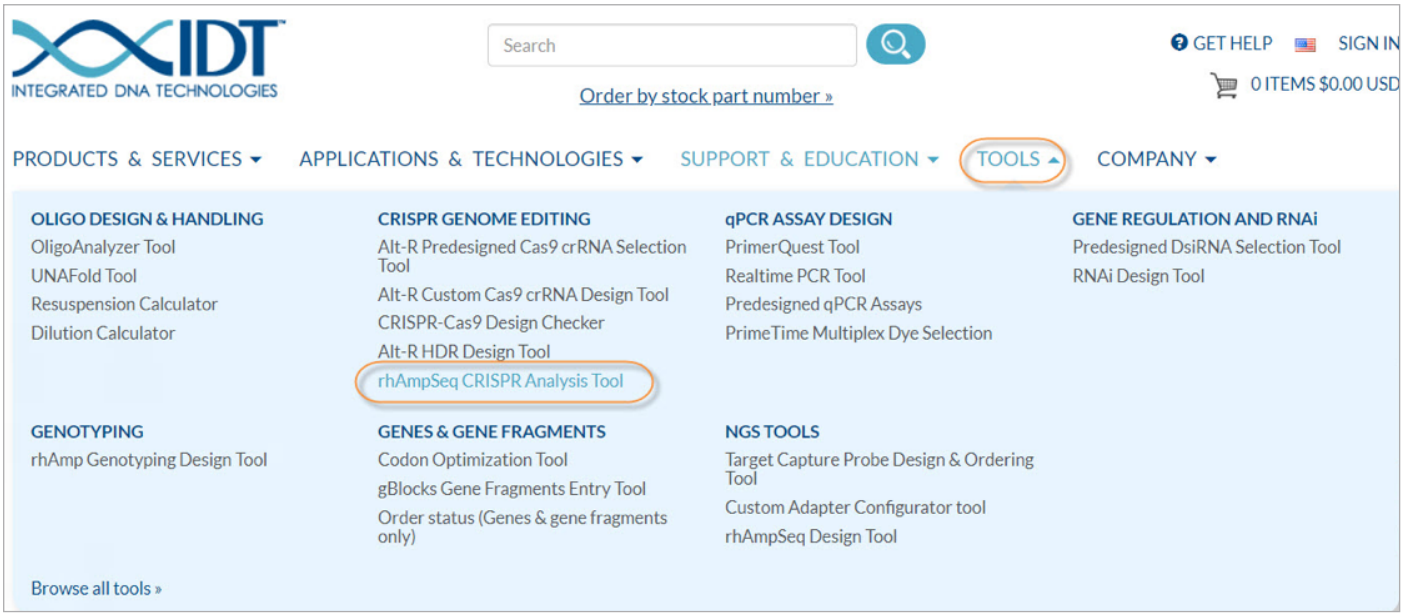
Using the rhAmpSeq™ CRISPR Analysis Tool

