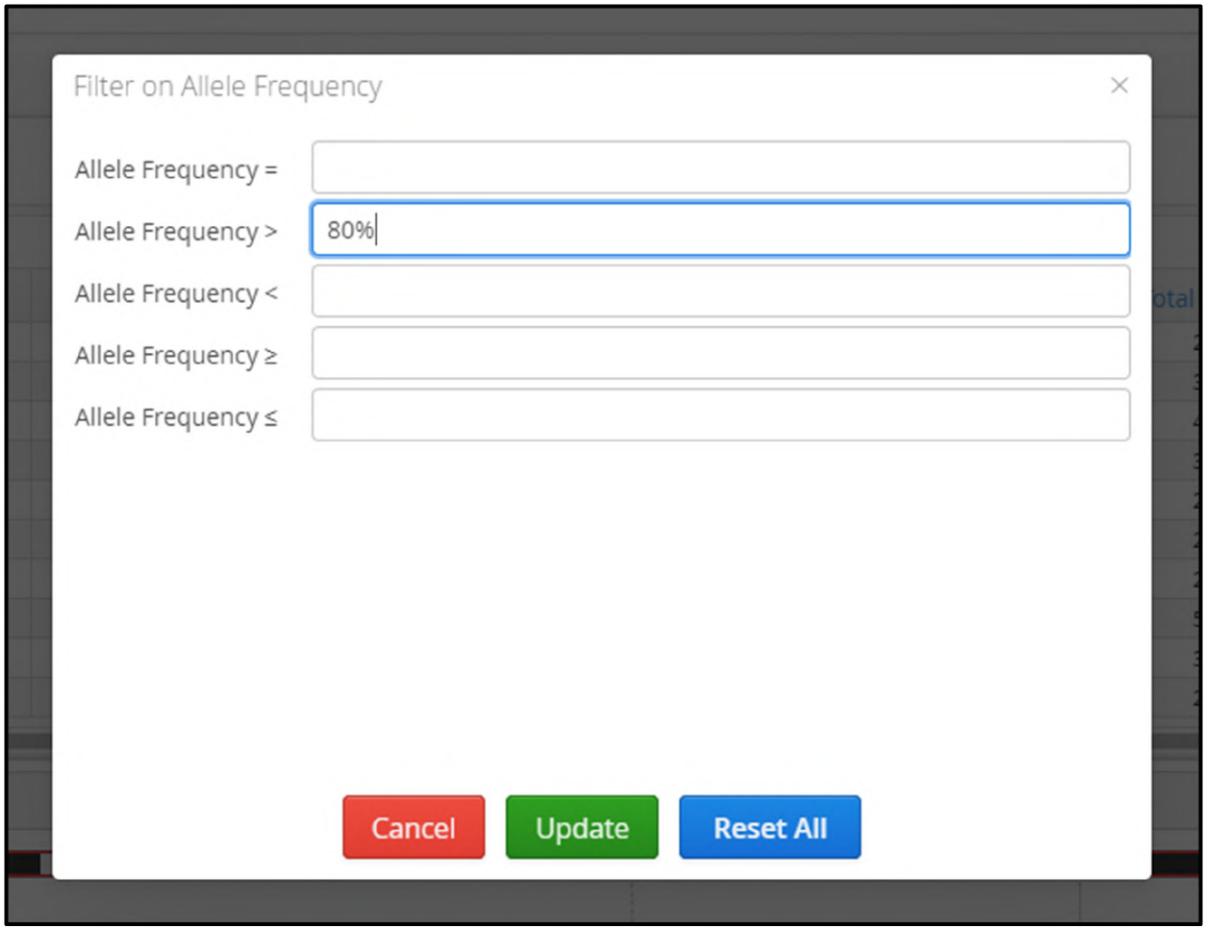


Figure: Selection of another dynamic filter to start creating combinations

Now the Dynamic Filter shows "Total Depth > 200 and Allele Frequency > 0.8" and there are now 1457 variants remaining from the input of 2754.

Figure: Variants being filtered by a compound dynamic filter

There is no requirement for the user to have to repeat the setting of dynamic filters every time they use the software, there is the option to name the filter and pressing Save Filter to retain for re-use.

Alternatively, all dynamic filters can be removed from the display by selecting to clear the filter Clear Filter

Viewing a Sample in IGV

Selection of a variant in the Variants Table causes it to be displayed in the embedded IGV.

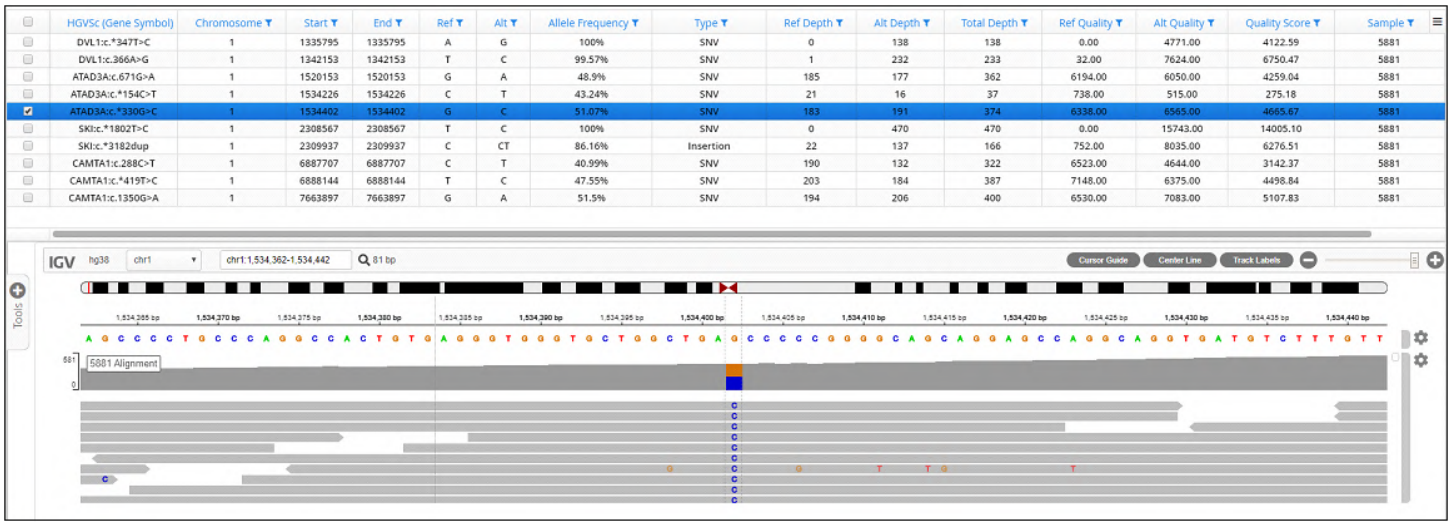


Figure: A variant selected and the aligned displayed in IGV

Within the IGV window there are several options for modifying the data being displayed.

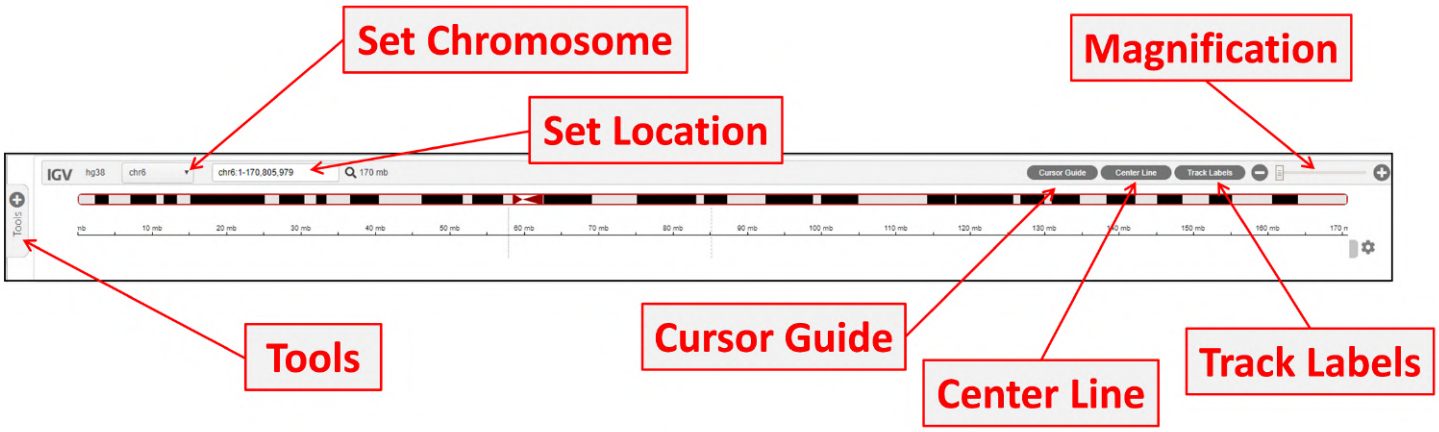


Figure: Display options for IGV

By default, the sequence viewer is centered upon the selected variant but users can drag the display upstream and downstream of the variant position. Also, it possible to zoom in and out via the magnification slider at the top of the window.

Additionally, the tracks displayed can also be modified via the setting options available on right hand side of the viewer  .

For example:

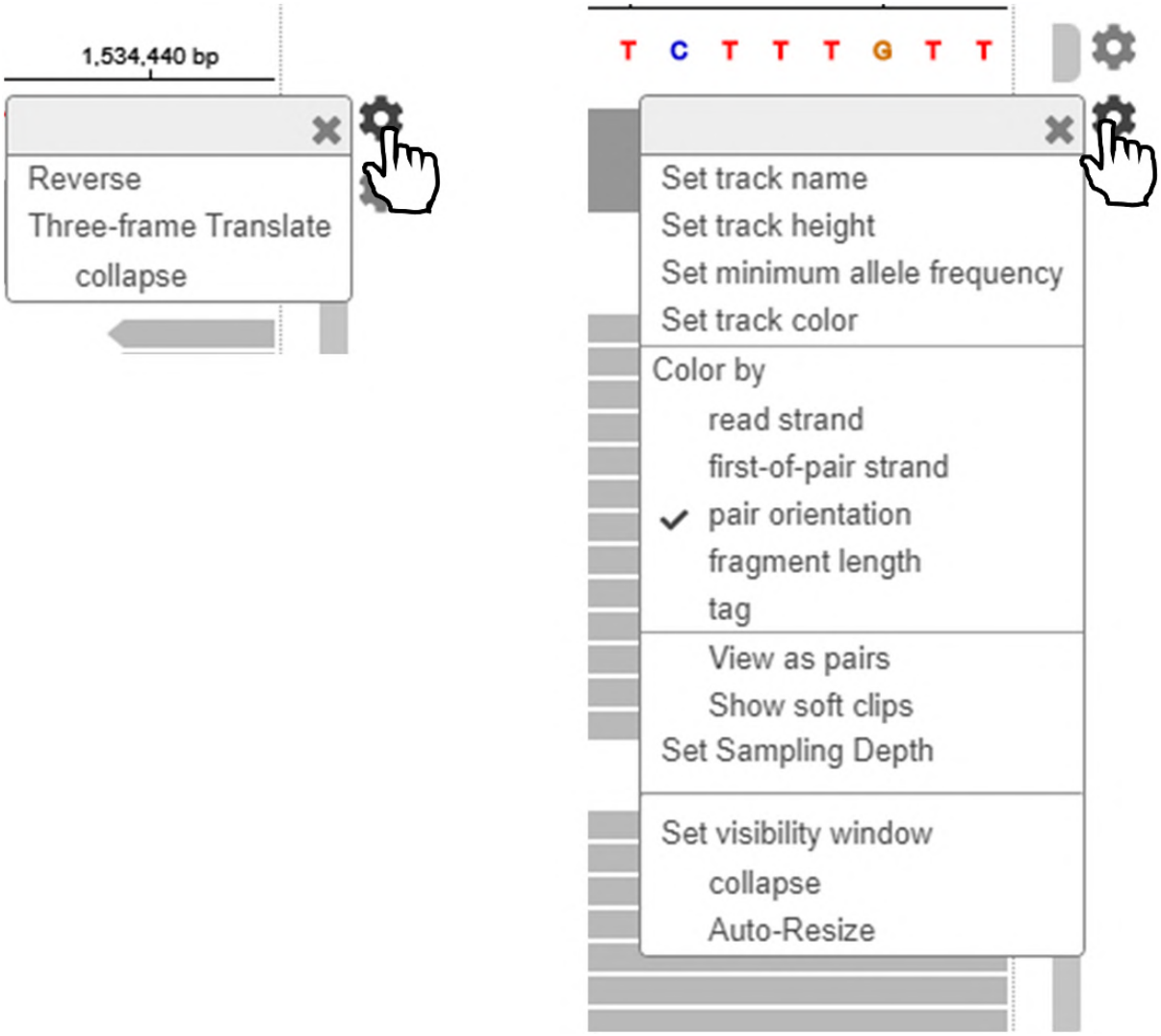


Figure: Display options for the integrated IGV browser, firstly using a left click on the mouse and secondly using a right click

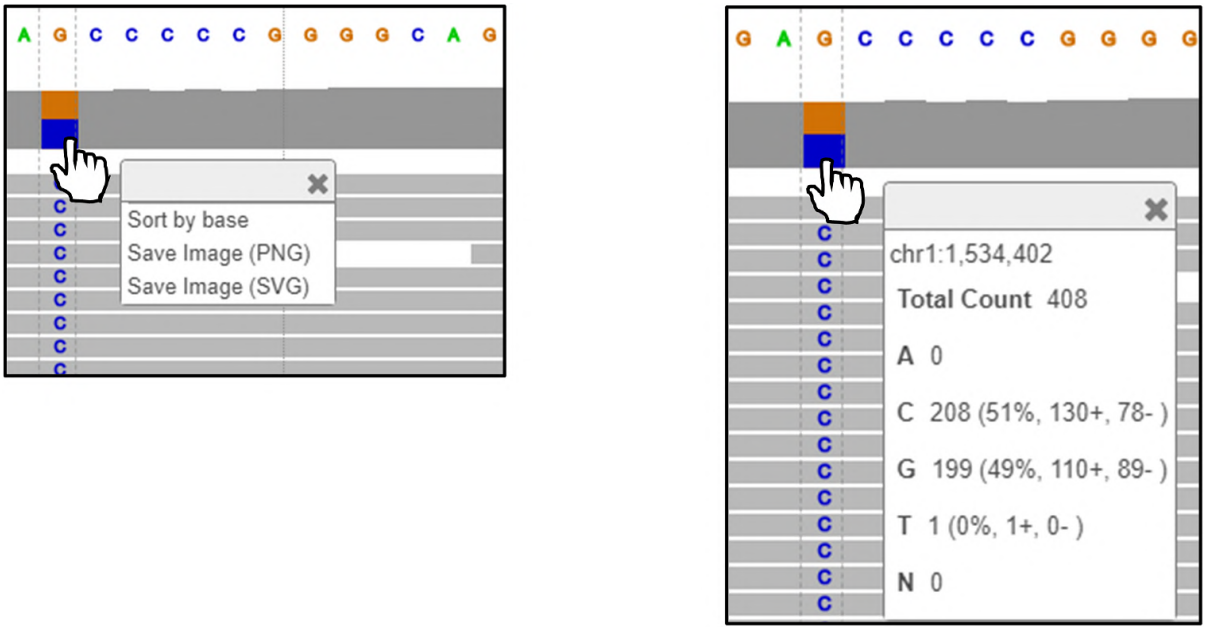


Figure: Display options available for sample reads, firstly with a left click and secondly with a right click

Using Tracks

Users can add or remove data tracks to the IGV view. This can be from publicly available sources or from proprietary internal or subscription-based sources.

Tracks can be added in the Software section of the Admin Controls (Admin Controls > Software > Annotation) and documentation of how to do this is in this section of the user guide.

To use this functionality, users need to access the Tools tab of IGV

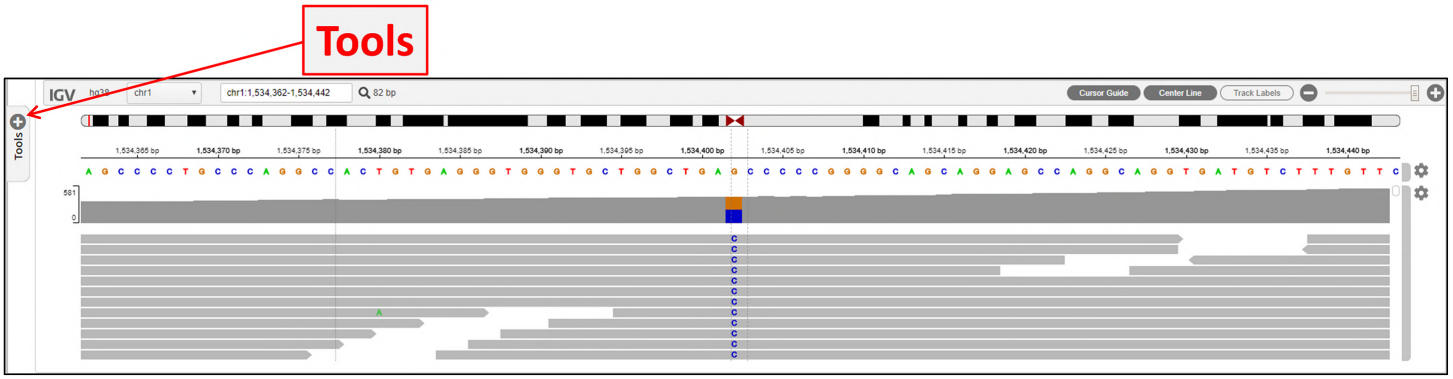


Figure: The Tools tab for adding data tracks to annotate an alignment displayed in IGV

Once accessed selecting the drop-down arrow will list the available tracks.

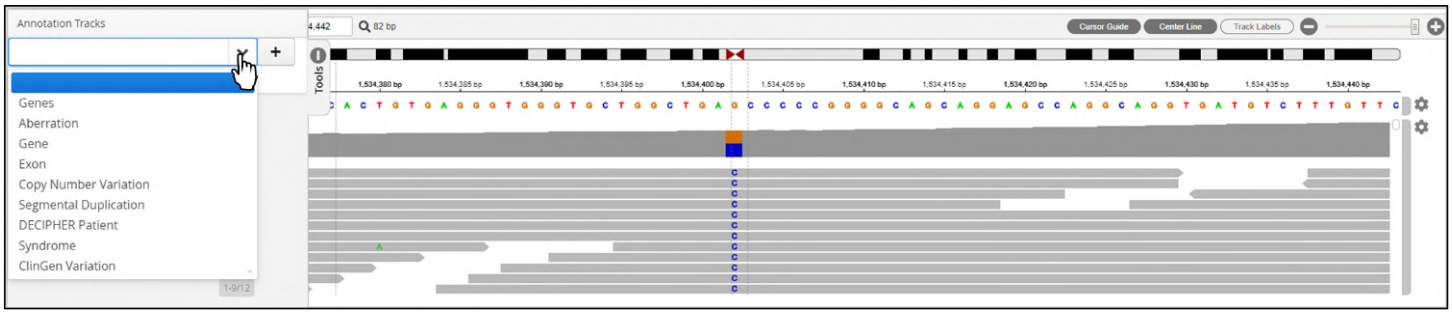


Figure: The drop-down list of data tracks available

Select the data track to be added.

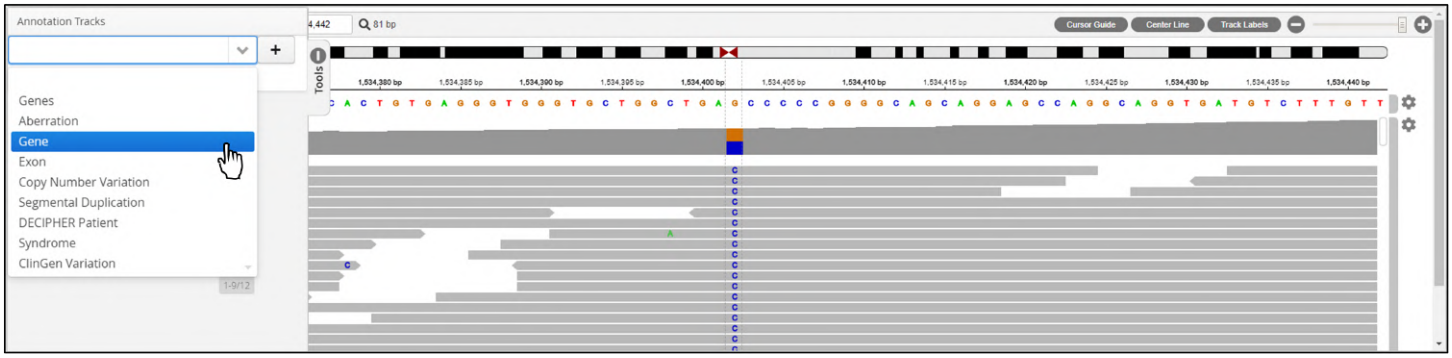


Figure: Selection of a track

And then click on the + icon to add it to the set of tracks for the software to display

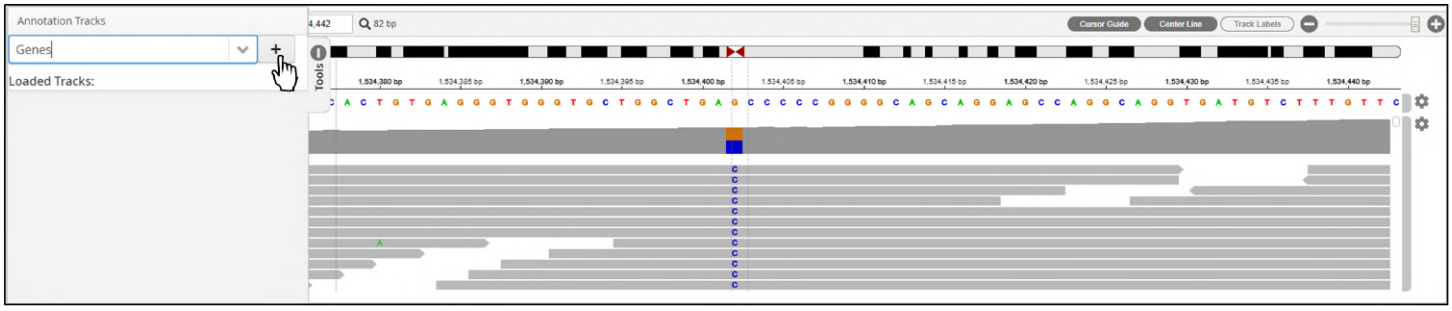


Figure: Click on the + icon to add the data track to the display

The selected track will be displayed. It can be removed by clicking on the minus icon −

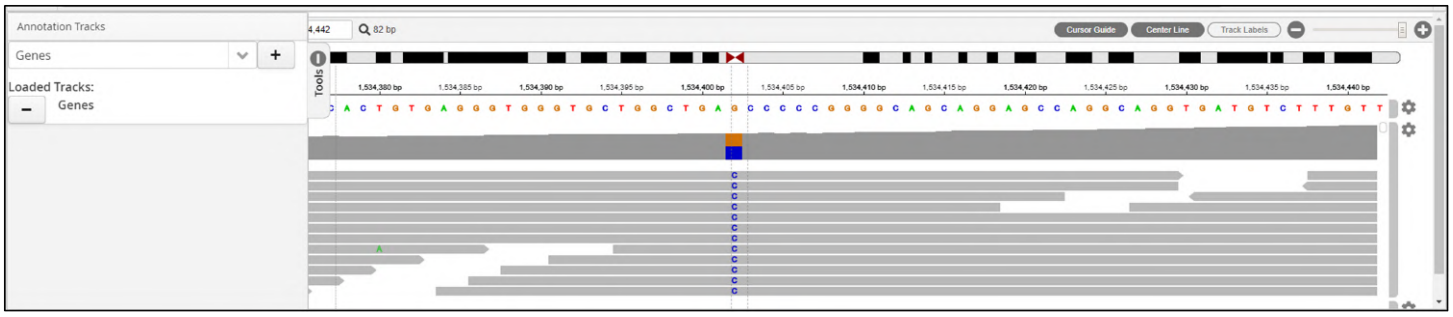


Figure: The new track is now loaded

Select any further tracks to add to the view

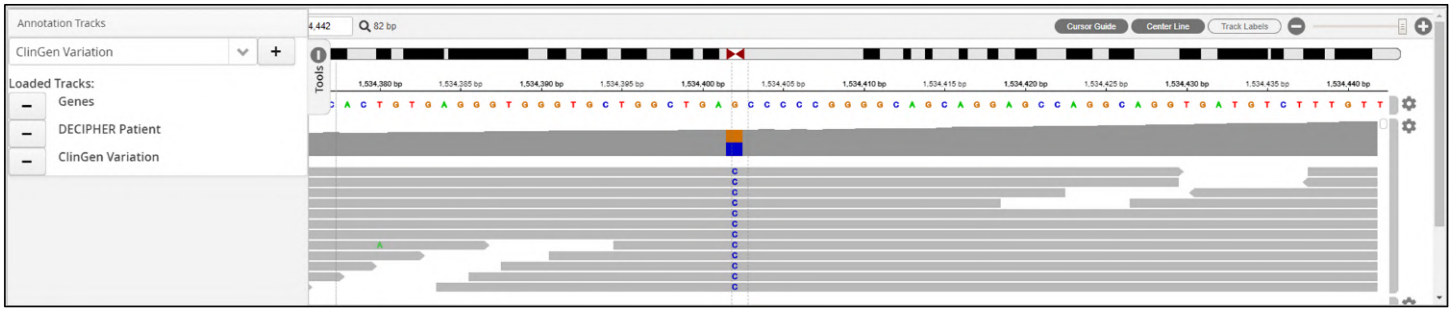


Figure: Addition of the required tracks

Finally, close the Tools tab and the data tracks will be displayed.

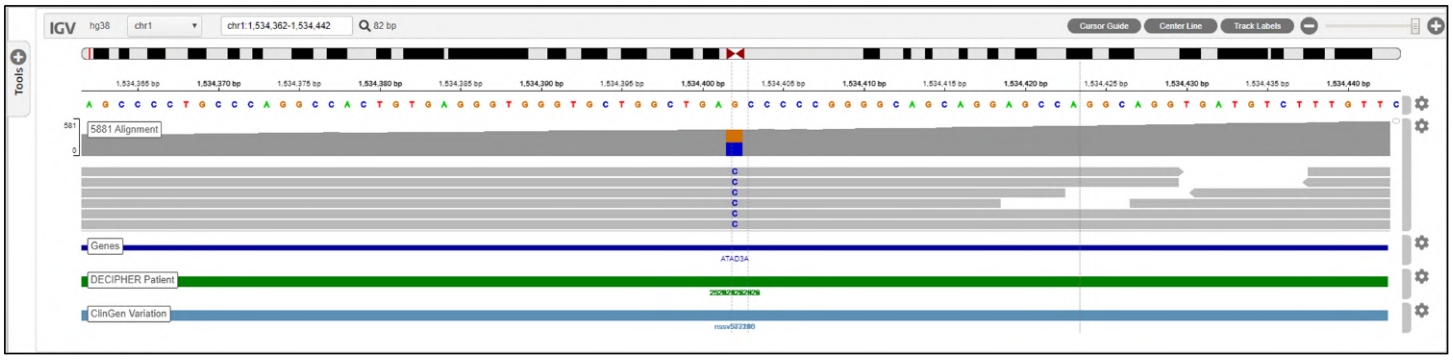


Figure: Display of the selected data tracks following closure of the tools tab

'Popping out' of the IGV display

There is, potentially, a substantial amount of information that can be displayed in the IGV view. To accommodate the information and make it easier for the user it is possible to 'pop out' the IGV view into a new browser tab.

This is accomplished using the button in the display

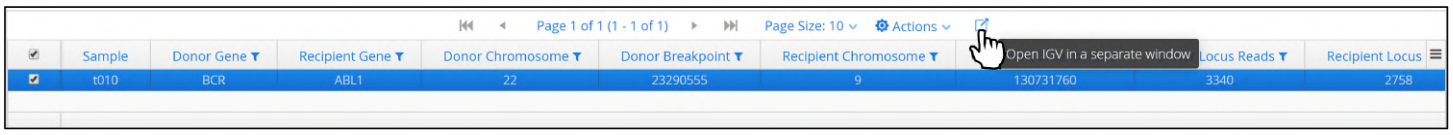


Figure: Button to allow display of IGV into a new tab in the browser

Selecting Multiple Samples

As discussed above users can opt to view multiple samples simultaneously by selecting them in the batch view.

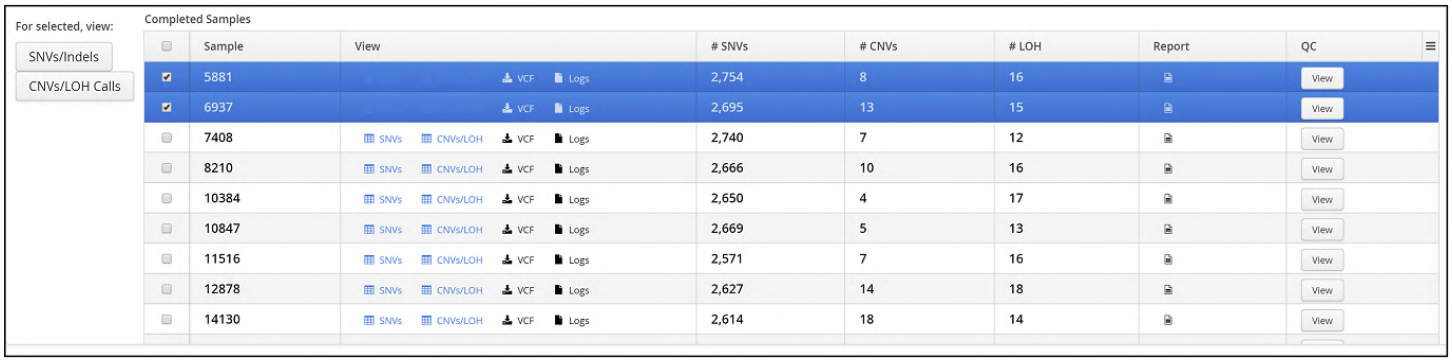


Figure: Selecting multiple samples to view in the Variants page

When multiple samples are selected there will be separate tracks for each sample in IGV. This makes it possible to compare the same variant in different samples.

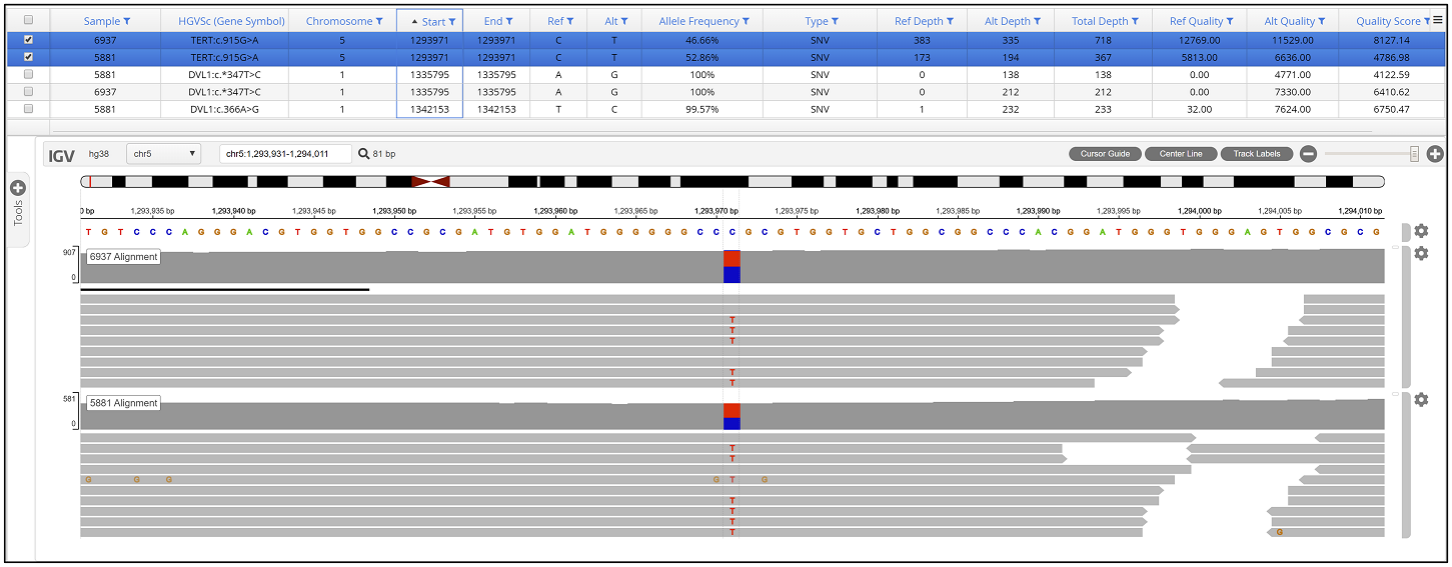


Figure: An example of two samples sharing a variant as displayed in the integrated IGV browser

Variant Table Options

There are options within the Variant table accessed using the Actions drop down menu.

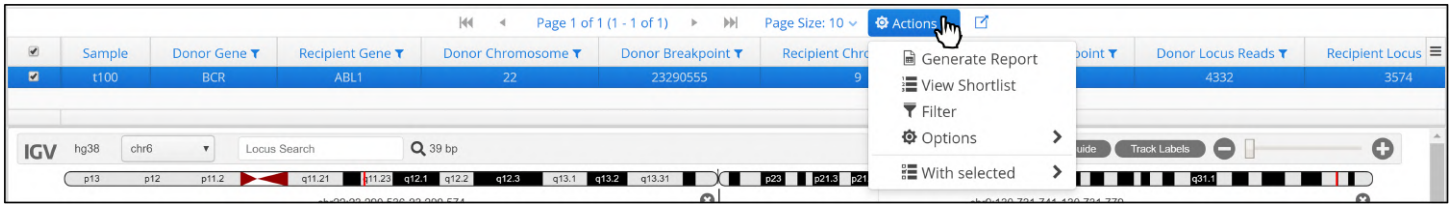


Figure: Accessing the options in the variant table

Reporting

Results can be exported by clicking on the Generate Report button below the variant table.

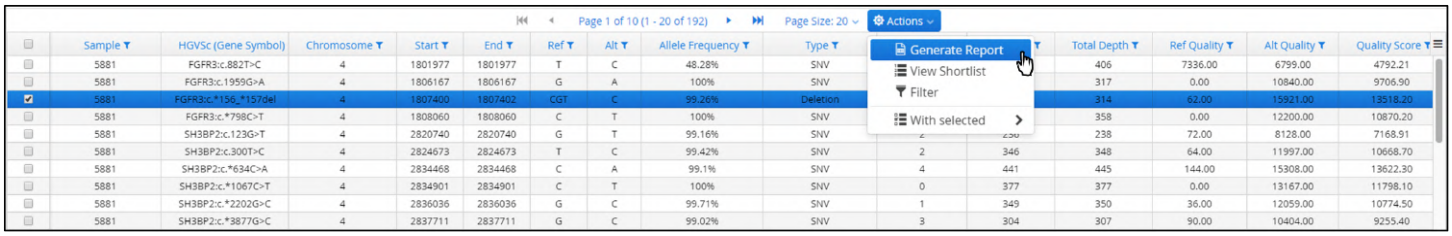


Figure: Selecting the Generate Report option from the variant table header menu

Interpret provides multiple types of report and for each of these types there are templates. These are highly customisable and updates can be easily applied in Admin Controls-Analysis-Reports.

When selecting to generate a report, the initial window allows users to select the type of report to generate.

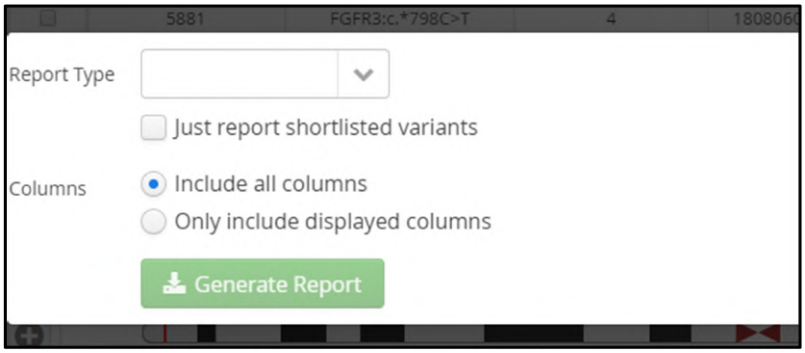


Figure: Initial report option

Default reports supplied with the software are listed.

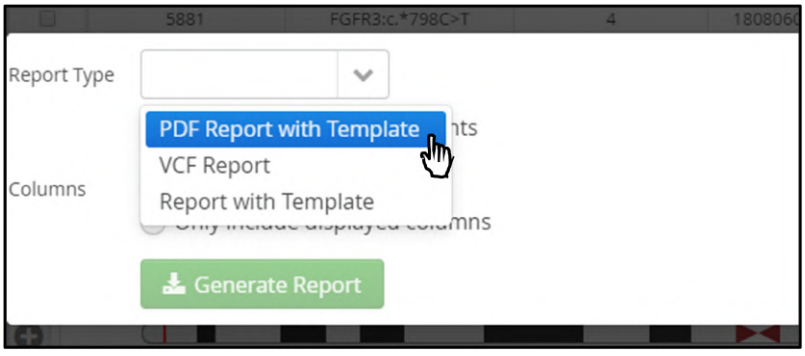


Figure: Selection of PDF report type

Once the report type has been selected the user needs to specify the template to use.

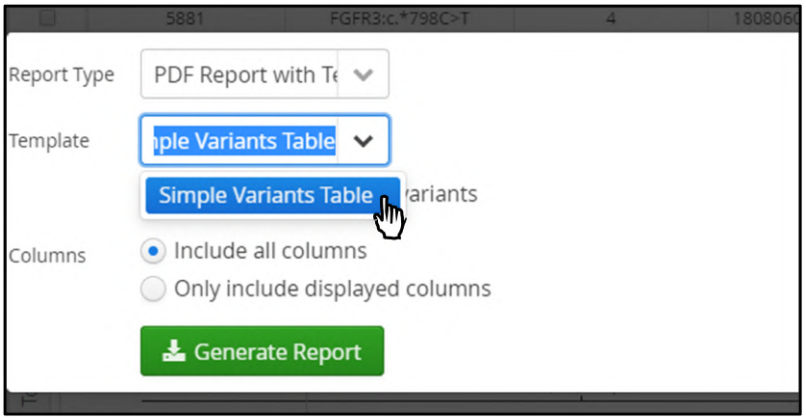


Figure: Selection of the template to use with the PDF report type

Once all options are chosen, pressing Generate Report will create the PDF file and the web browser will download it.

Sample	Chromosome	Start	End	Length	Genome Build	Ref	Alt	Type
5881	4	1801977	1801977	0b	GRCh38	T	C	SNV
5881	4	1806167	1806167	0b	GRCh38	G	A	SNV
5881	4	1807400	1807402	2b	GRCh38	CGT	C	Deletion
5881	4	1808060	1808060	0b	GRCh38	C	T	SNV
5881	4	2820740	2820740	0b	GRCh38	G	T	SNV
5881	4	2824673	2824673	0b	GRCh38	T	C	SNV
5881	4	2834468	2834468	0b	GRCh38	C	A	SNV

Figure: An example of the PDF report generated

Other options are included, for instance an HTML based report.

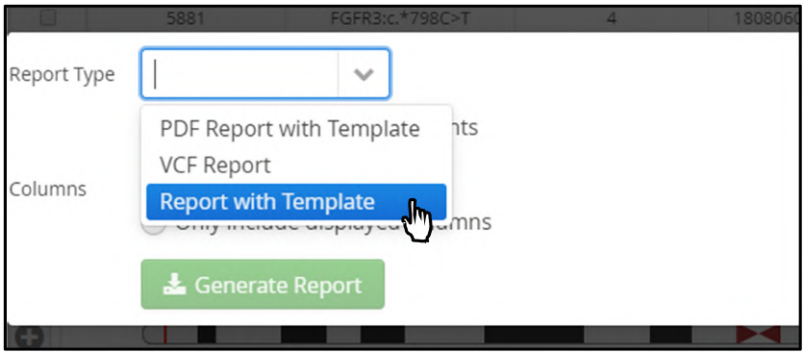


Figure: Selecting a template type report

Again, a template needs to be chosen

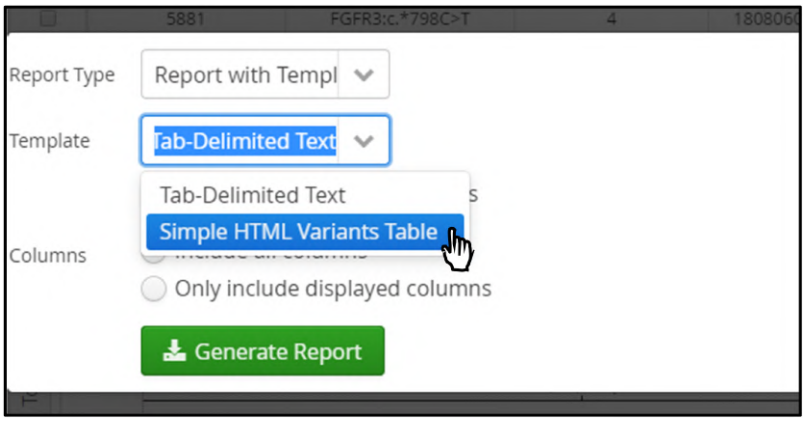
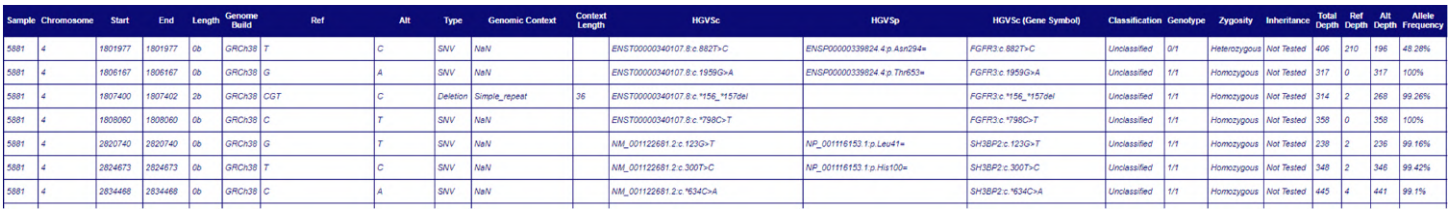


Figure: Selection of a HTML format report

The HTML formatted report is then generated and available,



Sample	Chromosome	Start	End	Length	Genome Build	Ref	Alt	Type	Genomic Context	Context Length	HGVSc	HGVSp	HGVSc (Gene Symbol)	Classification	Genotype	Zygosity	Inheritance	Total Depth	Ref Depth	Alt Depth	Allele Frequency
5881	4	1801977	1801977	0b	GRCh38	T	C	SNV	NaN		ENST00000340107.8:c.882T>C	ENSP00000339824.4:p.Asn294=	FGFR3:c.882T>C	Unclassified	0/1	Heterozygous	Not Tested	406	210	196	48.28%
5881	4	1806167	1806167	0b	GRCh38	G	A	SNV	NaN		ENST00000340107.8:c.1959G>A	ENSP00000339824.4:p.Thr653=	FGFR3:c.1959G>A	Unclassified	1/1	Homozygous	Not Tested	317	0	317	100%
5881	4	1807400	1807402	2b	GRCh38	CGT	C	Deletion	Simple_repeat	36	ENST00000340107.8:c.*156_*157del		FGFR3:c.*156_*157del	Unclassified	1/1	Homozygous	Not Tested	314	2	268	99.26%
5881	4	1808060	1808060	0b	GRCh38	C	T	SNV	NaN		ENST00000340107.8:c.*798C>T		FGFR3:c.*798C>T	Unclassified	1/1	Homozygous	Not Tested	358	0	358	100%
5881	4	2820740	2820740	0b	GRCh38	G	T	SNV	NaN		NM_001122681.2:c.123G>T	NP_001116153.1:p.Leu41=	SH3BP2:c.123G>T	Unclassified	1/1	Homozygous	Not Tested	238	2	236	99.16%
5881	4	2824673	2824673	0b	GRCh38	T	C	SNV	NaN		NM_001122681.2:c.300T>C	NP_001116153.1:p.His100=	SH3BP2:c.300T>C	Unclassified	1/1	Homozygous	Not Tested	348	2	346	99.42%
5881	4	2834468	2834468	0b	GRCh38	C	A	SNV	NaN		NM_001122681.2:c.*634C>A		SH3BP2:c.*634C>A	Unclassified	1/1	Homozygous	Not Tested	445	4	441	99.1%

Figure: An example of the HTML report

Actions

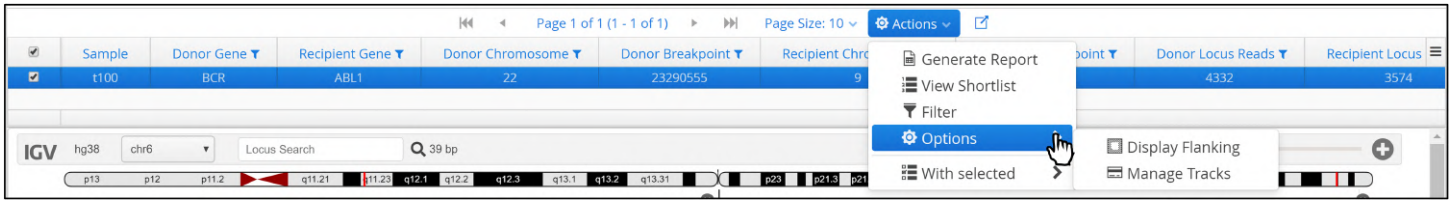


Figure: Options available for configuring the view in IGV

Display Flanking

Users can choose whether or not to display flanking sequence in the IGV display.

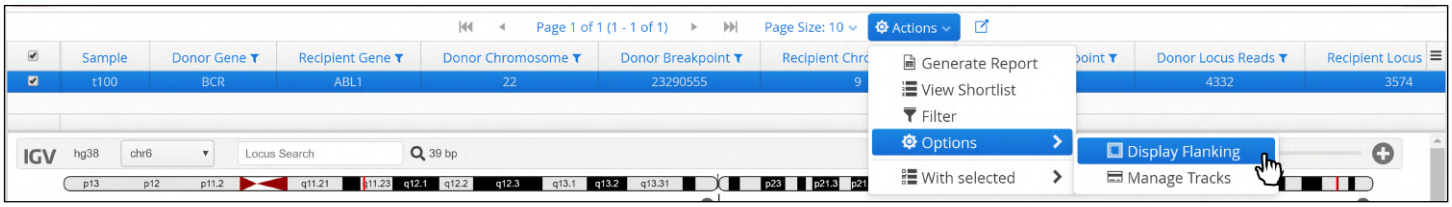


Figure: Selecting the Display flanking option in the Actions menu.

Manage Tracks

Users can add or remove data tracks to the IGV view. This can be from publicly available sources or from proprietary internal or subscription-based sources.

Tracks can be added in the Software section of the Admin Controls (Admin Controls > Software > Annotation) and documentation of how to do this is in this section of the user guide.

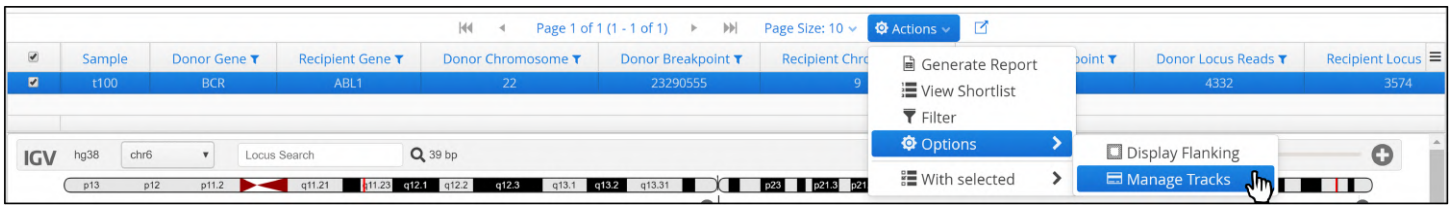


Figure: Selecting the manage tracks options

The available tracks will be displayed in a pop-up window and users can select the tracks that they want to add to the display.