For Research Use Only | Version 1

Table of contents

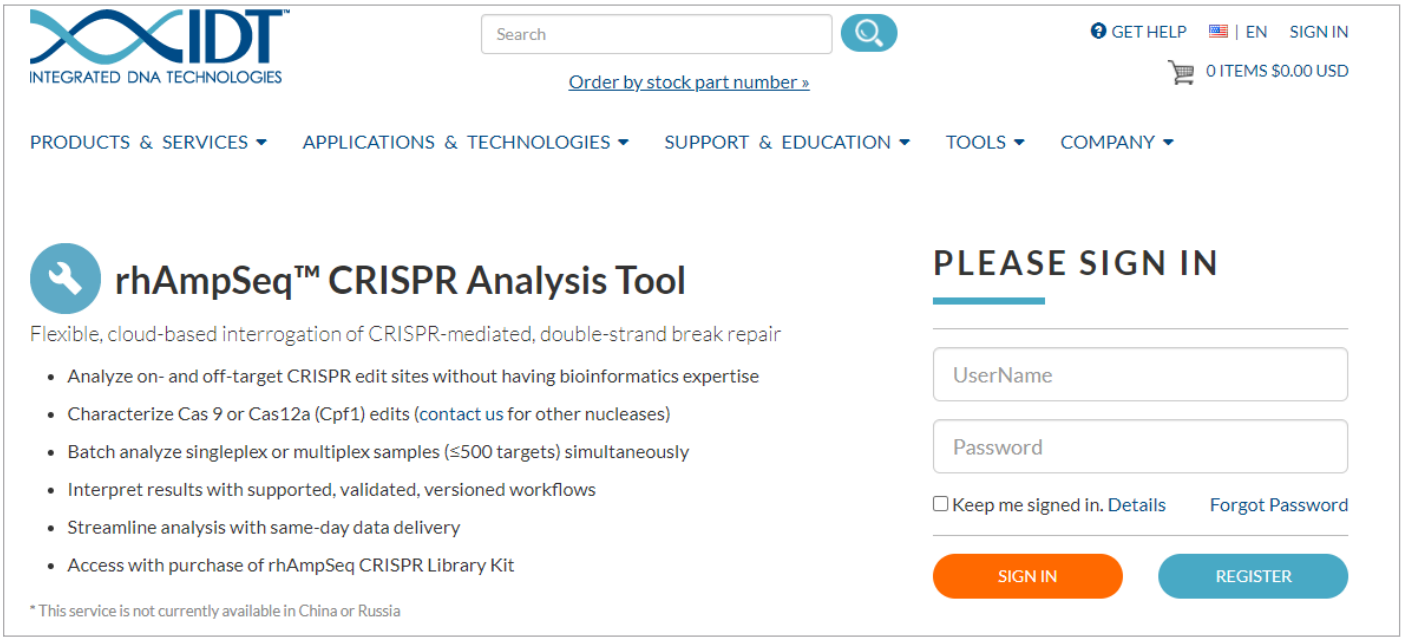
Access the rhAmpSeq CRISPR Analysis Tool

From the IDT website, go to Tools, CRISPR Genome Editing, then select the **rhAmpSeq CRISPR Analysis Tool**.