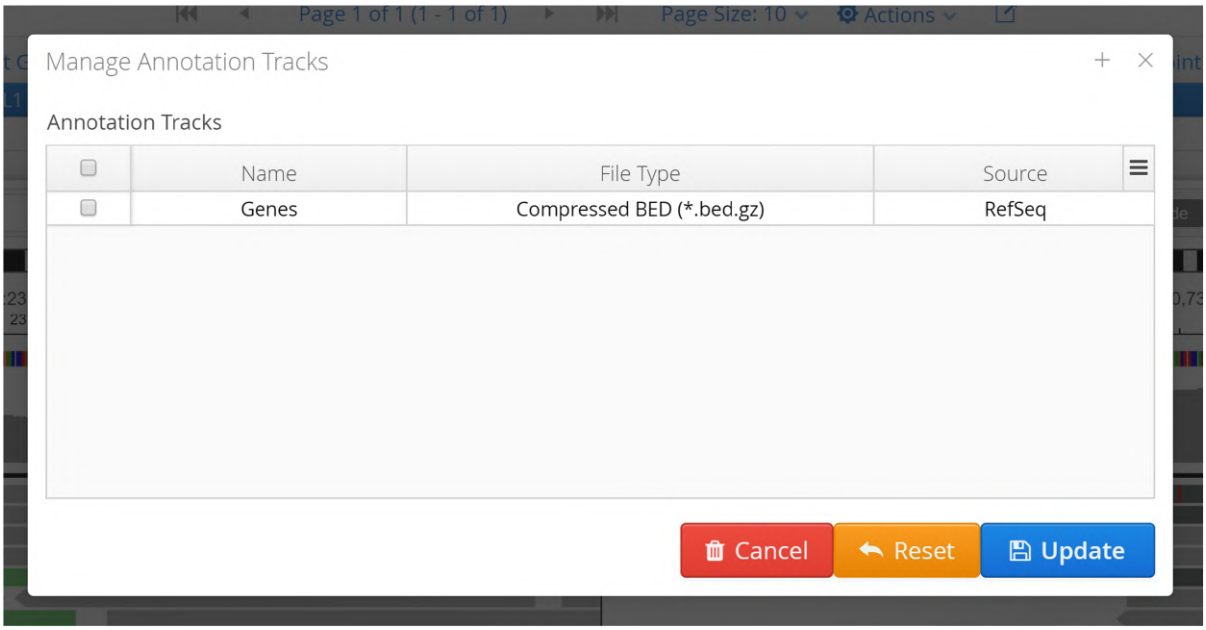


Figure: Tracks available to display

Once the required tracks are selected, users can press Update to update the IGV display.

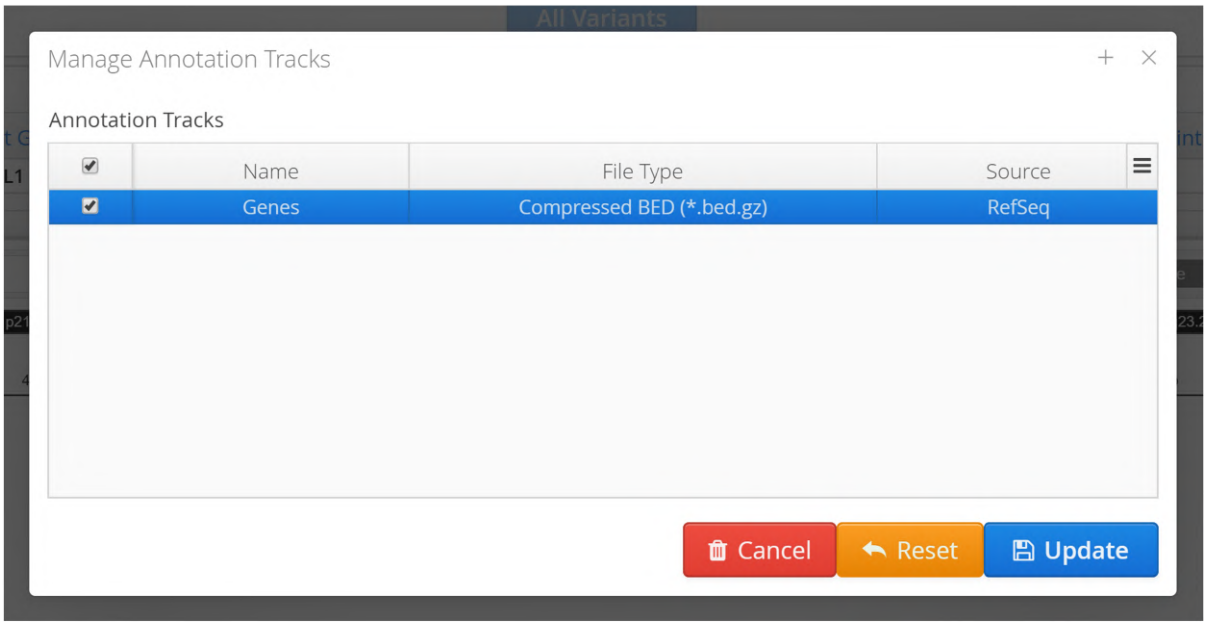


Figure: Selecting tracks to add to the IGV display

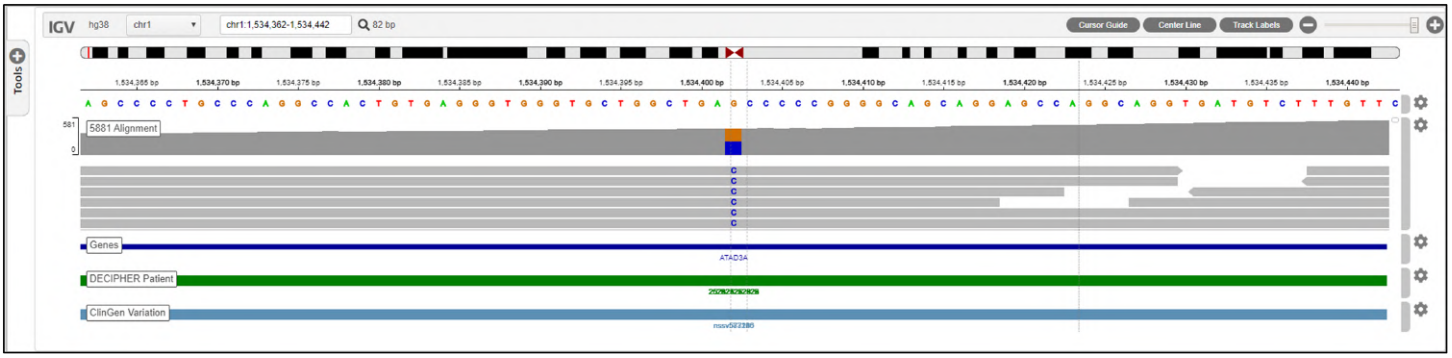


Figure: Displaying of tracks in the IGV display

12 Viewing Analysis Results By Variant

As results of samples are generated, they are stored in the Interpret database and can be analysed from a variant-centric point of view

Accessing of this viewpoint is via the Variants button on the dashboard menu bar shown in the figure below.

Figure: Selection of Variants from the Dashboard menu bar

Selecting the Variants tab in the menu bar opens up a new page to display all the variants recorded in the database.

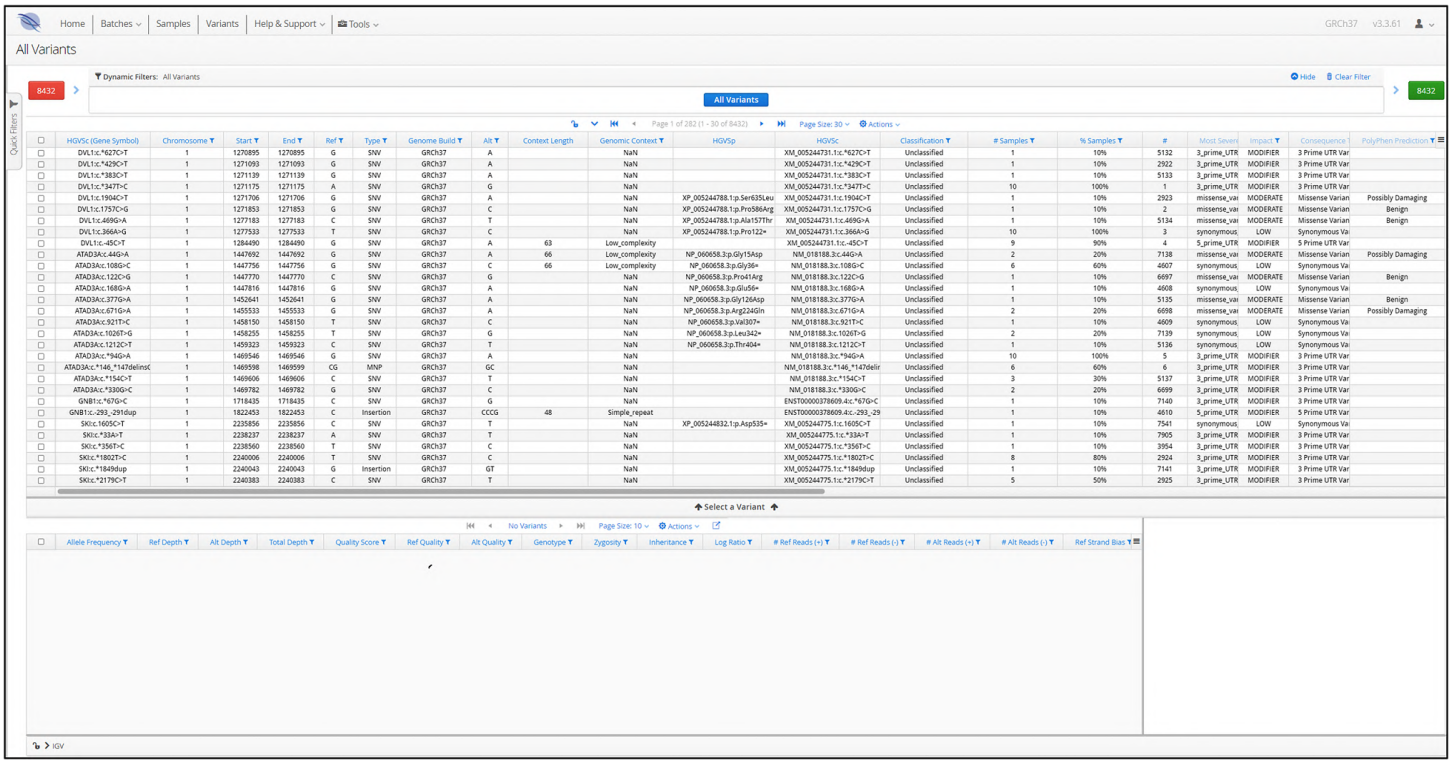


Figure: The start page for viewing variants

There is a substantial amount of information available in the variants page and the different sections are highlighted in the figure below.

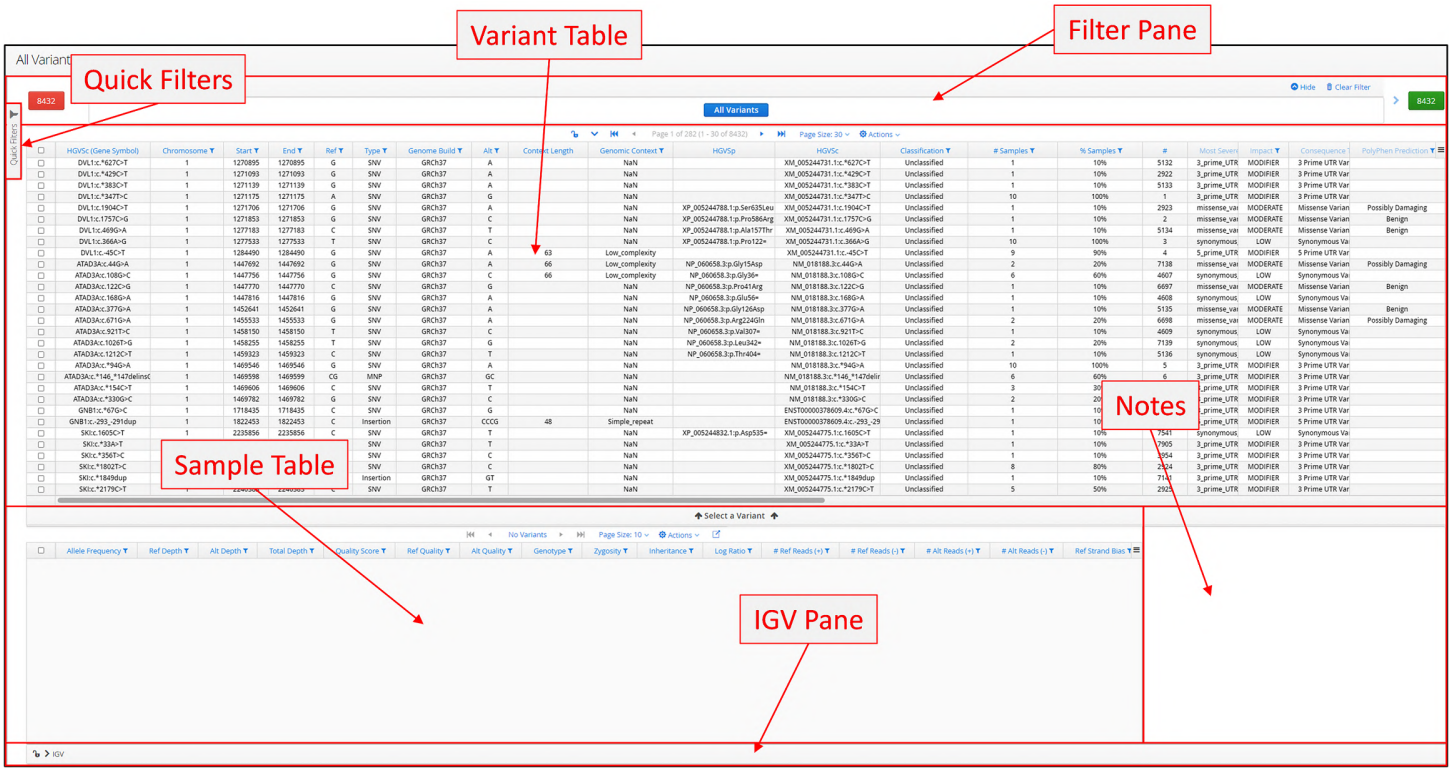


Figure: The different sections of the variant page

There a number of active regions

- Filter Pane

The filter pane allows for dynamic filtering of the variants. By default the filter is set to All Variants, so all variants are displayed in the variants table, however these can be refined according to your specific requirements.

- Variant Table

This displays all variants in the database that meet the filtering requirements of the dynamic filter. By default this is for displaying all variants.

- Sample Table

When a row in the variant table is selected all samples that contain the selected variant will be displayed in this table.

- IGV Pane

Selection of samples in the

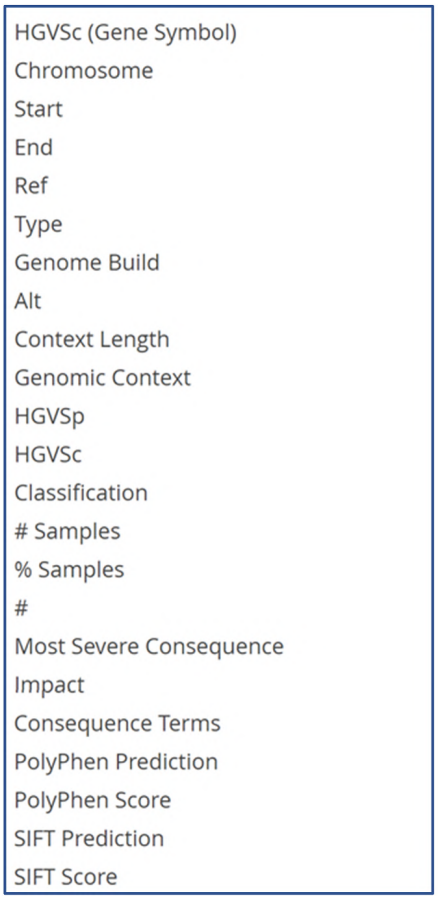
- Notes

Users can add notes to variants.

- Quick Filters

These are selection of options that allow users to rapidly filter variants on the basis of some general conditions.

Clicking on the ≡ icon on far right of the column headers in the variant table will display all the columns that can be selected for display in the variant table.



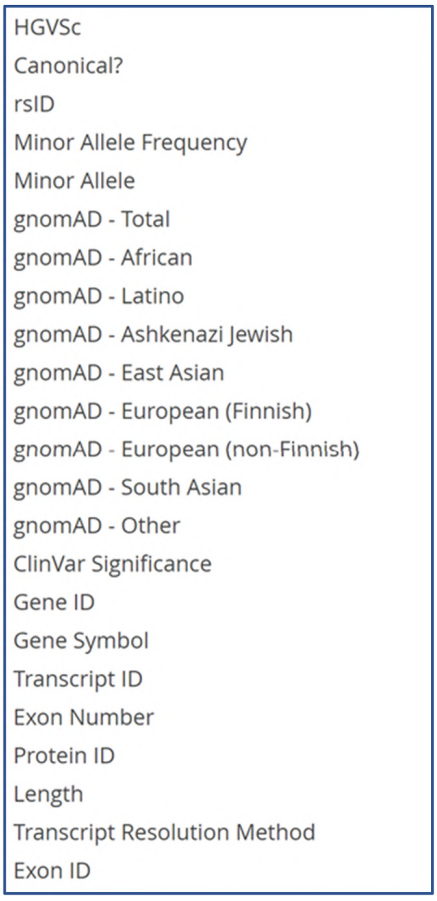


Figure: Columns available to select for display in the variant table

Similarly clicking on the same icon on the sample table provides a series of column options.

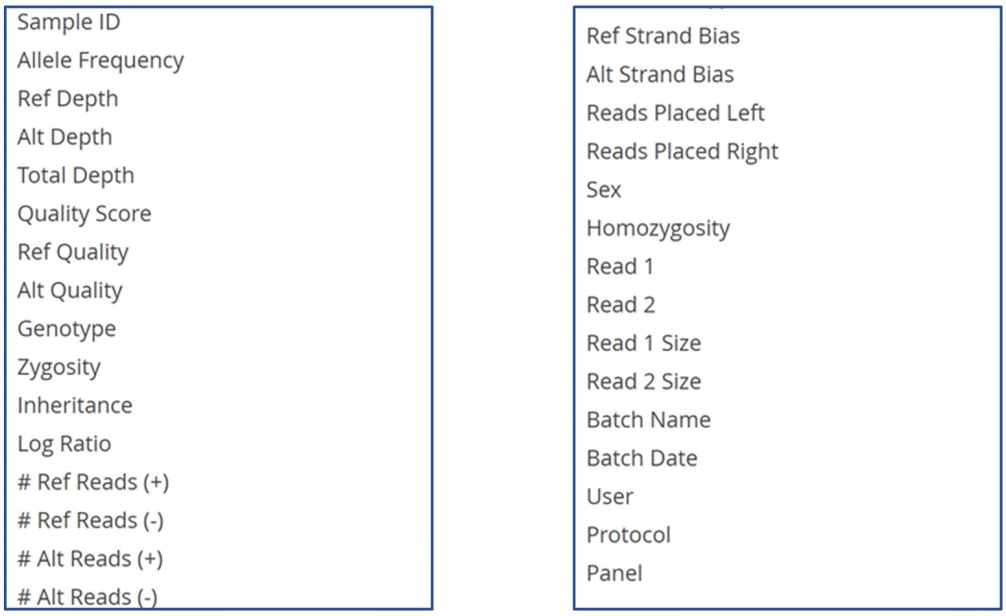


Figure: Columns available to select for display in the sample table

Filtering Variants

Add a section on filtering variants either using the dynamic filtering or the quick filter tools.

Dynamic Filtering

The dynamic filters are the same set of options discussed in the previous section. They can be accessed from the Actions drop down menu shown below.

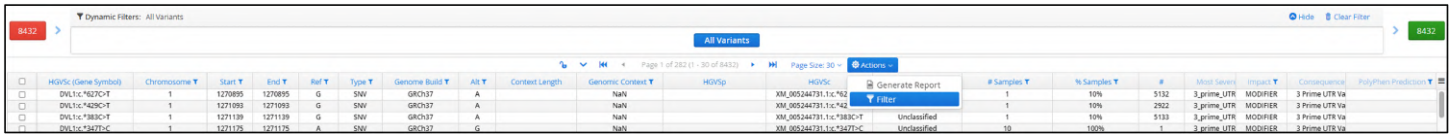


Figure: Accessing the dynamic filtering options

The dynamic filtering window provides a detailed set of options for investigating the variants stored within the Interpret database.

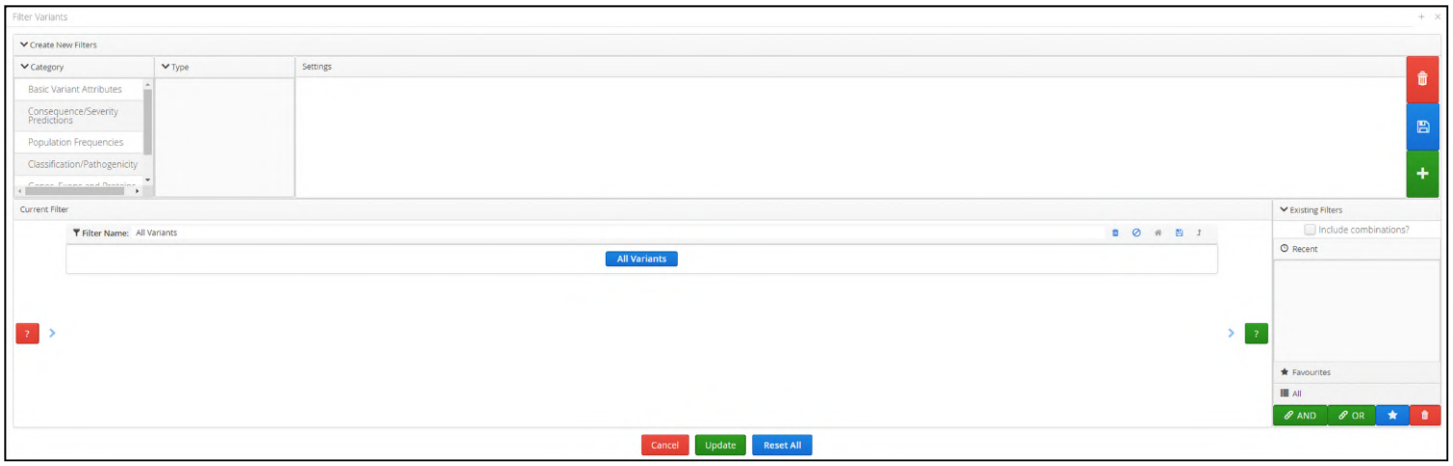


Figure: Full filtering options available to filter variants

Filtering by Quick Filters

Selecting the Quick Filters tab on the side of the variants page opens a tab that provides some options for quickly drilling down into variants of interest. The options currently available are to select based on a classification type, the NGS panel or gene.

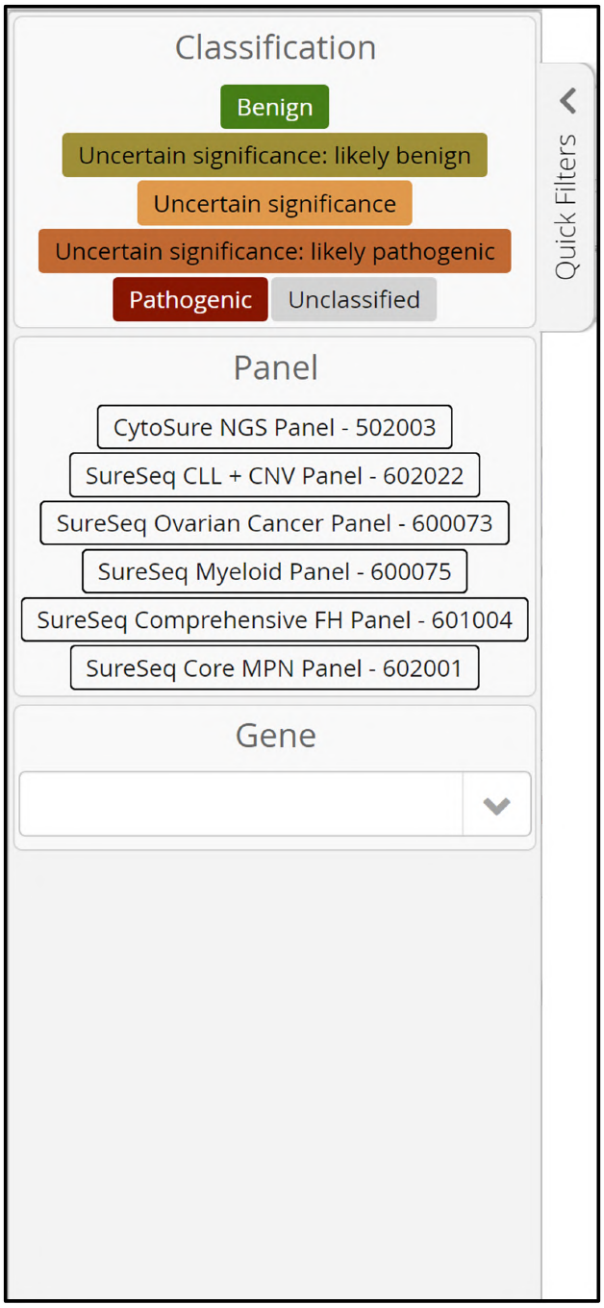


Figure: Quick filter options

Classification or Panel can be selected by pressing the corresponding buttons with multiple selections allowed. To filter by genes start typing the gene name in the text box matching values will be displayed.

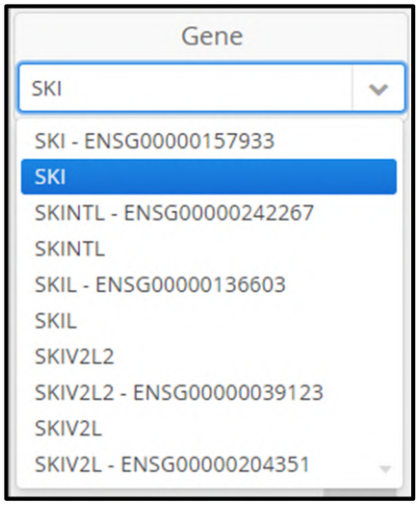


Figure: Using quick filters to select for the gene SKI

Once a gene is selected it will be displayed as below and can be removed by clicking on the x next to the gene name.

Figure: Using quick filters to filter by the SKI gene

The dynamic filter is now updated and shows that, from the input of 8432 variants, there are only 10 found within the SKI gene.

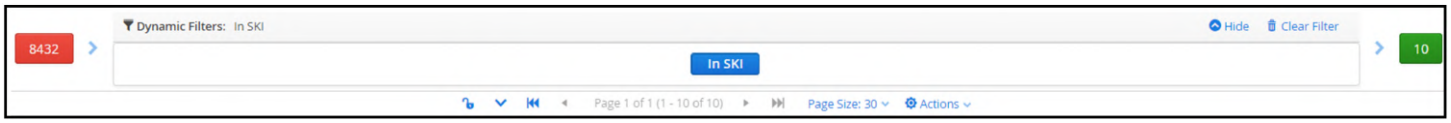


Figure: Displaying only variants in the SKI gene

Additional genes can be selected and these will be displayed in the same way.

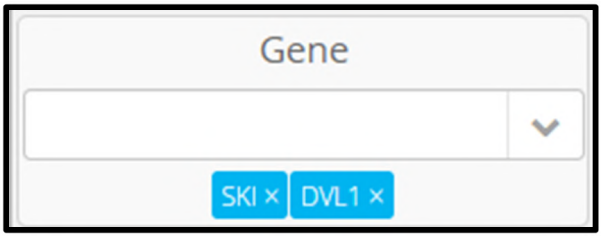


Figure: Using quick filters to select variants in the SKI or DVL1 genes

When the 2-gene filter is applied the output now increases to 19 variants being displayed.

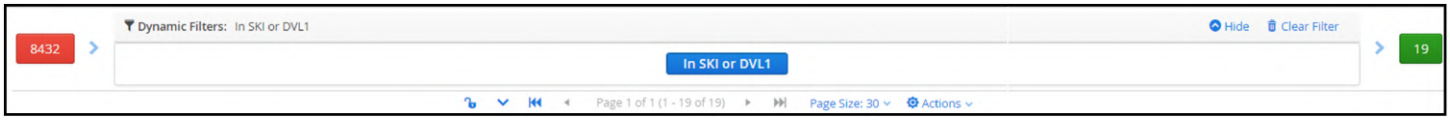


Figure: Displaying only variants in the SKI or DVL1 genes

Displaying Variants

Each row of the variant table represents a variant that has been detected in at least one sample. Selecting a variant displays all the samples in which the variant is present. In the figure below, the variant DVL1:c.*347T>C is present in 10 samples.

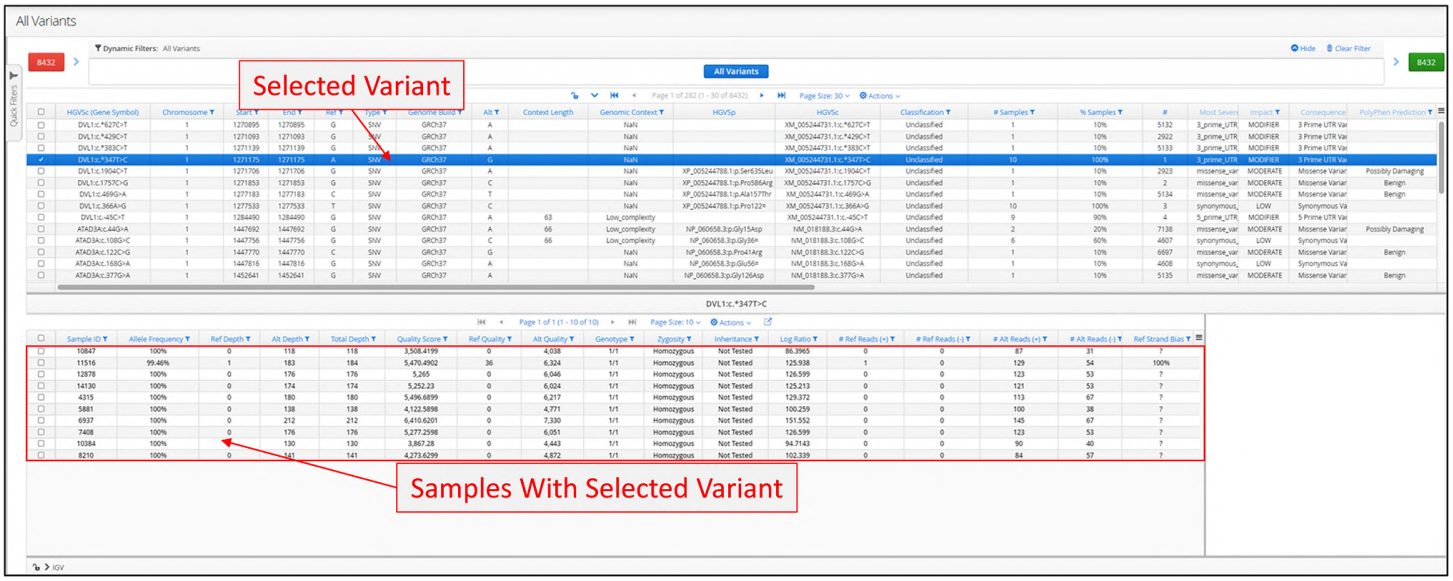


Figure: Selecting a variant displays all samples in which it is present

Subsequently, selecting any of the sample or samples rows will display the alignment for the variant in the corresponding sample.

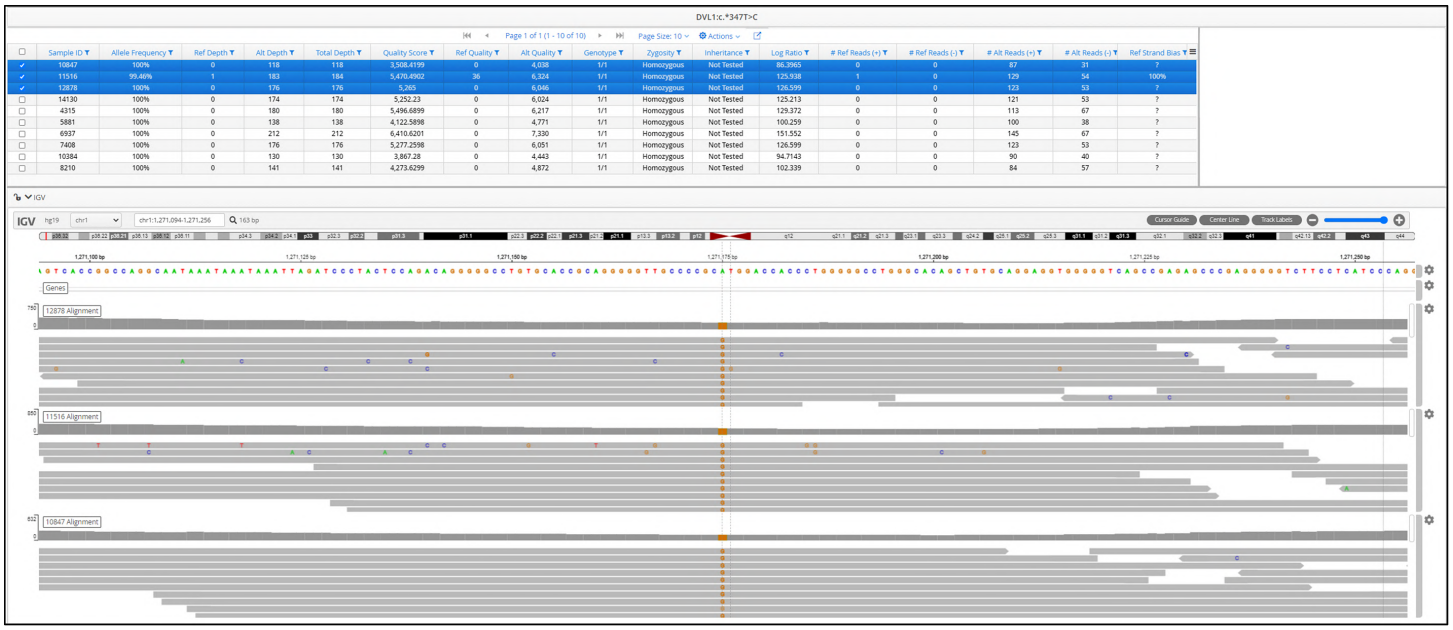


Figure: Display of the alignment for 3 samples in IGV

Adding Notes to Variants

It is possible to for users to add annotations to variants through the notes function. When a variant has been selected and there are rows populated in the sample table, the user can make a right click on one of these. From the popup menu select the Notes > Add Note options

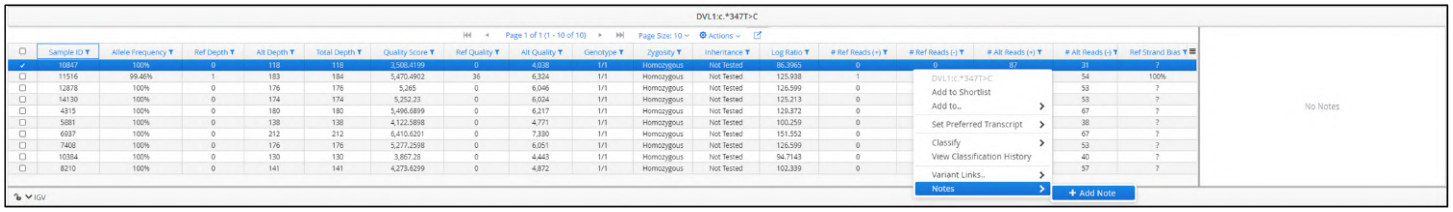


Figure: Selecting the Add Note option

A window is displayed with a text box where up to 250 characters can be used. Any other pre-existing notes will also be shown.

Figure: Note creation template

The user can enter the required notation.

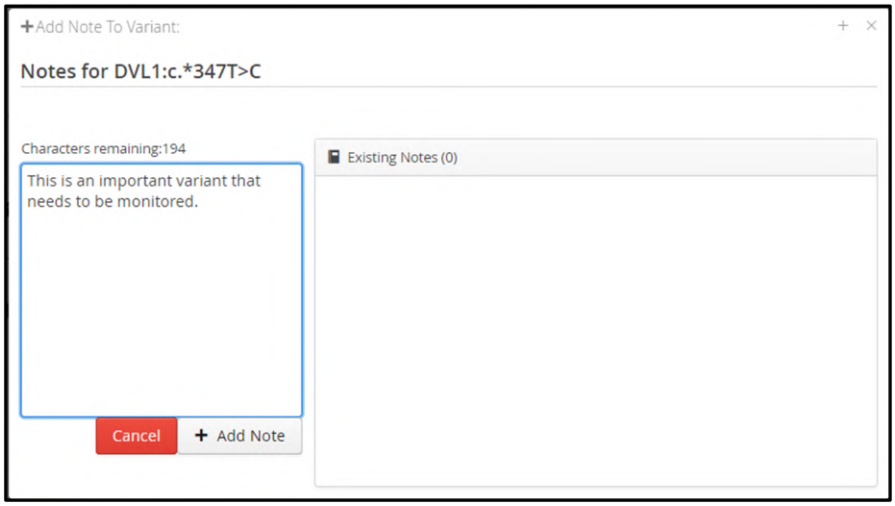


Figure: Note creation

Then selecting + Add Note completes the process and the existing notes section is updated to include the newly created note.

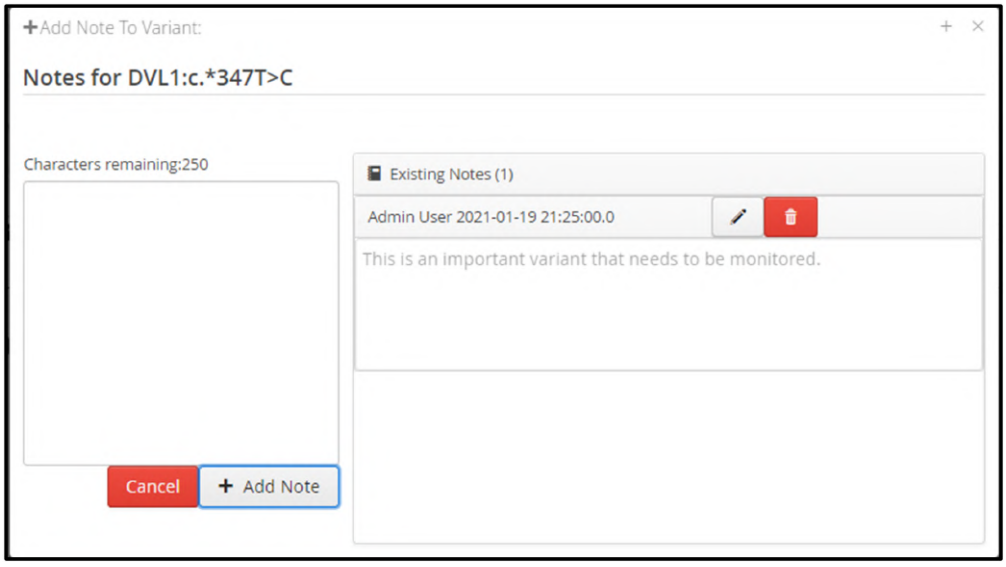


Figure: Note generation

Once all changes have been made, the notes window can be closed and the view will return the normal variant display with the note now being displayed in the Notes panel.

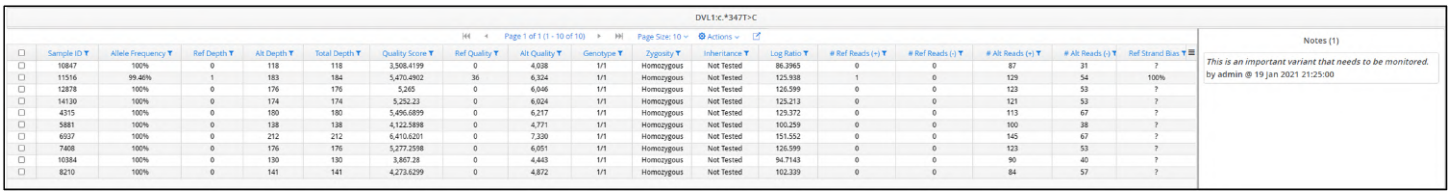


Figure: Displaying a note in the note panel

A note can be deleted by pressing the red rubbish bin icon. If the Confirm Delete option is then selected the note will be removed.

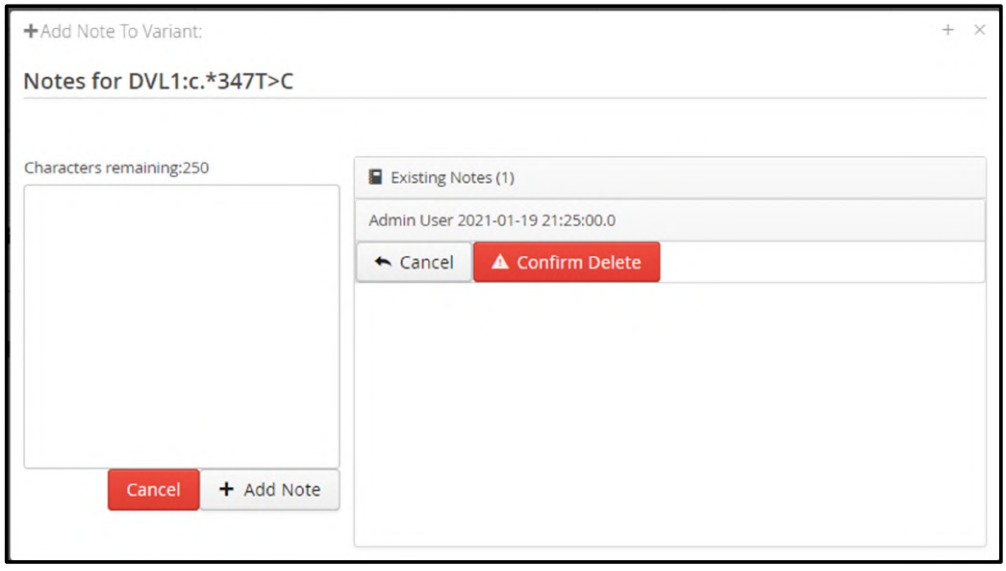


Figure: Deleting a note

Users can also edit a note by clicking on the pen icon; which will show the note in a text box where changes can be made. The update is confirmed by the pressing the Apply button.

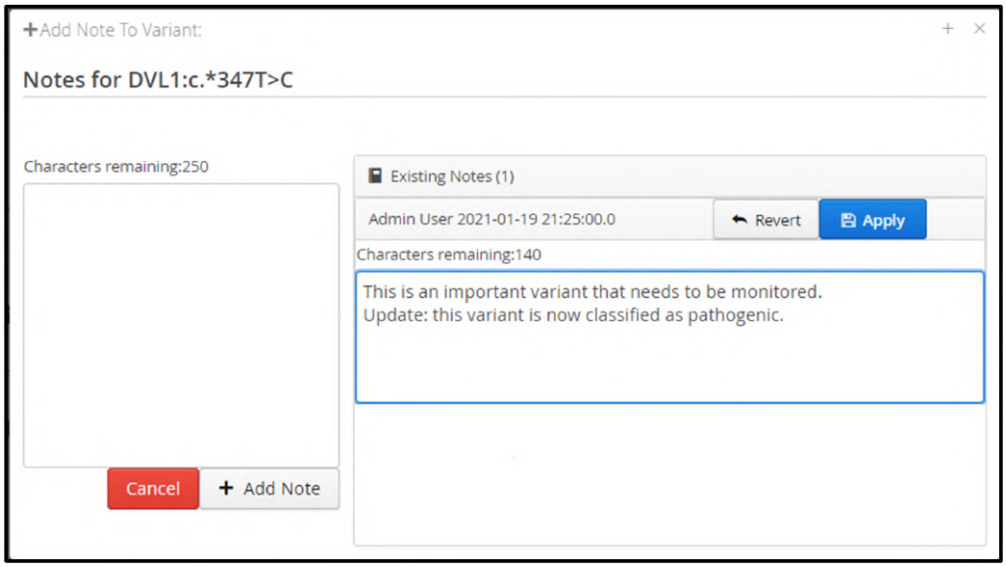


Figure: Update an existing note

The notes panel is subsequently updated to show the revised note.

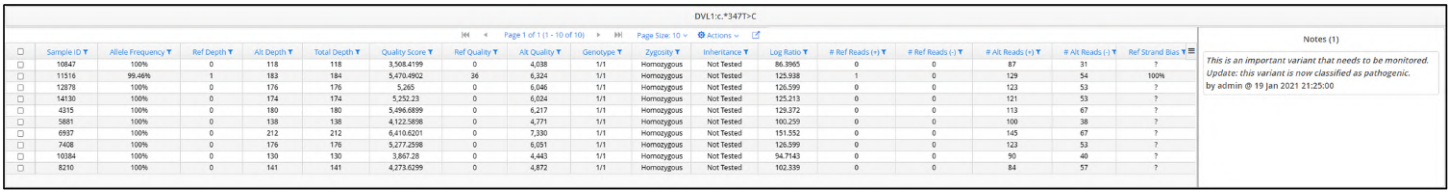


Figure: Display of an updated note

13 Product-specific Analysis

13.1 Minimal Residual Disease

Detection and monitoring of Minimal Residual Disease (MRD) with the SureSeq Myeloid MRD Panel is made possible in Interpret through:

1. The ability to specify "Hotspots" (variants) which should be specifically interrogated by the pipeline for their presence at very low frequency.
2. The ability to visualise the change in allele frequency of these hotspots in multiple sequencing runs over time.

13.1.1 Hotspot specification

Hotspots to be interrogated by the pipeline should be defined as part of the analysis protocol in Admin Controls -> Analysis -> Protocols:

1. **Hotspot Monitoring** must be enabled in the Optional Capabilities section of the protocol editing/creation screen.

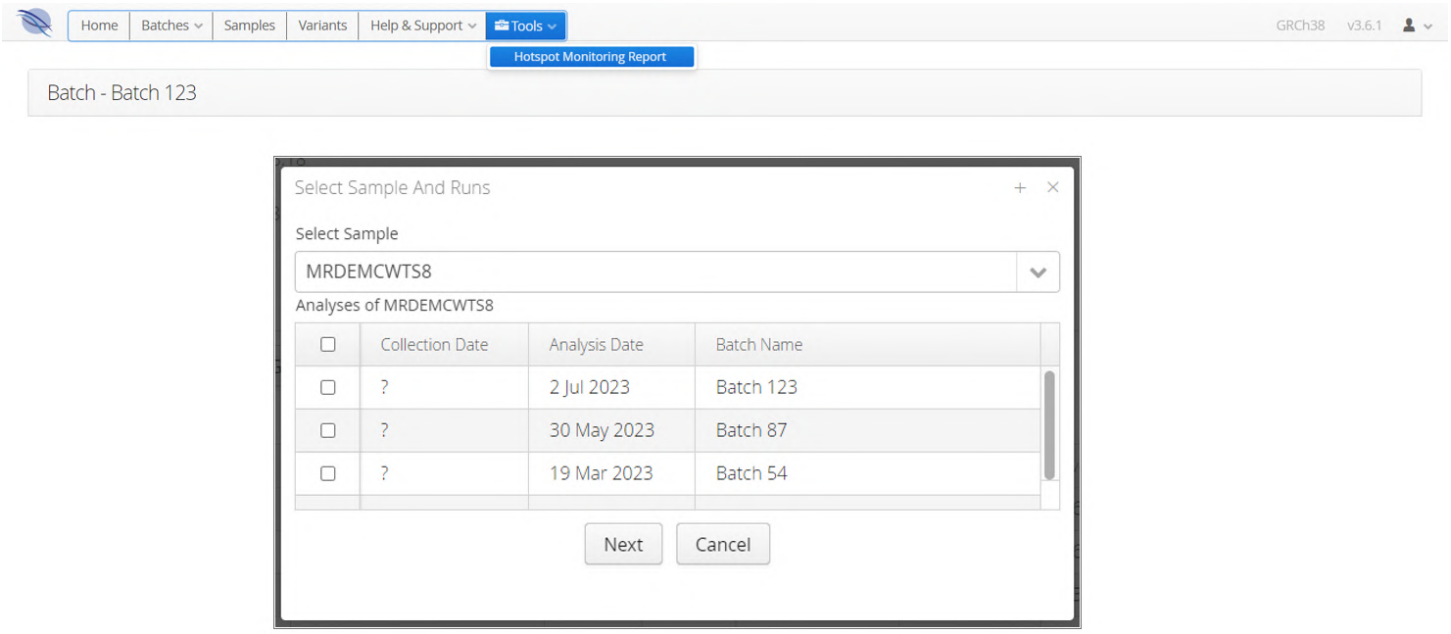
2. The user must also specify the hotspots/variants to be monitored via the Hotpots section of the Basic Pipeline Configuration by selecting from variants which have already been added to a Variant List .