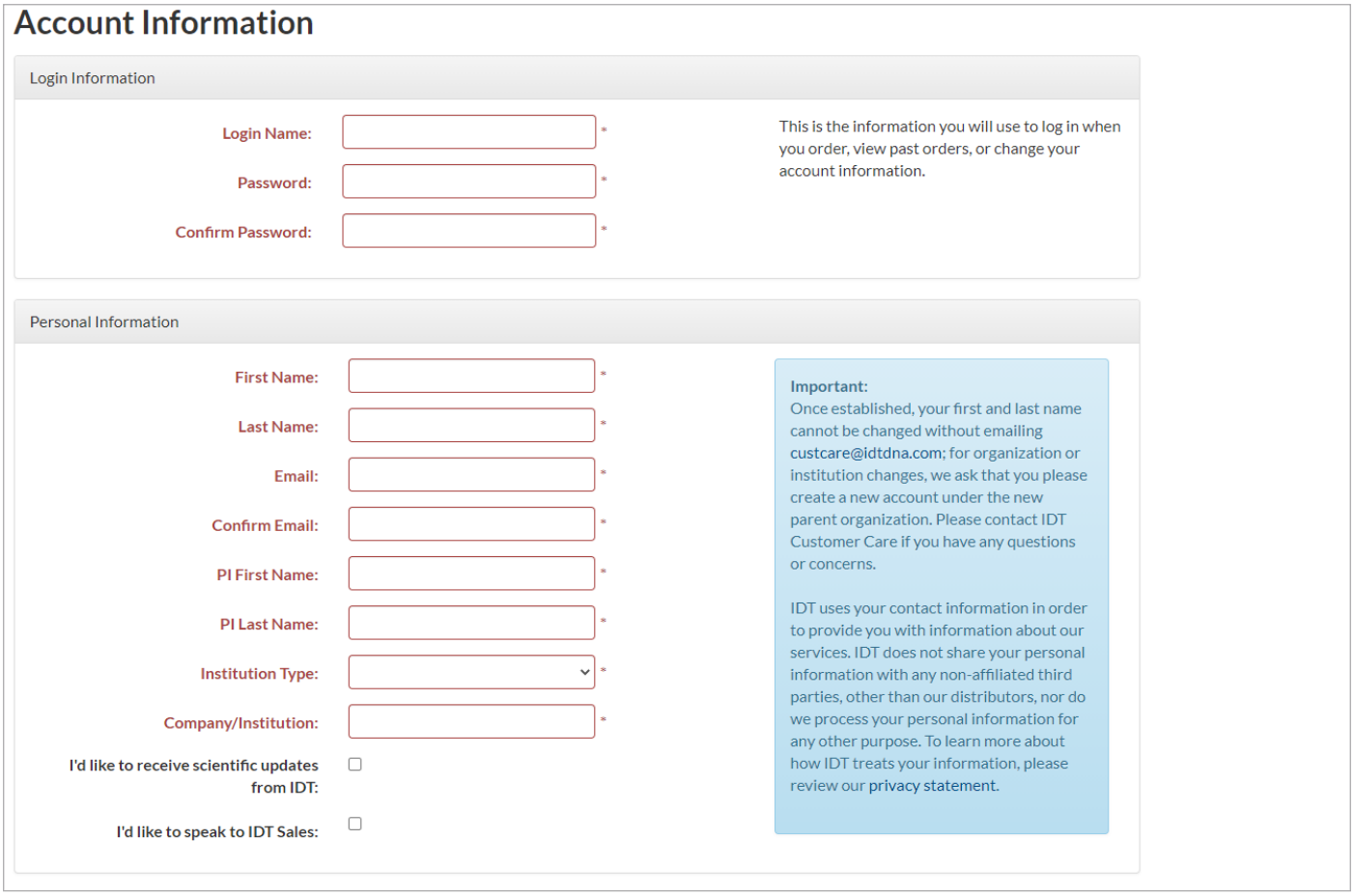


Create an account

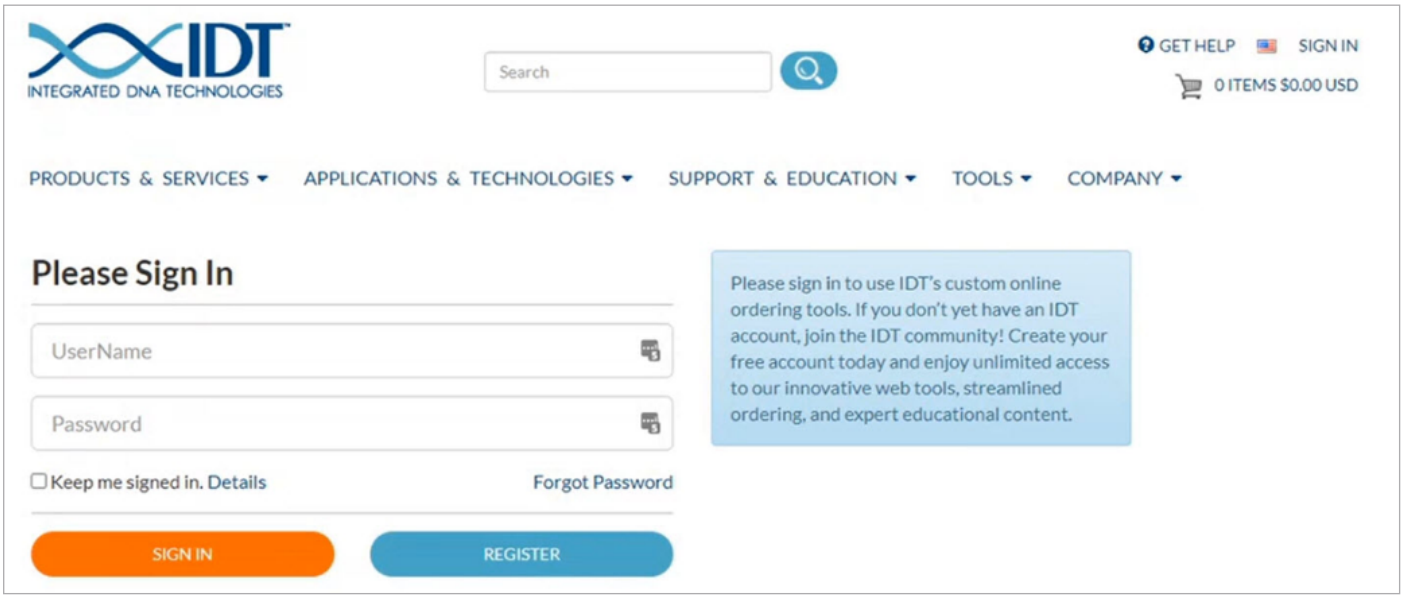
An IDT account provides access to the **Analysis Lab** page and the ability to use the rhAmpSeq CRISPR Analysis Tool.