A Variant List , from which hotspots may be selected, can be created in two ways:

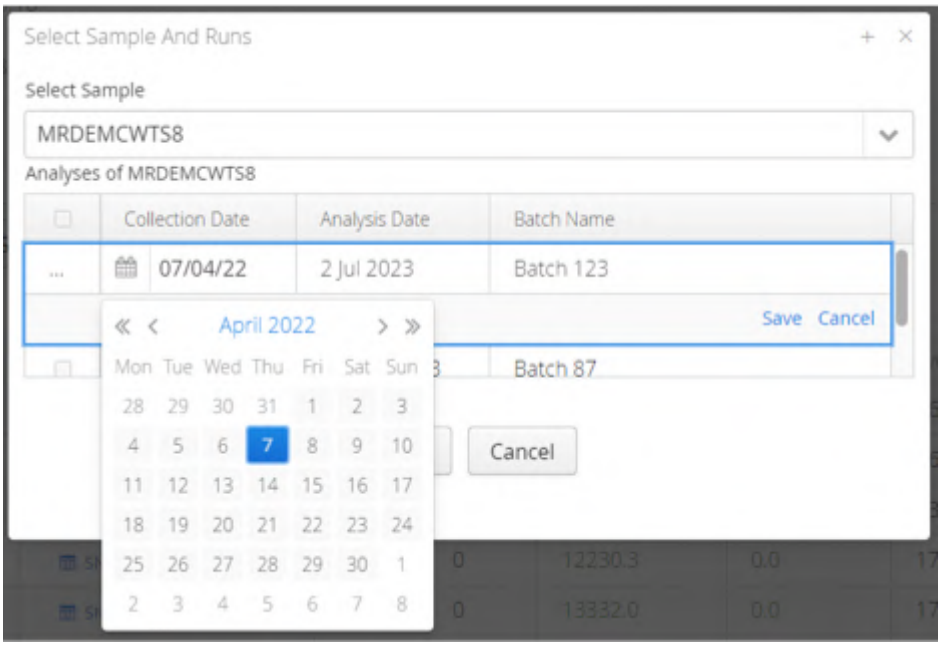
1. A variant which has already been detected by the software can be manually added to the list by right-clicking on the variant in the variants table, selecting Add to... and either selecting New List, to add the variant to a list variant list, or clicking on the name of an existing variant list.
2. A file containing a list of hotspots (either as a VCF or a line-separated list of HGVSc) can be uploaded via Admin Controls -> Analysis -> Variant Lists .

13.1.2 Hotspot Monitoring

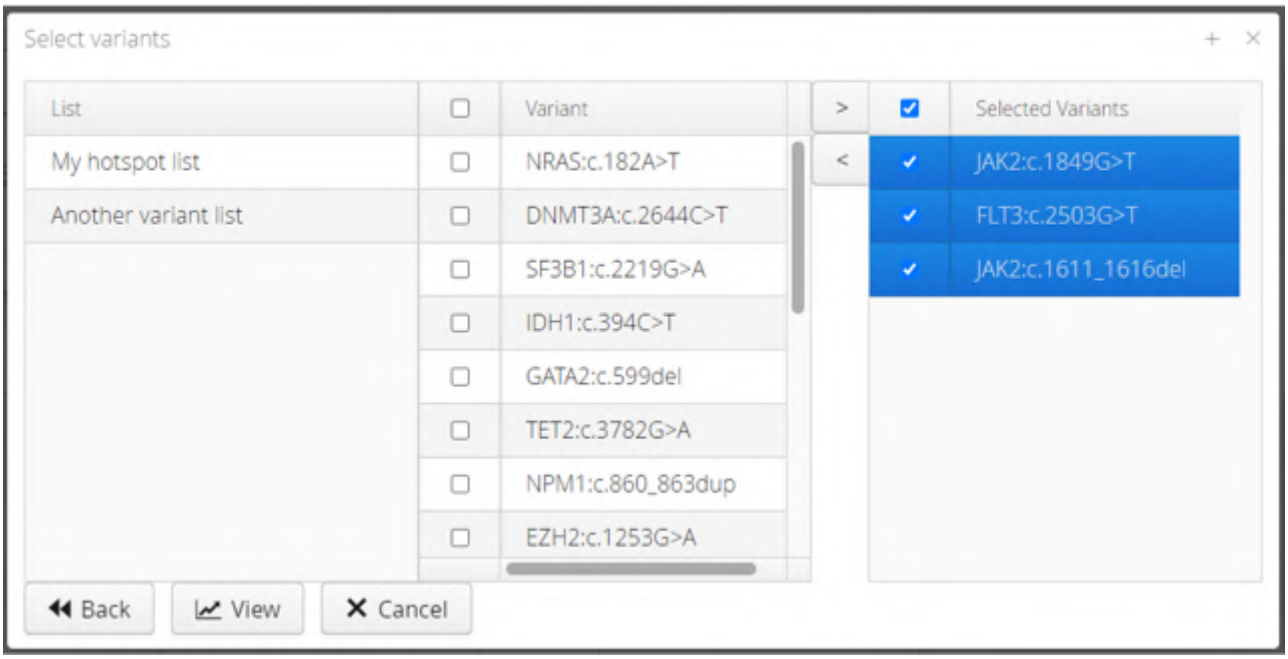
In order to visualise the results of hotspot monitoring, select Tools -> Hotspot Monitoring Report , and select the sample/patient to be reported.



If necessary, enter the collection date of the sample(s). This only needs to be carried out once for each sample and will be remembered for future reports.



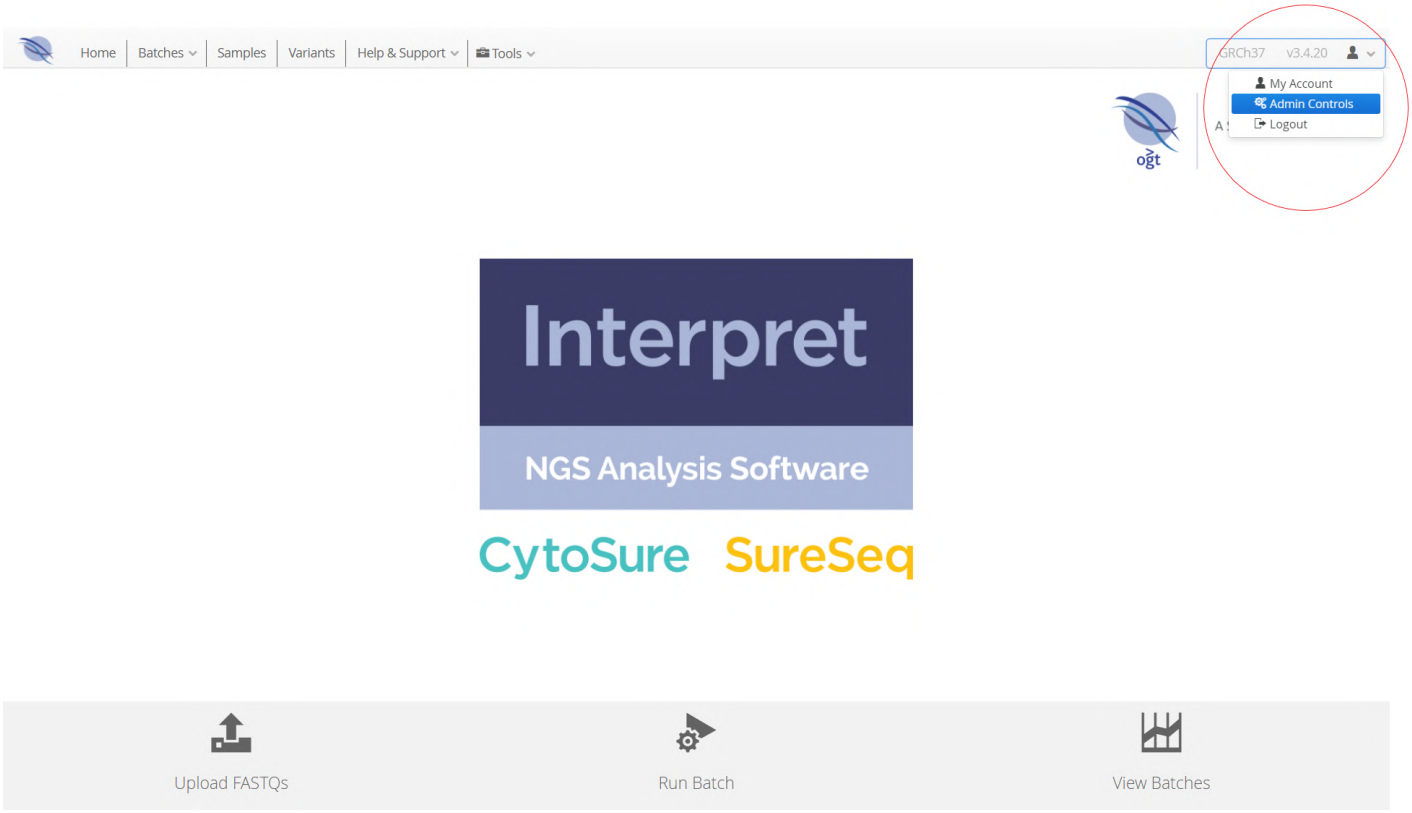
Select the hotspots to be reported.



A graph containing the allele frequencies of all selected hotspots in all selected sample runs will be displayed, along with tabs allowing the user to view the results for individual hotspots. Graph images may be exported by right-clicking of the graph and selecting “Save Image”. The table containing the data underlying the graph may be exported as a CSV via the Export Data button.

14 Administration Controls

These are accessed by selecting the Admin Controls options in the user drop down menu.



There are 4 main parts to the admin controls:

1. Overview
2. User Controls
    - Current Users
    - Add Users
3. Analysis
    - Manage Samples
    - Current Analyses
    - Protocols
    - Panels
    - Region Lists
    - Variant Lists
    - Classifications
    - Metric Sets
    - Manage Links
    - Filters
    - Preferred Transcripts
    - Reports
    - Guidelines
4. Software
    - Software Overview
    - Annotation
    - Advanced Settings
    - Plug-ins

14.1 Administration Overview

The administration overview provides a view of the latest activity in Interpret

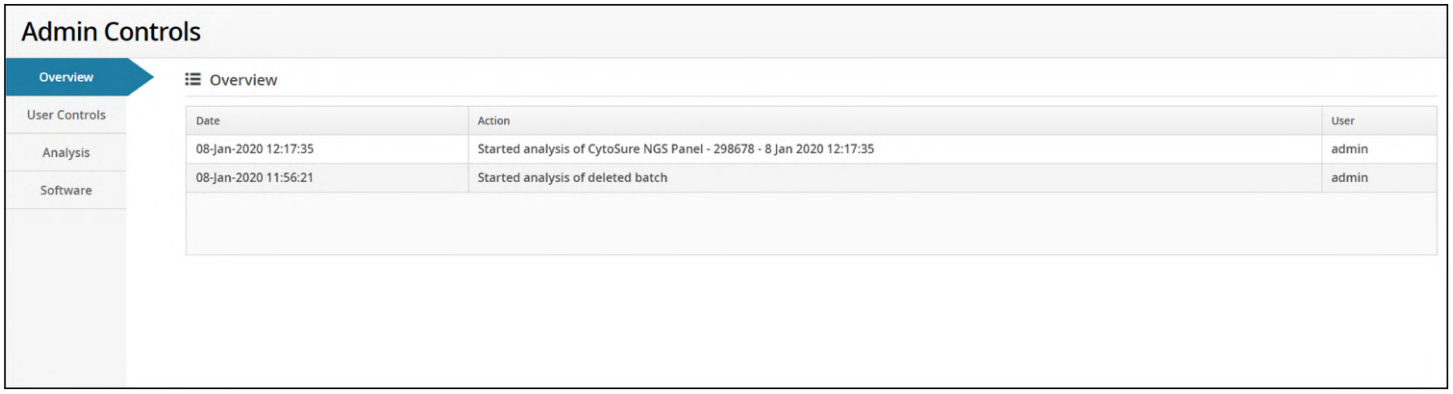


Figure: The overview window showing the latest activity in Interpret

14.2 User Controls

14.2.1 Current User

Current Users shows a list of all the current users.

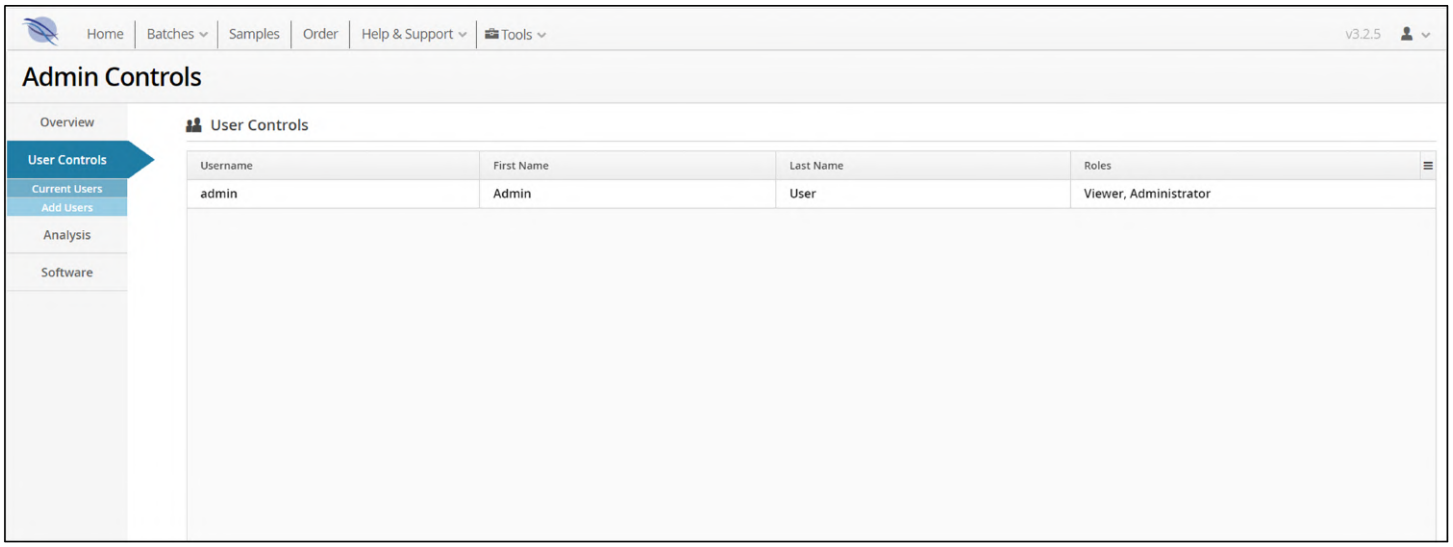


Figure: Front page for current users in user controls

Selecting a user provides an overview of the user account

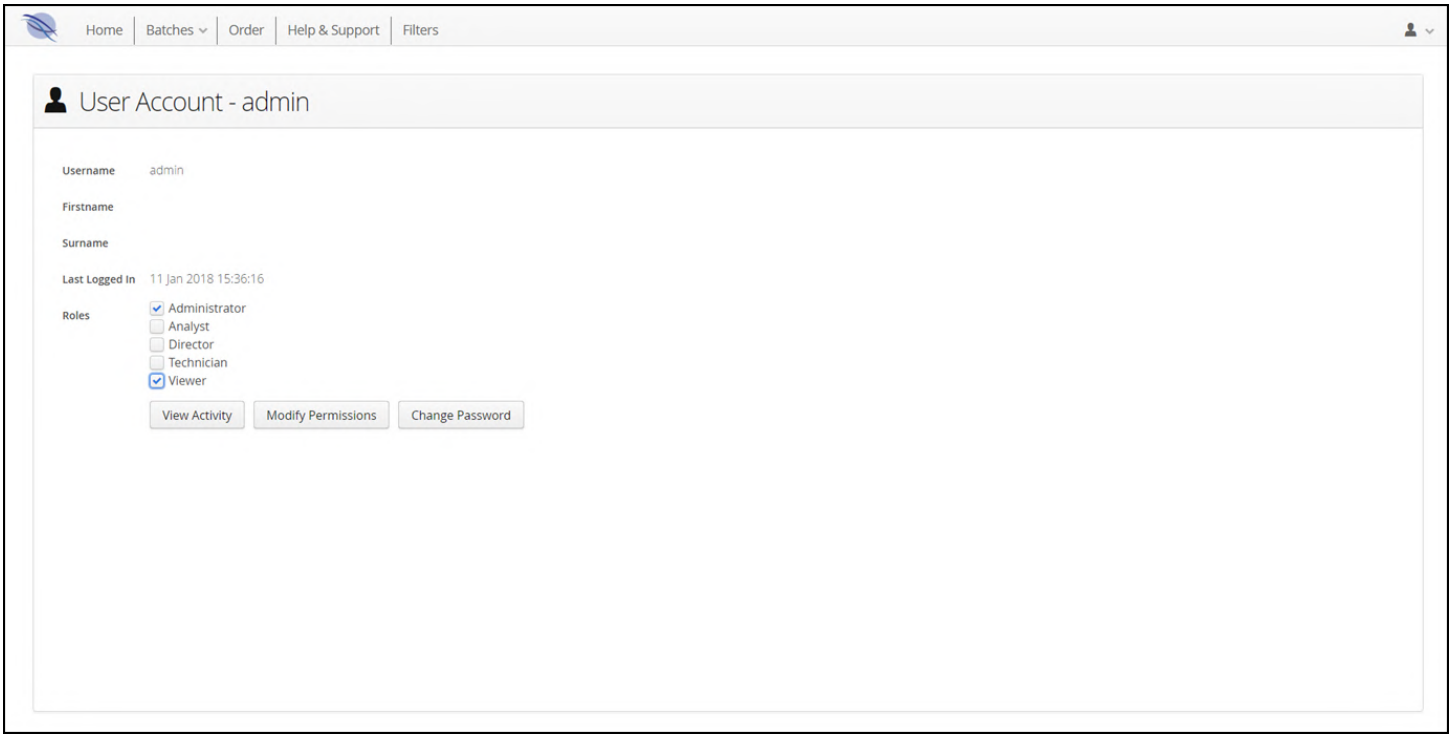


Figure: An overview of a user account

Selecting the "View Activity" button will give an overview of what the user has done and when.

Figure: Viewing a users activity

Selecting "Modify Permissions" will display the current permissions available and those with a tick in the adjacent checkbox have been enabled for the user.

As the current user is admin, as expected, all permissions are enabled.

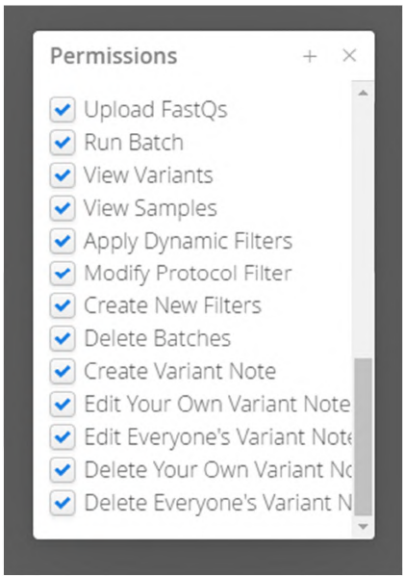


Figure: Permission options that can be modified

Finally, selecting "Change Password" gives a pop-up window that allows the user to change their password.

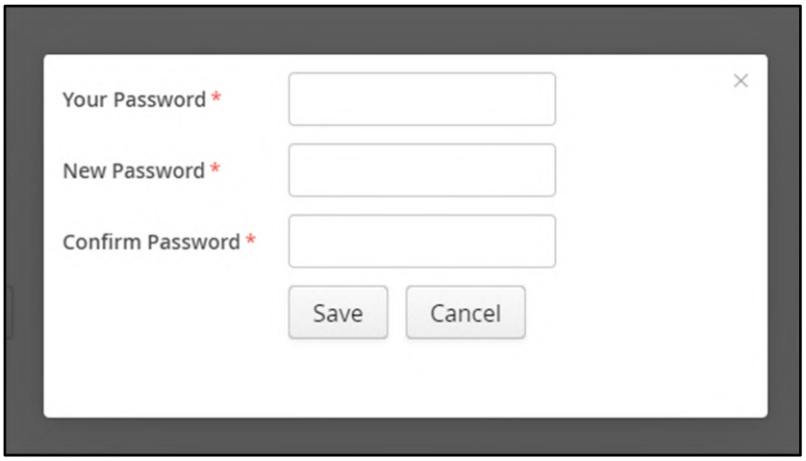


Figure: Pop-up menu to change a user's password

14.2.2 Add User

Add User provides a means to add new users. The following information needs to be supplied or selected:

- User details
- Login details
- Roles

Following correct completion of the form the 'Create User' button becomes active allowing the process of adding a user to be completed.

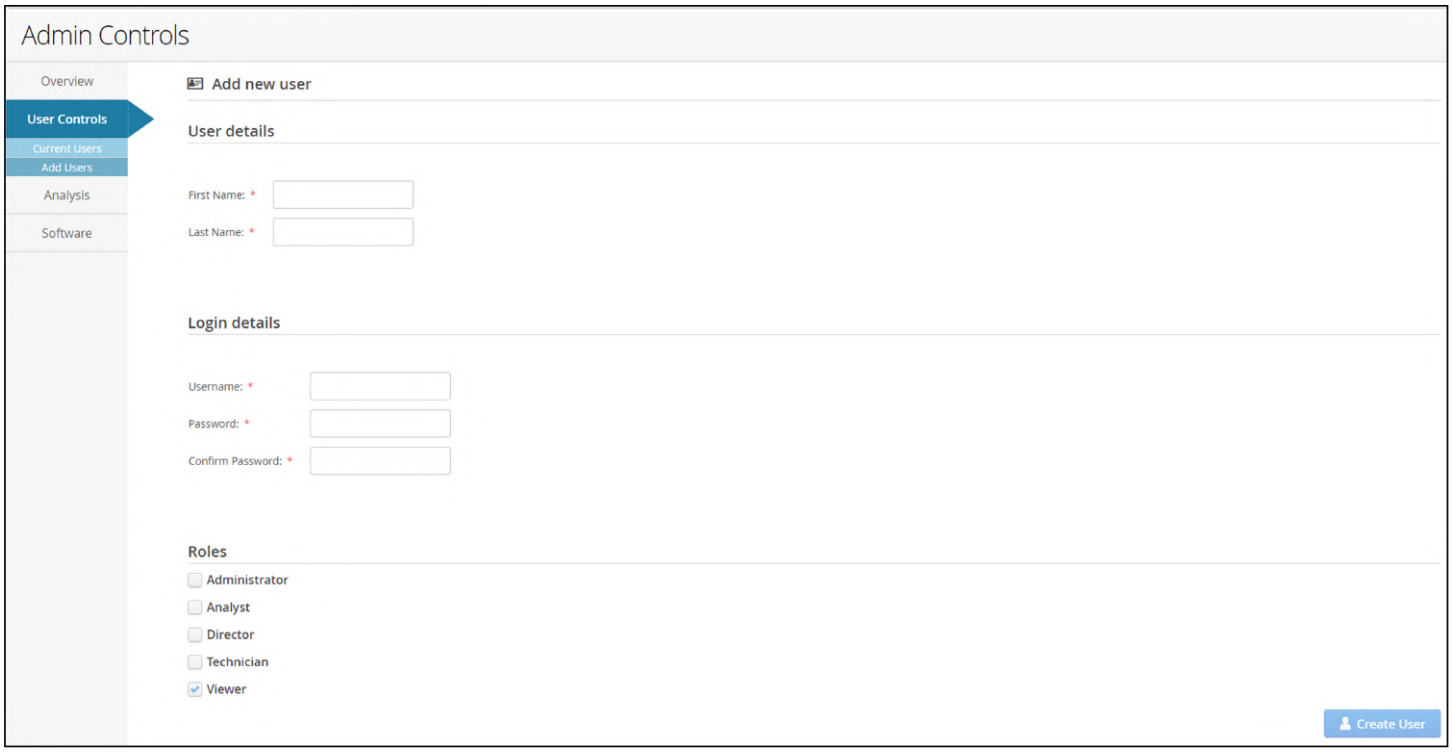


Figure: The add user start page

When the details are syntactically correct there will be a tick in the corner of the window.

Figure: In form validation of the new user details

Similarly, there are controls in place for the login details. A username that already exists cannot be used.

Figure: In form validation of the new user login details

An existing user name cannot be used.

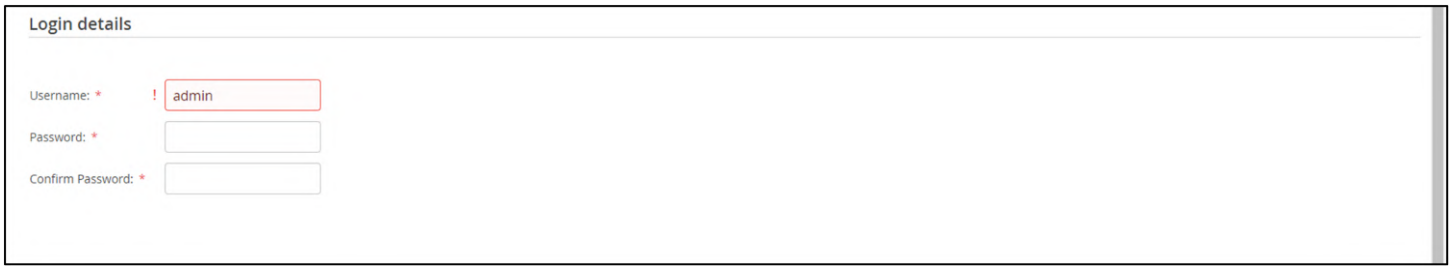


Figure: Example of using a pre-existing username

The form also checks to ensure the password is entered correctly.

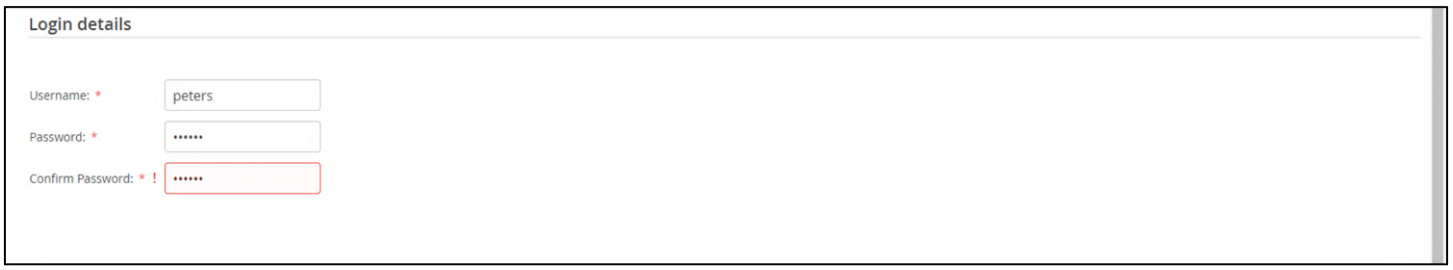


Figure: Example of entering non-matching passwords

Lastly, roles are assigned from the list using the checkboxes.

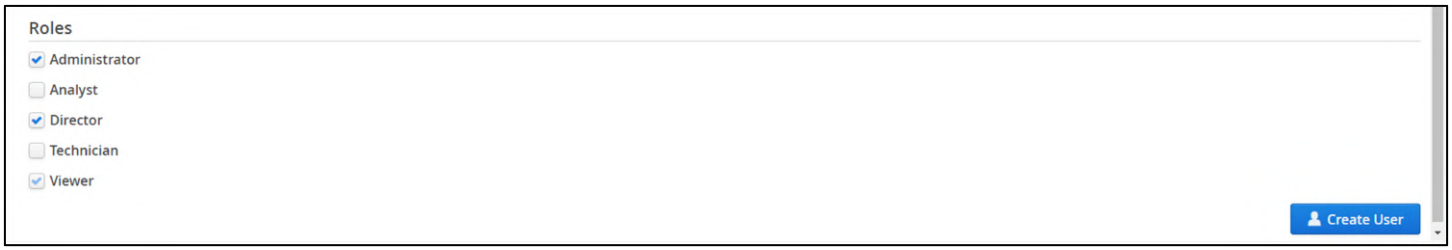


Figure: Selection of the roles for the new user

There are currently 5 different roles defined within the software and these are described in the following table.

Role	Abilities
Administrator	All
Director	All
Analyst	Starting an analysis and viewing results
Technician	Loading FASTQs and Starting an analysis
Viewer	Basic permissions required for all users

Table: Roles available to assign

In order to login to Interpret, all users are required to be a Viewer. Other roles can be assigned as required.

Selecting "Create User" processes the form and there is a popup to confirm that the new user has been created.

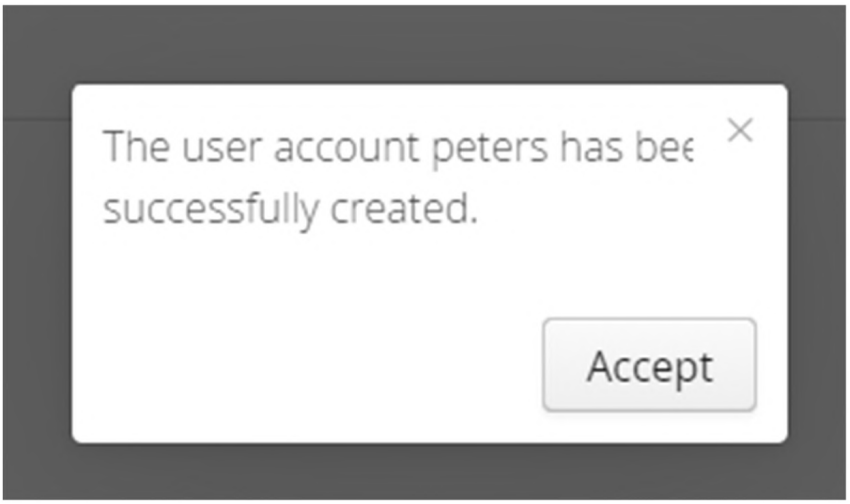


Figure: Popup menu confirming creating of the new user

Now looking at the Current Users in the User Controls you can see that the new user is listed.

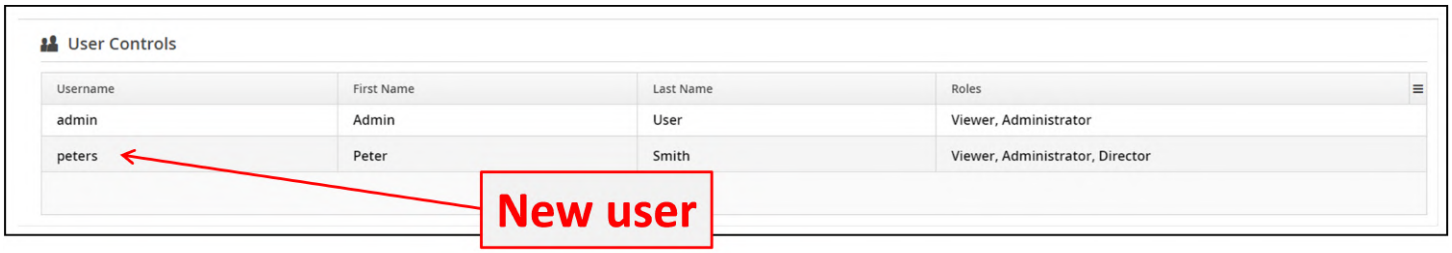


Figure: Display of the newly created user in the user table

14.3 Analysis

14.3.1 Manage Samples

Interpret allows users to manage samples and the data associated with them.

Figure: The Manage Samples start page

There are two tabbed panes displayed, an overview and a variables table.

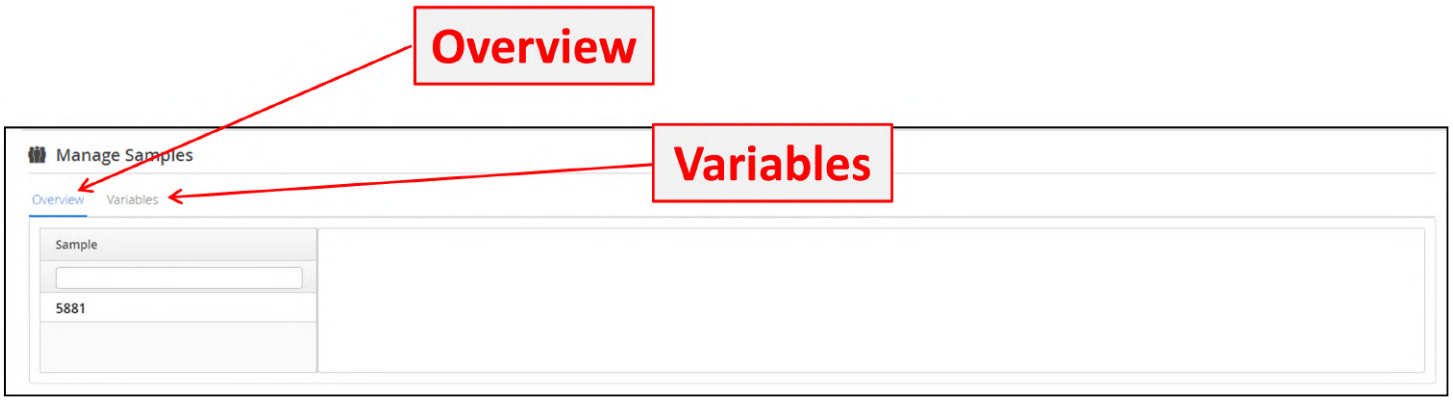


Figure: The two sections of the manage samples page

Selecting a sample in the overview tab brings up a series of sub-pages.

There is a batch history showing when the sample had been used in an analysis.

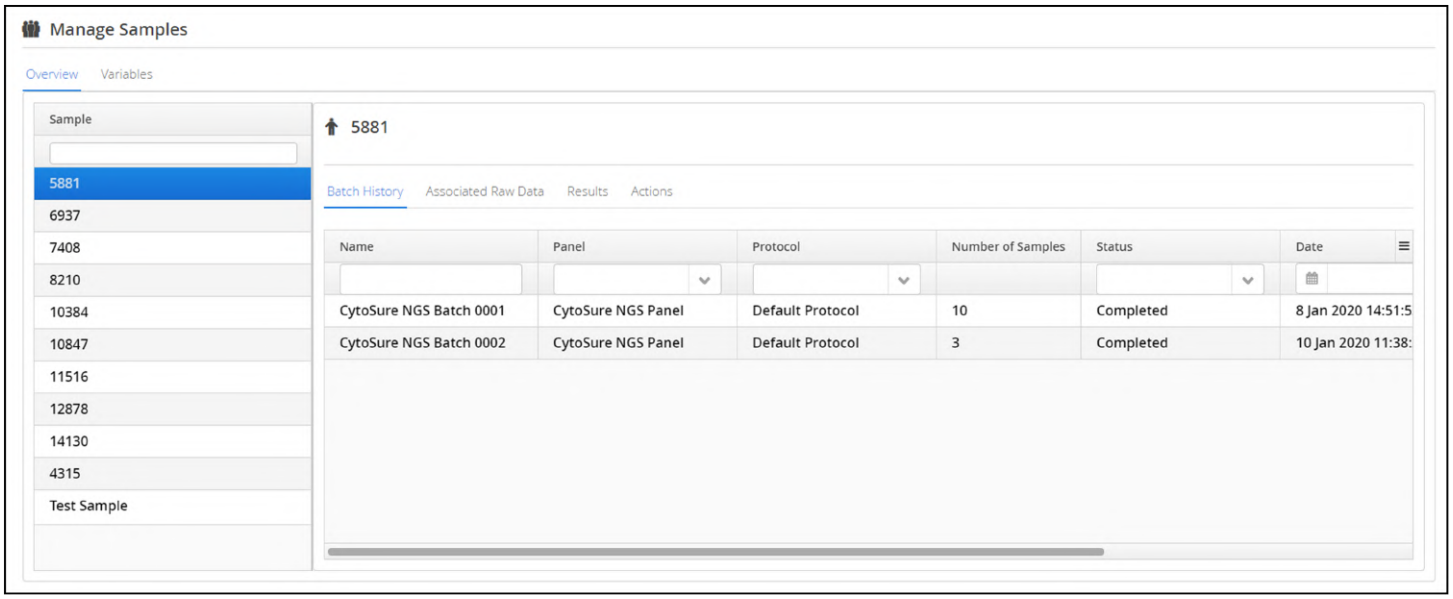


Figure: The batch history for a sample

There is a table showing all data associated with a sample name, including duplicate data if the sample data has been uploaded more than once.

These data can be deleted by selecting the Delete Data Files option.

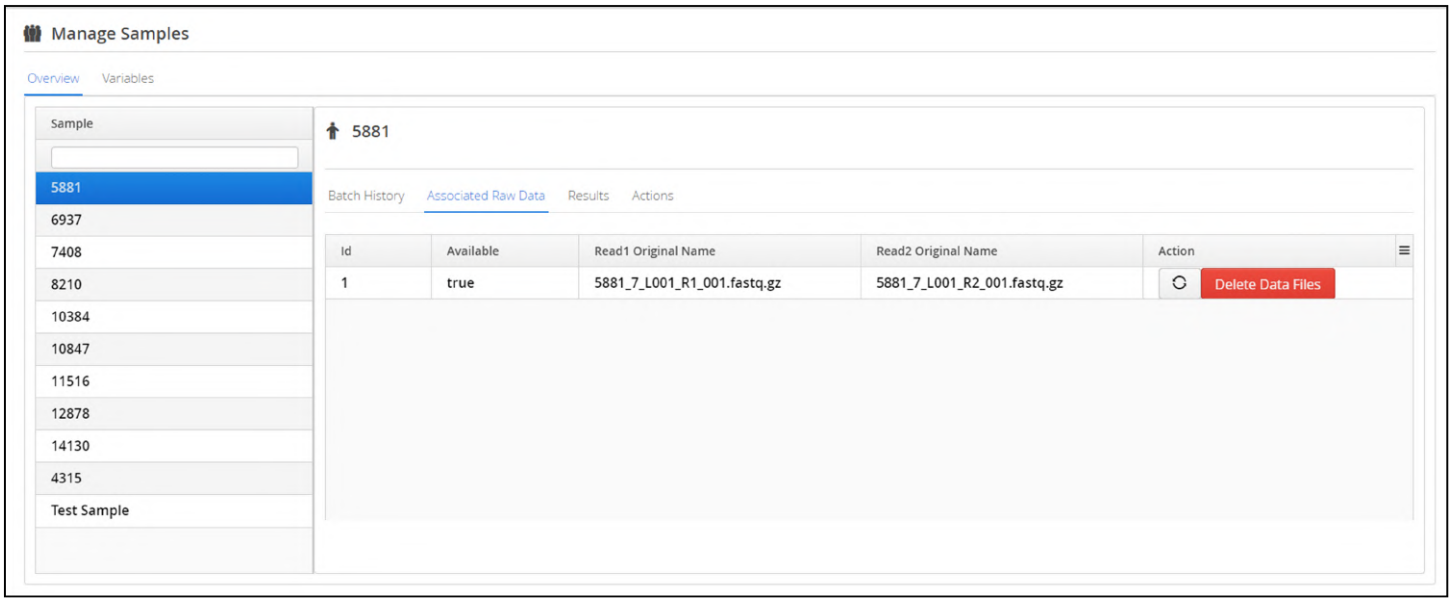


Figure: The associated raw data for a sample

If a user tries to delete data files, there will be a popup menu asking for confirmation prior to any data being deleted.

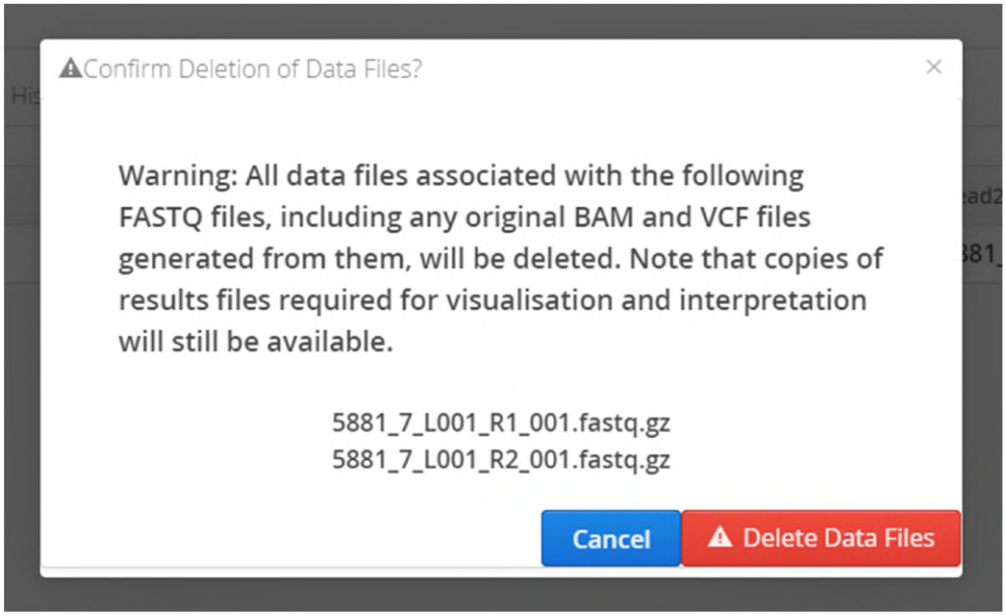


Figure: Popup menu asking for deletion confirmation

The user can see the variant and QC results associated with a sample.

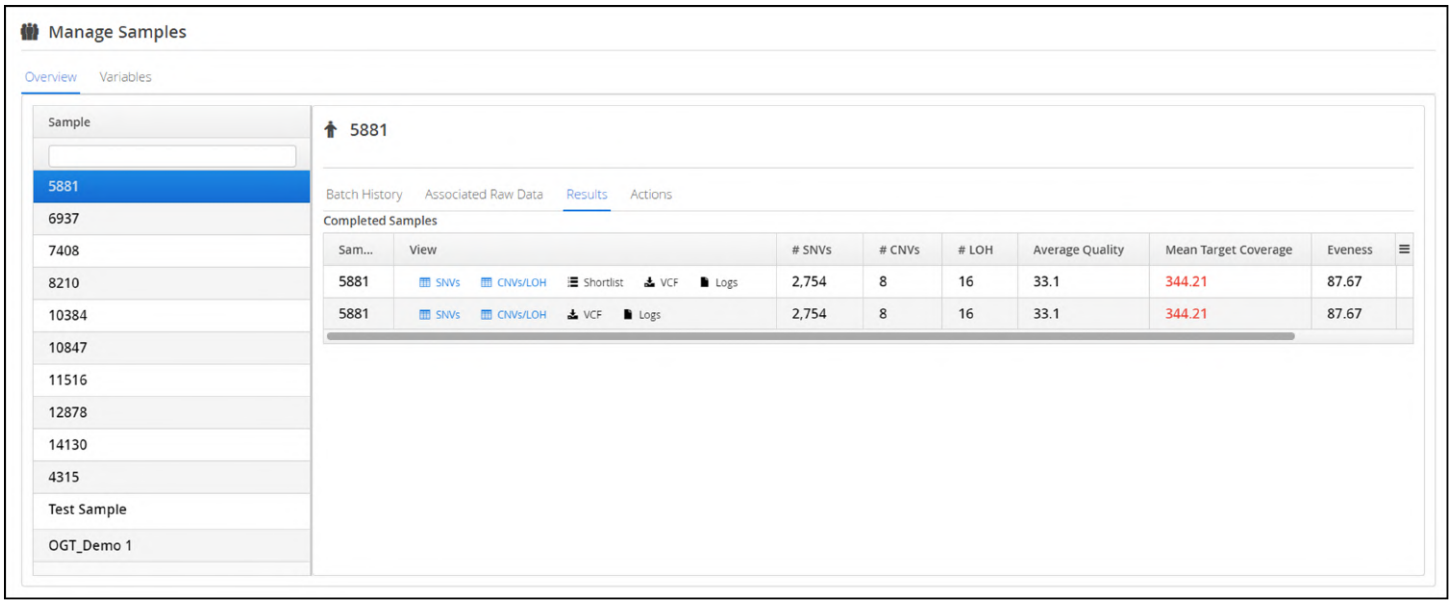


Figure: The results generated for a sample

Lastly, there are actions available for a sample. Currently, this is limited to updating a sample name.

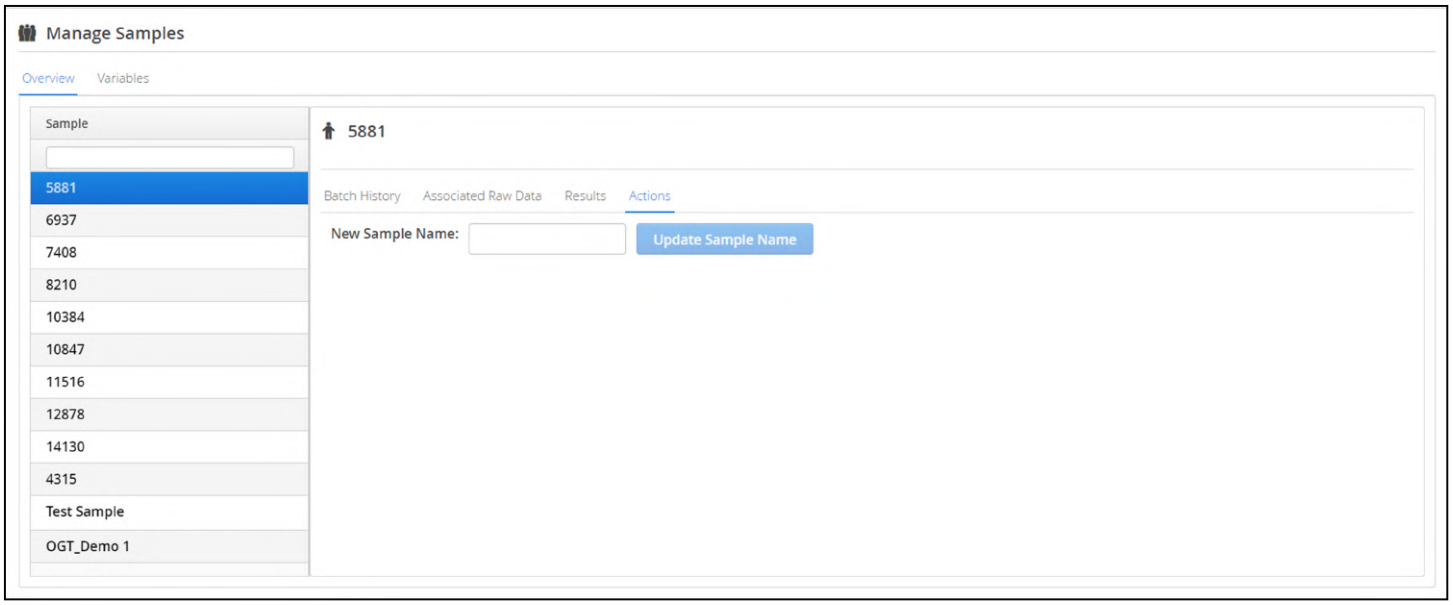


Figure: Actions available for a sample

Entering a new name and selecting Update Sample Name will update the sample name.

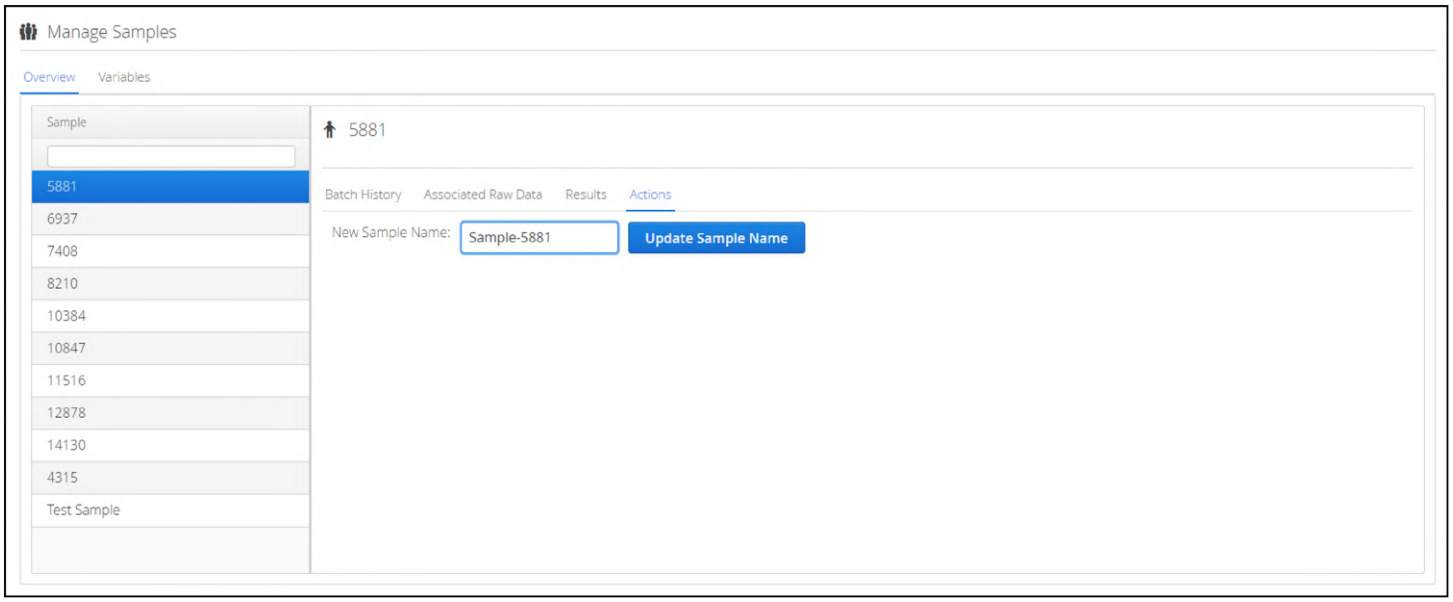


Figure: Updating a sample name

In the variables tab, users can modify existing variables are create new ones.

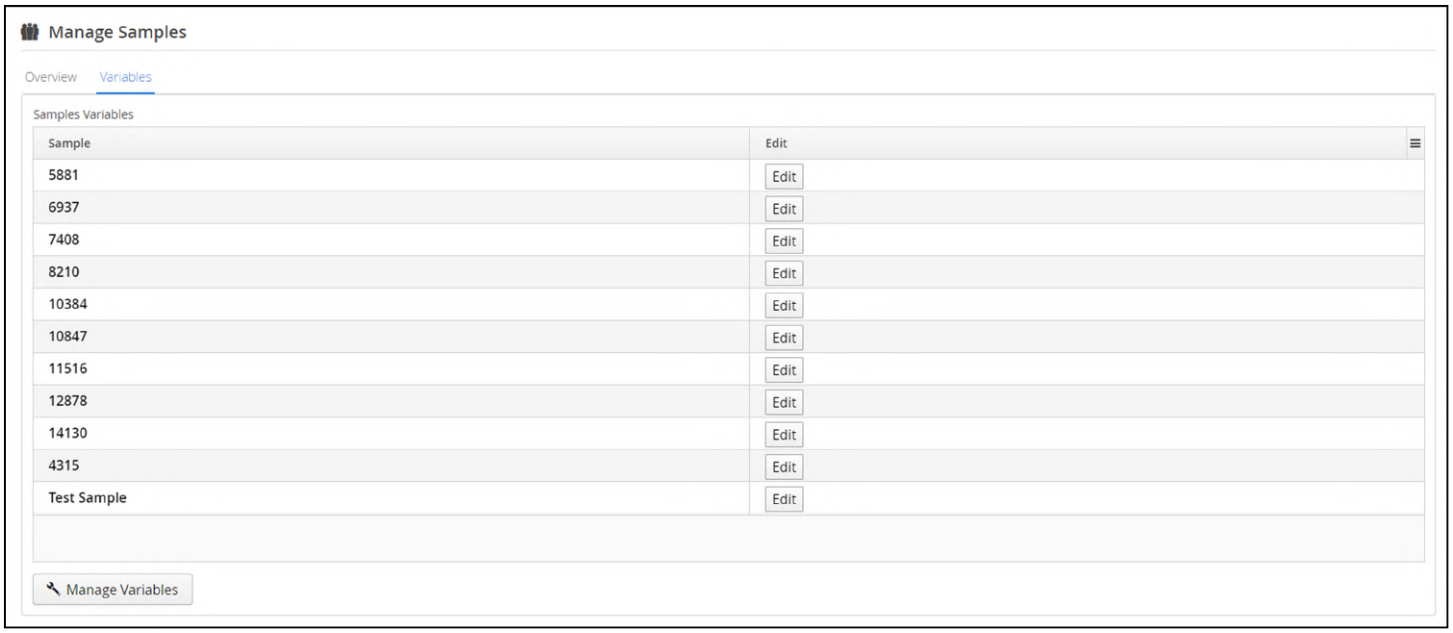


Figure: Sample variables

Selecting Edit allows the user to add variables for a sample or a new variable category.

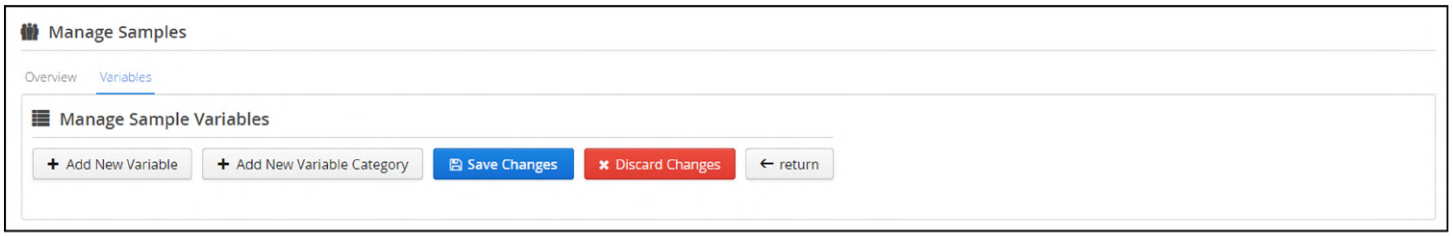


Figure: Managing variables for a sample

14.3.2 Current Analyses

This provides a means to view all current analyses.

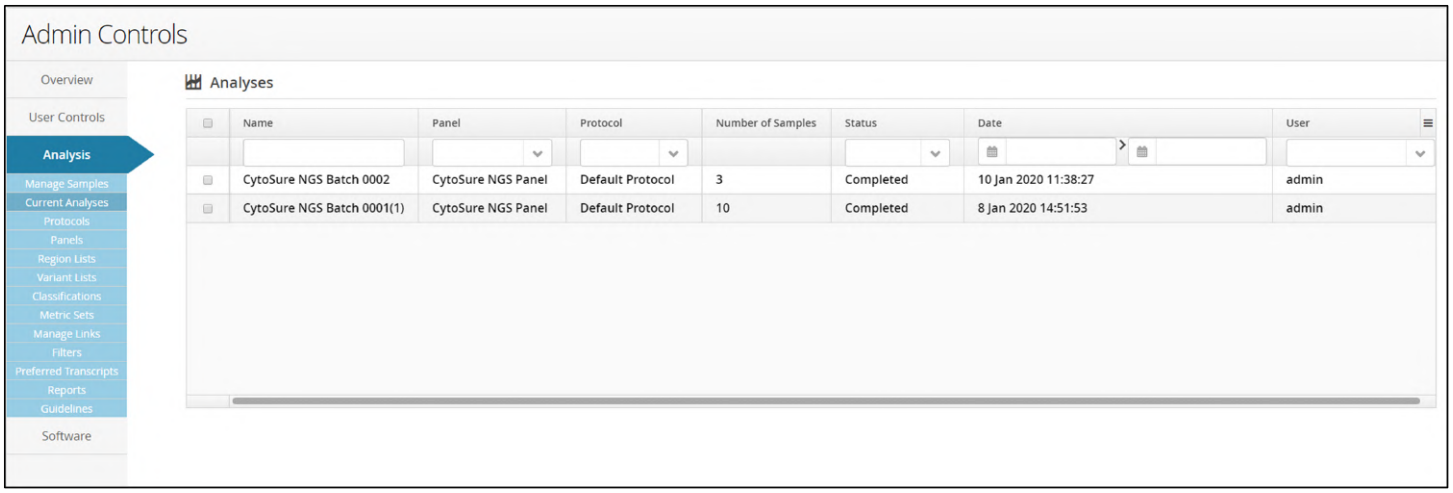


Figure: The Current Analyses start page

As with other pages in the software where there is a column selector icon  ,

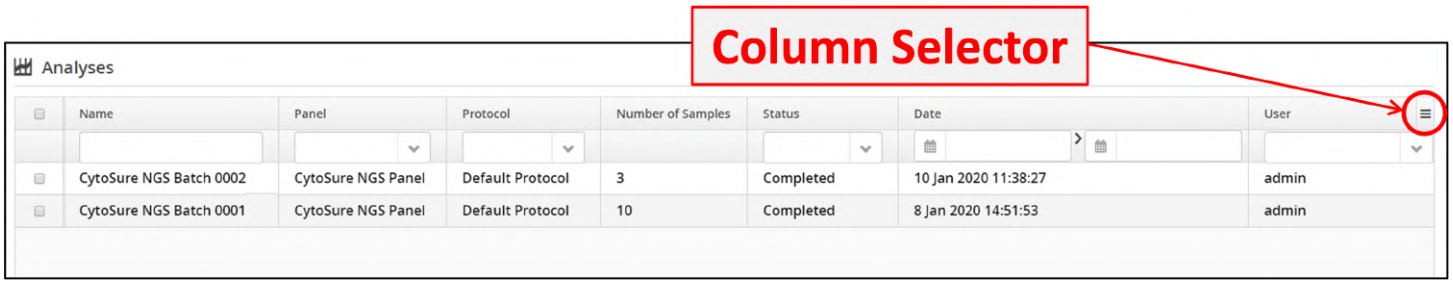


Figure: Current analyses with the column selector highlighted

Columns can be added or removed as required from the popup menu.

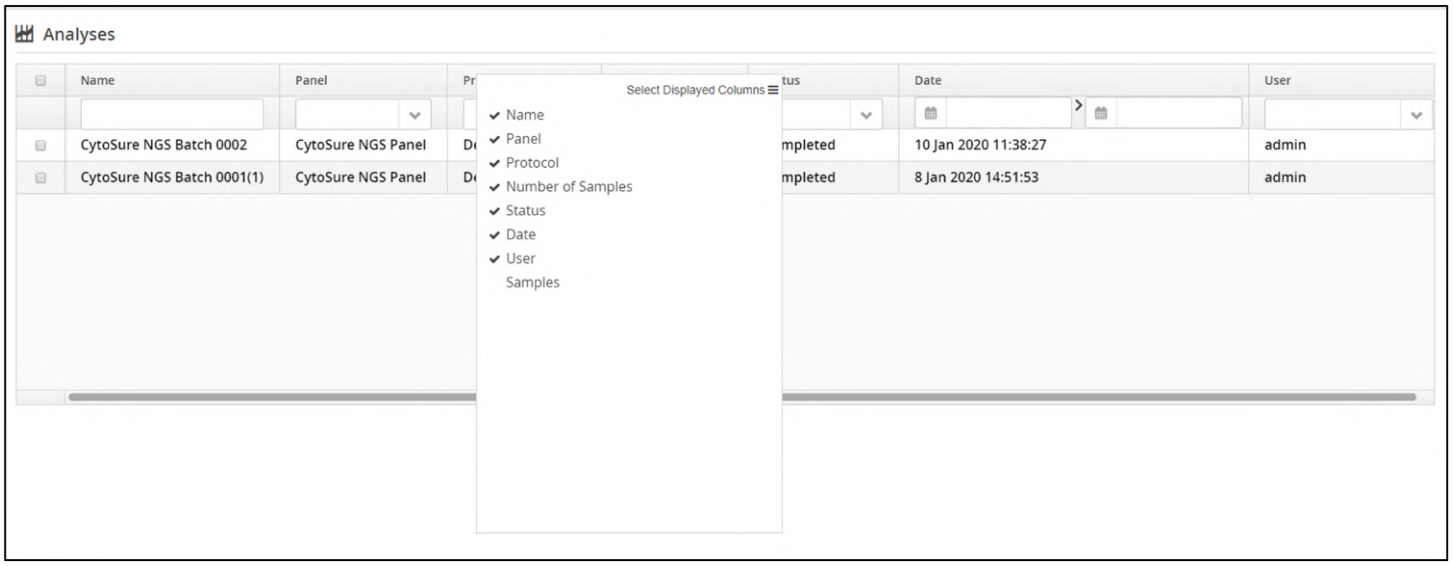


Figure: Columns available to select for display

Additionally columns can be sorted using column filters.

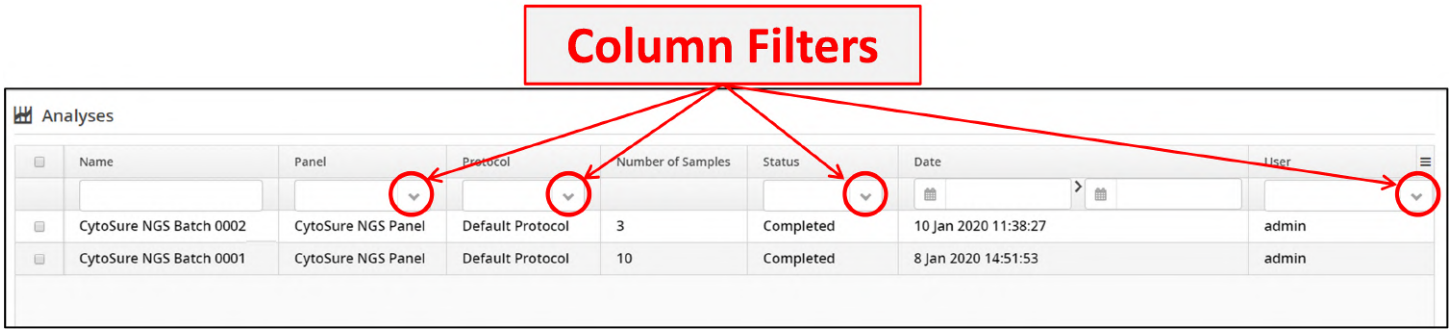


Figure: Column filtering options available

For example, in order to view all analysis with the Default Protocol this can be selected from the drop down in the protocol column

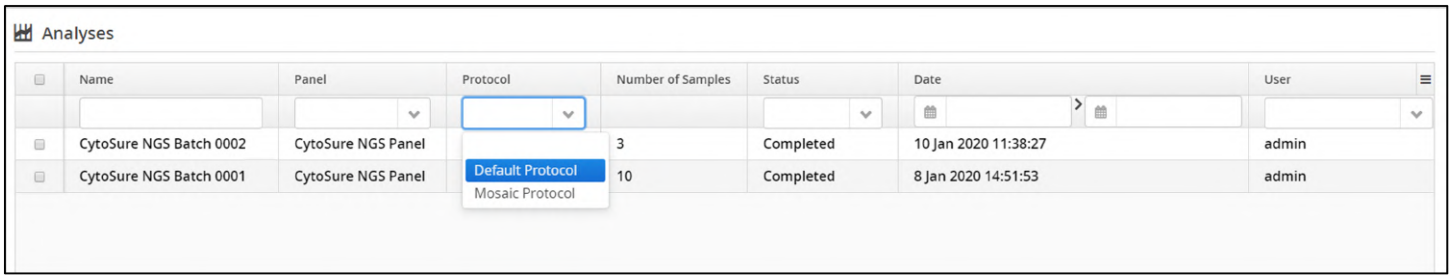


Figure: Selection of analyses only performed using the Default Protocol

14.3.3 Protocols

A Protocol defines how a sample is analysed.

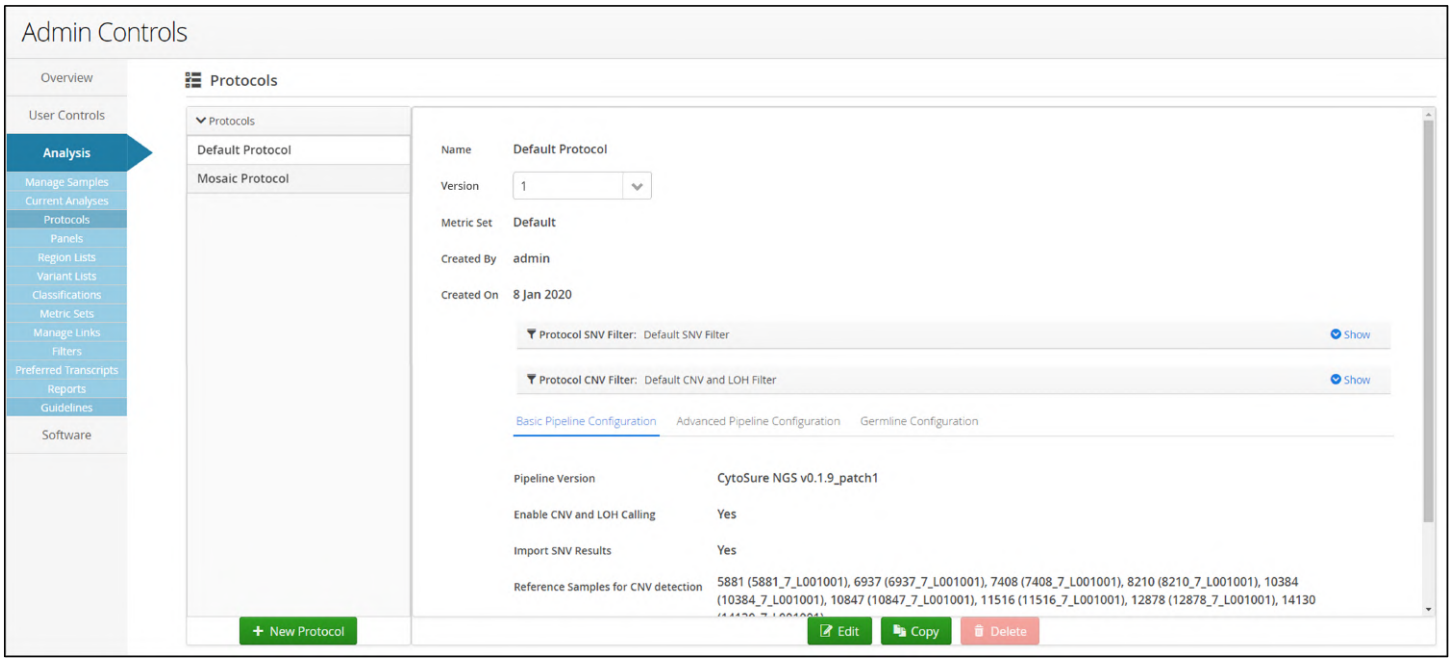
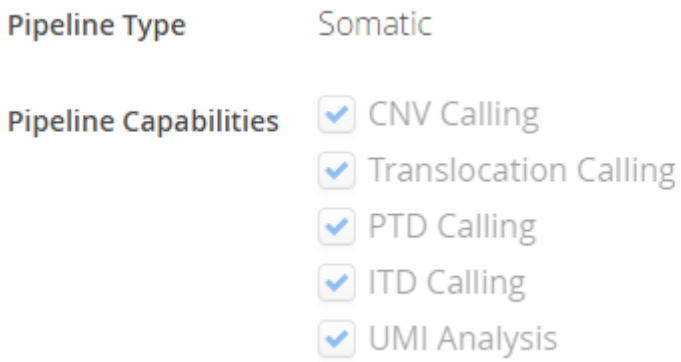


Figure: The Protocols start page

There are 4 components in a Protocol:

1. Pipeline Type (and Pipeline Capabilities)- the analysis pipeline type with which the protocol is compatible, and the functions of that pipeline that should be run as part of the analysis of samples in batches which use the protocol (its "capabilities"). When a batch is created, Interpret matches pipeline type and capabilities with those supported by the selected panel.



2. Metric Set - settings with which to qualitatively assess the run data.
3. Protocol Filter(s) - a filter with which to process all variants produced by the analysis pipeline. Currently there are filters for the following:
    a. SNV filter
    b. CNV & LOH filter
    c. Translocation filter
4. Pipeline Configurations - specific configurable settings used in the pipeline.
    a. Basic pipeline configuration
    b. Advanced pipeline configuration
    c. Germline configuration - if germline mode is selected in the basic configuration settings
    d. Somatic configuration - if somatic mode is selected in the basic configuration settings

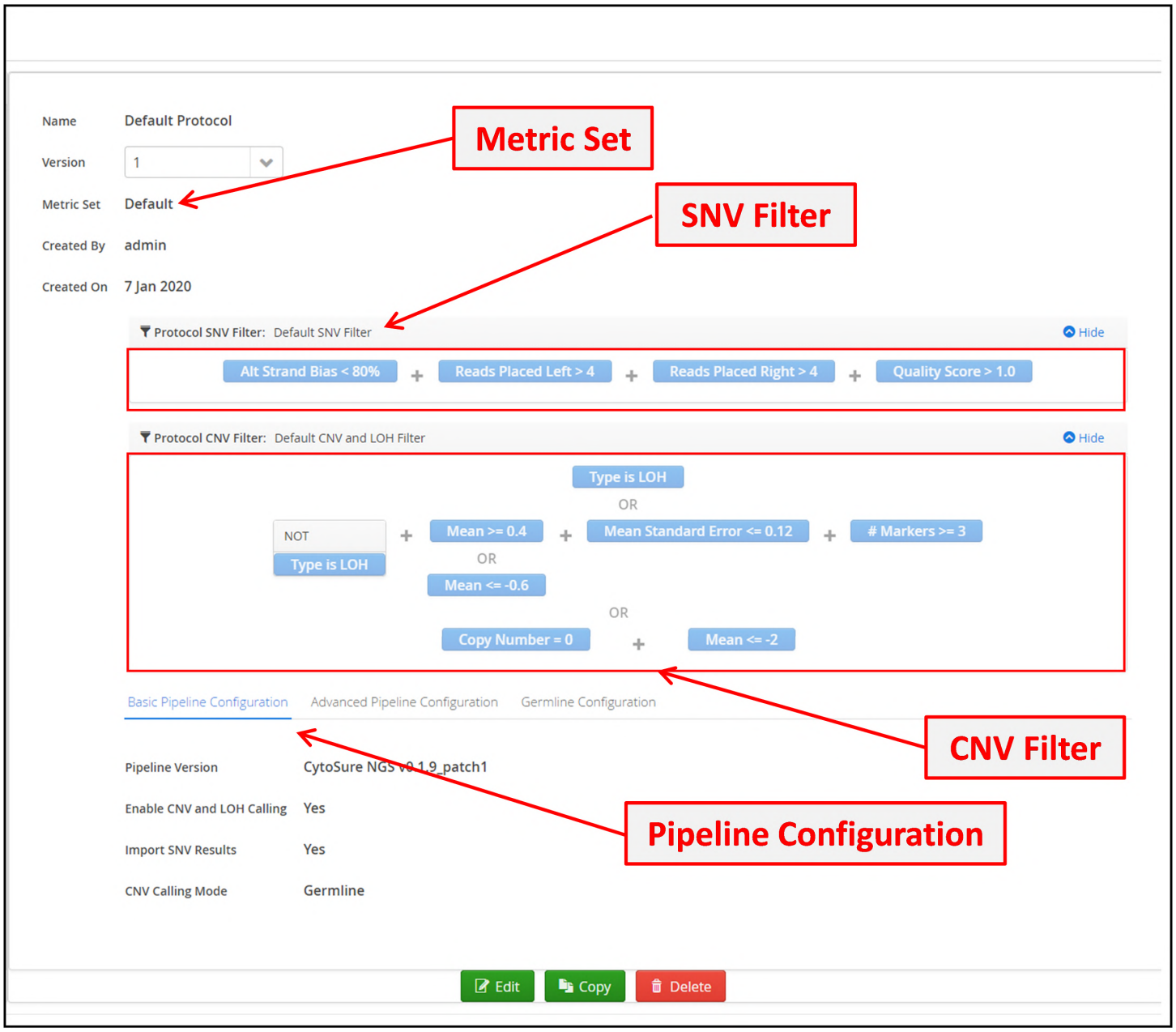


Figure: The sections of an analysis protocol

At the bottom of the protocol are a set of tabbed pages with different configuration settings.

There are basic pipeline settings:

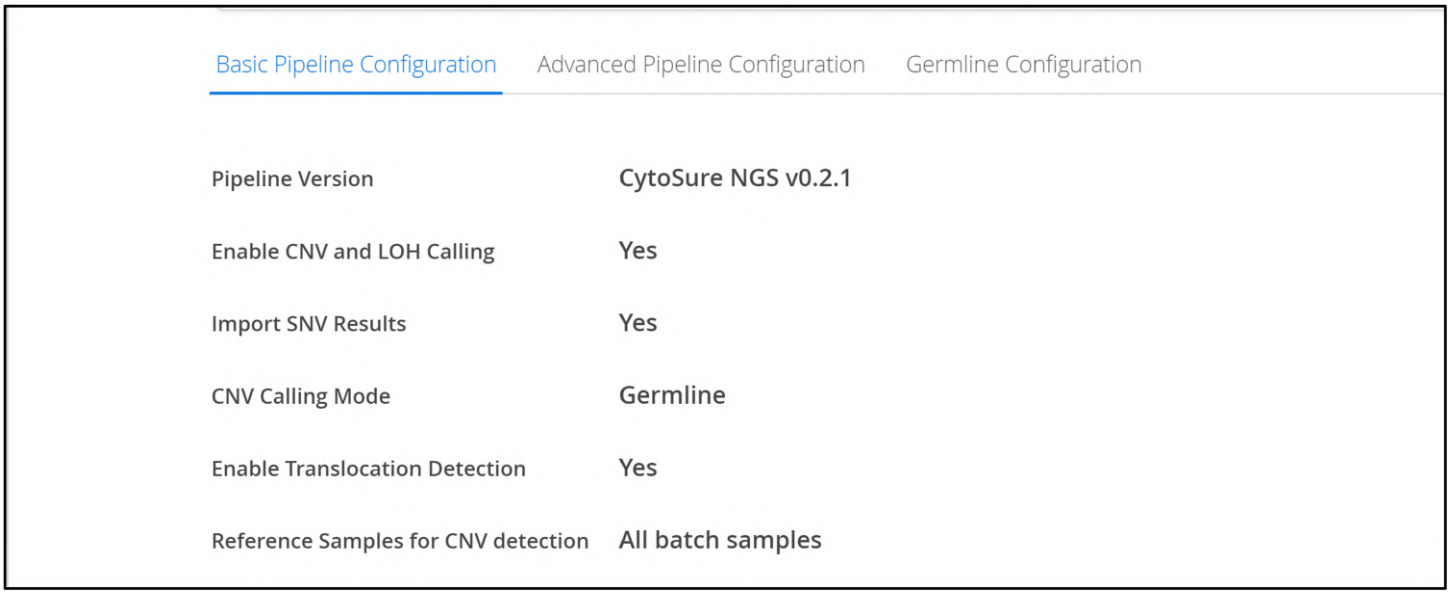


Figure: Default basic pipeline configurations

There are advanced pipeline settings:

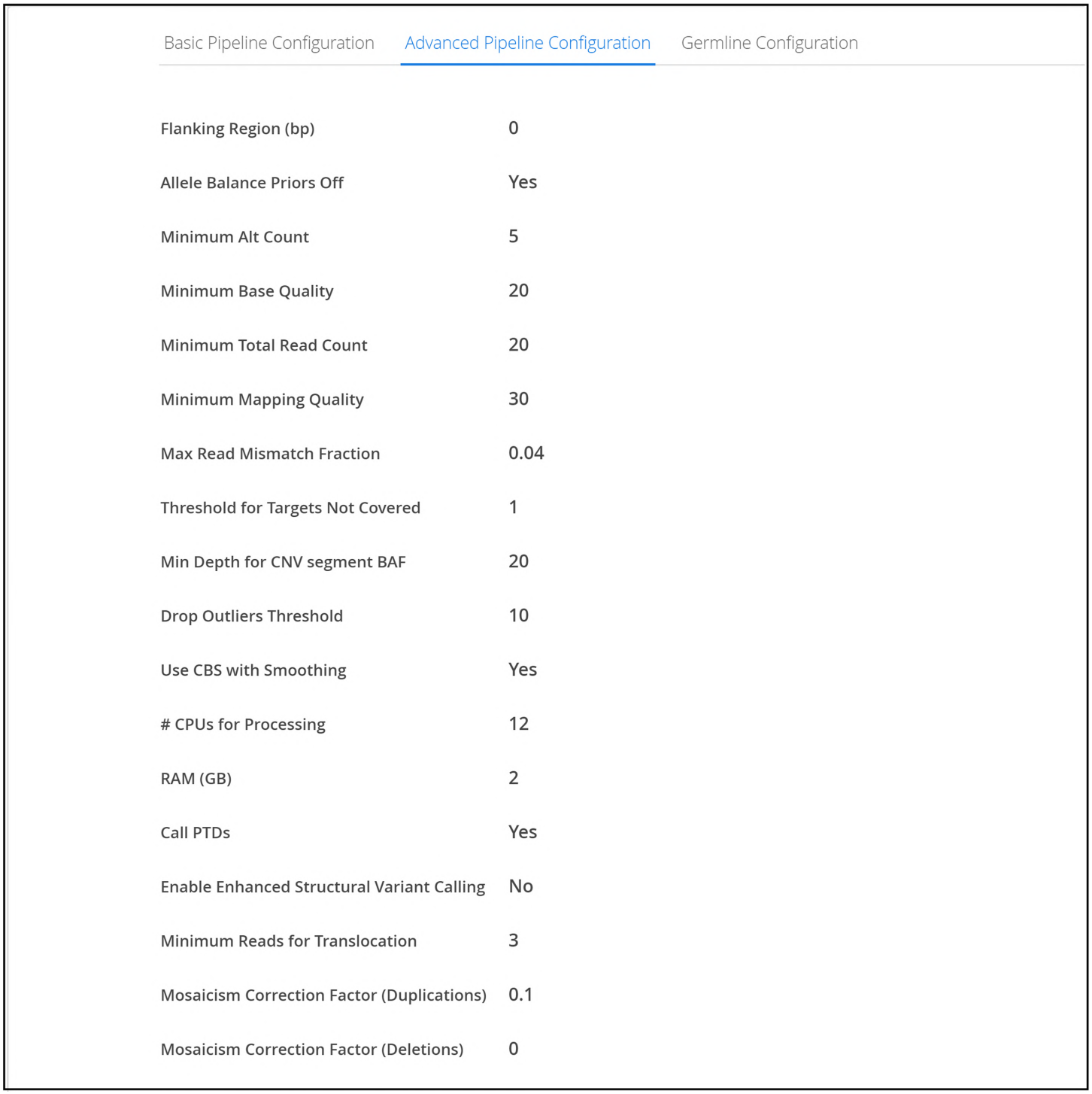


Figure: Default advanced pipeline configurations

Depending on the CNV calling mode selected in the basic pipeline configuration there will be either a Germline Configuration tab

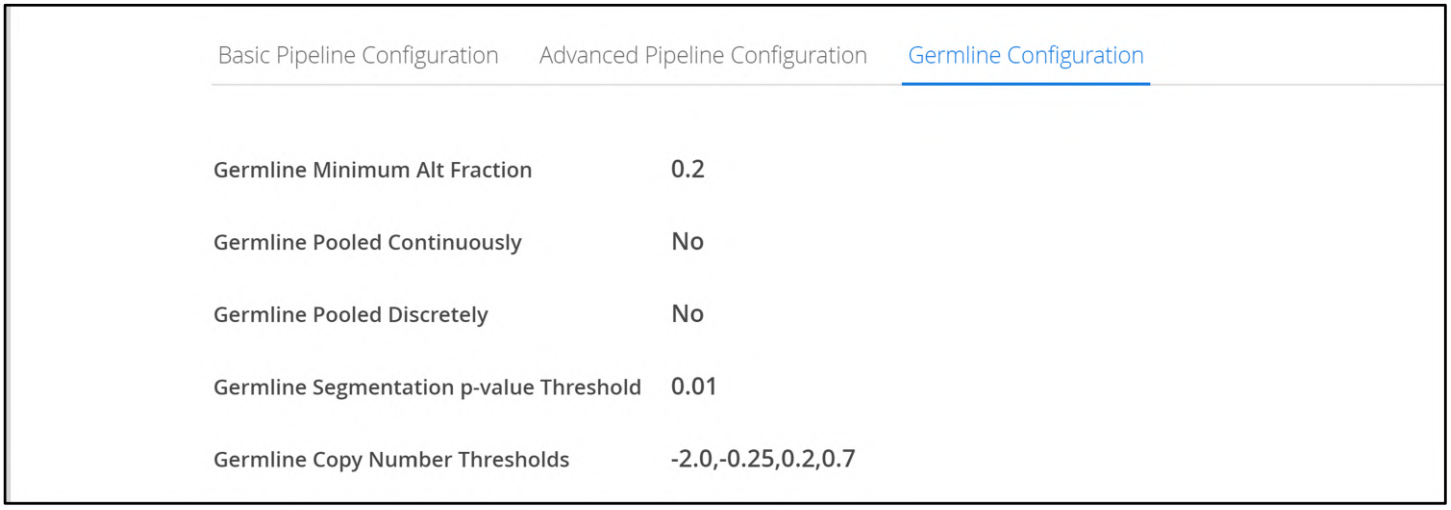


Figure: Germline configuration settings

Or a Somatic Configuration tab.

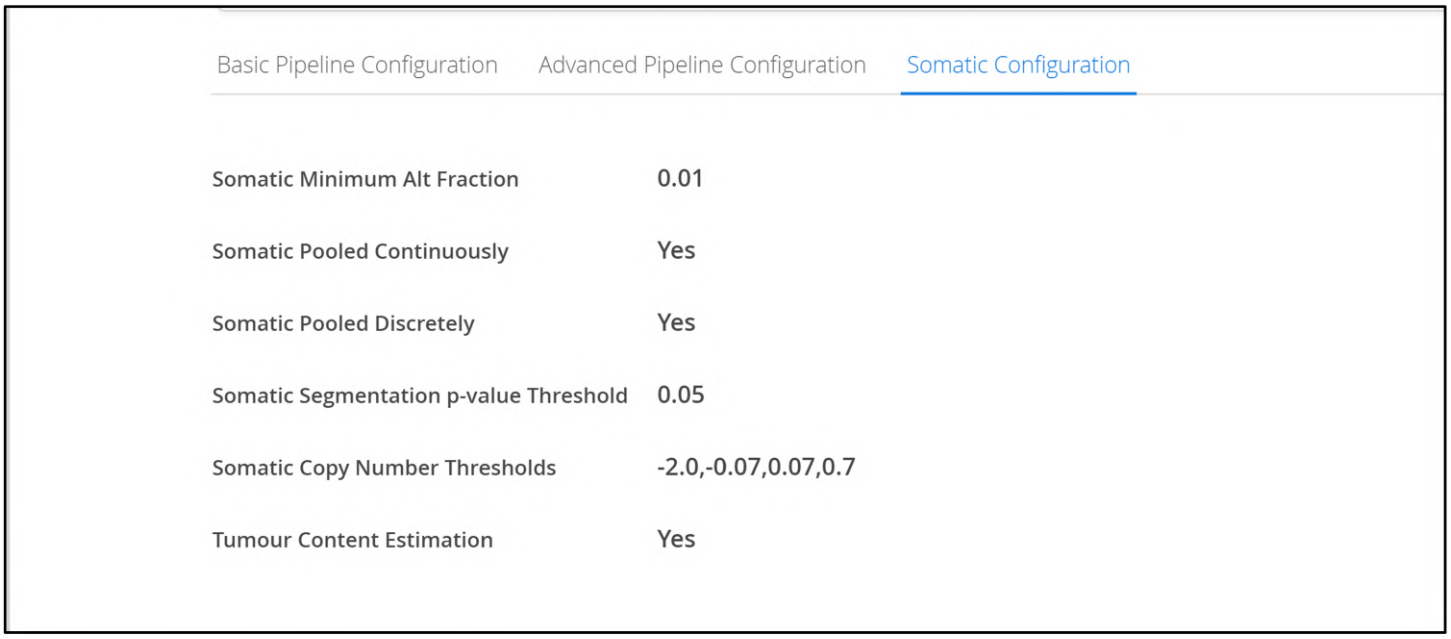


Figure: Somatic configuration settings

Modifying a Protocol

To modify a protocol users select the Edit button at the bottom of the protocol. A new version of the protocol is displayed which is the same, with the exception of the Version number that is incremented.

The user can retain the same metric set and filters or select alternatives from the drop-down lists.

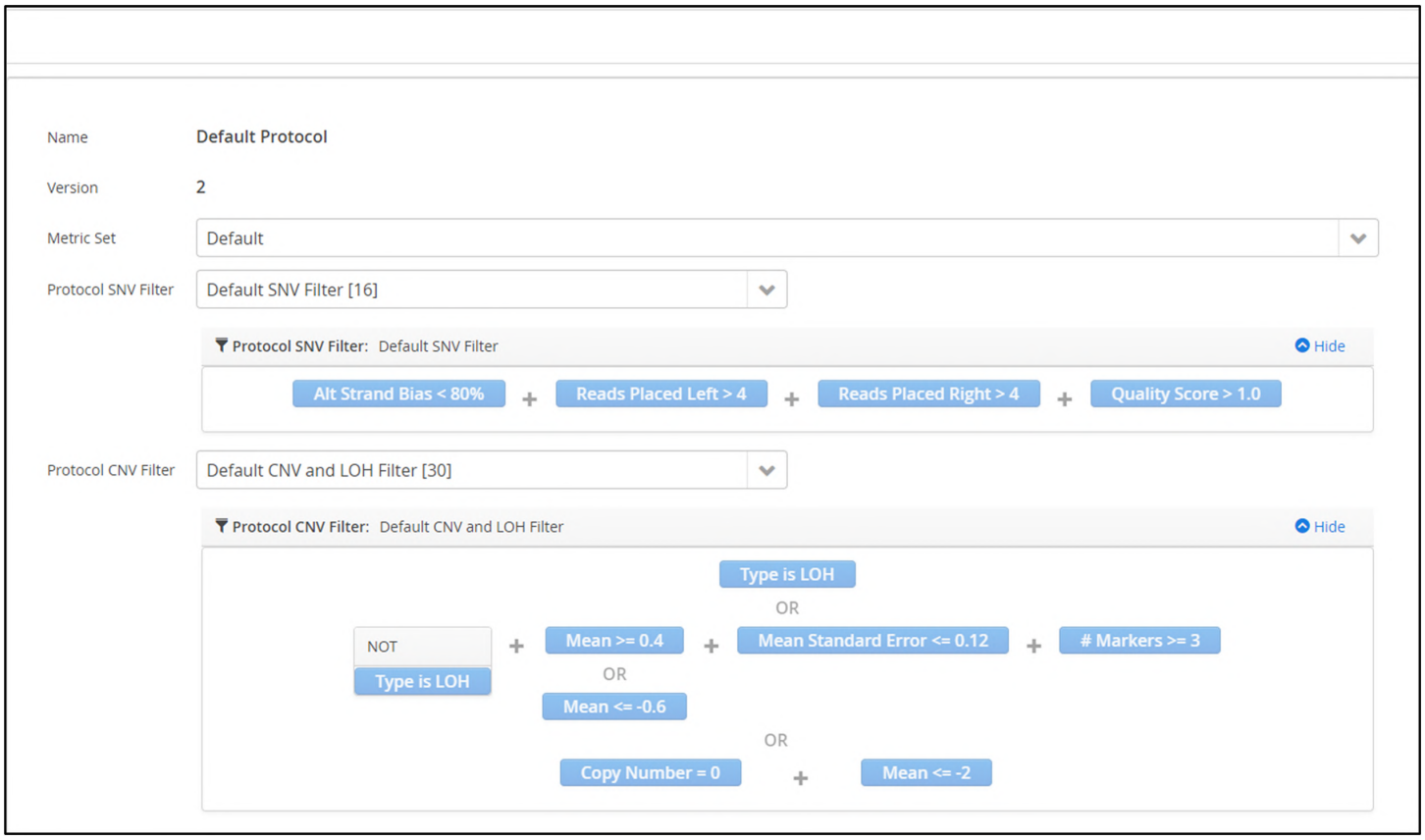


Figure: Editing an existing protocol

Protocol configurations can also be set; with the default values being updated. Software update does not automatically change the pipeline version for the protocol. User must modify protocol to modify desired pipeline version.

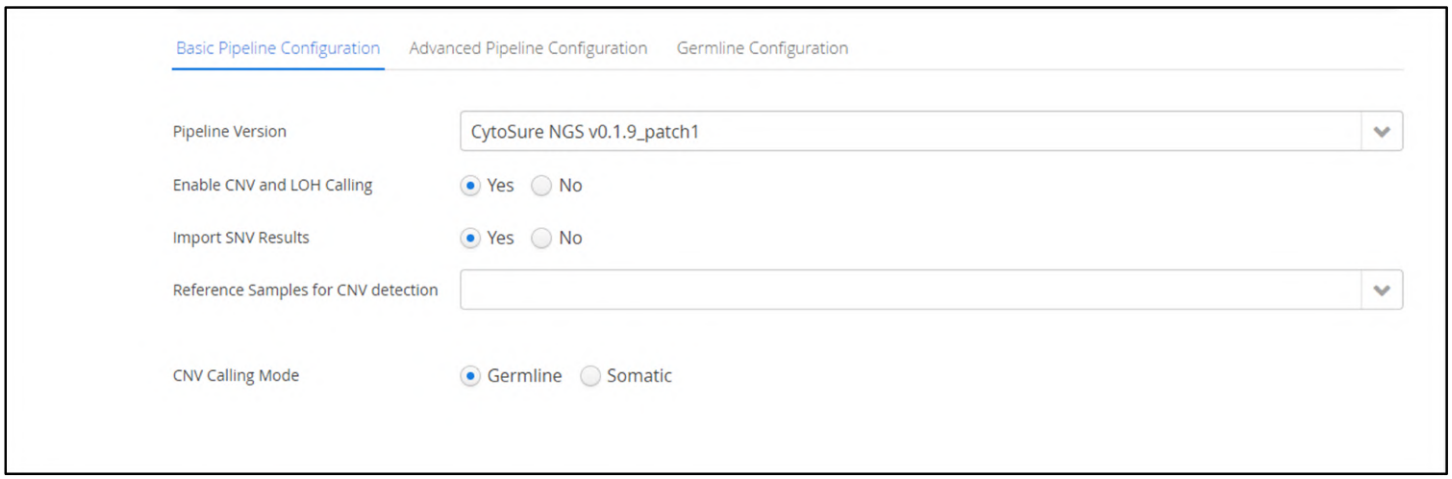


Figure: Setting basic pipeline configurations

Similarly with advanced pipeline configurations.

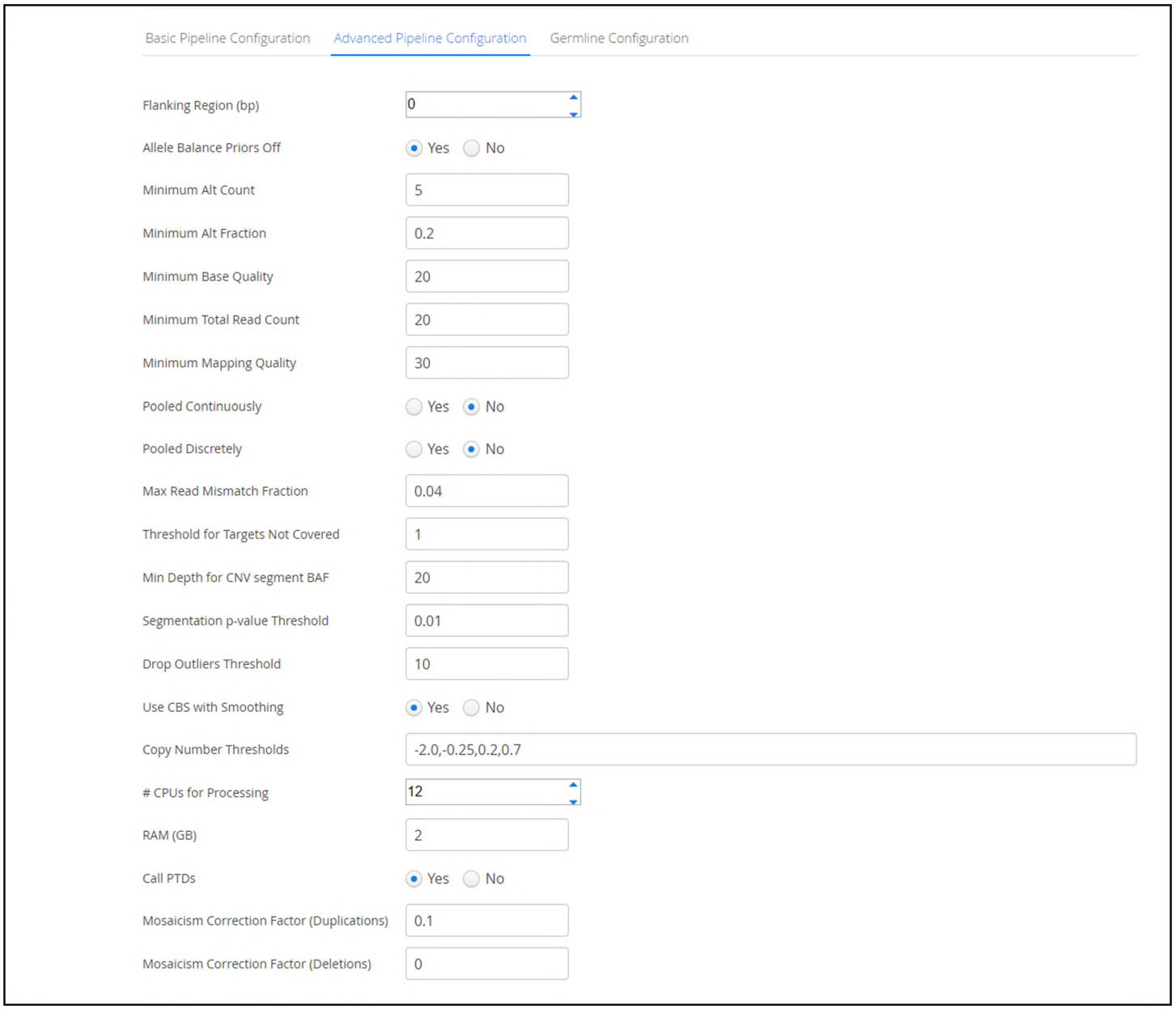


Figure: Setting advanced pipeline configurations

And lastly with the germline configurations.

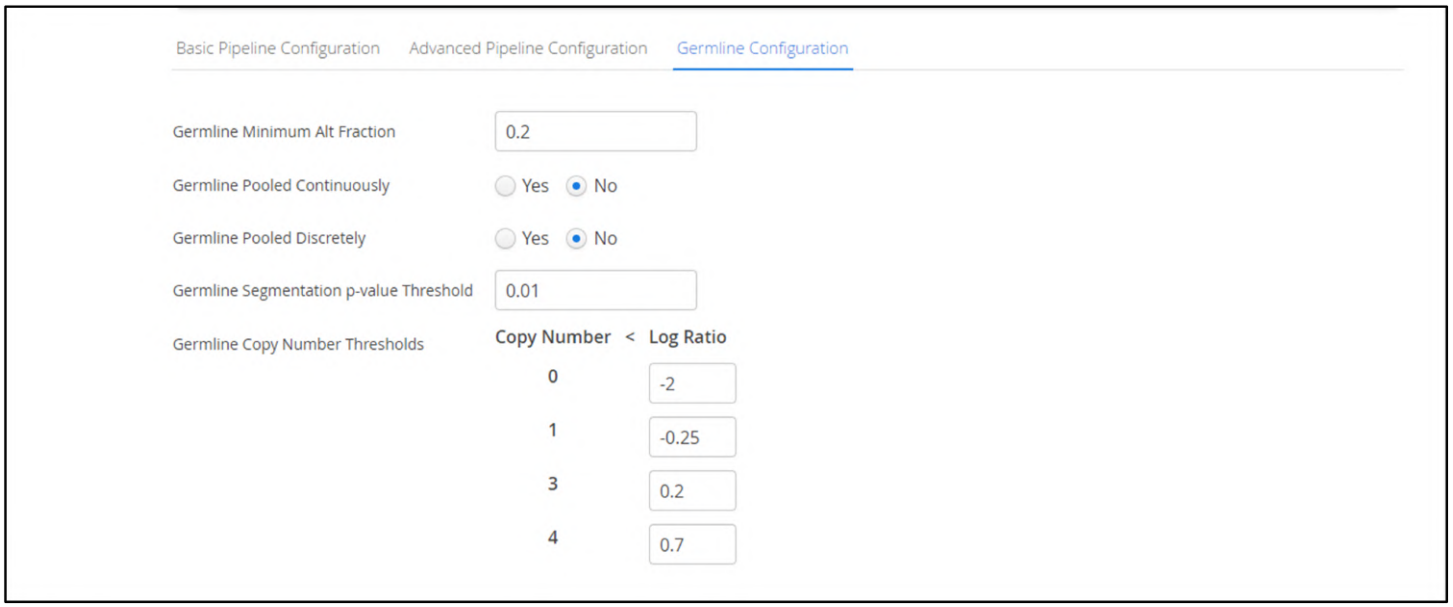


Figure: Setting germline configurations

Creating a Reference Pool for calling CNVs

In order to call CNVs, the software will use a set of samples to curate a pool of references against which variants can be detected.