Note: If you do not already have an IDT account, click **Register** to create your own account.

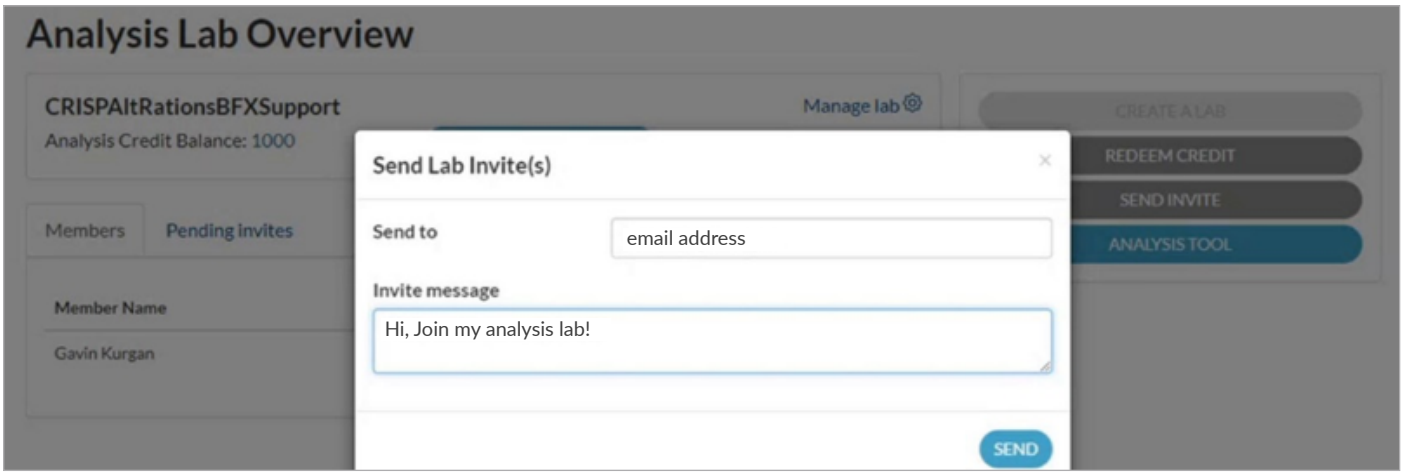


Once established, you will use this account login to sign in and use the tool.



After you log in, the **Analysis Lab** landing page opens. If you are not already part of an analysis lab, click the **Create a Lab** button and give a unique name to your lab.

After creating a lab, you can invite collaborators, colleagues, and students to join your analysis lab by clicking **Send Invite**. Enter a list of those you want to collaborate with, and a custom message to accompany your invitation.

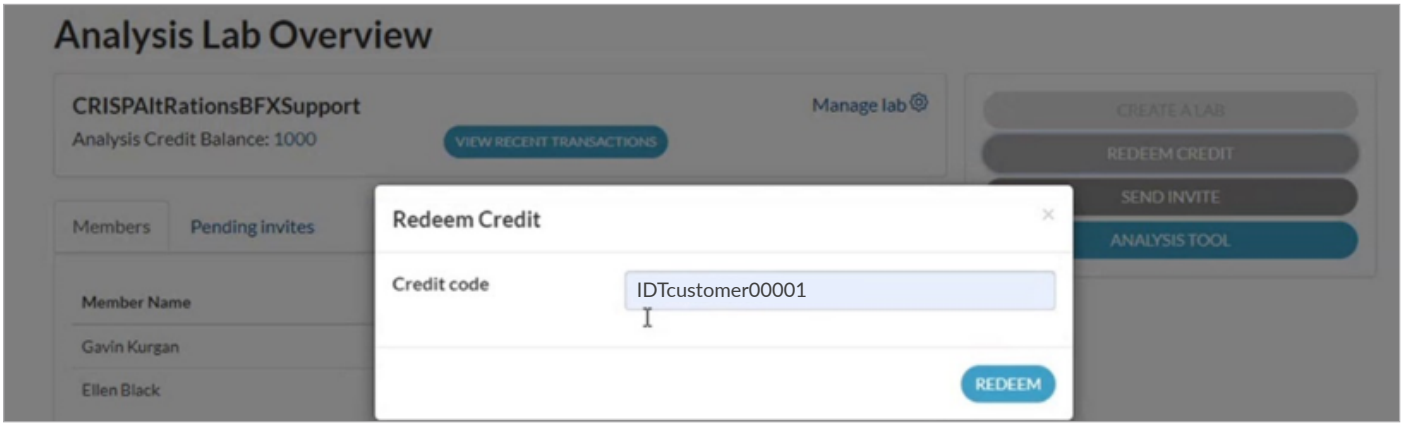


Tip: You can control user privileges by setting permission levels for each invited member.