Samples that will be part of the reference pool to be used for calling CNVs can be selected from all the samples that have previously been loaded in the software.

 It is important that samples used for the reference have been processed using the NGS panel with baits from the same lot that has been used for processing the test samples.

OGT can provide a set of suitable reference sample data for each of the bait lots that can be purchased. These can be used by the user as a starting point for building a reference of control samples suitable for CNV calling.

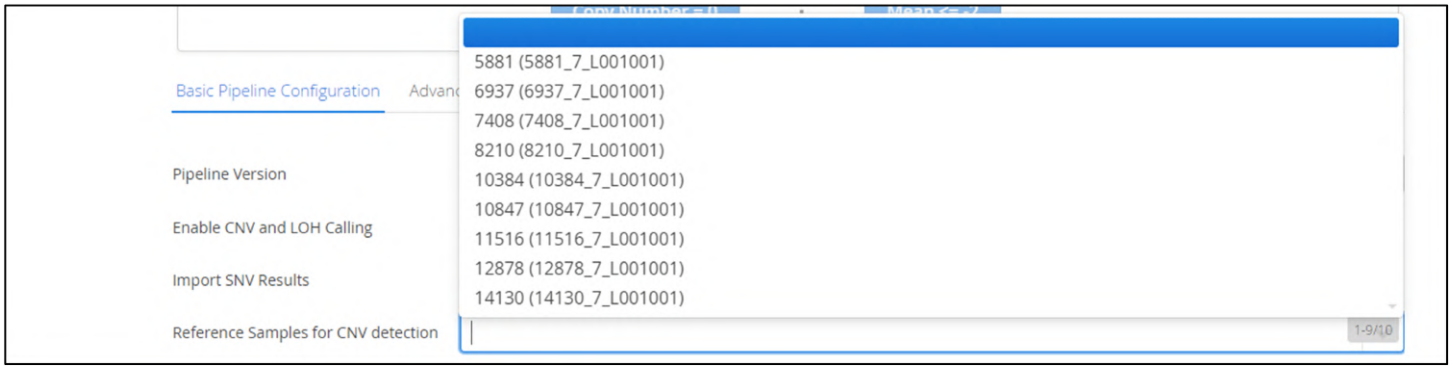


Figure: Selecting samples for the reference pool

Once selected the sample will be displayed, highlighted in blue to denote pending status.

Figure: A sample selected for the reference pool

Samples can be added as required.

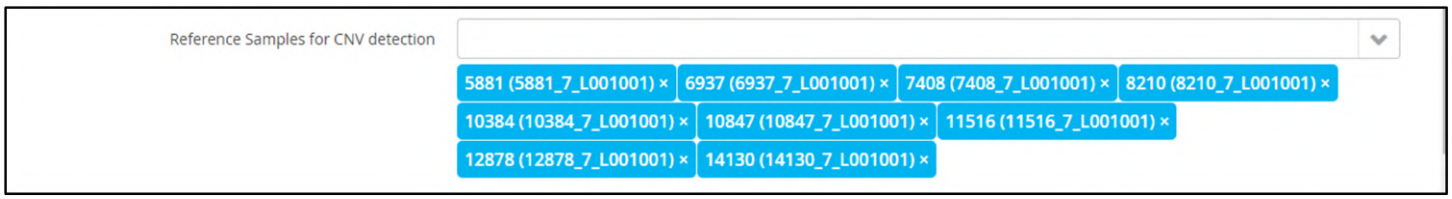


Figure: A reference pool of samples for a protocol

Once all required samples have been selected for the pool, the configuration can be saved.

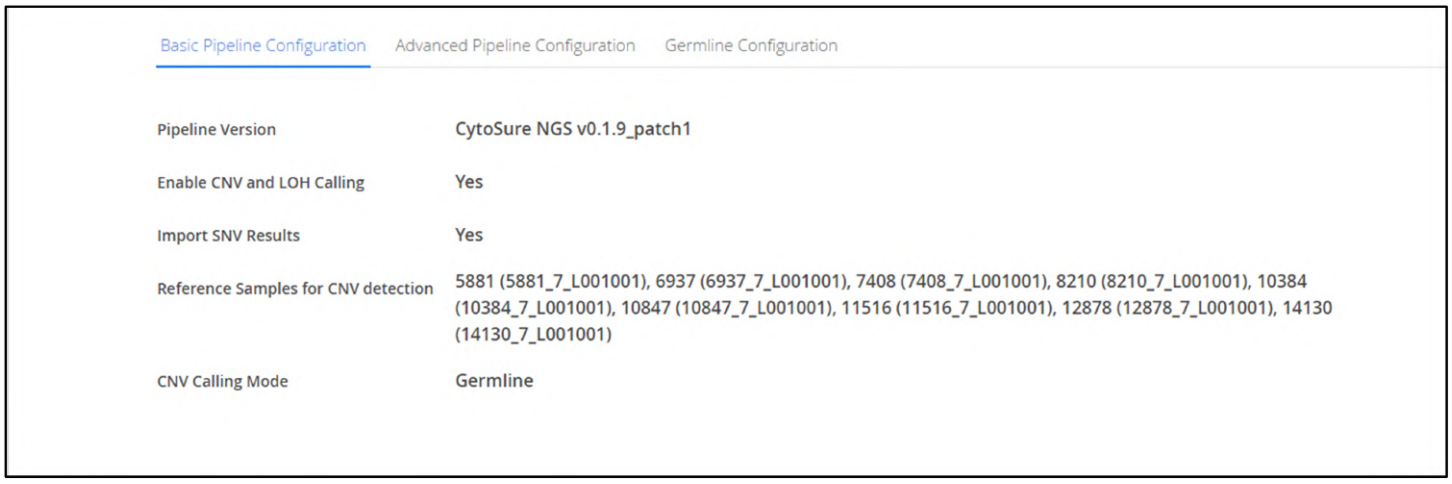
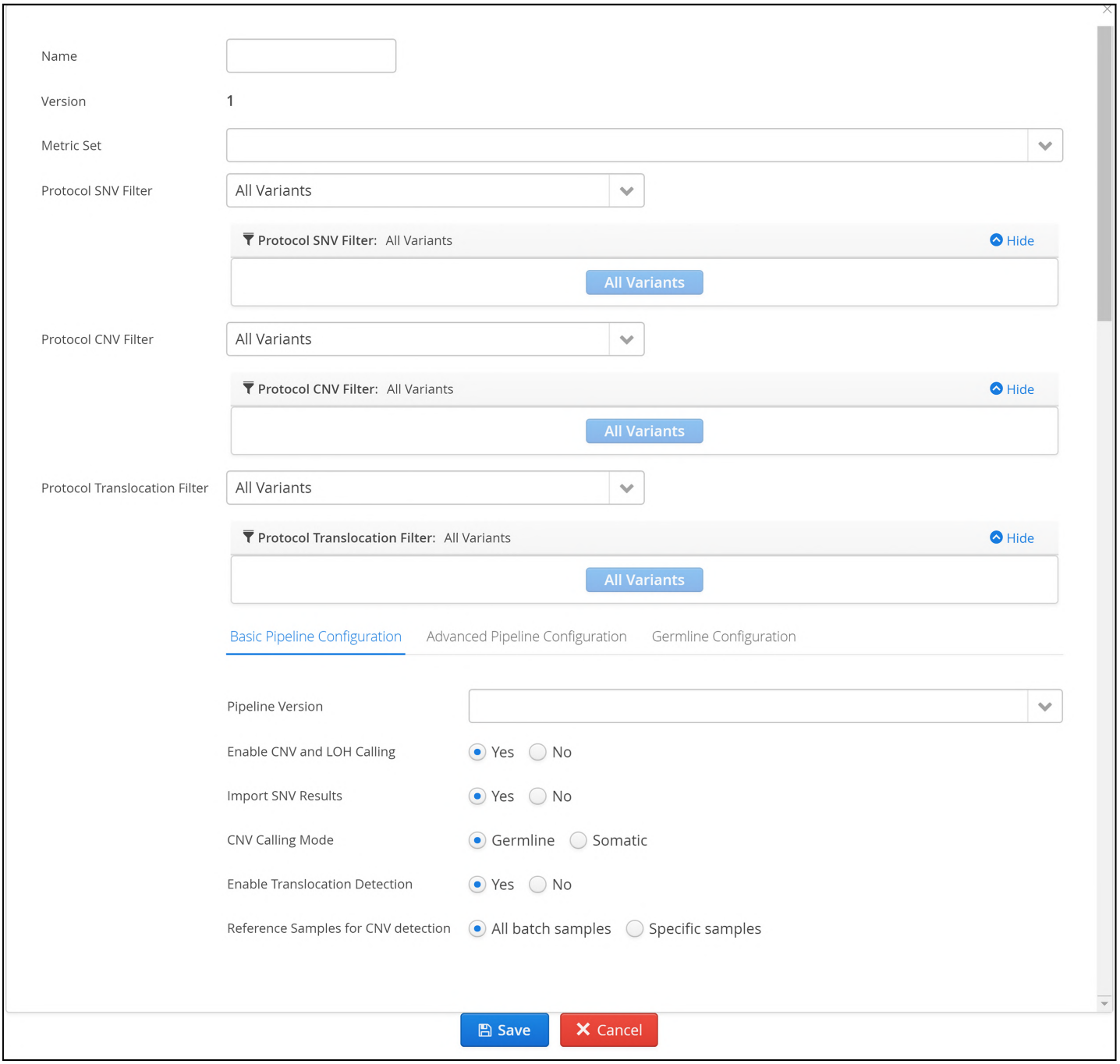


Figure: The basic pipeline configuration showing a a reference pool

Creating a New Protocol



Basic Pipeline Configuration

Setting	Description	Default Setting	Other Options
Pipeline Version	Select the version of the pipeline to use in the protocol.		
Enable CNV and LOH Calling	Enable or disable calling of CNVs and LOH by the protocol.	Yes	No

Setting	Description	Default Setting	Other Options
Import SNV Results	Enable importing of the SNV calls into the database.	Yes	No
CNV Calling Mode (CytoSure NGS (< 0.2.13) only)	Select the calling mode to either germline or somatic variants.	Germline	Somatic
Enable Translocation Detection	Enable or disable detection of translocations.	Yes	No
Reference Samples for CNV Detection	Specify to use all samples in the batch, select specific samples from all samples in the system, or indicate that the user should select one or more of the samples in the batch when the batch is started.	All batch samples	Specific samples, Samples from batch
Call PTDs	Enable calling of partial tandem duplications	Yes	No
Call ITDs	Enable calling of internal tandem duplications	Yes	No
Run UMI Processing (Somatic only)	Enable analysis with unique molecular identifiers (UMIs)	Yes	No
Hotspots (Somatic only)	Indicate which variants should be specifically analysed by the pipeline for monitoring purposes	(empty)	

Advanced Pipeline Configuration

Category	Setting	Description	Default Setting	Other Options
SNV Detection	Flanking Region (bp)	The amount of flanking sequence to include in the analysis.	0	Any value from 0 to 60
SNV Detection	Allele Balance Priors Off	Disable use of aggregate probability of observation balance between alleles as a component of the priors.	Yes	No
SNV Detection	Minimum Alt Count	The minimum alternative allele read count	5	
SNV Detection	Minimum Alt Fraction	The minimum alternative allele read fraction	0.2	
SNV Detection	Minimum Base Quality	The minimum allowed base quality	20	
SNV Detection	Minimum Total Read Count	The minimum total read count	20	

Category	Setting	Description	Default Setting	Other Options
SNV Detection	Minimum Mapping Quality	The minimum mapping quality of the reads	30	
SNV Detection	Pooled Continuously	Output all alleles which pass input filters, regardless of genotyping outcome or model.	No	Yes
SNV Detection	Pooled Discretely	Assume that samples result from pooled sequencing and model pooled samples using discrete genotypes across pools.	No	Yes
SNV Detection	Max Read Mismatch Fraction	The maximum fraction of mismatches in the read	0.04	
Quality Control	Threshold for Targets Not Covered	The coverage threshold to consider a target as not covered	1	
CNV Detection	Min Depth for CNV Segment BAF	The minimum depth required calculating the CNV segment B-allele frequency	20	
CNV Detection	Segmentation p-value Threshold	The significance threshold to accept segment	0.01	
CNV Detection	Drop Outliers Threshold	Drop bins that lie more than this many multiples of the 95th quantile away from the average within a rolling window as they are considered as outliers.	10	
CNV Detection	Use CBS with Smoothing	Use smoothing with the circular binary segmentation algorithm.	Yes	No
CNV Detection	Copy Number Thresholds	Thresholds for calling each integer copy number, i.e. 0,1,3,4	-2.0, -0.25, 0.2, 0.7	
Hardware Resources	# CPUs for Processing	The number of CPUs to be used by the protocol	12	
Hardware Resources	RAM(GB)	The amount of memory to be used by the protocol	2	
Structural Variant Detection	Threshold for PTDs	Exon-to-control ratio threshold to indicate a PTD	1.9	
Structural Variant Detection	Minimum Reads for Translocation	The minimum number of reads required to call translocations	3	

Category	Setting	Description	Default Setting	Other Options
CNV Detection	Mosaicism Correction Factor (Duplications)	The mosaicism correction factor for duplications	0.1	
CNV Detection	Mosaicism Correction Factor (Deletions)	The mosaicism correction factor for deletions	0	
UMI Processing	Minimum Input Base Quality	Minimum base quality to be used to call molecular consensus reads	10	
UMI Processing	Minimum Read Mapping Quality	Minimum read mapping quality to be used in reads grouping by UMI	30	
UMI Processing	Minimum Reads Supporting Consensus	Minimum number of reads supporting a consensus base/read. Used in consensus reads filtering. 0 indicates no filtering.	0	
Fusion Detection	Supporting Threshold	Minimum number of reads in support of the fusion call	5	

Germline Configuration

Setting	Description	Default Setting	Other Options
Germline Minimum Alt Fraction	The minimum alternate allele fraction	0.2	
Germline Pooled Continuously	Output all alleles which pass input filters, regardless of genotyping outcome or model.	No	Yes
Germline Pooled Discretely	Assume that samples result from pooled sequencing and model pooled samples using discrete genotypes across pools.	No	Yes
Germline Segmentation p-value Threshold	The significance threshold to accept segment	0.01	

Setting	Description	Default Setting	Other Options
Germline Copy Number Thresholds	Set the log ratio thresholds for each of the copy number variations.	<table><tr><th>Copy Number</th><th>Log Ratio</th></tr><tr><td>0</td><td>-2</td></tr><tr><td>1</td><td>-0.25</td></tr><tr><td>3</td><td>0.2</td></tr><tr><td>4</td><td>0.7</td></tr></table>	

Somatic Configuration

Setting	Description	Default Setting	Other Options
Somatic Minimum Alt Fraction	The minimum alternate allele fraction	0.01	
Somatic Pooled Continuously	Output all alleles which pass input filters, regardless of genotyping outcome or model.	Yes	No
Somatic Pooled Discretely	Assume that samples result from pooled sequencing and model pooled samples using discrete genotypes across pools.	Yes	No
Somatic Segmentation p-value Threshold	The significance threshold to accept segment	0.05	

<table>
<tr><th>Setting</th><th>Description</th><th>Default Setting</th><th>Other Options</th></tr>
<tr><td>Somatic Copy Number Thresholds</td><td>Set the log ratio thresholds for each of the copy number variations.</td><td><table><tr><th>Copy Number</th><th>Log Ratio</th></tr><tr><td>0</td><td>-2</td></tr><tr><td>1</td><td>-0.07</td></tr><tr><td>3</td><td>0.07</td></tr><tr><td>4</td><td>0.7</td></tr></table></td><td></td></tr>
<tr><td>Tumour Content Estimation</td><td>Enable the software to estimate the percentage tumour content in the sample</td><td>Yes</td><td>No</td></tr>
</table>

14.3.4 Panels

The Panels page displays information about the panels currently loaded into the software.

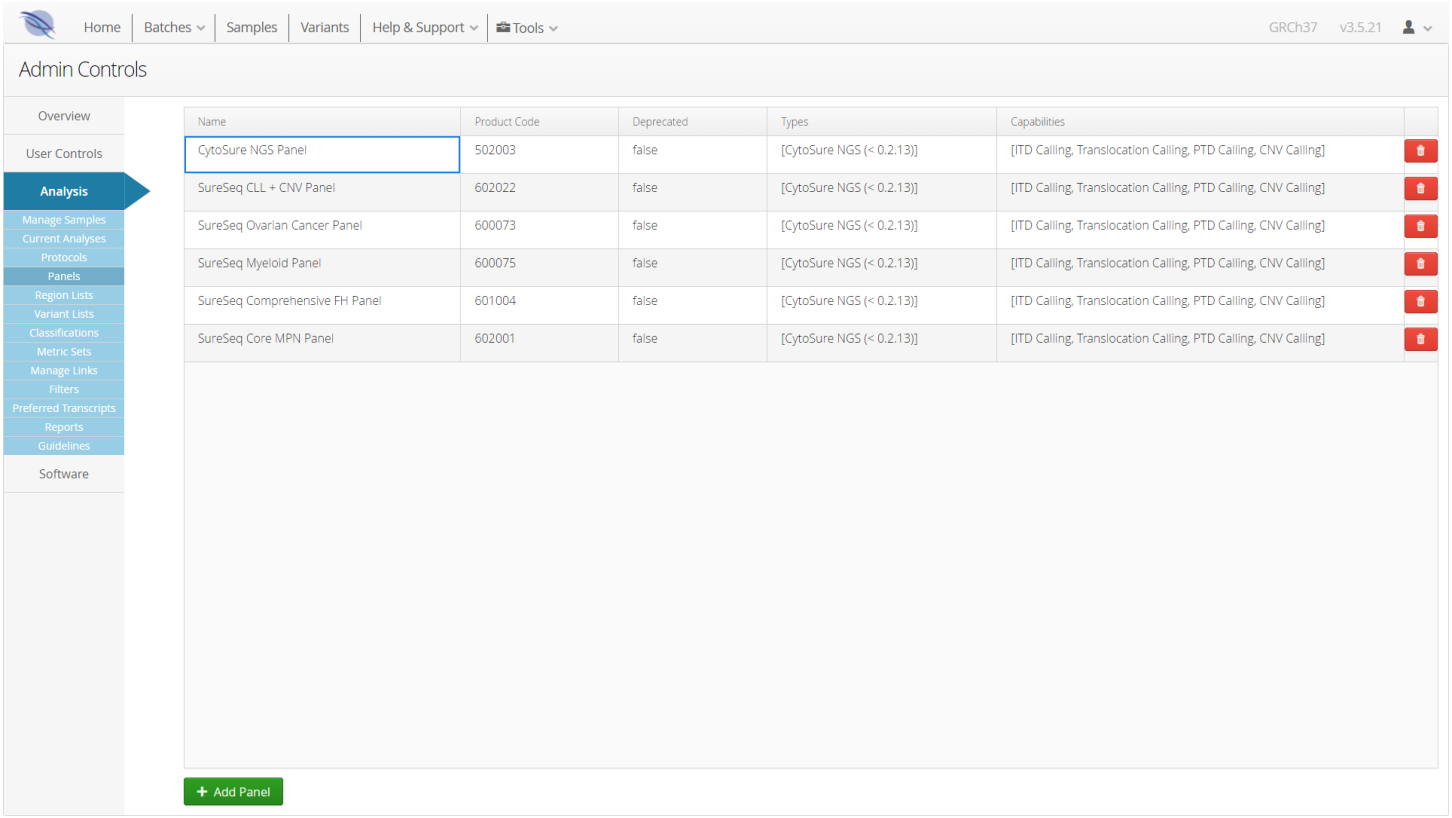


Figure: The panel start page

Adding New Panels

New panels can be added via the Add Panel button.

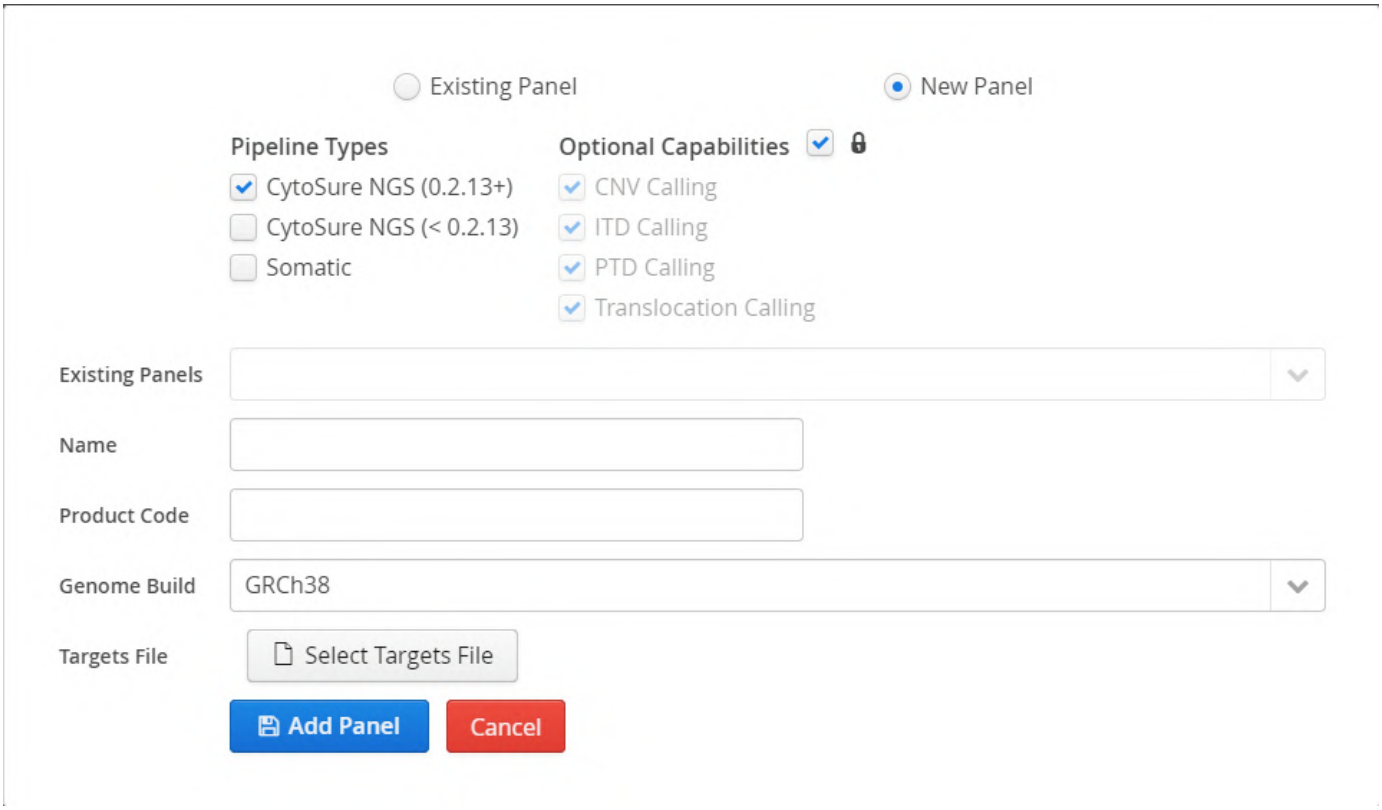


Figure: The Add Panel window.

To add a new panel to the system, users must indicate with which of the available Pipeline Types, and which of the Optional Capabilities of those pipeline types, the panel is compatible (contact OGT support if unsure). Assign a unique Name and Product Code, select the correct Genome Build for the Targets File, then select the file from the file system by clicking the Select Targets File button, and finally click Add Panel.

> Targets File Format
> Targets Files should be provided by OGT. Standard BED files are not compatible with Interpret and their use will result in pipeline failure. Contact OGT for the correct Targets File for your panel and build.

The Add Panel window may also be used to add a targets file to an existing panel for different genome build - select the Existing Panel radio button and the select the appropriate panel from the Existing Panels drop-down list in order to do this.

Modifying Existing Panels

Attributes of existing panels may be modified by double-clicking on the appropriate row of the table, making the required changes to the Name, Product Code, Types and/or Capabilities, and clicking Save.

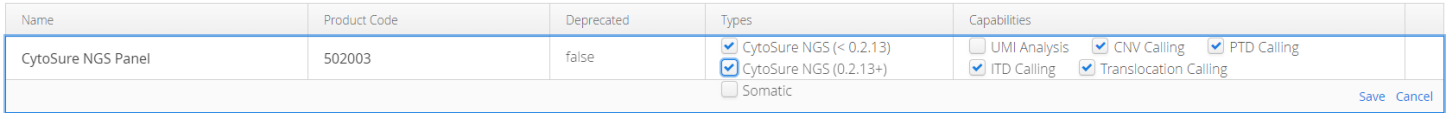


Figure: Editing attributes of a panel.

14.3.5 Region Lists

Region lists are a set of defined genomic regions

Figure: The Region Lists start page

To add a region list users select the 'Add Region List' button which provides a form.

Figure: Form for adding a Region List

Regions can be defined using three parameters:

1. Regions can be imported from a file.
2. Regions can be defined as genomic features
3. Regions can be defined as chr:start-finish

The first step is defining the name of the region list

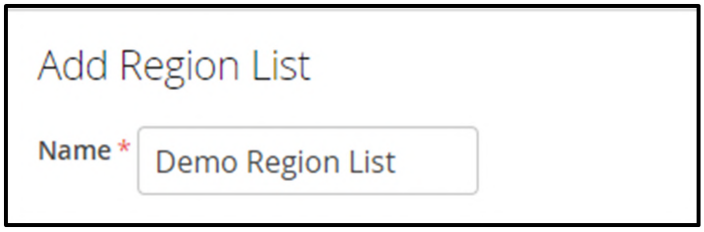


Figure: Setting the region list name

Importing Regions from a File

Importing a file of defined regions requires the file be in BED or zipped BED formats.

Figure: The file chooser for importing regions from a file

Selecting 'Choose File' opens a file browser and when a file is selected it is shown in the window, as below.

Figure: Selection of a file to be imported

When the Import button is pressed the selected file is checked and if there is a problem with the file format the following is displayed.

Figure: File import error message

Clicking on the error will remove it.

If the file format is correct then the user needs to select the genome build from the drop-down menu.

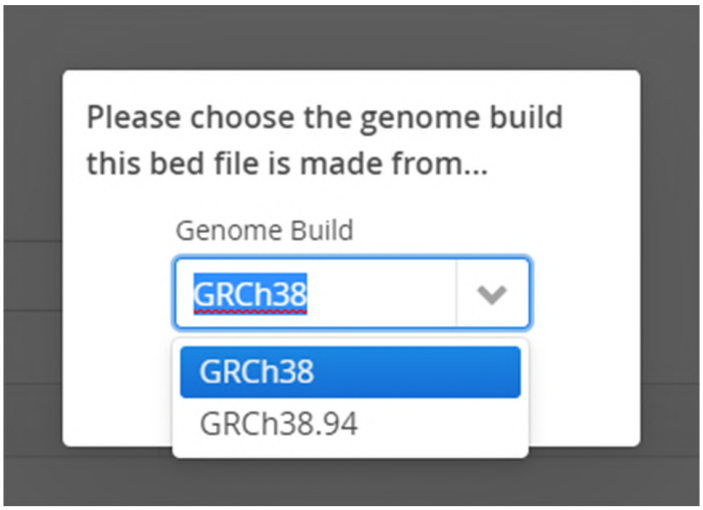


Figure: Selecting the genome build

Finally, selecting Import completes the process

Figure: Select import to complete the upload

And the regions are now displayed in the table.

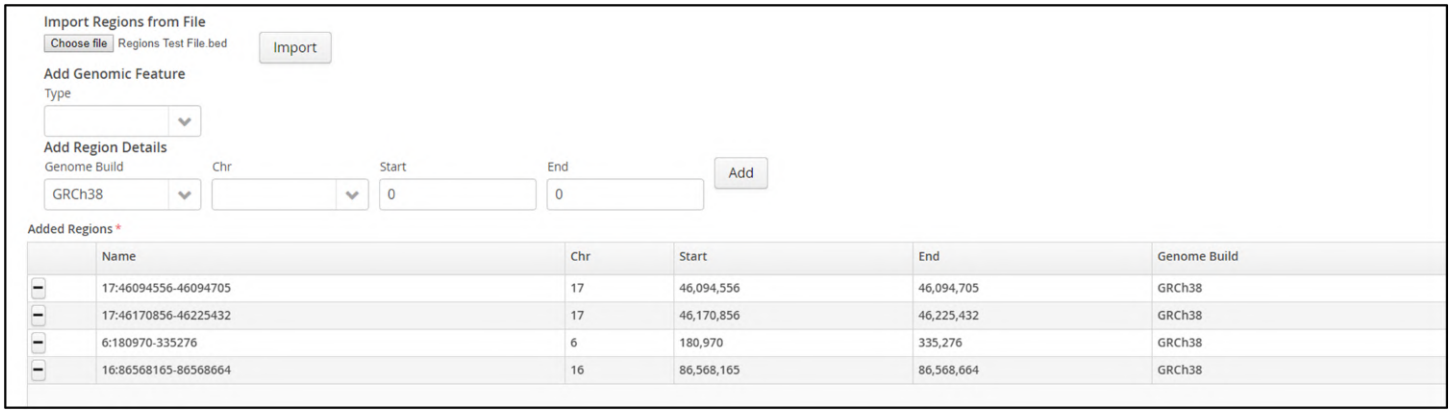


Figure: Imported regions displayed

The upload of regions can be edited by selecting the − icon on the side to remove the region on that row.

Defining a Region as a Genomic Feature

Regions can be added as genomic features selected from the drop-down menus.

Currently available features are exons, genes, proteins and transcripts.

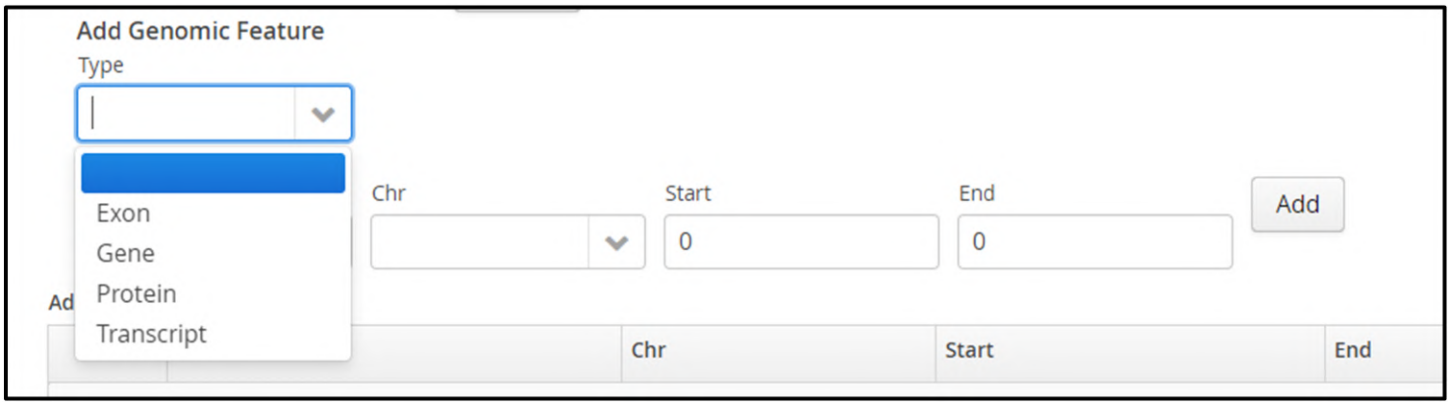


Figure: Types of genomic feature available

When a type has been chosen a second menu is provided containing the features of the chosen type that can be selected.

Users can scroll up and down the list to find the required feature

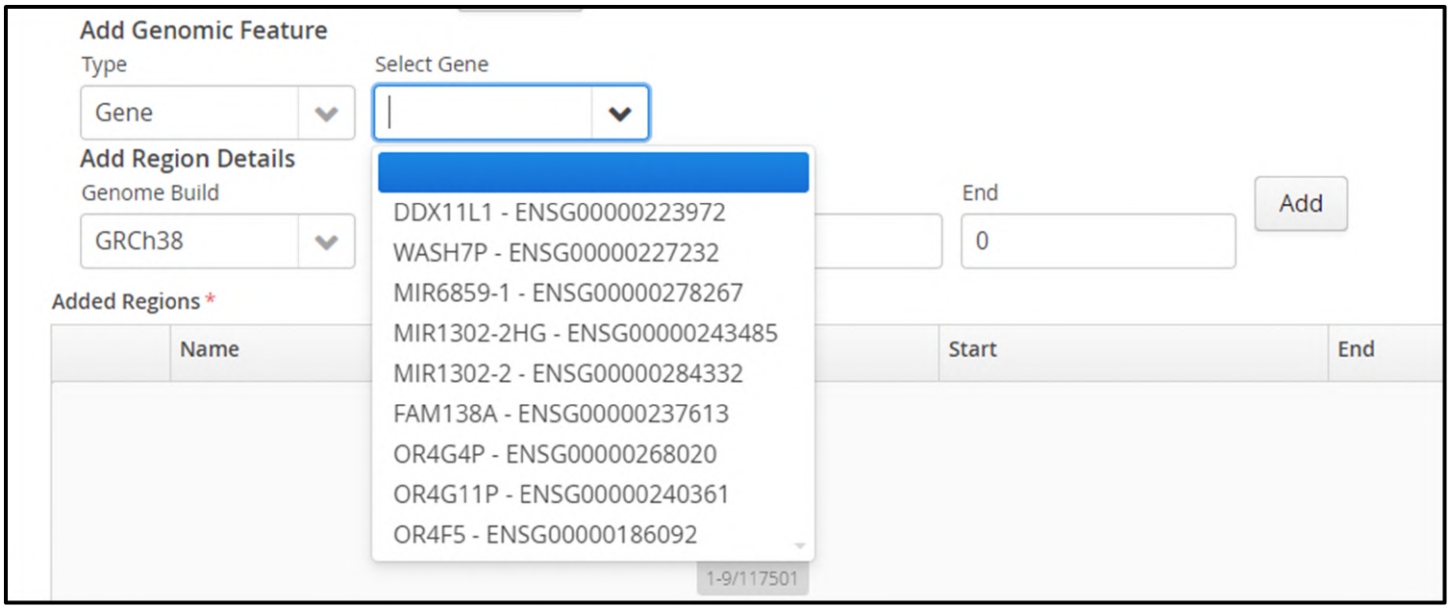


Figure: List of genes available to add as a region

Alternatively, typing text in the text field will automatically subset those features that contain the text.

In the figure below the user has typed DMT and the menu lists all features in which this text is found.

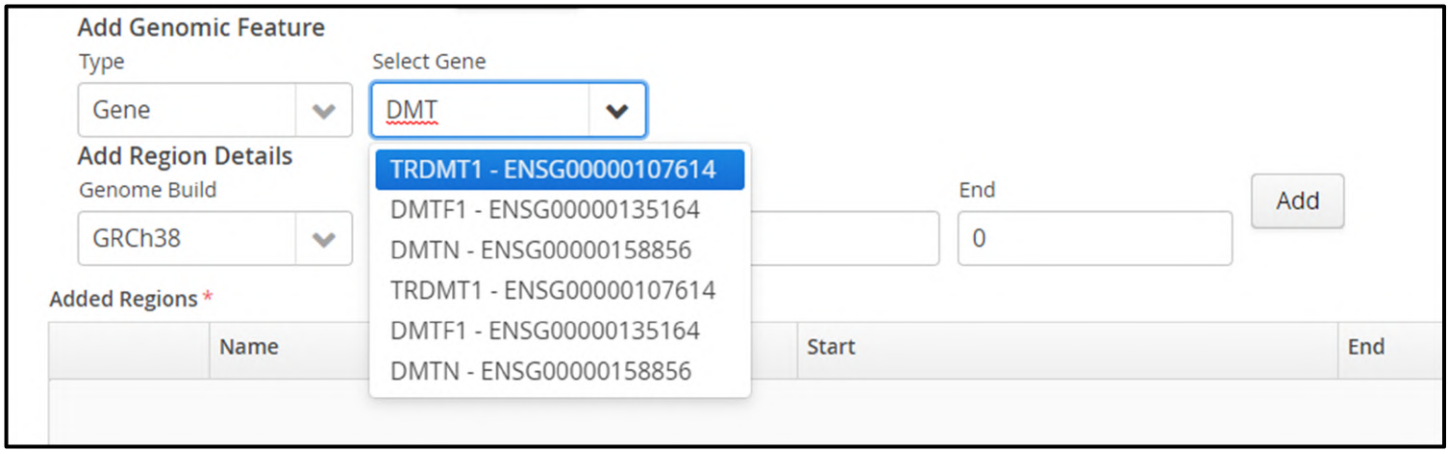


Figure: Filtering genes available by

Pressing return selects the feature and adds it to the region list.

Figure: Addition of a genomic feature to the region list

When selecting features, the drop-down lists are populated with the following information:

- Exon - the EnsEMBL ENSE number
- Gene - the gene name
- Protein - the EnsEMBL ENSP number
- Transcript - the EnsEMBL ENST number

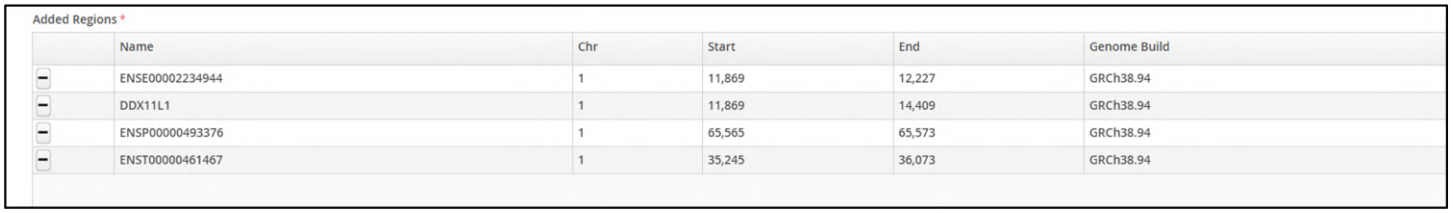


Figure: A region list showing different genomic features

Regions can be defined as chromosome, start and end

The final mechanism by which to create a region is to manually specify the build, chromosome, start and end positions.

Figure: The add region detail section of the form

The drop-down menu provides a list if available genome builds.

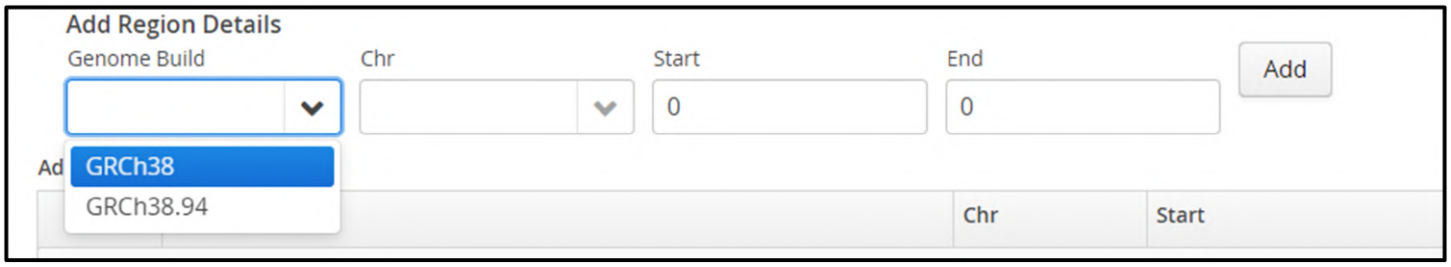


Figure: Specification of the required genome build

Subsequently, the chromosome can be selected from he populated list.

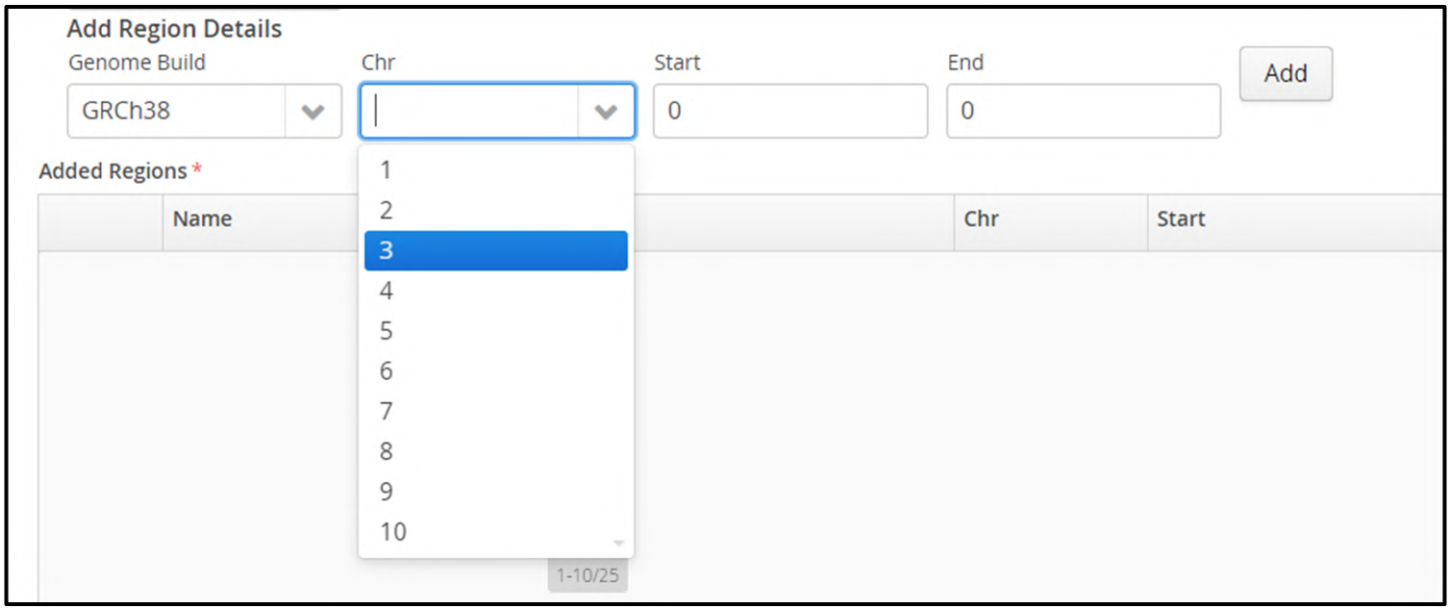


Figure: Selection of the chromosome from the drop-down menu

Lastly, the user can enter the start and end coordinates.

Figure: selection of region details

Pressing Add appends the region to the table

Figure: Addition of the region details to the new region list

Pressing save completes the process and you now see that there is a new Region List called Demo Region List that comprises 4 defined regions.

Figure: Addition of the new region list

Using a Region List

When a Region List has been defined it can be used to filter variants.

Region Lists can be found in the 'Region/Variant List' category and then the 'Region List' type.

All available Region Lists will be presented in the drop down menu.

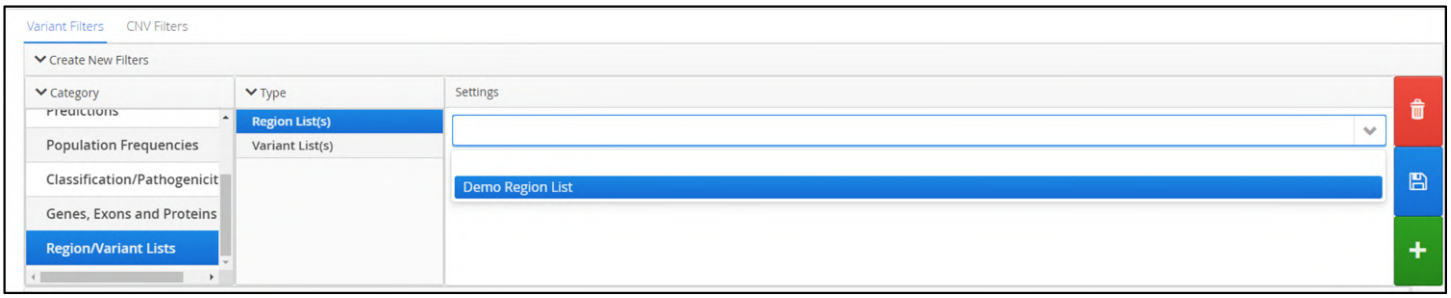


Figure: Availability of the new region list in the Filters

When it has been selected a Region List is added to the filter builder.

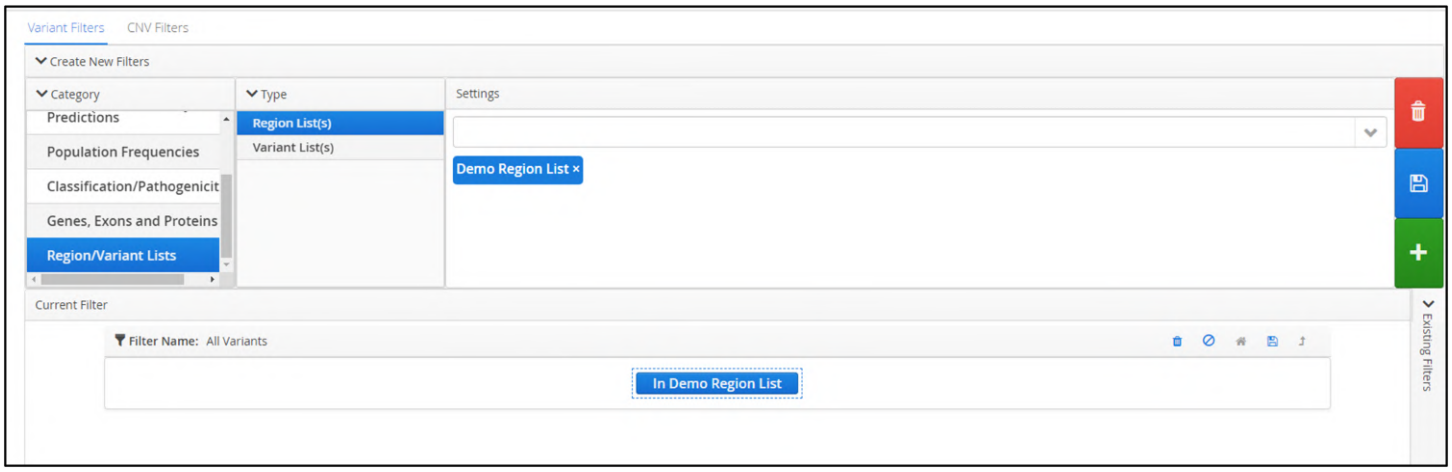


Figure: Addition of the Demo Region List to the filter

Alternatively, the newly created Region List can be used to dynamically filter variants.

In the figure below the Demo Region List has been deployed dynamically with the result that 3449 SNPs are filtered down to just 1 SNV.

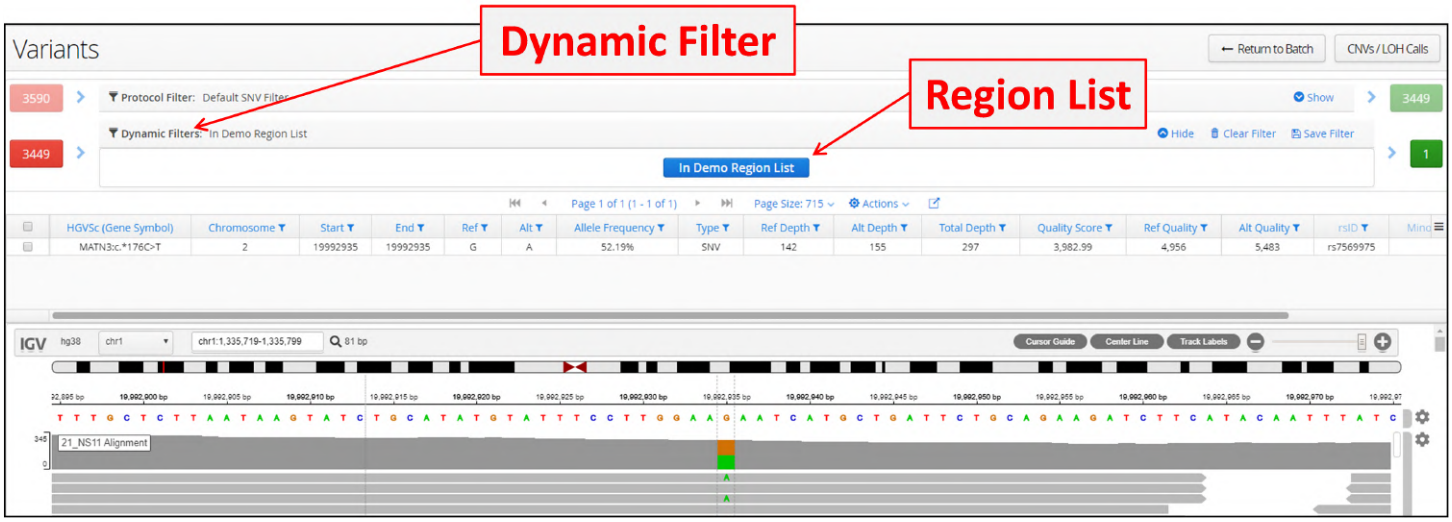


Figure: Deployment of a region list in a dynamic filter

14.3.6 Variant Lists

Interpret allows users to generate variant lists that can be used as a filter in analysis of data generated for samples.

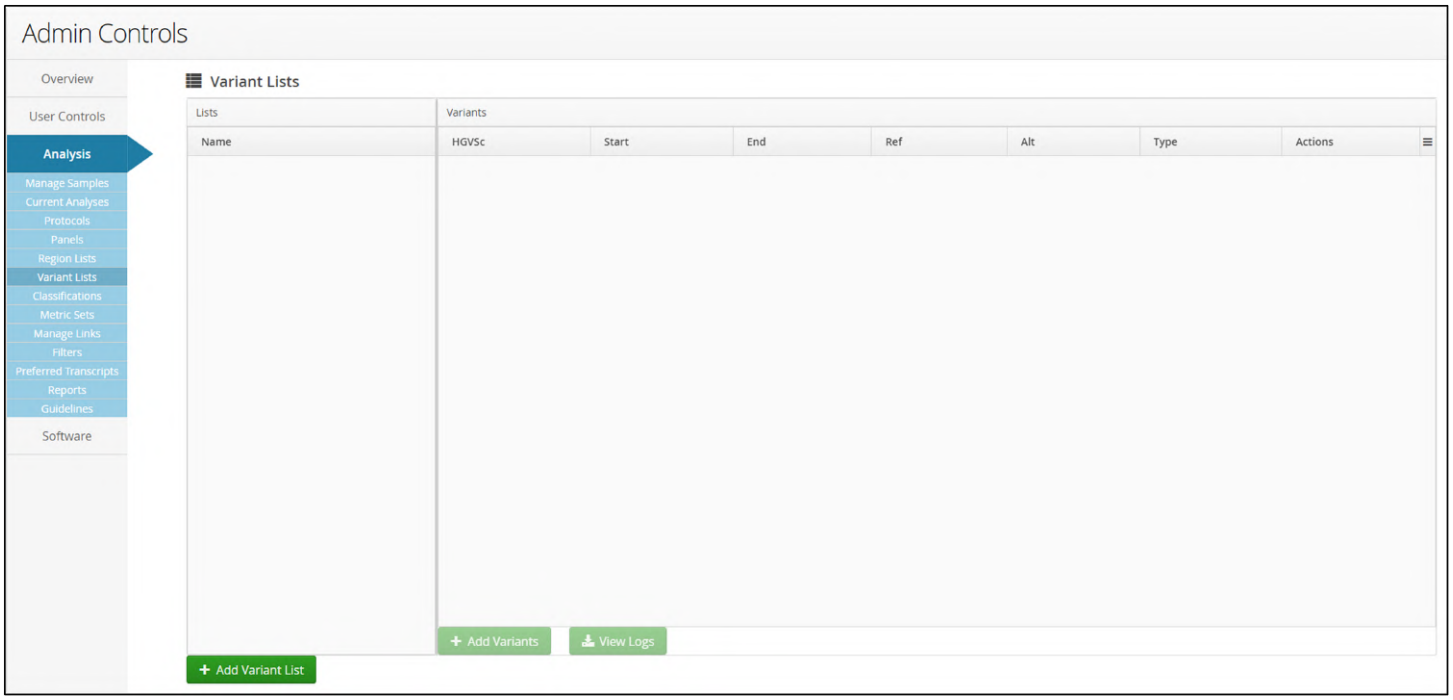


Figure: The Variant List start page

Selecting Add Variant List provides a popup where the user can enter the name of the new variant list.

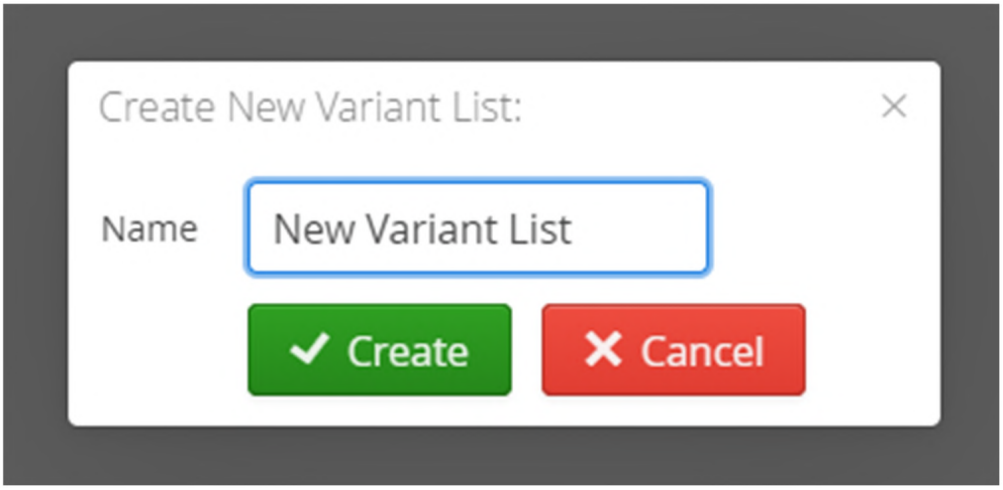


Figure: Creating a new variant list

Once a variant list has been created, variants can be added from files.

The format for uploading variants is either VCF or HGVS formats.

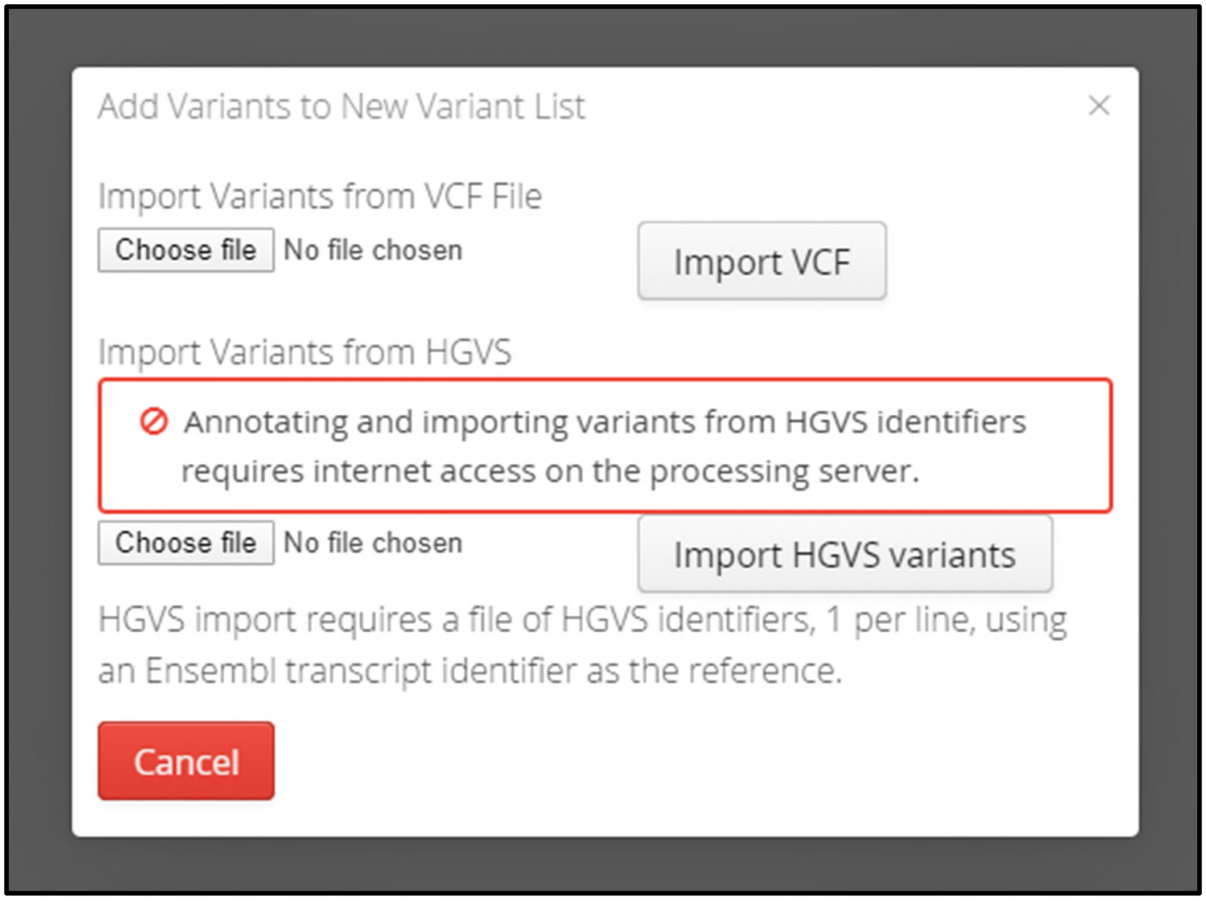


Figure: Uploading variants to a variant list

Once created and populated with variants the new variants list is available as filter.

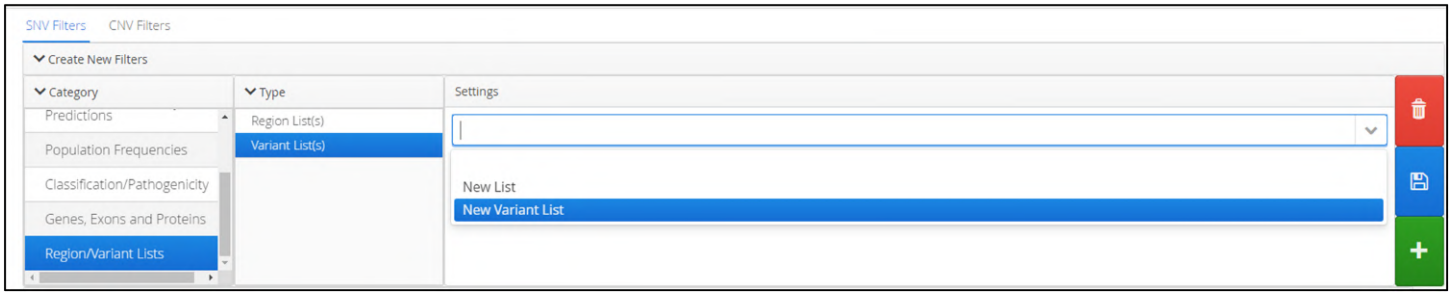


Figure: Selecting the new variant list

The variant list can also be appended to within the results page. Right clicking on a variant provides an option to add a variant to an existing variant list.

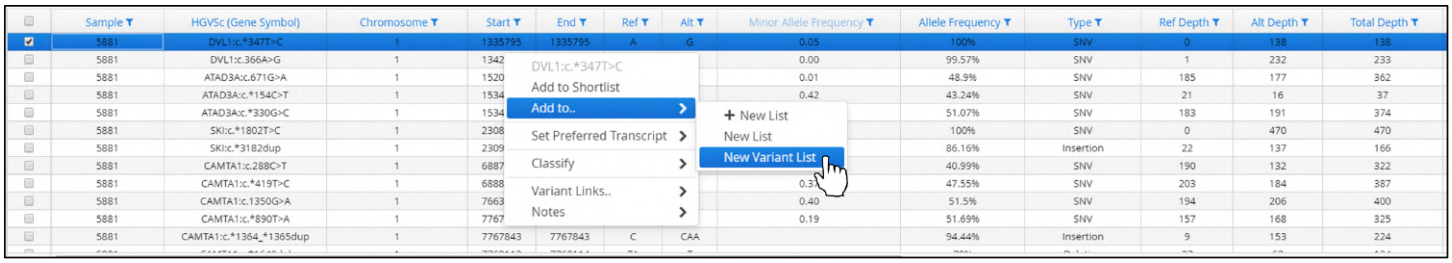


Figure: Adding a variant to the new variant list

14.3.7 Classifications

It can be helpful to assign colours to particular classifications and in this page users are able to create new and modify existing assignments.

By default the software will ship with 5 classifications already created. These are listed below and displayed in Figure XX.

1. Benign
2. Uncertain significance (likely benign)
3. Uncertain significance
4. Uncertain significance (likely pathogenic)
5. Pathogenic

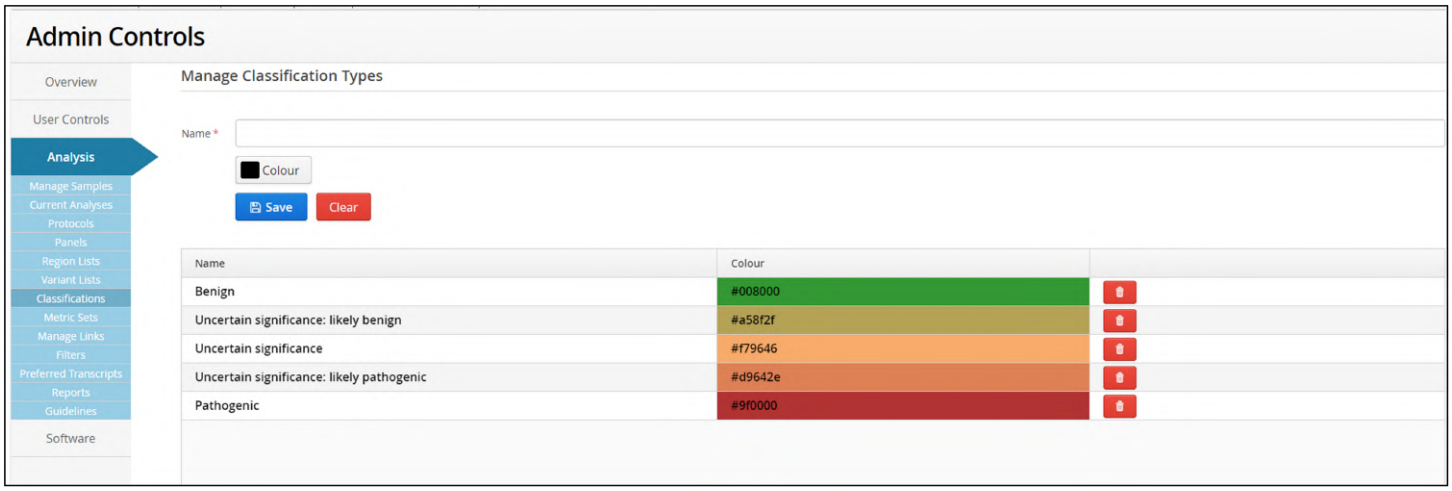


Figure: View of the default settings in Classifications

Adding a Classification

Figure: Setting a new classification name

By default the colour to be assigned is black but users can select a different colour by pressing on the Colour button.

This will produce a sub-window containing 3 tabs of different colour palettes called RGB, HSV and Swatches, each of which is displayed below.

Selecting on a name will display the tab or that palette.

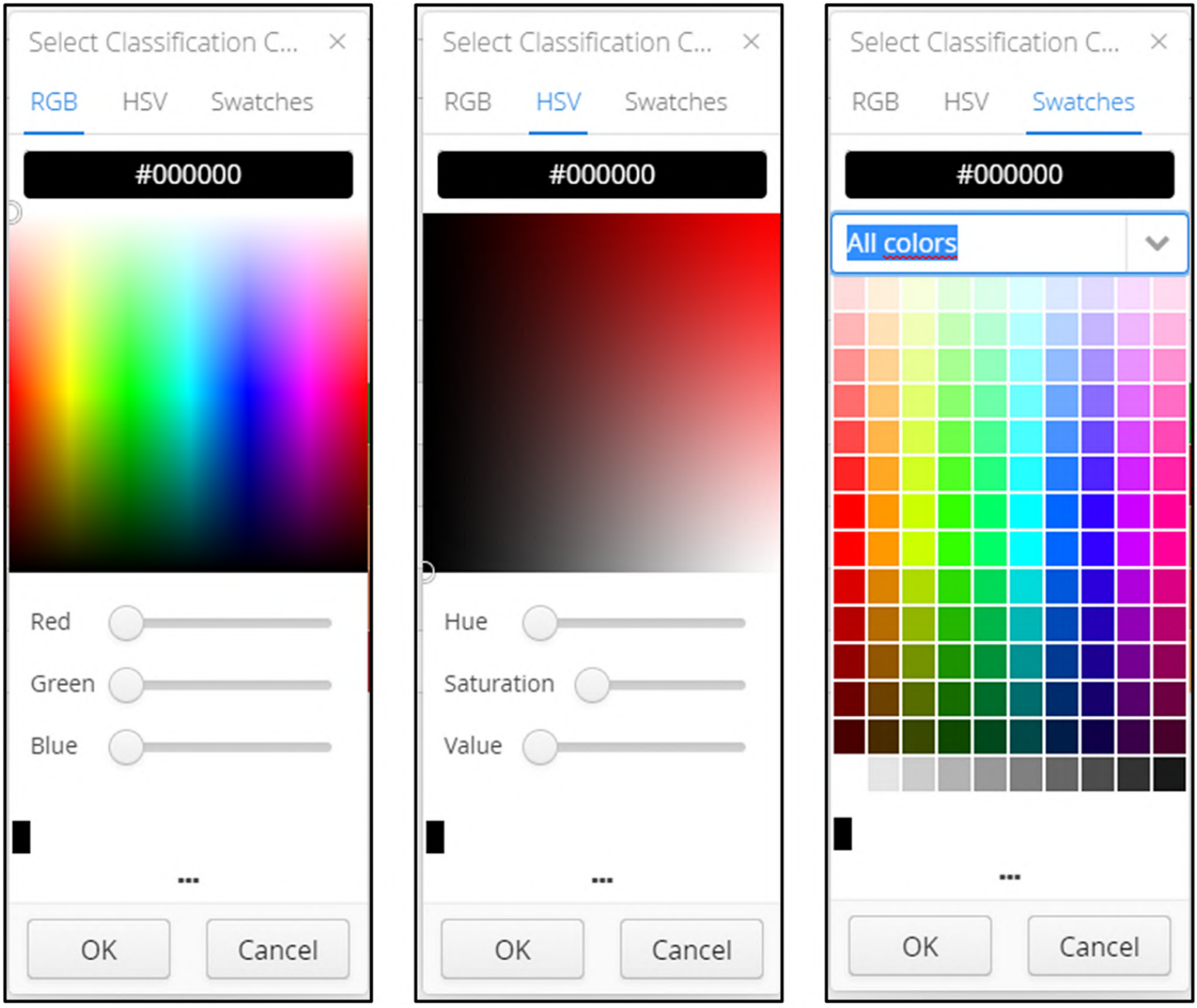


Figure: Colour palettes available to setting a classification colour

Once the colour has been selected, you can click on OK to add the classification to the software

Figure: A new classification ready to add

And the new classification is displayed, as shown below:

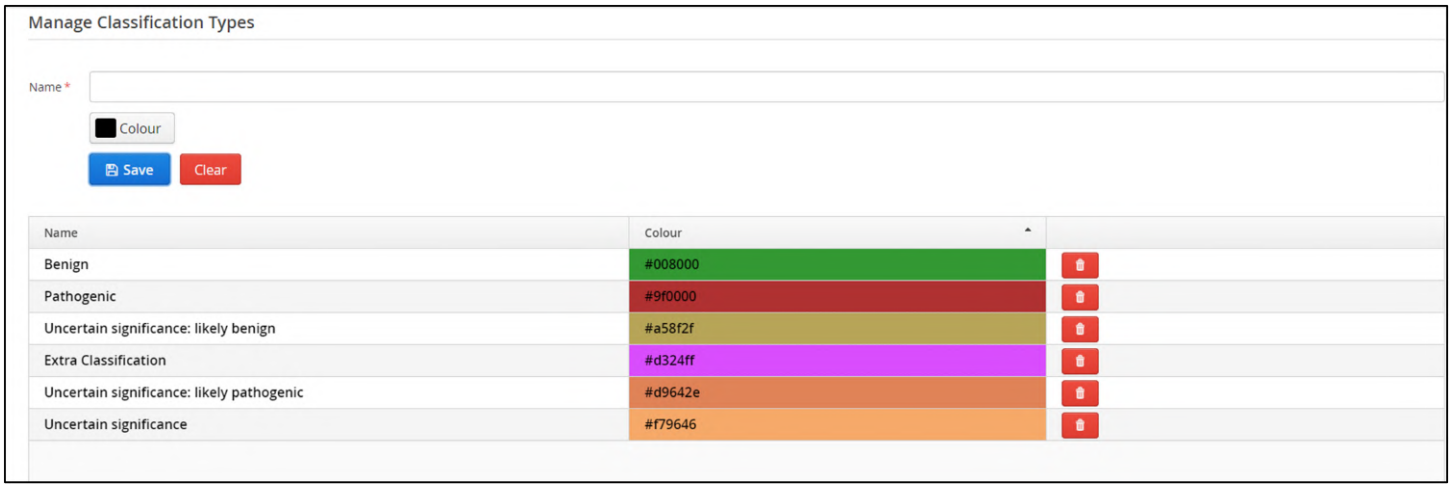


Figure: New classification added

A classification can be removed by clicking on the wasted bin icon on the row of the classification that needs to be removed.

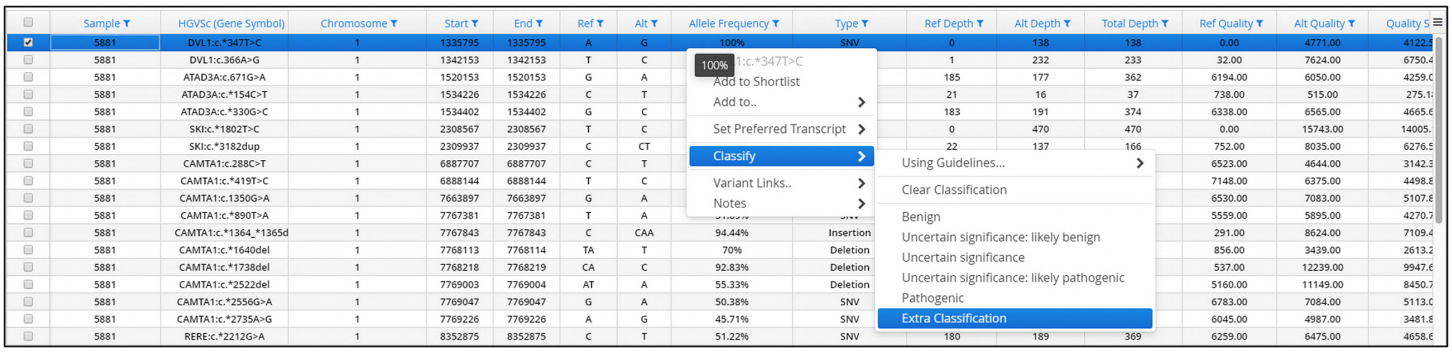


Figure: The new classification can now be used to apply to a variant

The variants table is then updated to show the colour of the new classification.

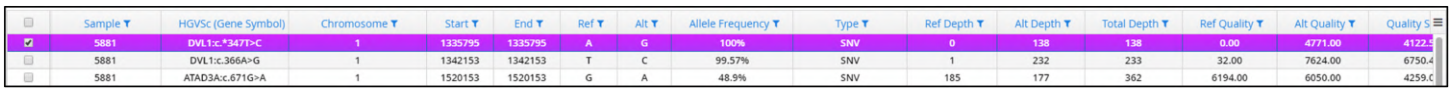


Figure: A variant annotated with the newly generated classification type

14.3.8 Metric Sets

The metric sets define parameters to qualitatively assess an analysis. Selecting the Metric Sets page will display all the metric sets that have been created in Interpret.

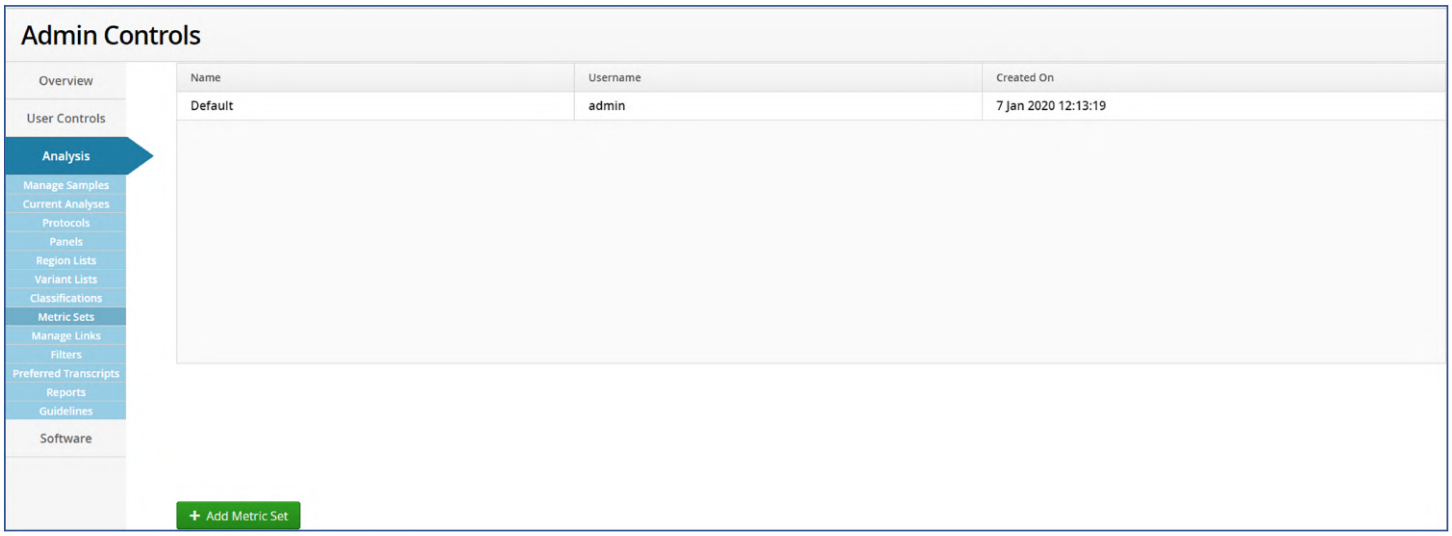


Figure: The metric sets start page

The settings in the metric set are used in the batch reports for samples that have been analysed. Each metric reported is coloured red, blue or green depending on the values set as Excellent, Pass or Failed accordingly.

Completed Samples

Sam...	View	# SNVs	# CNVs	# LOH	Mean Target Coverage	% Reads Aligned	% Duplication	Targets Not Cove...	% Usable On Target Reads	Report
5881	SNVs CNVs/LOH	2,754	8	16	344.21	99.55	9.08	138	71.95	
6937	SNVs CNVs/LOH	2,695	13	15	556.51	99.63	10.75	122	72.18	
7408	SNVs CNVs/LOH	2,740	7	12	381.59	99.64	8.6	142	72.41	
8210	SNVs CNVs/LOH	2,666	10	16	356.72	99.58	8.48	137	73.48	
10384	SNVs CNVs/LOH	2,650	4	17	313.08	99.63	8.05	150	73.06	
10847	SNVs CNVs/LOH	2,669	5	13	314.99	99.57	8.53	140	72.82	
11516	SNVs CNVs/LOH	2,571	7	16	407.31	99.63	9.29	136	72.07	
12878	SNVs CNVs/LOH	2,627	14	18	388.99	99.61	8.47	129	72.9	
14130	SNVs CNVs/LOH	2,614	18	14	396.32	99.61	8.41	150	72.49	

Figure: A table of completed samples showing metrics in different colours

For example the default metric set that is included in Interpret has the settings shown in the figure below.

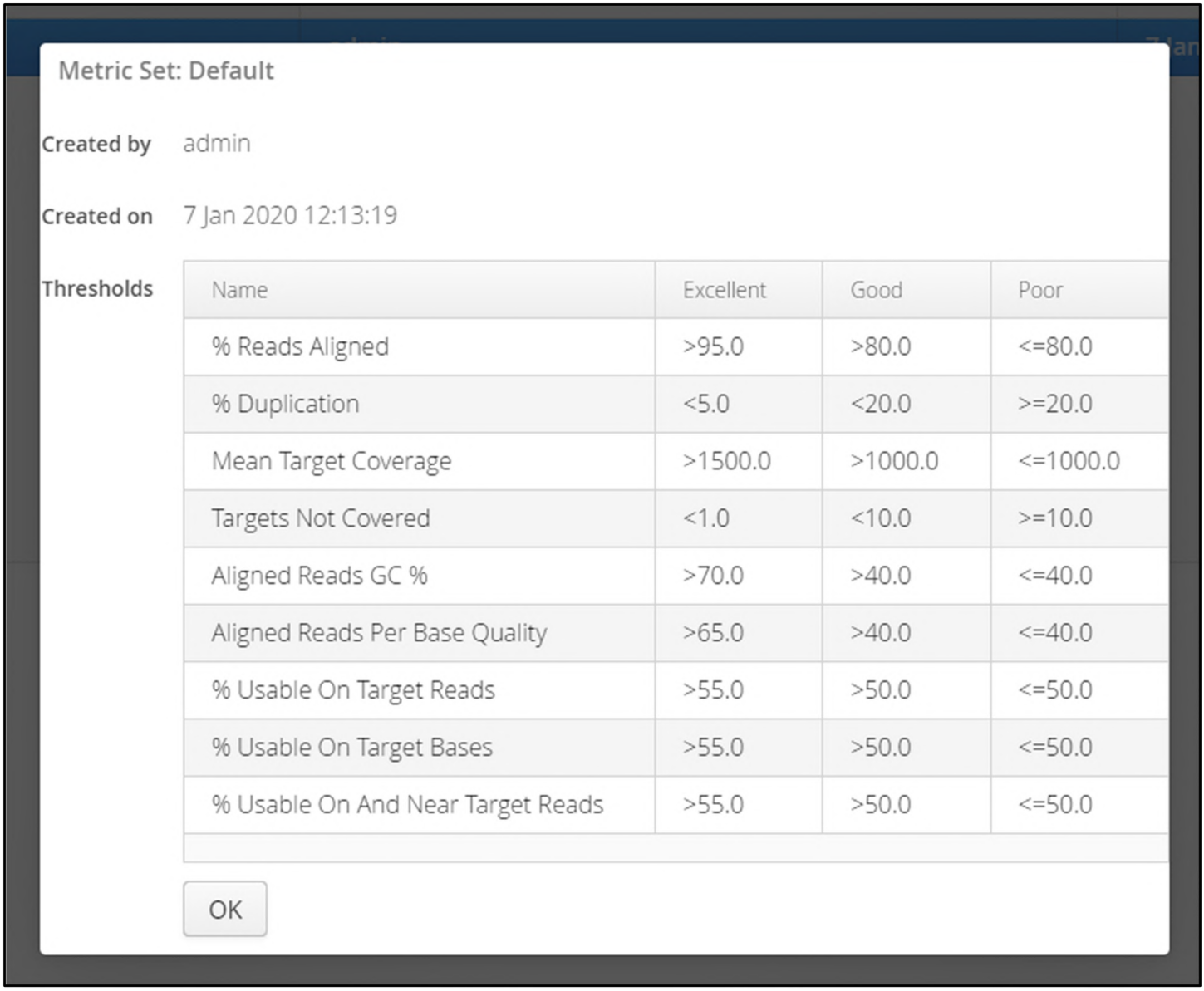


Figure: Settings in the default metric set

Selecting 'Add Metric Set' on the start page opens a window that allows users to set their own definitions.

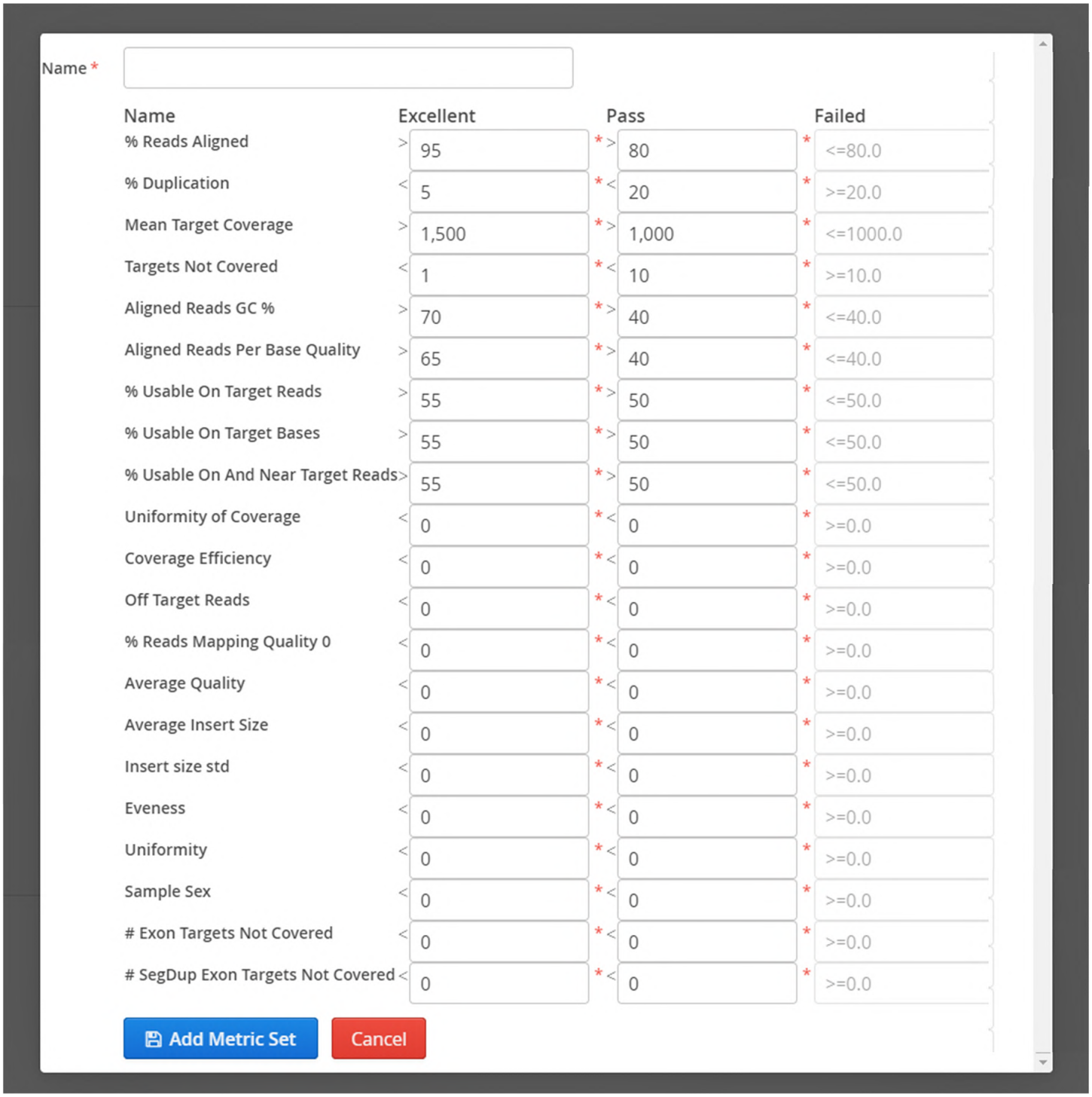


Figure: Settings available for metric sets

Once the settings have been made, select 'Add metric Set' to complete and the new metric set will be displayed in the Metric Sets start page.

Name	Username	Created On
Default	admin	7 Jan 2020 12:13:19
New Metric Set	admin	27 Feb 2020 10:57:46

Figure: New metric set displayed in start page

Once a metric set has been created it is available to use in a protocol.

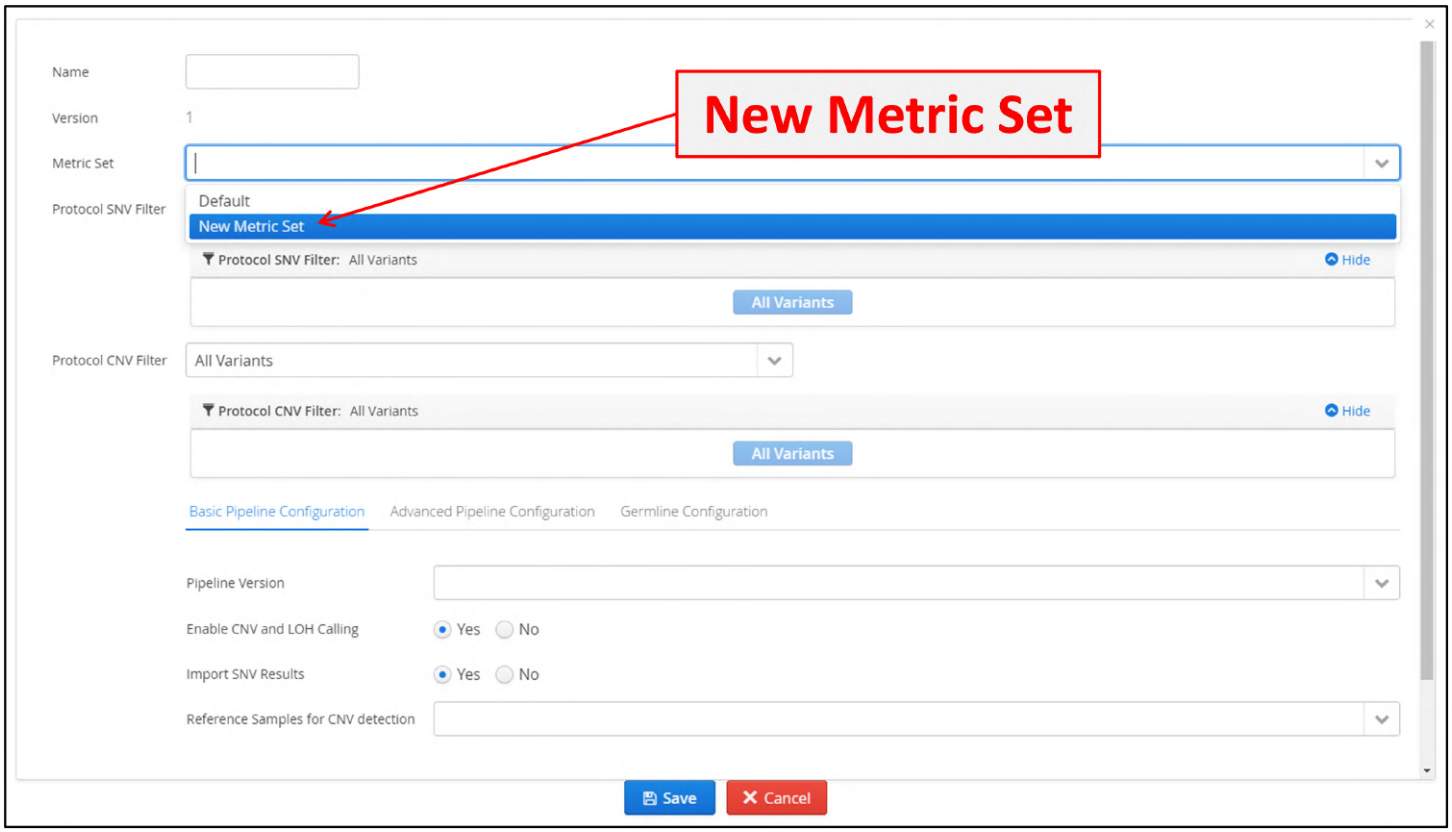


Figure: New metric set available for use in protocols

14.3.9 Manage Links

Manage links provides a means to define links to external resources such as ClinVar.

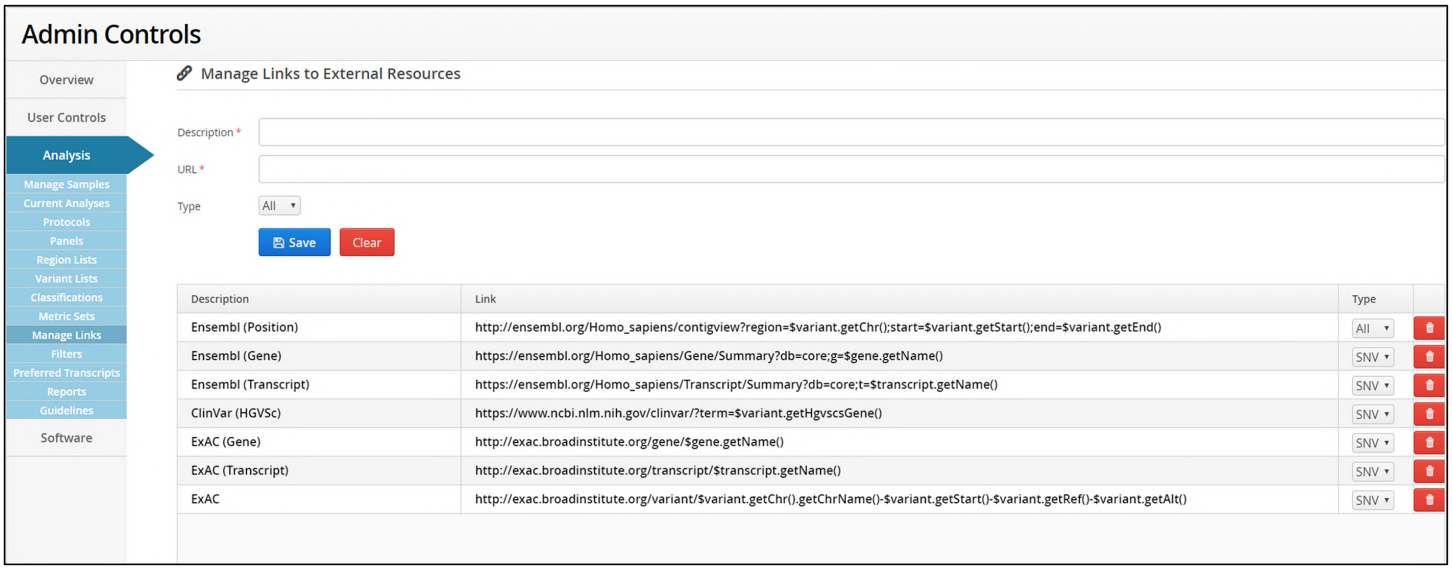


Figure: The Manage Links start page

Existing links can be removed by clicking on the waste bin icon.

New links can be generated by adding a description and URL in the table. Please contact OGT support to assist with this.

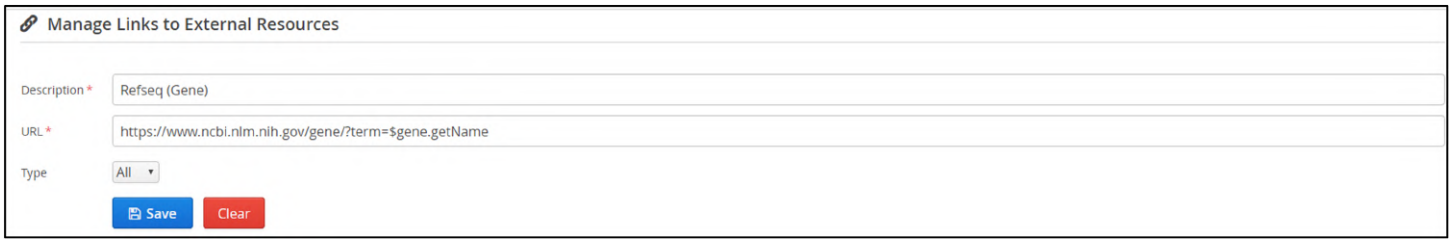


Figure: Defining a new link

A link type can be for only SNVs, only CNVs or both

Figure: Setting the link type

Once saved the link new link is displayed in the table on the start page.

Figure: New link displayed

The new link is now available to be used in the results page for variants.

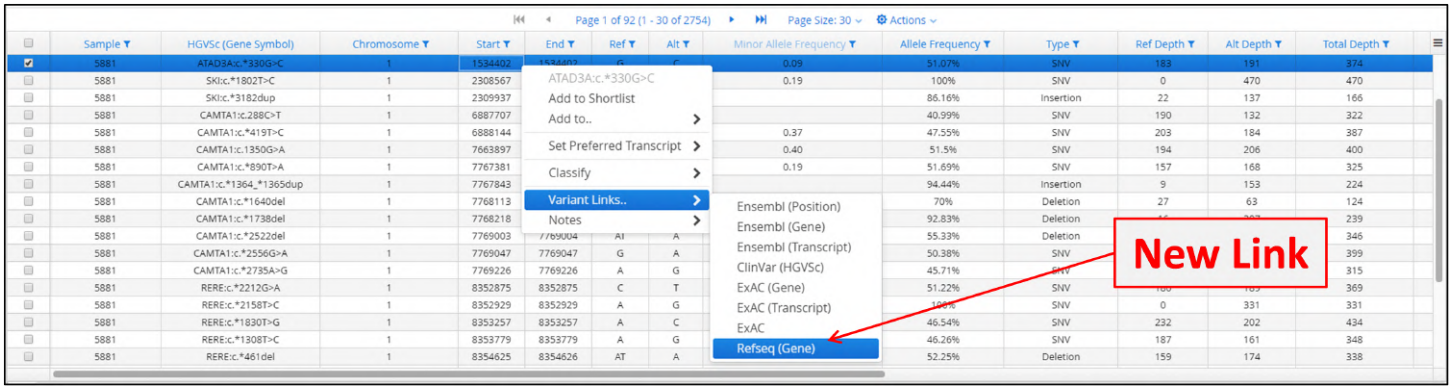


Figure: New link available for use in results tables

14.3.10 Filters

Filters provide the means to control which variants are displayed in the SNV and CNV variant tables.