Manage your account

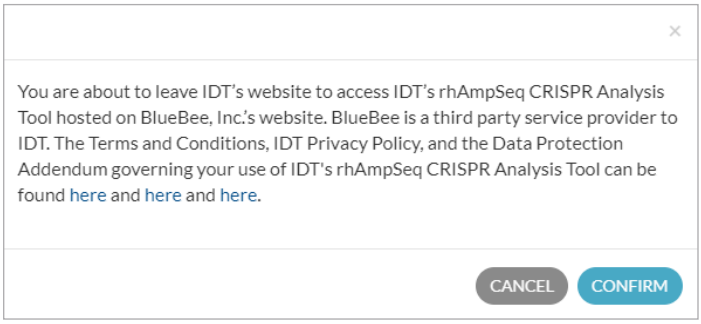
To use the rhAmpSeq CRISPR Analysis Tool, you will need Analysis Credits. You received a code for these credits when the order for your purchased product was invoiced. To redeem your Analysis Credits, click the gray **Redeem Credit** button, then enter the code. View your current balance and transaction history in the top left corner.

Tip: Refresh the page to see updated information on your account balance.



Start analyzing

Open the rhAmpSeq CRISPR Analysis Tool by clicking the blue **Analysis Tool** button. Select **Confirm** when prompted.



A new login page will open where you will enter your IDT account credentials again for security purposes. Now you are ready to begin analyzing data!

Tip: Since you have learned how to create and manage your account, you are ready to explore the other tutorial videos that further describe how to use the IDT rhAmpSeq CRISPR Analysis Tool.

At any point if you need help or have questions, let us know. We are here to assist you. Reach us at applicationsupport@idtdna.com.

Upload data

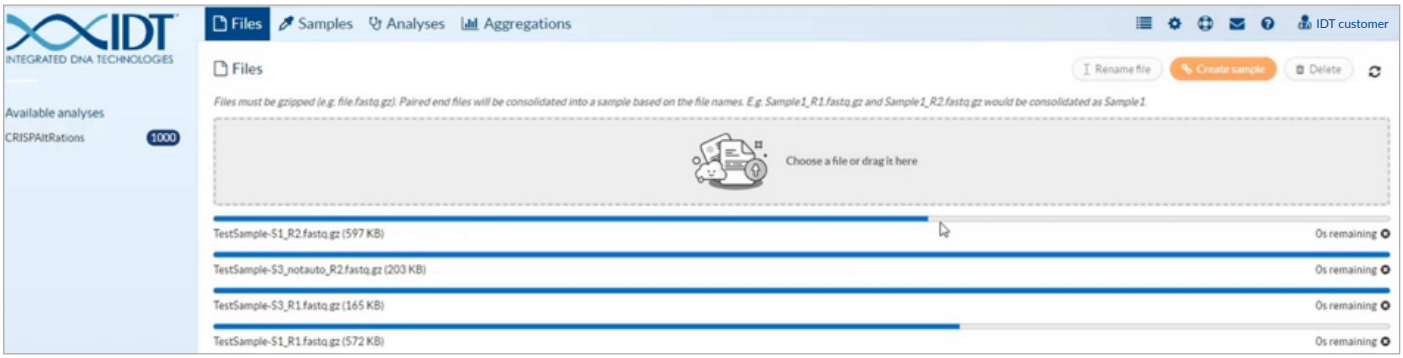
Before we delve into the process of uploading data, let us review the supported data formats. Specifically, you will have to provide demultiplexed sequencing data, usually in the form of .fastq or .fastq.gz files as shown in the below image. These files contain the following attributes for each read in the sequenced sample:

- read name
- sequence
- associated quality score

There are a few different ways to upload your data to the rhAmpSeq CRISPR Analysis Tool.
You can:

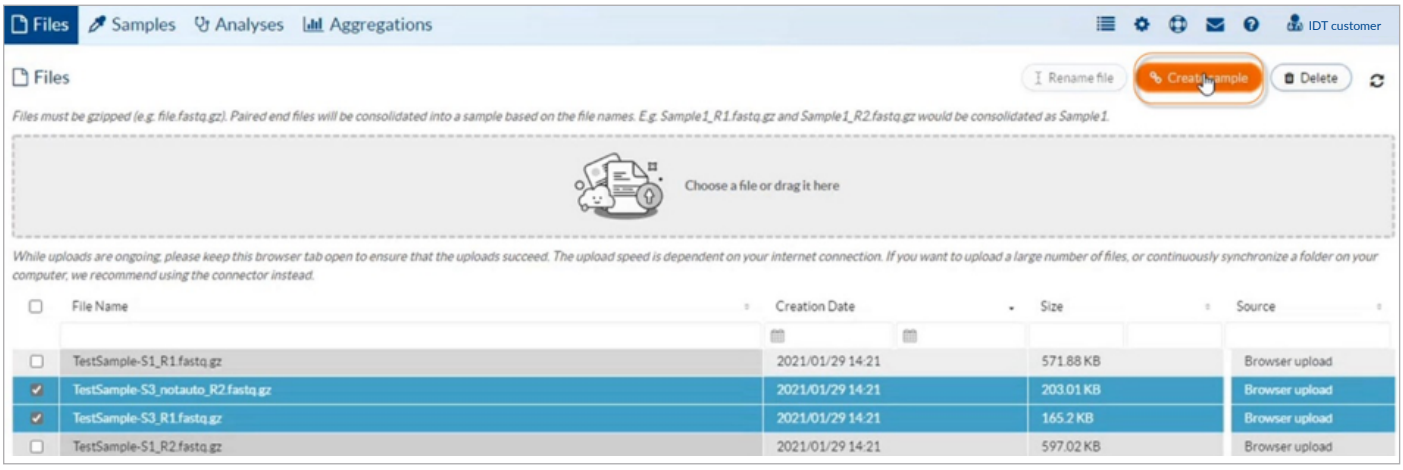
- drag and drop files into the application
- upload files from a local drive
- upload files from the cloud using the Connector

For smaller file sizes, using drag-and-drop is the fastest and easiest method for uploading sequence data.
Simply navigate to the **Files** tab, then drag the samples you want analyzed into the open gray box in the interface as shown below.
Files are uploaded in parallel, and upon successful upload, R1/R2 files are automatically paired to create a "Sample."



Note: For drag-and-drop uploads you must stay on the upload page until the transfer is completed. Wait for the progress bars to indicate when the uploads are complete before moving to the next step.

Note: Alternatively, you can manually pair read data into a sample by selecting the uploaded files, then clicking the orange **Create sample** button.