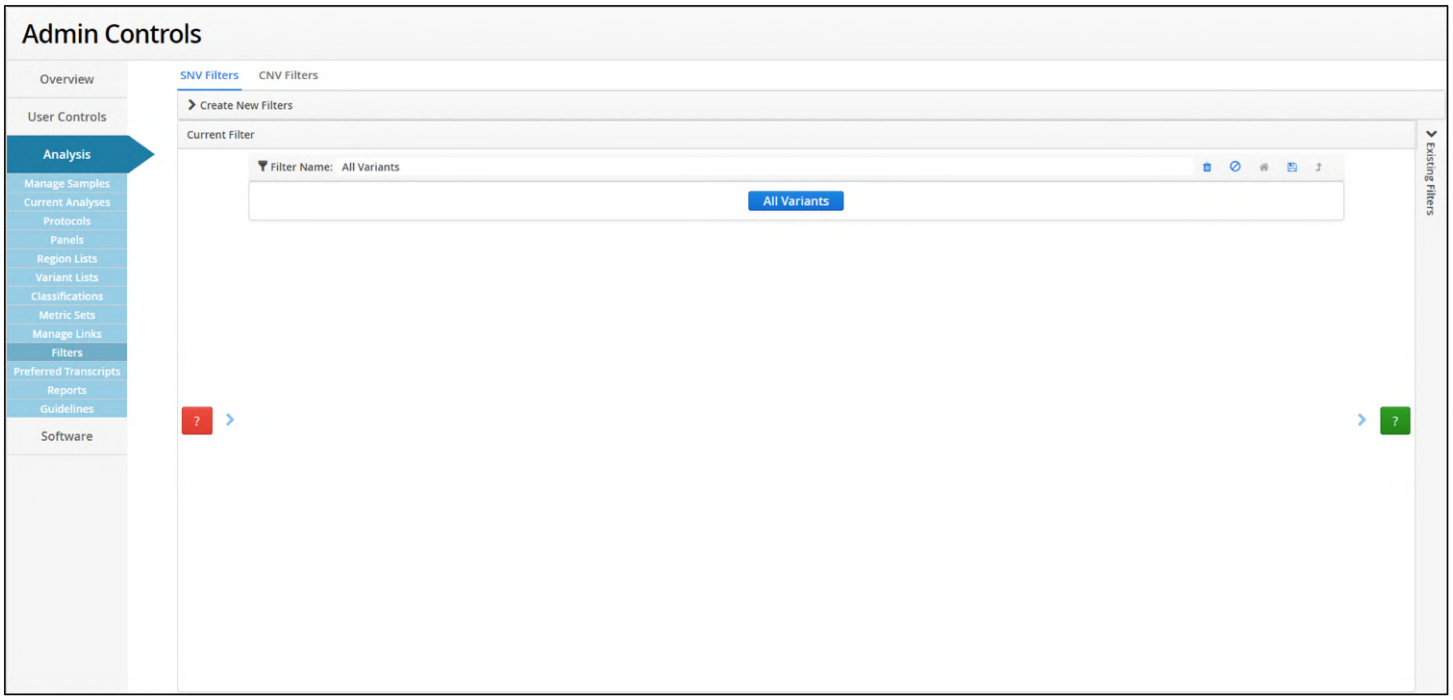


Figure: The start window for the filters

The filters page is divided into 3 parts

1. Creation of new filters
2. The current filter
3. Existing filters

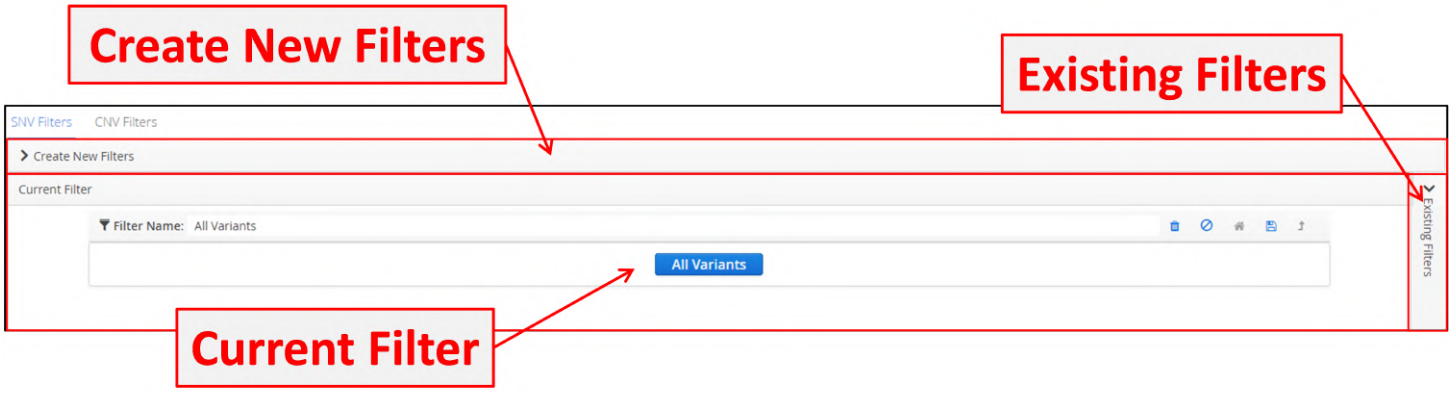
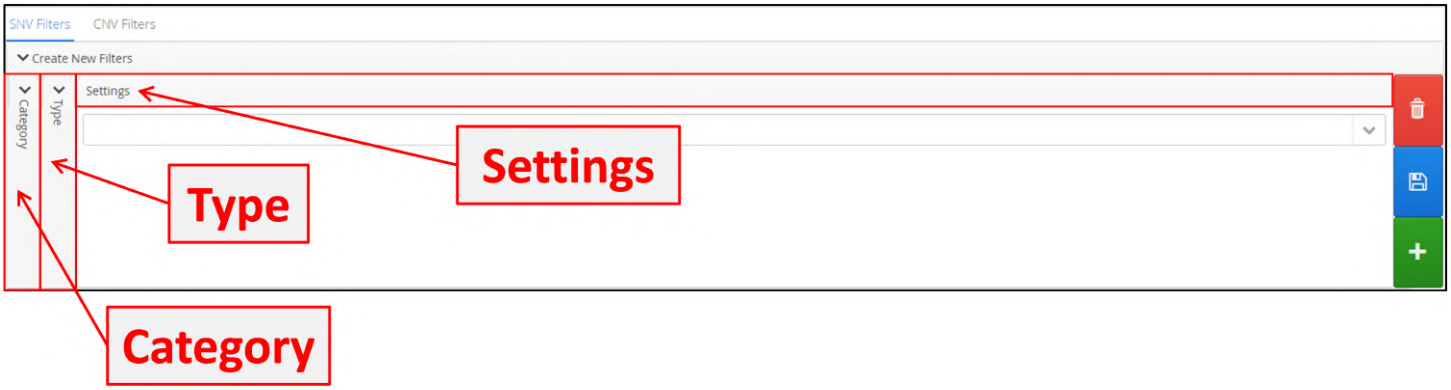


Figure: The different section of the filters page

New filters are created by selecting criteria and combining to generate a chain.

Opening the Create New Filters tab shows category, type and settings.



Within Category are a range of options and these are different for SNVs and CNVs.

Figure: SNV filter categories

Figure: CNV filter categories

Selecting one will then populate Type with the different values, for example below are some of the types of basic variant categories.

Figure: Example of displaying type available for a selected category

Selecting a type will then populate Settings with specific options for that type; for example below users can set a filter on selected chromosomes including mitochondrial (MT).

Figure: The setting available for the selected category type

To create a filter, users make the necessary selections and click on the on white plus in the green box to the side of the window.

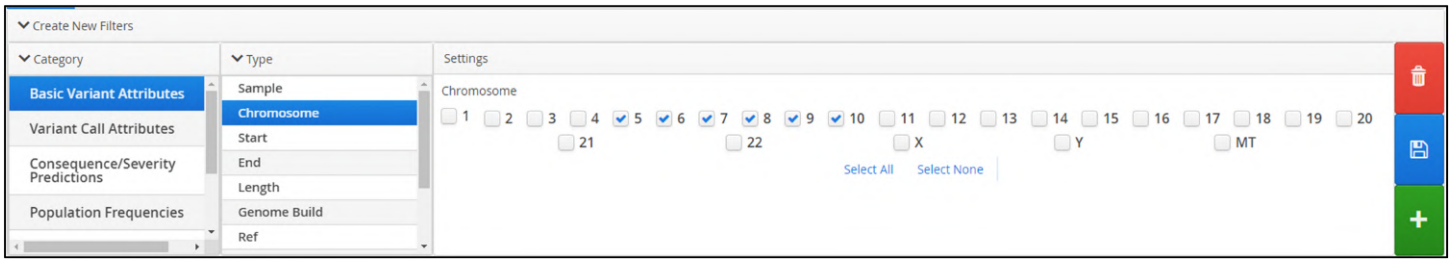


Figure: Selection of settings to use in a filter

This will add the filter to the "Current Filter" window.

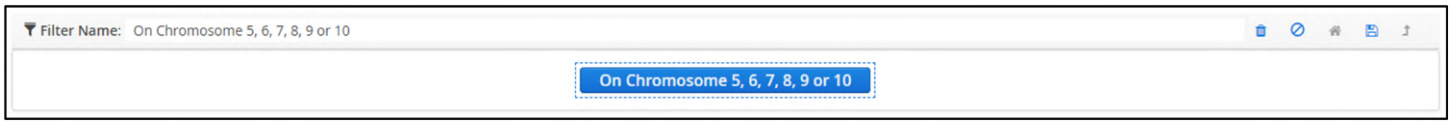


Figure: Display of the newly created filter

The filter options are numerous and very flexible and software allows users to chain filters together in order to tailor to a users needs.

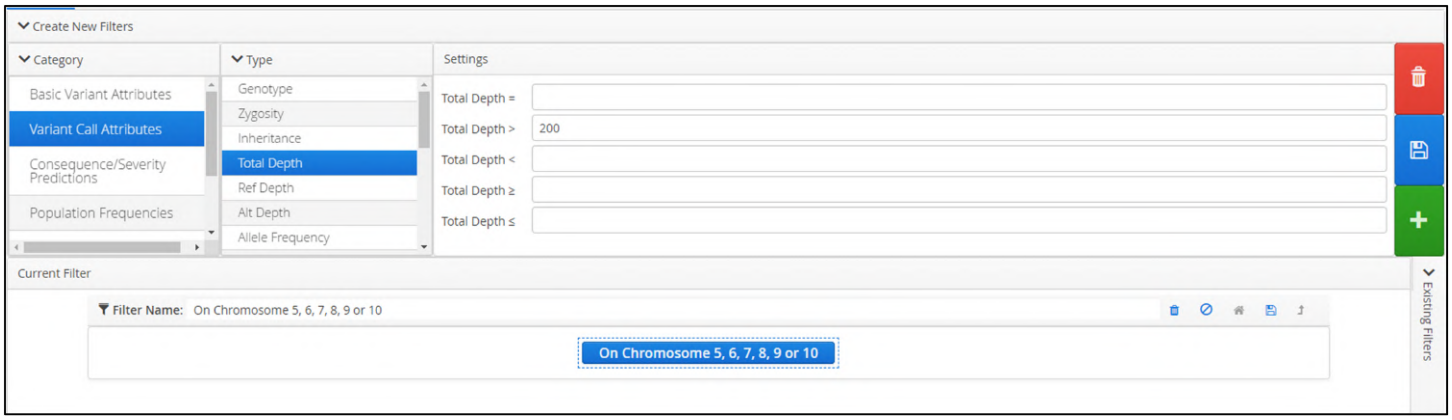


Figure: Specifying a second filter

Adding another filter provides the user with the option how the second filter will be combined with the first filter.

The user must now decide to specify whether the options are to satisfy either filter with an OR statement

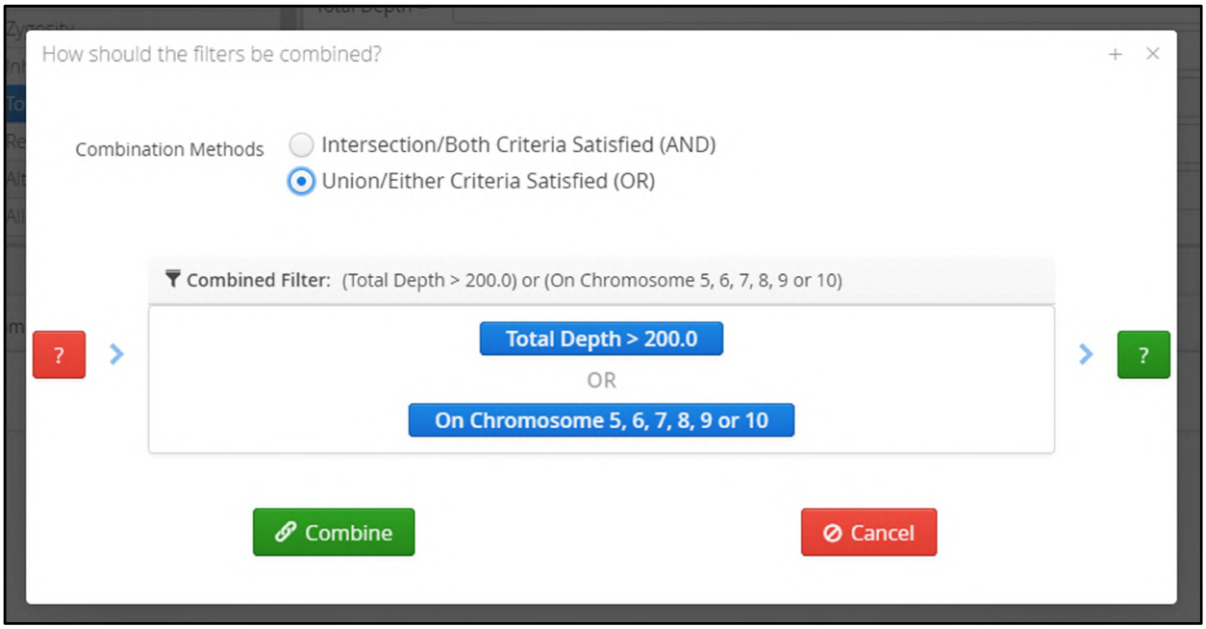


Figure: Combining filters with an OR statement

whether the options are to satisfy both filters with an AND statement.

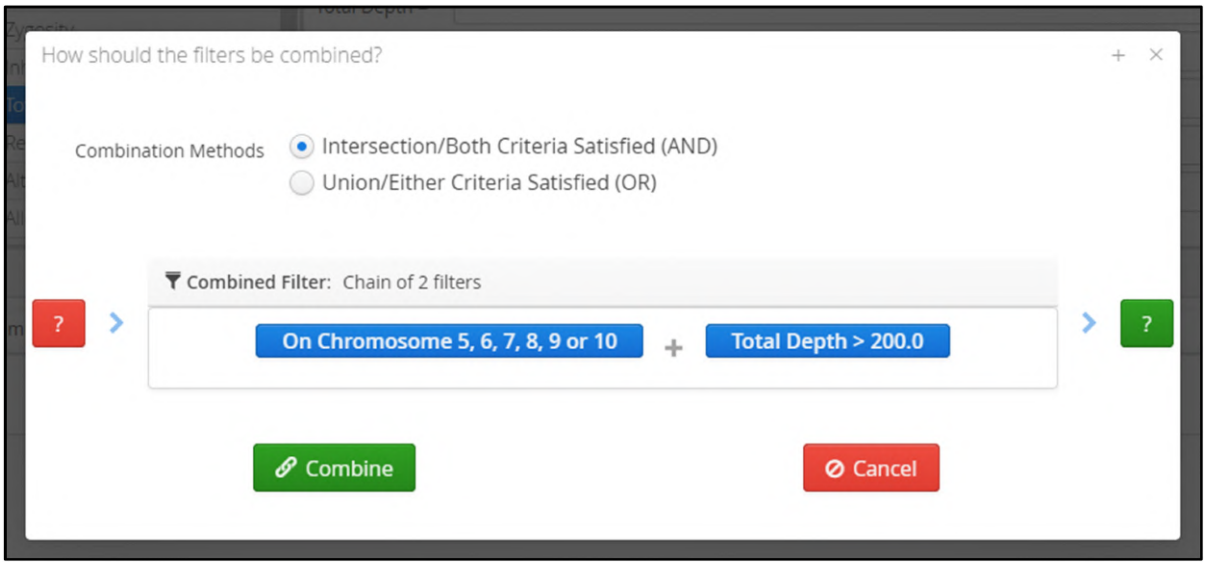


Figure: Combining filters with an AND statement

In this case AND was selected and the current filter shows the chain of 2 filters.

Figure: A filter showing 2 settings combined with an AND statement

Additional filters can be added, for example adding a filter on Allele Frequency.

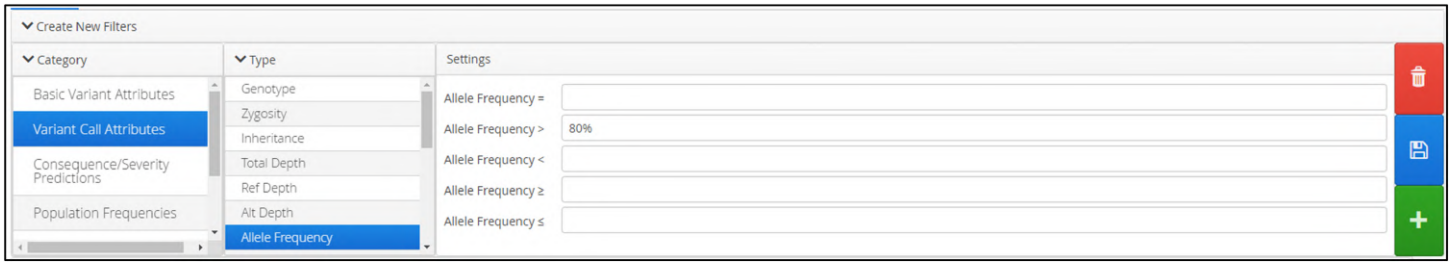


Figure: Adding an allele frequency filter

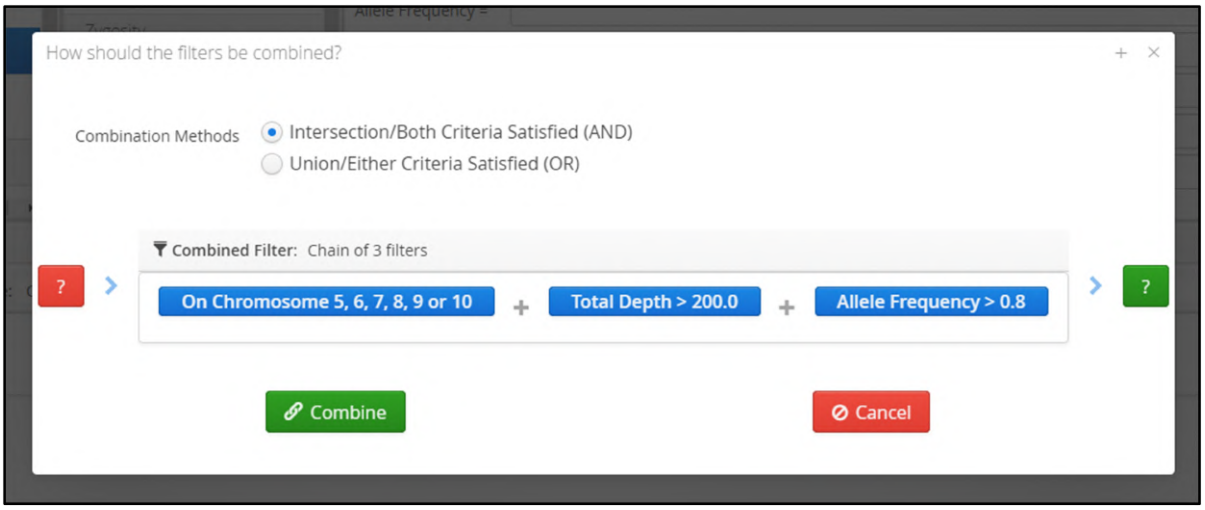


Figure: Adding the allele frequency filter with an AND statement

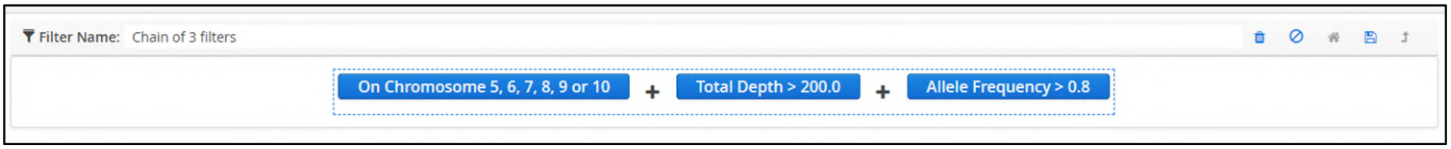


Figure: The updated filter is now displayed

As more filters are added they are grouped and this is reflected in the display by the drawing of dashed boxes around filter groups. Thus far there is a single filter contained within a single dashed box.

More complicated arrangements are possible whereby individual filters can selected and modified. In this chain of 3 filters the Allele Frequency > 80.0 has been selected and when selected the colour changes from blue to green.

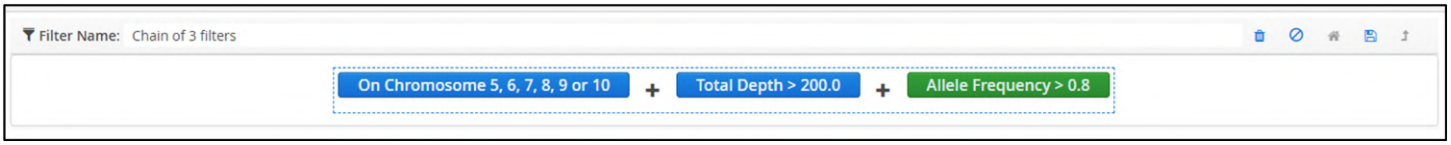


Figure: Selection of a part of an existing filter

As such an additional filter, in this case for Allele Frequency < 20.0, will be added to the selected filter

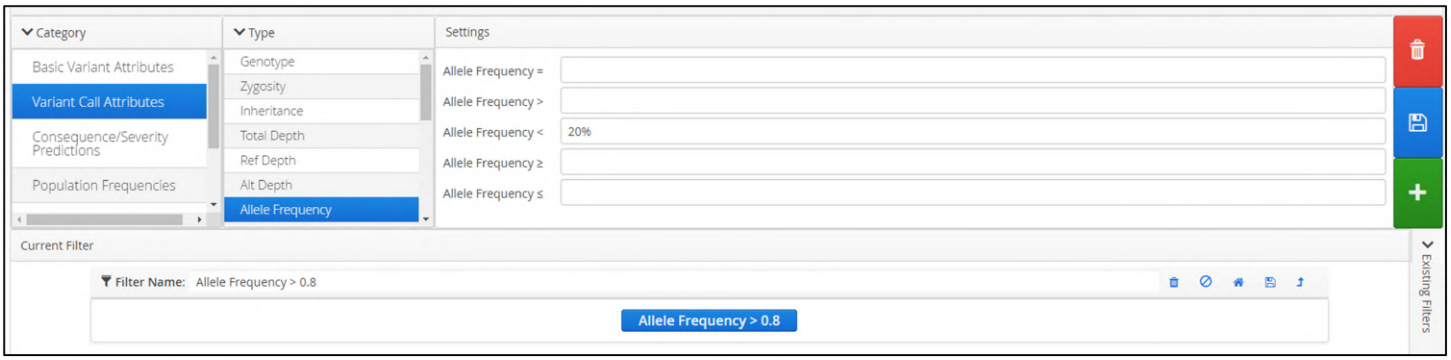


Figure: Modifying part of an existing filter

When presented with the choice of AND or OR the new filter is only shown with the previously selected filter.

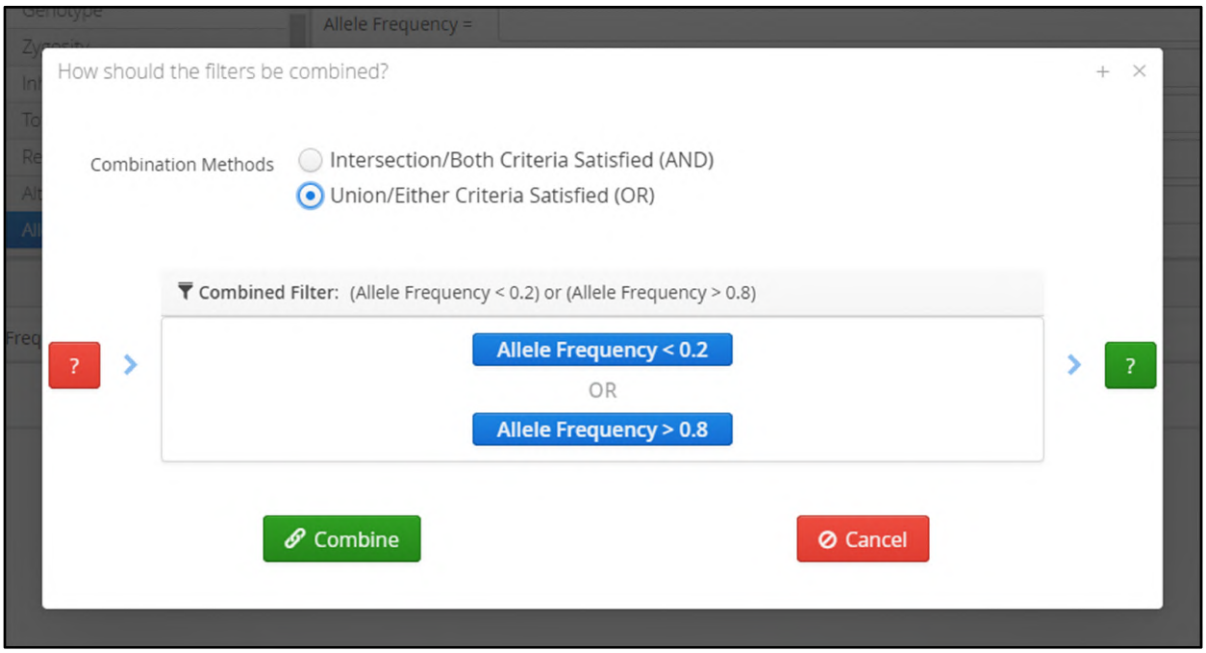


Figure: Updating part of an existing filter

The current filter will be updated to show the OR combination of the two filters.

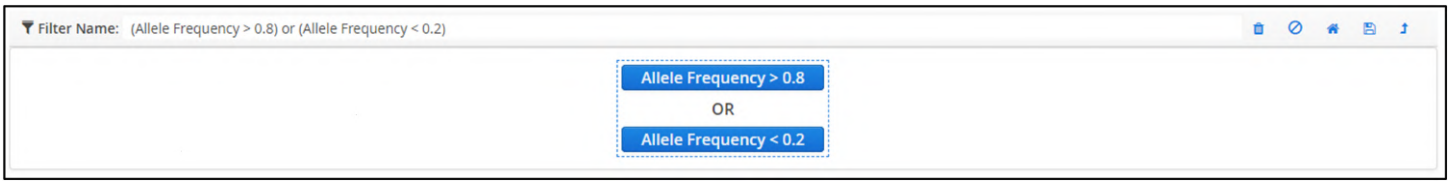


Figure: Addition of the new filter

Clicking on the home or house icon returns the current filter to the full view with the hierarchy of filters defined by dashed boxes.

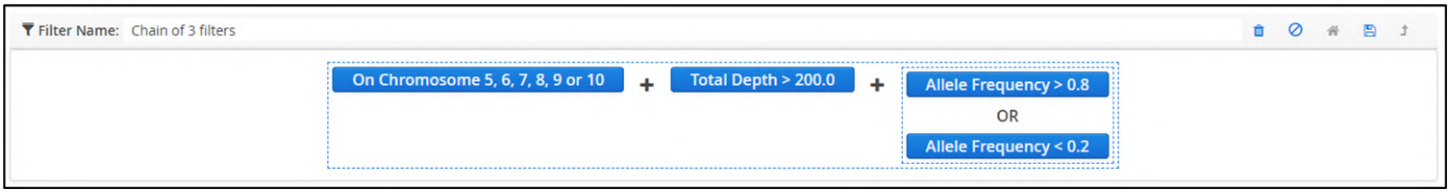


Figure: Updating of the display to show the new filter

As the filter builds the sections of dashed boxes increase.

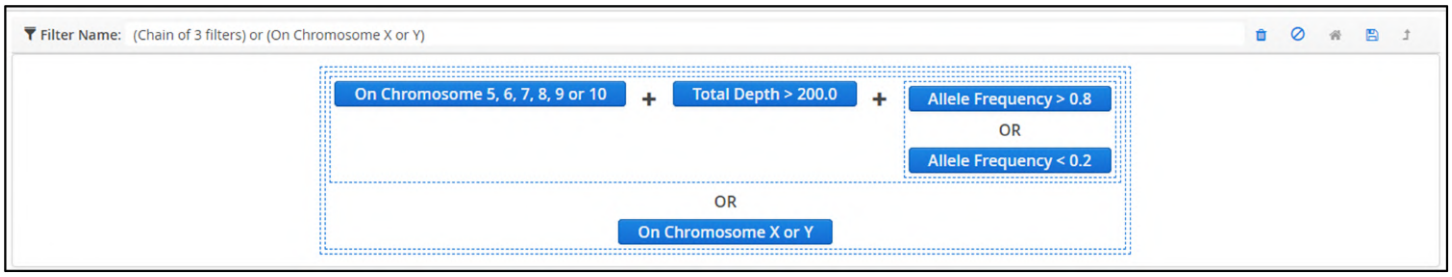


Figure: An example of a filter with multiple sections

The user to modify individual parts or groups of filters by using the mouse.

As the mouse moves over the filter the different parts will be highlighted in a different colour.

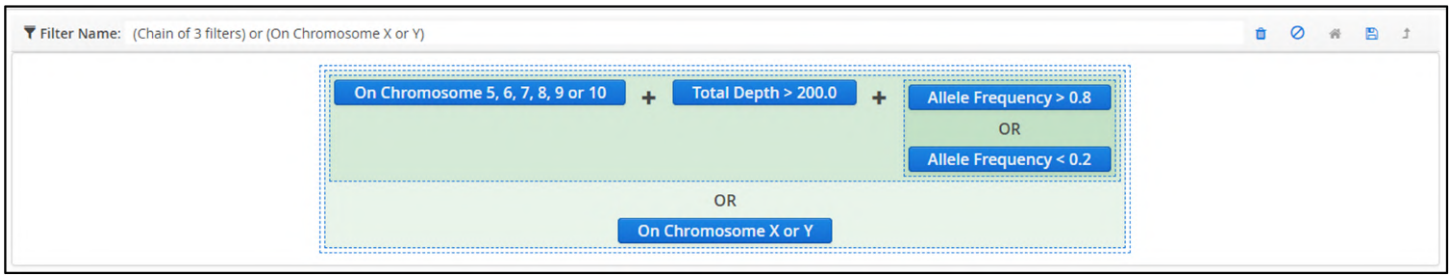


Figure: Highlighting different parts of a filter

If there is a modification to make to a part of a filter, then highlight the area to be updated.

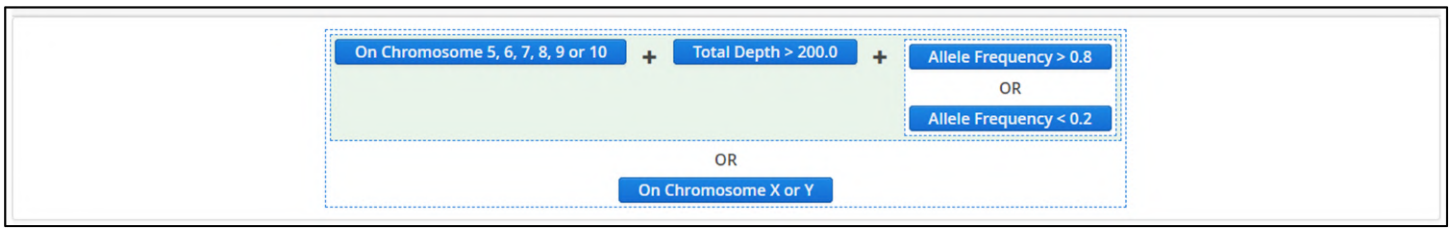


Figure: Selection of part of the filter

And clicking on it will show the sub-section chosen.

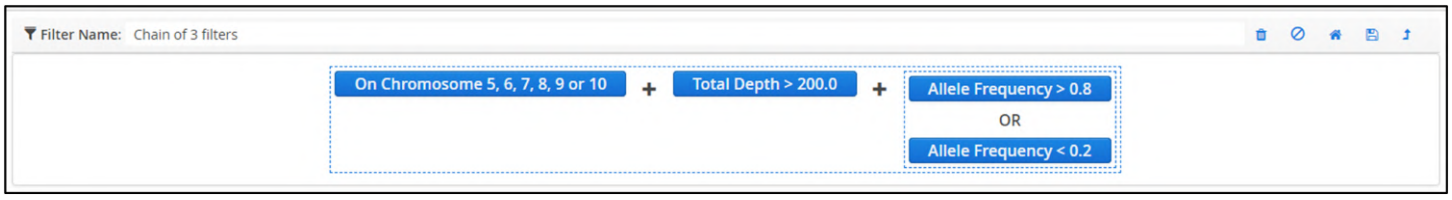


Figure: A section of the filter

This sub-section can be be modified

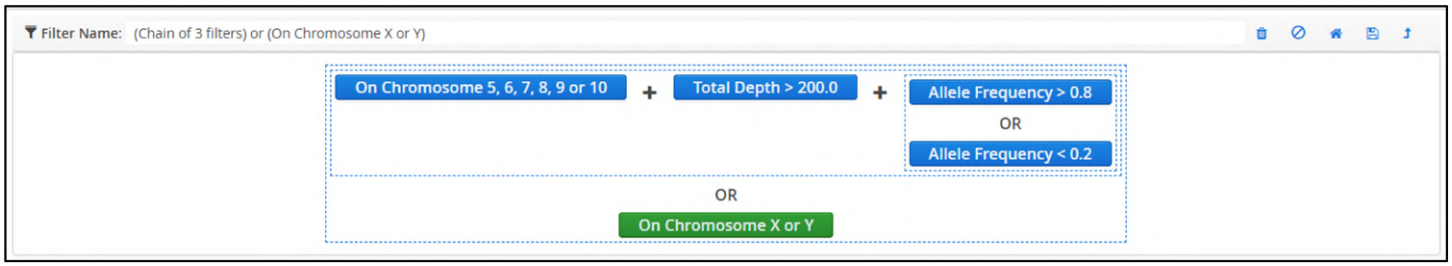


Figure: Modification of the selected part of a filter

Clicking on the home icon returns the view to the whole filter.

This can be continued until the required filter has been generated. The flexibility in Interpret allows a user to build highly complex queries.

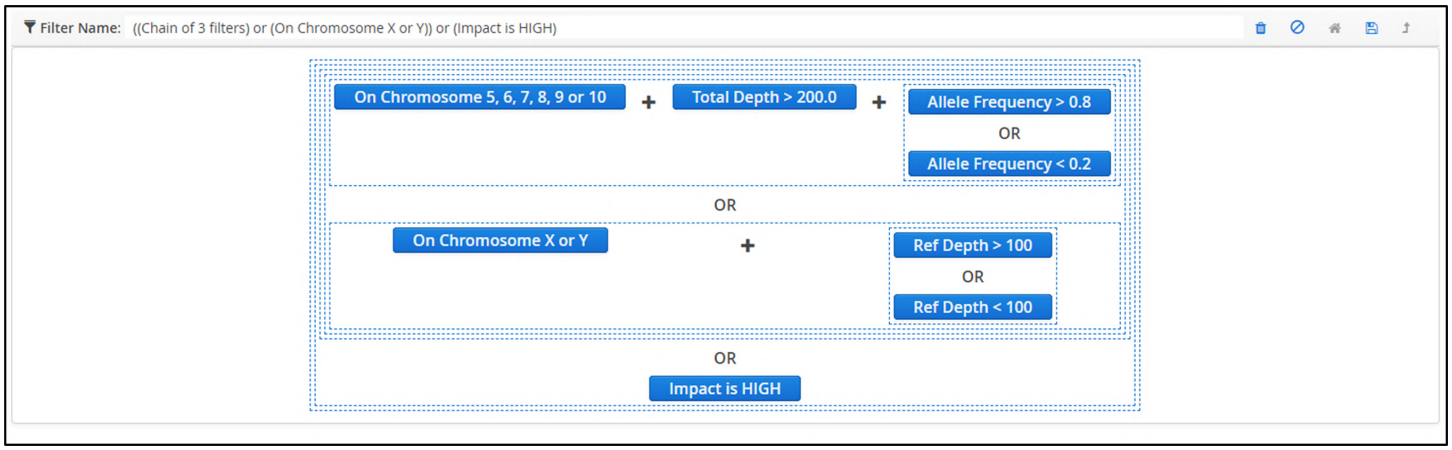


Figure: An example of a complex filter generated in Interpret

14.3.11 Preferred Transcripts

It is usual for genes to have multiple transcripts. As users sometimes have specific transcripts of interest, Interpret provides the means to define a preferred transcript for each gene.

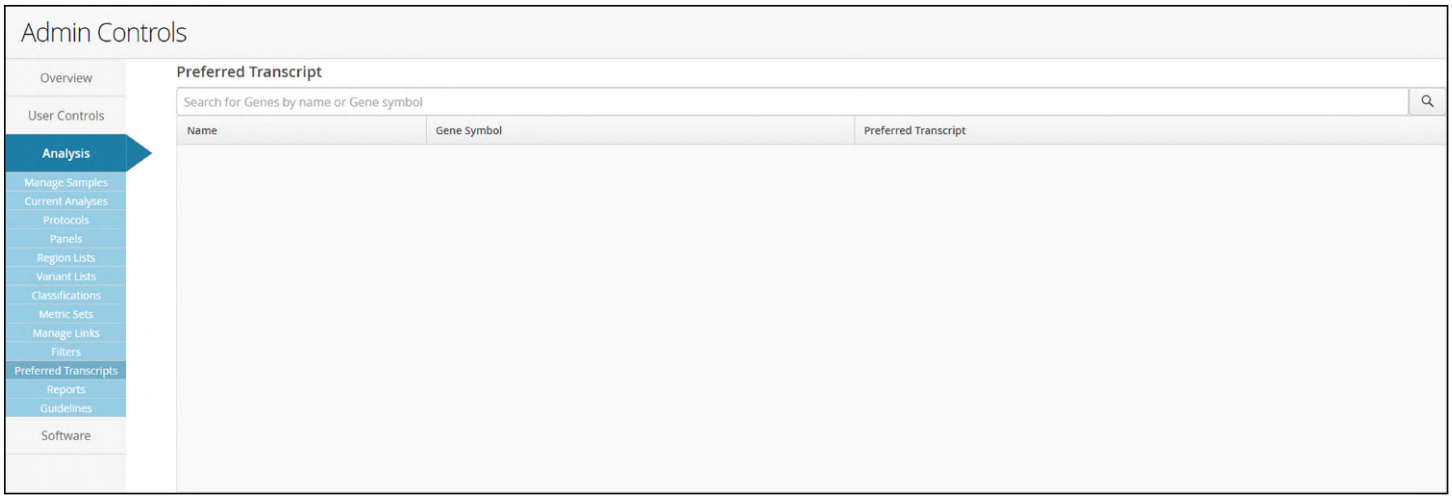


Figure: The Preferred Transcript start page

When no preferred transcript has been assigned the default behaviour is for the software to select the longest canonical transcript and report the annotation for that transcript alone.

To see a different annotation the user therefore needs to set their preferred transcript using either the admin controls prior to running an analysis, or from within the variants table post-analysis. The table will update with the different annotation once this is done.

To set a preferred transcript the user needs to enter a search term as in the figure below.

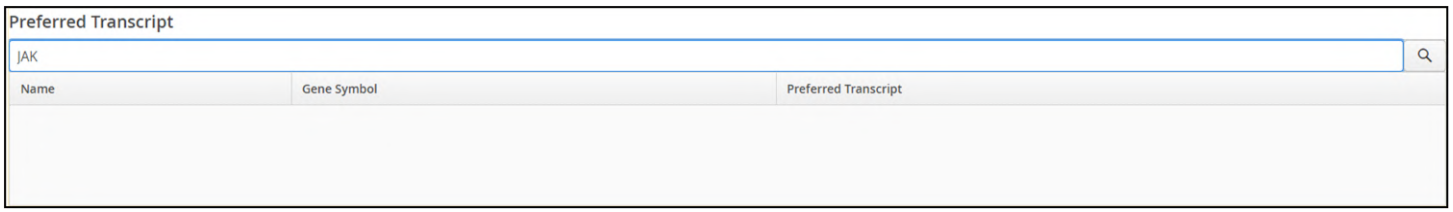


Figure: Adding a search term

The text entered in the field is matched across all the genes in the database. The search term must simply be present somewhere in the gene name. All matches will be displayed.

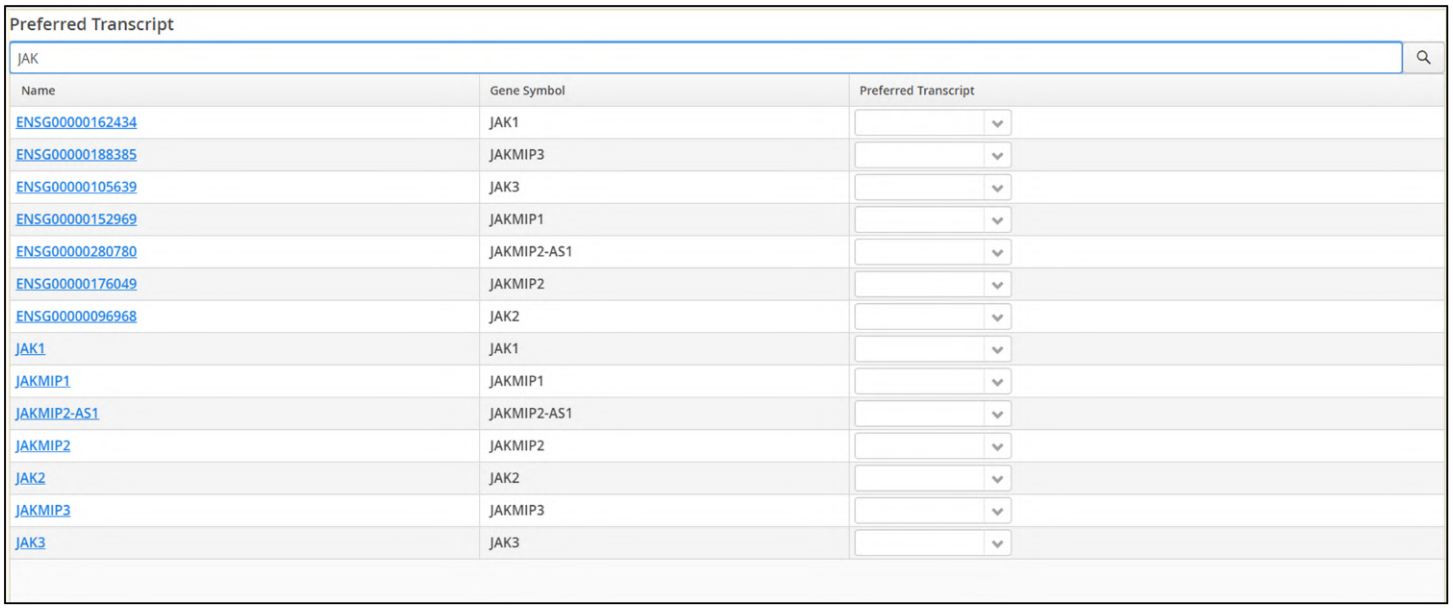


Figure: Search results

For each gene returned by the search term there will be a drop down menu of transcripts available.

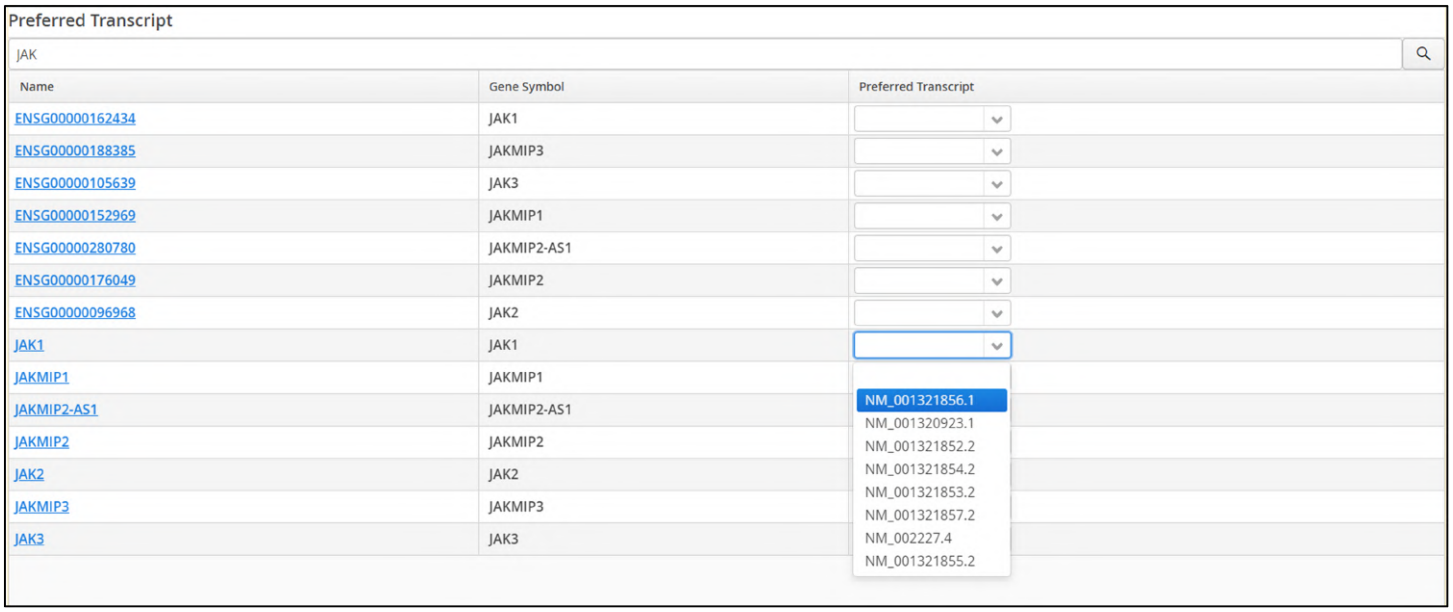


Figure: Selecting a transcript form the list available

When this is done the preferred transcripts page will update to reflect the selection.

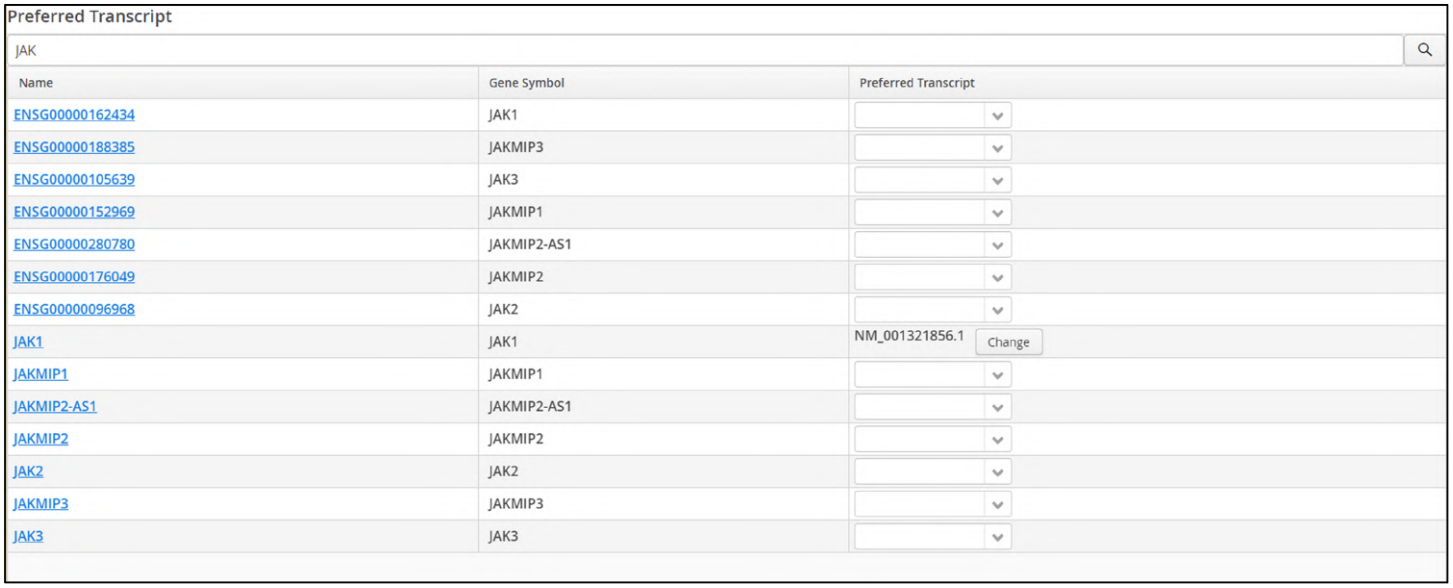


Figure: The selected transcript for a gene is displayed

14.3.12 Reports

The reporting section displays the reports and their templates currently loaded in the software; each type of report can have different templates associated with it,

OGT is able to assist with the creation of customer specific reports and templates.

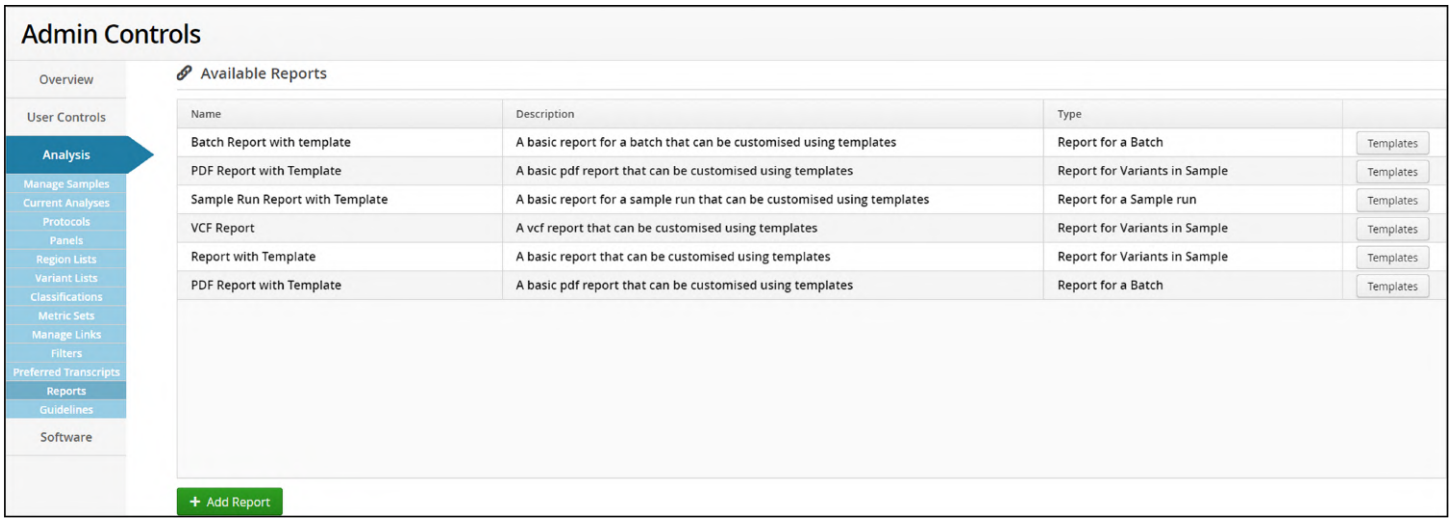


Figure: Initial view of the currently loaded reporting templates

New reports can be added by selecting the Add Report option and following the instructions in the popup window.

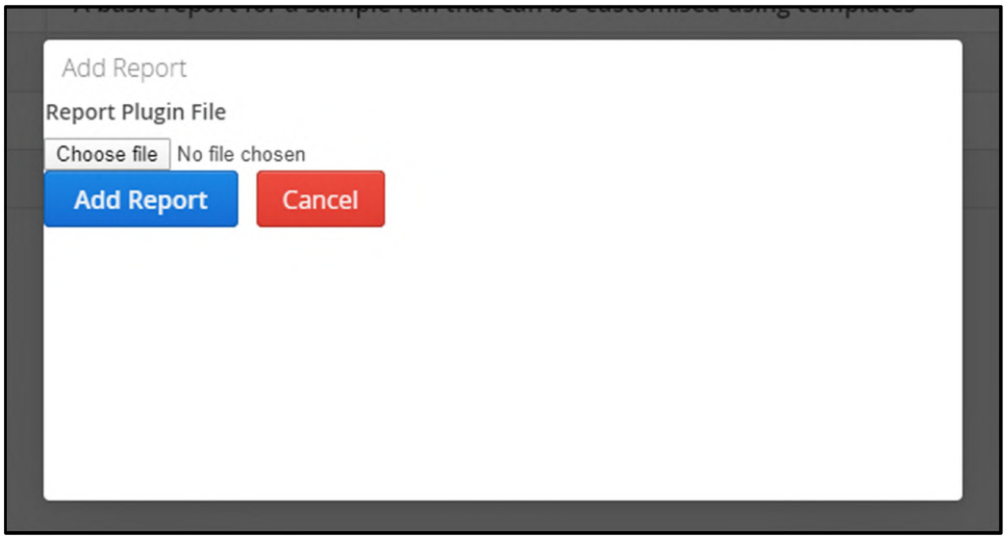


Figure: The Add Report popup window

Selecting the templates button for report shows the templates that are available to that particular report type

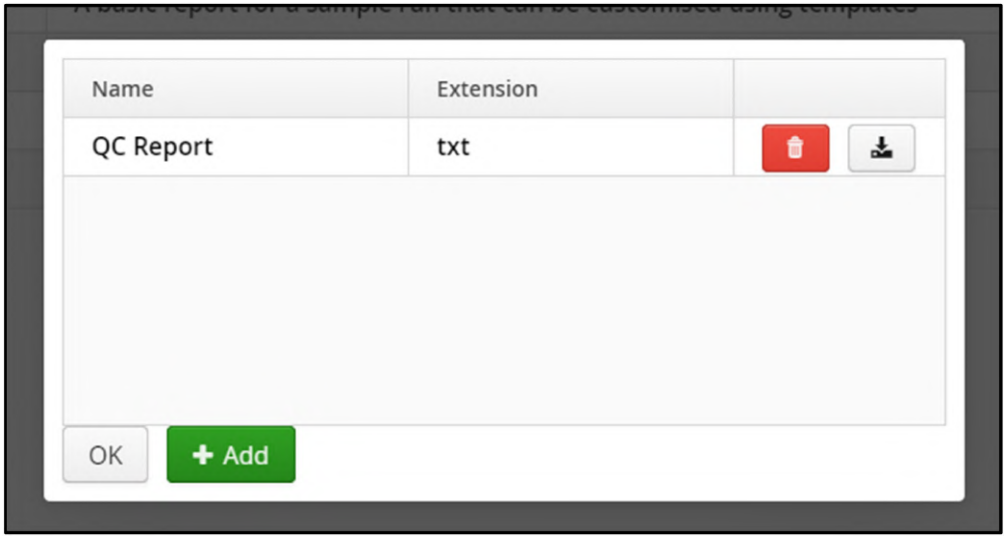


Figure: Templates associated with a report type

New templates can be added using the +Add button, selecting a template file and adding.

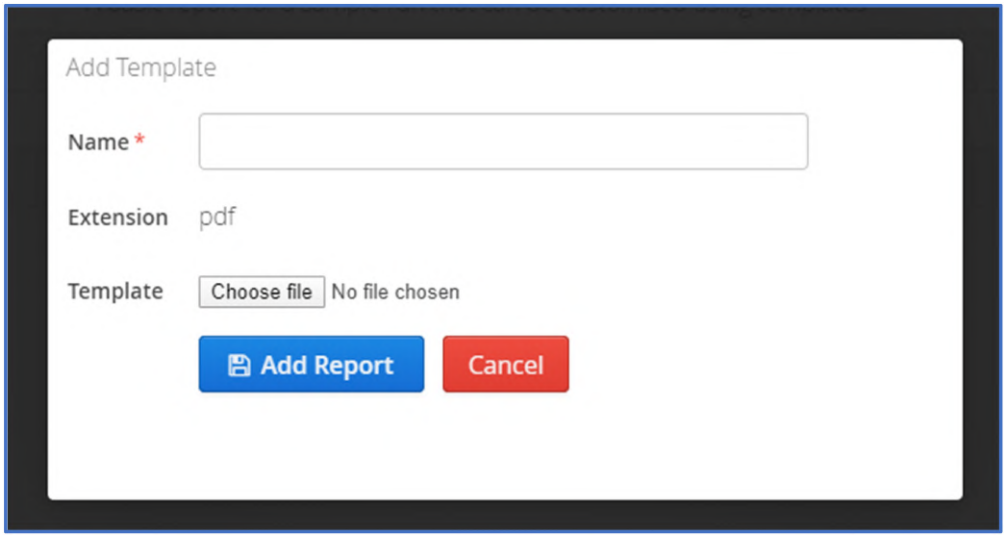


Figure: Add a template to a report

14.3.13 Guidelines

Guidelines provide a means to regulate classifications of variants by forcing the user to follow a question tree that will lead to a classification being generated. It is recommended to assess the template with reference to a user's own laboratory guidelines.

Using such a tree should remove any user variation with the result a classification will be generated consistently.

When first selected the guideline page is as follows:

Figure: The Guidelines start page

The guidelines page has two sections showing existing guidelines and the questions and transitions within them.

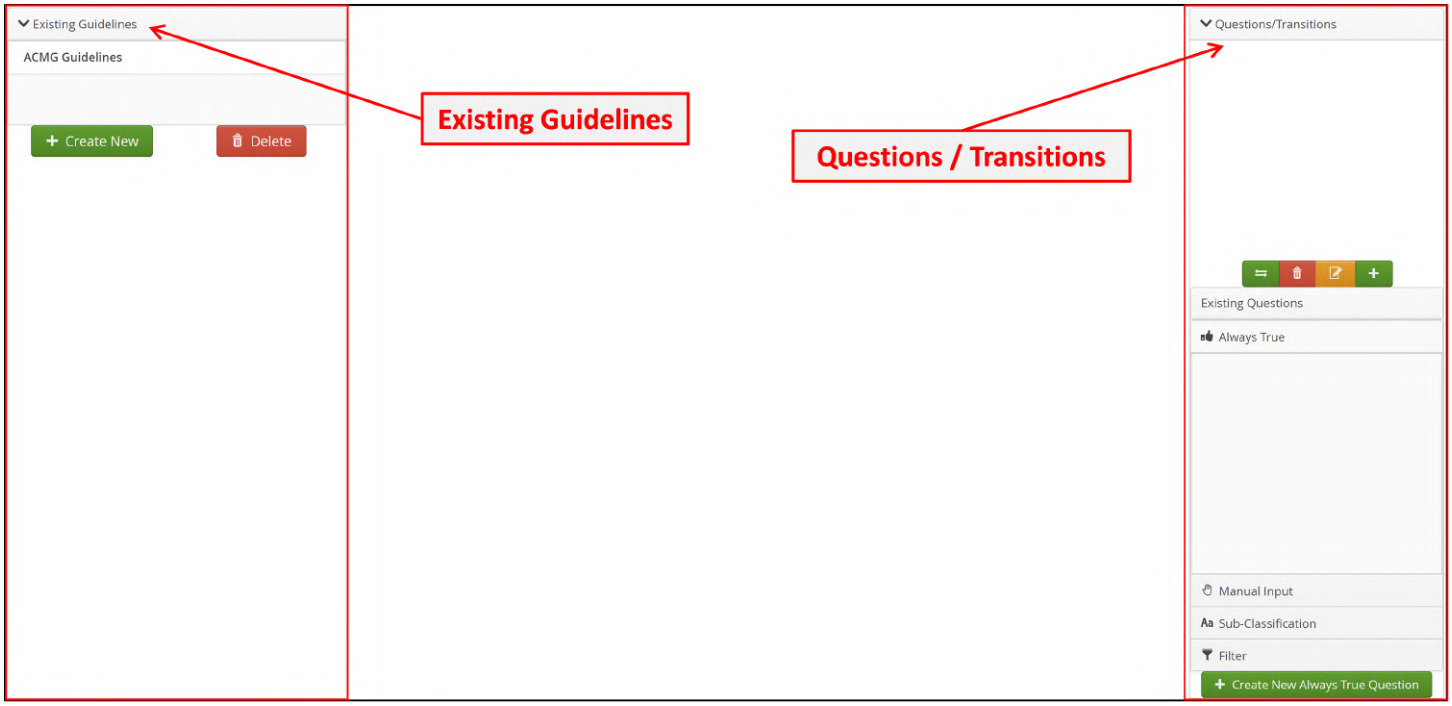


Figure: Existing guidelines and questions/transitions sections

Selecting ACMG guidelines will show a schematic of the questions and transitions. This is a complex decision tree as shown by the schematic below.

Selecting any of the questions, that are in the boxes, will show in the question in the box to the side as well as the rules associated with the question.

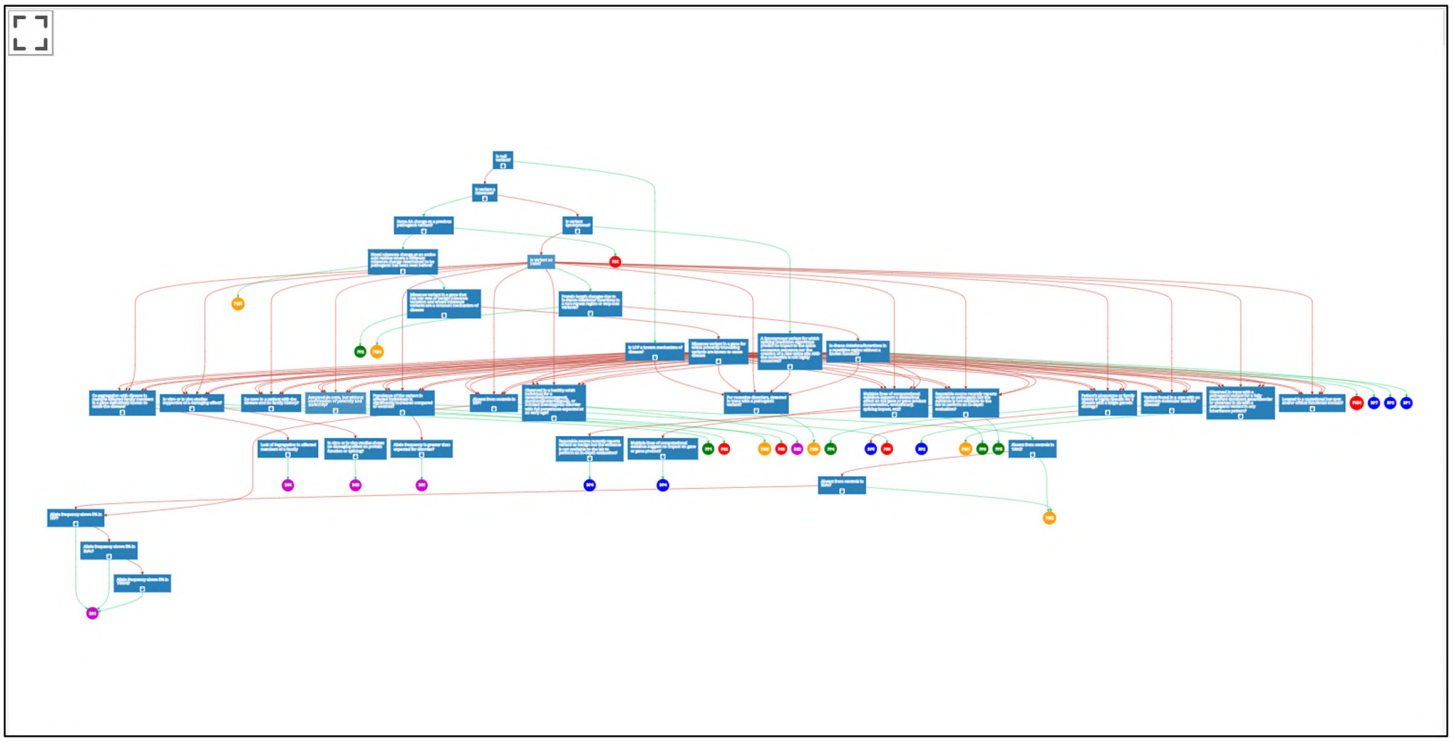


Figure: Schematic of the ACMG guidelines included with Interpret

Generating New Guidelines

Users can generate new guidelines to meet their own requirements. This is, currently, beyond the scope of this user manual. If this is required please contact OGT for further assistance.

14.4 Software

14.4.1 Advanced Settings

The advanced settings within Interpret are primarily for use by OGT support.

The default view shown in the figure; if any changes are required it is import to contact OGT in the first instance.

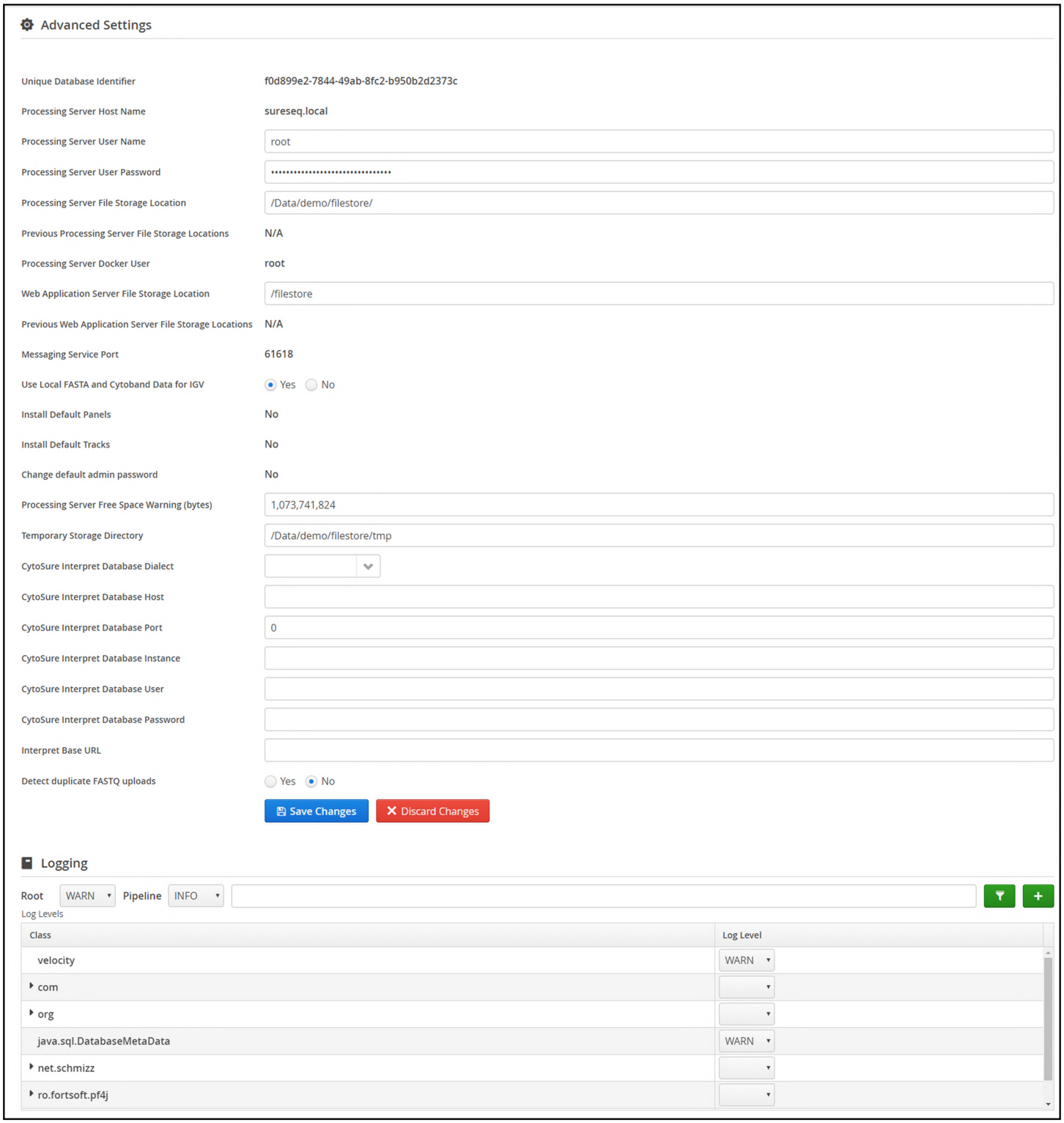


Figure: Default advanced settings

14.4.2 Annotation

Annotation used within Interpret are managed on this page.

The initial table shows all the annotations loaded in the software.

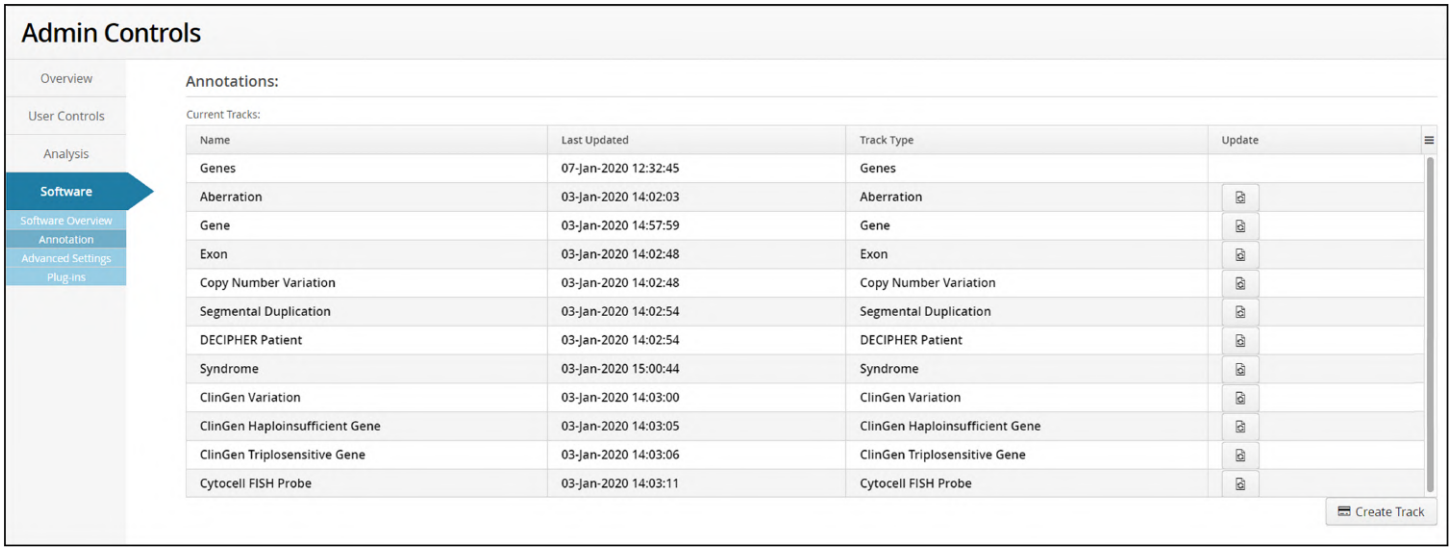


Figure: The Annotation start page

Selecting Create Track provides a popup menu which allows a track to be created either from a file or by importing from an existing CytoSure Interpret installation.

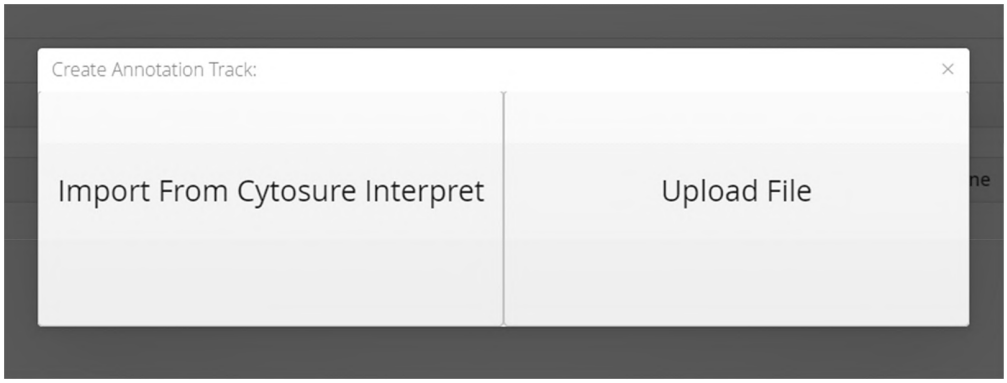


Figure: Create track options

14.4.3 Plug-ins

The use of plug-ins is a mechanism whereby additional functionality can be provided.

In situations where a user requests some specific functionality it may be easier to generate a plug-in rather than a new version of the software.

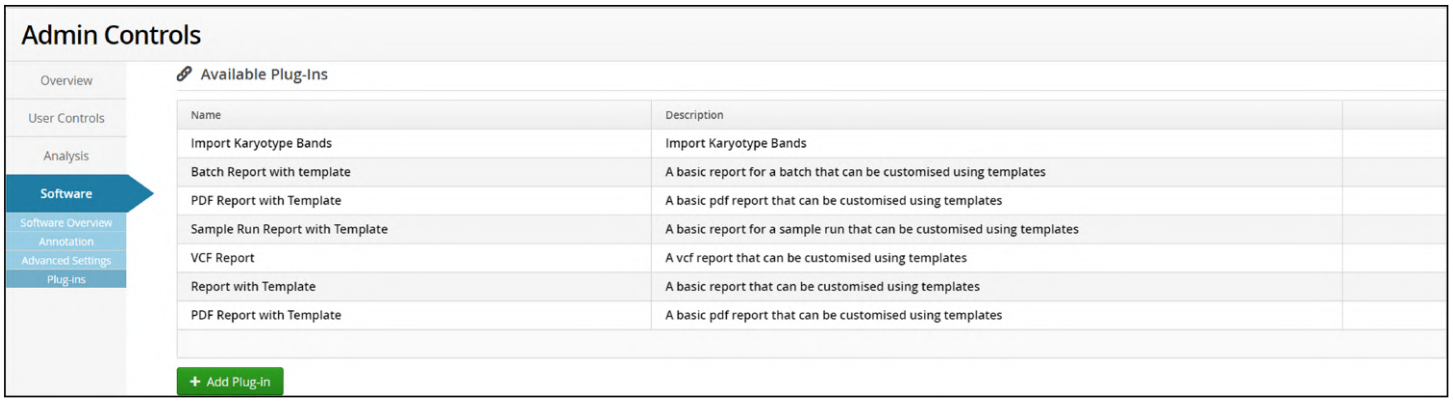


Figure: The Plug-ins start page

Plug-ins can only be generated by OGT and if a plug-in provided then it can be loaded by selecting Add Plug-in and then using the file browser to select the file.

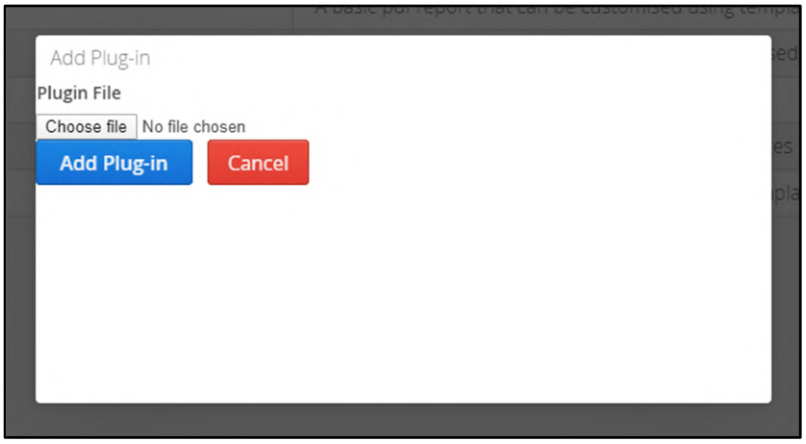


Figure: Load plug-in popup menu

14.4.4 Software Overview

The Software Overview provides version detail of the software installed.

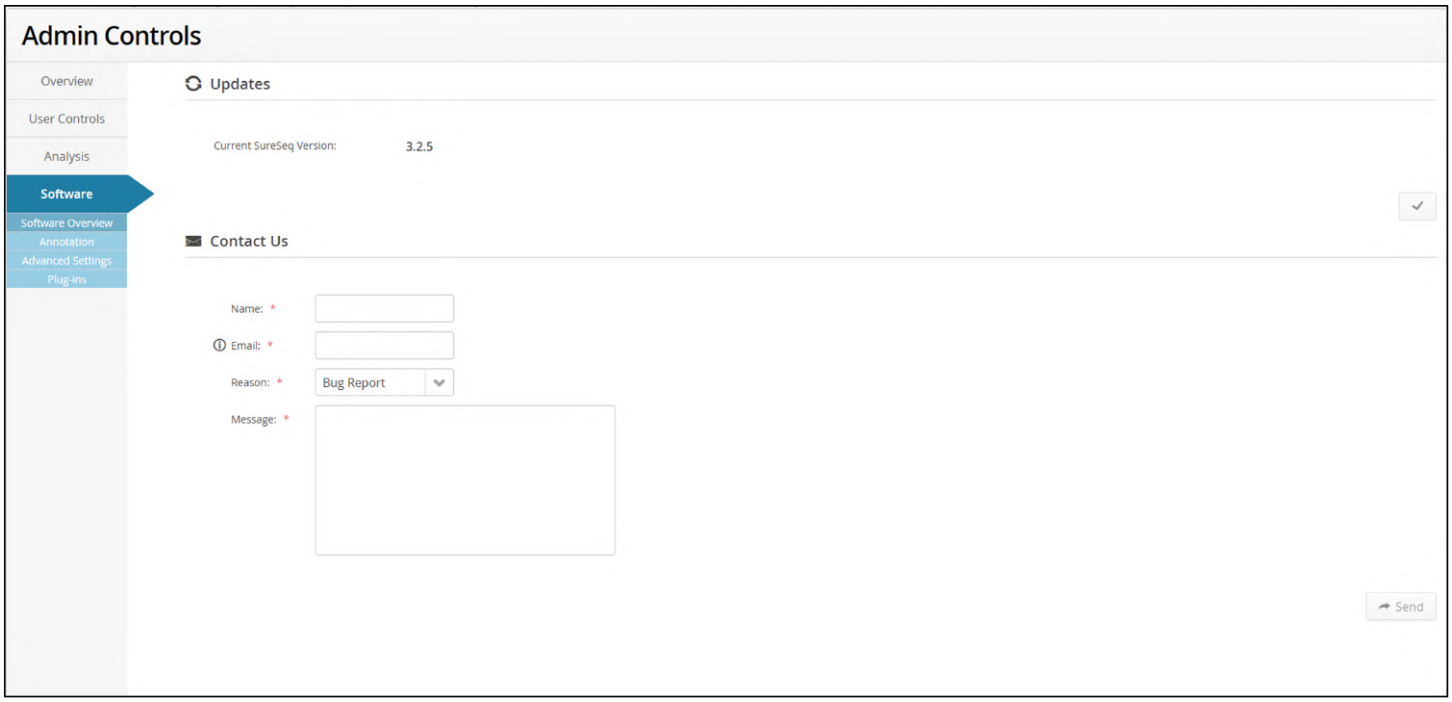


Figure: The Software Overview start page

Users can use the form to contact OGT.

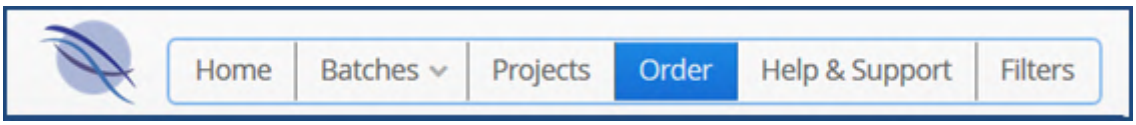
Figure: Options for contacting OGT

14.5 Reporting

Reporting has been discussed previously in the sections title Viewing Analysis Batches and Viewing Analysis Results where the details can be found.

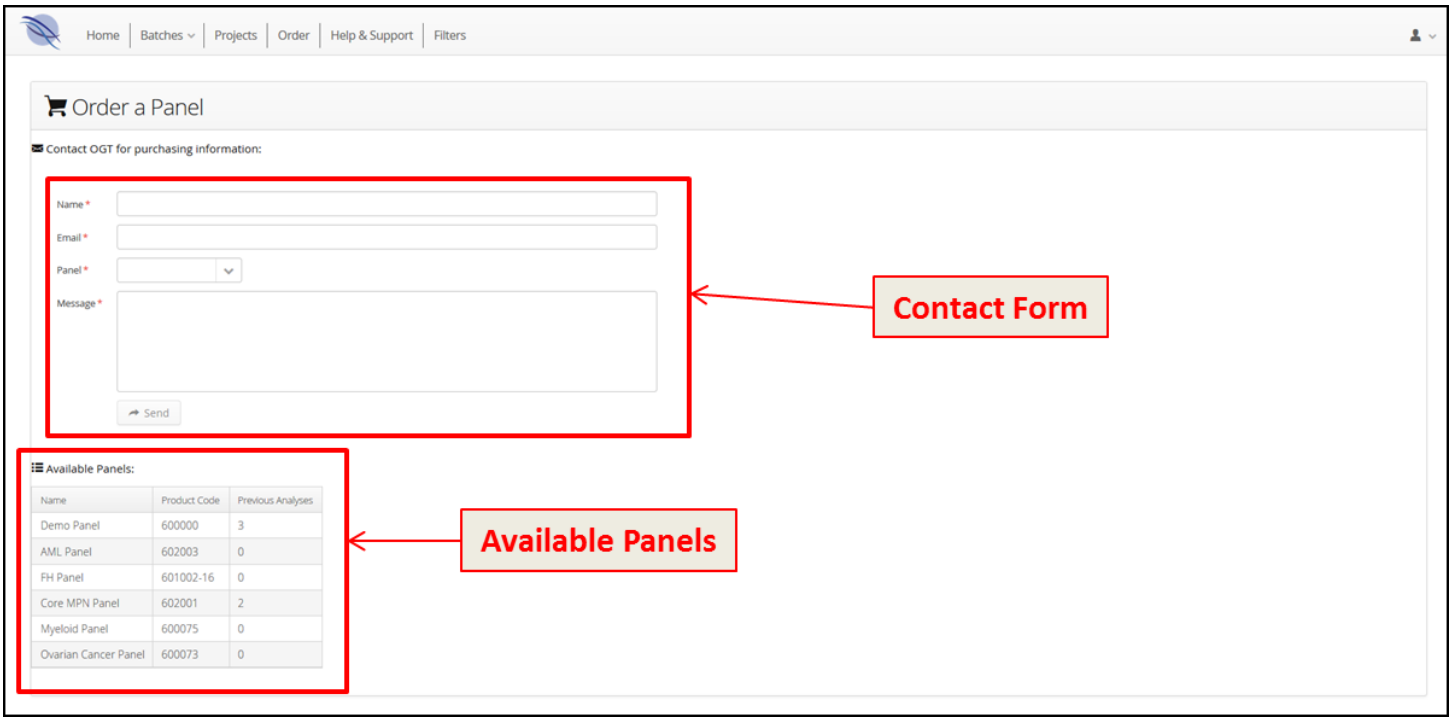
15 Ordering Panels

Users are able to request additional reagents through Interpret. This is accessed through the 'Order' menu button displayed in the software dashboard



The order panel page provides a form for the user to complete and submit to OGT.

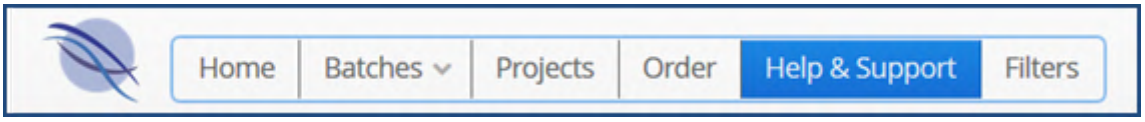
There is also information concerning the range of available panels and a record of which panels have previously been ordered.



16 Help and Support

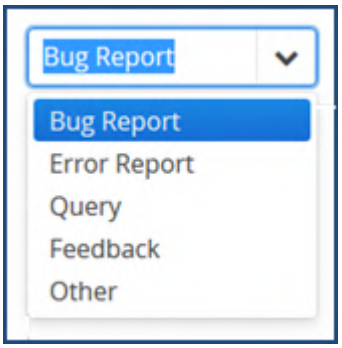
If you have requirement for support of any kind there is a form included in the software to allow requests for assistance to be processed.

This is accessed by clicking on the 'Help & Support' menu item on main dashboard.

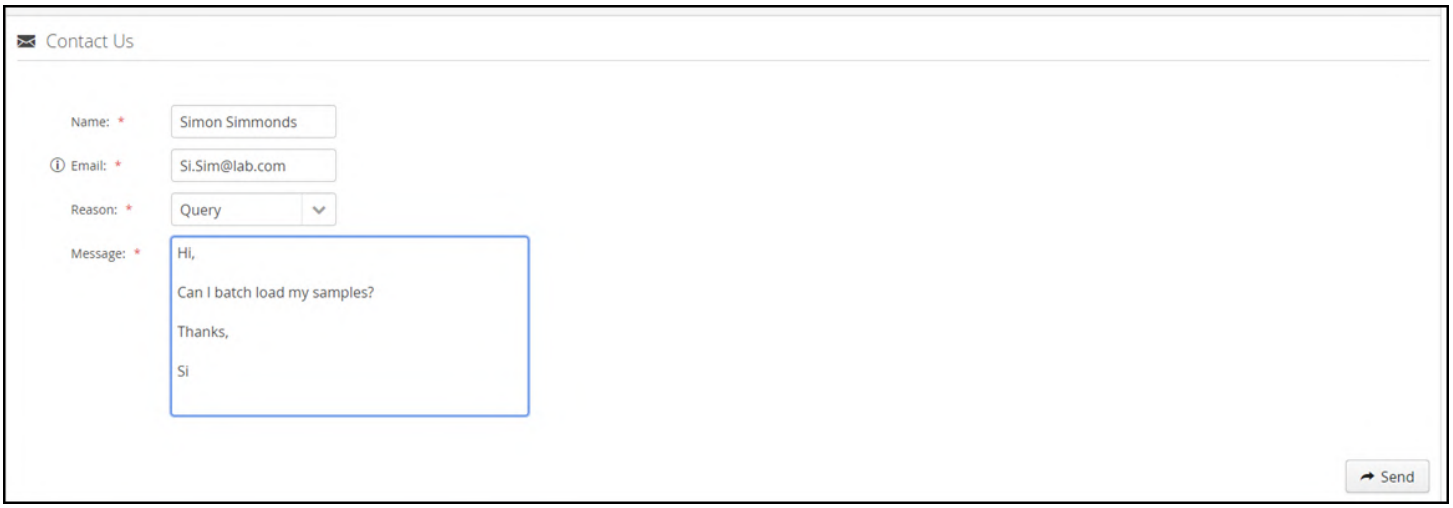


Users can then complete the form by supplying contact details and a description of the issue.

Additionally it is possible to specify the type of communication being made by the user by choosing from the drop down list of options.



Once completed, selecting send will deliver the message to OGT for processing.



17 Appendix

17.1 Attribute Definitions

Attribute	Variant Type	Description
#	SNV/Indel	Variant Database Identifier
# Alt Reads (-)	SNV/Indel	Number of alternative alleles on negative strand

Attribute	Variant Type	Description
# Alt Reads (+)	SNV/Indel	Number of alternative alleles on positive strand
# Markers	CNV/LOH	Number of bins (markers) used to identify CNVs
# Ref Reads (-)	SNV/Indel	Number of reference alleles on negative strand
# Ref Reads (+)	SNV/Indel	Number of reference alleles on positive strand
# Samples	SNV/Indel	Number of samples the variant is present in the database
% Samples	CNV/LOH	Percentage of samples the variant is present in the database
% Samples (Similar CNVs)	CNV/LOH	Number of samples in the database with an overlapping CNV
Allele Frequency	SNV/Indel	Allele Frequency
Alt	SNV/Indel	Alternate allele
Alt Depth	SNV/Indel	Number of reads supporting the alternative allele at the position
Alt Quality	SNV/Indel	Sum of alternative base qualities at the position
Alt Strand Bias	SNV/Indel	Sequencing bias in which one DNA strand is favoured over the other in the reads containing the alternative allele (Percentage)
Bands	CNV/LOH	Location of the variant on the chromosome
Batch Date	SNV/Indel	Date the batch was performed
Batch Name	SNV/Indel	Name of the batch containing the run
Canonical?	SNV/Indel	A flag indicating if the transcript is denoted as the canonical transcript for this gene
Chromosome	CNV/LOH	Chromosome of the CNV/LOH
Chromosome	SNV/Indel	Chromosome of the variant
Classification	CNV/LOH	User-assigned classification of the variant
ClinVar Significance	SNV/Indel	Clinical significance of variant according to ClinVar (e.g. benign, pathogenic, uncertain significance etc.)
Confidence	CNV/LOH	Confidence that the call is correct (e.g. High, Low) dependant on the Standard Error of Mean

Attribute	Variant Type	Description
Consequence Terms	SNV/Indel	Most severe outcome caused by the specific variant (e.g Frameshift variant, Stop gained, Synonymous variant etc.)
Context Length	SNV/Indel	Length of the genomic context overlapping the variant
Copy Number	CNV/LOH	Number of copies of the CNV event
Depth	CNV/LOH	Depth of coverage of the sample at position of the CNV
Description	Translocation	Donor and recipient gene symbol pair
Donor Breakpoint	Translocation	Position on the donor chromosome of the translocation
Donor Chromosome	Translocation	The chromosome number of the donor gene
Donor Gene	Translocation	Donor gene symbol where the translocation originated
Donor Locus Reads	Translocation	Read depth at the donor breakpoint position
Donor Orientation	Translocation	Orientation (strand) of the donor gene
Donor Reads Position	Translocation	Which side of the breakpoint the donor read lies (left or right)
End	CNV/LOH	Genomic position of end of CNV
End	SNV/Indel	Genomic position of end of variant
Estimated Tumour Content	CNV/LOH	Estimated tumour content (only used in cancer panels)
Exon ID	SNV/Indel	Unique ID for the exon
Exon Number	SNV/Indel	The number of the exon in the gene the variant is present
Frequency	CNV/LOH	Average Allele Frequency of common SNP's overlapping the CNV
Fusion Type	Translocation	The method of detection that highlighted the fusion, either: Over-expression of the gene (Expression), detection of a known fusion by sufficient supporting reads (Canonical), detection of an unknown fusion by sufficient supporting reads (Non-canonical)
Gene ID	SNV/Indel	Unique ID of the gene where the variant is
Gene Symbol	SNV/Indel	Gene symbol where the variant is
Genes	CNV/LOH	List of genes overlapping the CNV

Attribute	Variant Type	Description
Genome Build	CNV/LOH	Genome assembly version
Genomic Context	SNV/Indel	Genomic context the variant is overlapping (Low Complexity, Homopolymer, Simple Repeat)
Genotype	SNV/Indel	Genotype (Heterozygous/Homozygous)
gnomAD - African	SNV/Indel	The frequency variant appears on the Genome Aggregation Database from people of African decent
gnomAD - Ashkenazi Jewish	SNV/Indel	The frequency variant appears on the Genome Aggregation Database from people of Ashkenazi Jewish decent
gnomAD - East Asian	SNV/Indel	The frequency the variant appears on the Genome Aggregation Database from people of East Asian decent
gnomAD - European (Finnish)	SNV/Indel	The frequency the variant appears on the Genome Aggregation Database from people of Finnish decent
gnomAD - European (non-Finnish)	SNV/Indel	The frequency the variant appears on the Genome Aggregation Database from people of European (Non-Finnish) decent
gnomAD - Latino	SNV/Indel	The frequency variant appears on the Genome Aggregation Database from people of Latino decent
gnomAD - Other	SNV/Indel	The frequency the variant appears on the Genome Aggregation Database from people of another decent
gnomAD - South Asian	SNV/Indel	The frequency the variant appears on the Genome Aggregation Database from people of South Asian decent
gnomAD - Total	SNV/Indel	The frequency variant appears on the Genome Aggregation Database from all reference genomes
HGVSc	SNV/Indel	The HGVS coding sequence name
HGVSc (Gene Symbol)	SNV/Indel	The HGVS coding sequence name with the Transcript identifier replaced with its Gene Symbol
HGVSp	SNV/Indel	The HGVS protein sequence name
Homozygosity	SNV/Indel	Proportion of the genome covered by LOH regions larger than 5Mb
Impact	SNV/Indel	The Impact score according to Ensembl VEP of the genetic variation in the genetic sequence (e.g. LOW, MODERATE, HIGH etc.)

Attribute	Variant Type	Description
Inheritance	CNV/LOH	Estimated inheritance of the variant based on the presence of the variant in parental results, if available.
Inheritance	SNV/Indel	Estimated inheritance of the variant based on the presence of the variant in parental results, if available.
Inheritance	Translocation	Estimated inheritance of the variant based on the presence of the variant in parental results, if available.
ISCN	CNV/LOH	CNV/LOH variant encoded according to ISCN (International System for Human Cytogenomic Nomenclature)
ISCN	Translocation	Translocation variant encoded according to ISCN (International System for Human Cytogenomic Nomenclature)
Length	CNV/LOH	Length of CNV
Log Ratio	CNV/LOH	Mean log2 ratio of sample/reference of the CNV
Mean	CNV/LOH	Rescaled mean log2 of sample/reference of the CNV (only used in cancer panels)
Mean Standard Error	CNV/LOH	Standard Error of the Mean
Minor Allele	SNV/Indel	Base of the minor allele
Minor Allele Frequency	SNV/Indel	Rate at which the second most common allele occurs
Mosaicism	CNV/LOH	Estimate of the percentage of mosaicism observed in CNV region
Mosaicism Lower Bound	CNV/LOH	Estimate of the lower bound of mosaicism observed in the CNV region
Mosaicism Range	CNV/LOH	Estimate of the range of mosaicism observed in the sample
Mosaicism Upper Bound	CNV/LOH	Estimate of the upper bound of mosaicism observed in the CNV region
Most Severe Consequence	SNV/Indel	Most severe outcome caused by the specific variant (e.g Frameshift variant, Stop gained, Synonymous variant etc.)
Normalised Expression	Translocation	The expression of the baited gene relative to the housekeeping genes and normalised by total read count
Overlap	CNV/LOH	Genomic context of the CNV

Attribute	Variant Type	Description
P Value	Translocation	Probability of observing the translocation
Panel	SNV/Indel	Panel used for the analysis
PolyPhen Prediction	SNV/Indel	The prediction of how damaging a variant will be, based off the PolyPhen Score
PolyPhen Score	SNV/Indel	The probability that a substitution is damaging (e.g. 0.25 benign, 0.5 possibly damaging, 0.95 probably damaging)
Proportion	Translocation	Proportion of split reads over total reads at the donor breakpoint
Protein ID	SNV/Indel	Unique ID for the protein
Protocol	SNV/Indel	OGT Interpret software protocol used to analyse the run
Quality	CNV/LOH	(Not implemented)
Quality Score	SNV/Indel	Phred Quality score of the variant
Ratio	SNV/Indel	Ratio of depth observed in duplicated PTD exons compared to the exons in the rest of the gene
Read 1	SNV/Indel	File name of the FASTQ from R1 reads
Read 1 Size	SNV/Indel	Size of the FASTQ file from R1 reads
Read 2	SNV/Indel	File name of the FASTQ from R2 reads
Read 2 Size	SNV/Indel	Size of the FASTQ file from R2 reads
Reads Placed Left	SNV/Indel	Number of reads with supporting evidence to the left of the variant
Reads Placed Right	SNV/Indel	Number of reads with supporting evidence to the right of the variant
Recipient Breakpoint	Translocation	Position on the recipient chromosome of the translocation
Recipient Chromosome	Translocation	The chromosome of the recipient gene
Recipient Gene	Translocation	Recipient gene symbol where the translocation ended up
Recipient Locus Reads	Translocation	Read depth at the recipient breakpoint position
Recipient Orientation	Translocation	Orientation (strand) of the recipient gene

Attribute	Variant Type	Description
Recipient Reads Position	Translocation	Which side of the variant the donor read lies (left or right)
Ref	SNV/Indel	Reference nucleotide base
Ref Depth	SNV/Indel	Number of reads supporting the alternative allele at the position
Ref Quality	SNV/Indel	Sum of alternative reference qualities at the position
Ref Strand Bias	SNV/Indel	Sequencing bias in which one DNA strand is favoured over the other in the reads containing the reference allele (Percentage)
rsID	SNV/Indel	SNP id from NCBI dbSNP
Sample	CNV/LOH	ID of the sample containing the CNV
Sample	SNV/Indel	ID of the sample containing this variant
Sample	Translocation	ID of the sample containing this variant
Sample ID	SNV/Indel	ID of the sample containing this variant
Score	CNV/LOH	LOH score (Higher scores >30 indicate a higher confidence in the call)
Sex	SNV/Indel	Inferred chromosomal sex of the sample (Male, Female, Unknown)
SIFT Prediction	SNV/Indel	Prediction of how detrimental a variant will be to protein function (The opposite of Polyphen in terms of numbering)
SIFT Score	SNV/Indel	A score that predicts whether a variant will affect protein function (0 = deleterious , 1 = tolerated)
Source	CNV/LOH	Tool used for CNV identification
Start	CNV/LOH	Genomic position of start of CNV
Start	SNV/Indel	Genomic start position of variant
Supporting Reads	Translocation	The sum of split and discordant reads in support of the fusion call
Total Depth	SNV/Indel	Depth of coverage at the position
Transcript ID	SNV/Indel	Unique ID of the specific selected transcript
Transcript Resolution Method	SNV/Indel	Method used to determine which transcript to use

Attribute	Variant Type	Description
Type	CNV/LOH	Variant type (e.g. CNV, LOH)
Type	SNV/Indel	Variant type (e.g. SNV, ITD, PTD, etc.)
Type	Translocation	Variant type (e.g. Translocation)
User	SNV/Indel	Login name of user which ran the batch
VEP Version	SNV/Indel	Version of Ensembl Variant Effect Predictor
Zygosity	SNV/Indel	The degree at which both copies of the chromosome have the same genetic sequence (e.g. Homozygous or Heterozygous)

Legal Information

Interpret Software

Interpret software was developed by Oxford Gene Technology. The purchaser has the non-transferable right to use the product for RESEARCH USE ONLY AND NOT FOR DIAGNOSTICS PROCEDURES. It is not intended for use, and should not be used, for the diagnosis, prevention, monitoring, treatment or alleviation of any disease or condition, or for the investigation of any physiological process, in any identifiable human, or for any other medical purpose.

Customer's obligations

The Customer acknowledges that Oxford Gene Technology (Operations) Limited (or its group companies) owns all intellectual property rights in the design of the Product, including the choice and configuration of the oligonucleotide sequences used in the Product. The Product may only be reproduced or manufactured by Oxford Gene Technology (Operations) Limited or with its permission.

Contact us

Oxford Gene Technology,
Begbroke Hill,
Woodstock Road,
Begbroke,
Oxfordshire
OX5 1PF, UK

T: +44(0)1865 856800; (US: 914-467-5285)
E: contact@ogt.com
W: www.ogt.com

Technical support: E: support@ogt.com
T: +44(0)1865 856800