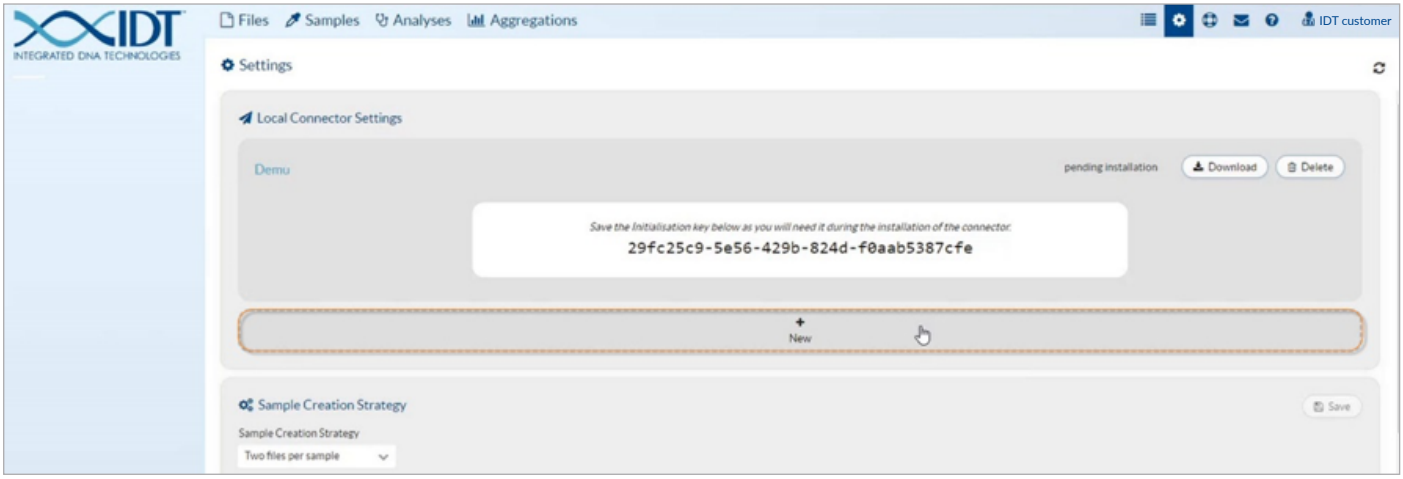
Upload large files via Connector setup

For larger files, or high-throughput applications, the drag-and-drop method may not be fast enough. Consider using the Connector setup upload feature. Navigate to the **Settings** page, indicated by a gear icon.

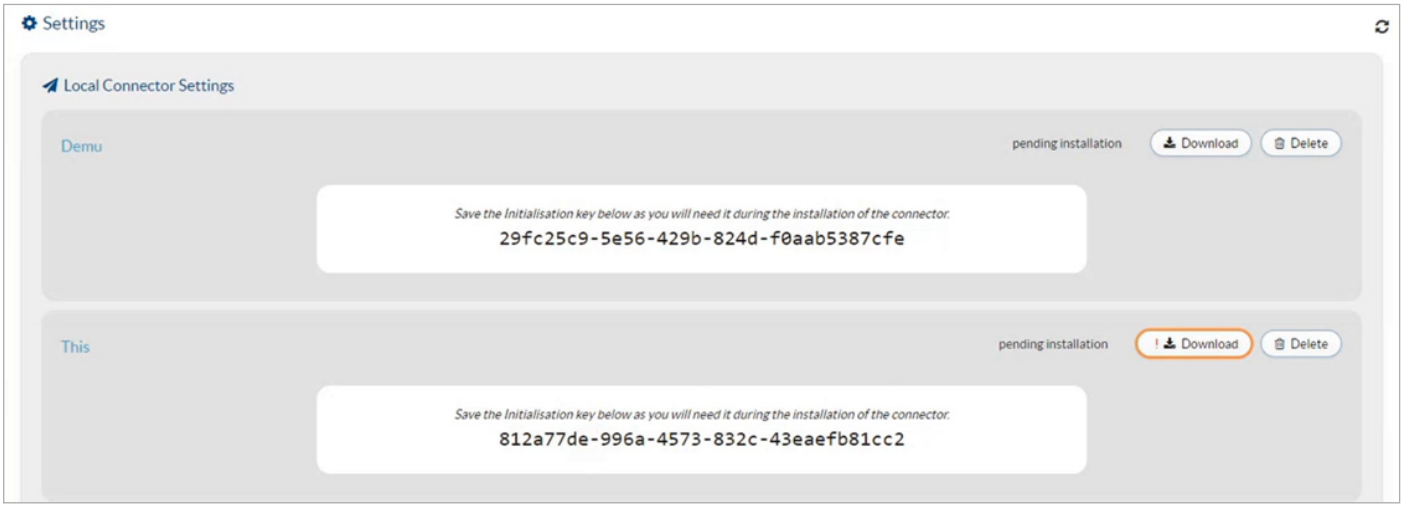
Here, you can set up the uploading of data from your local drive, or cloud storage systems. This guide will give you a high level overview of how to set up the Connector; for a more in-depth review, see the rhAmpSeq CRISPR Analysis Tool Data Transfer Guide.

To start uploads from your local drive to the analysis platform, click the **New** button under Local Connector Settings, then specify the following:

- name of the connector
- computer's operating system
- upload/download location of your data



Save your inputs, then download and install the connector software.



Follow the on-screen installation prompts.

To start uploads from a cloud streaming scenario, click the **New** button under the appropriate cloud service provider that you use. The setup is the same as above (uploads from a local drive), except that some specifics for each cloud provider may differ. Reference the rhAmpSeq CRISPR Analysis Tool Data Transfer Guide for detailed instructions based on your cloud provider.

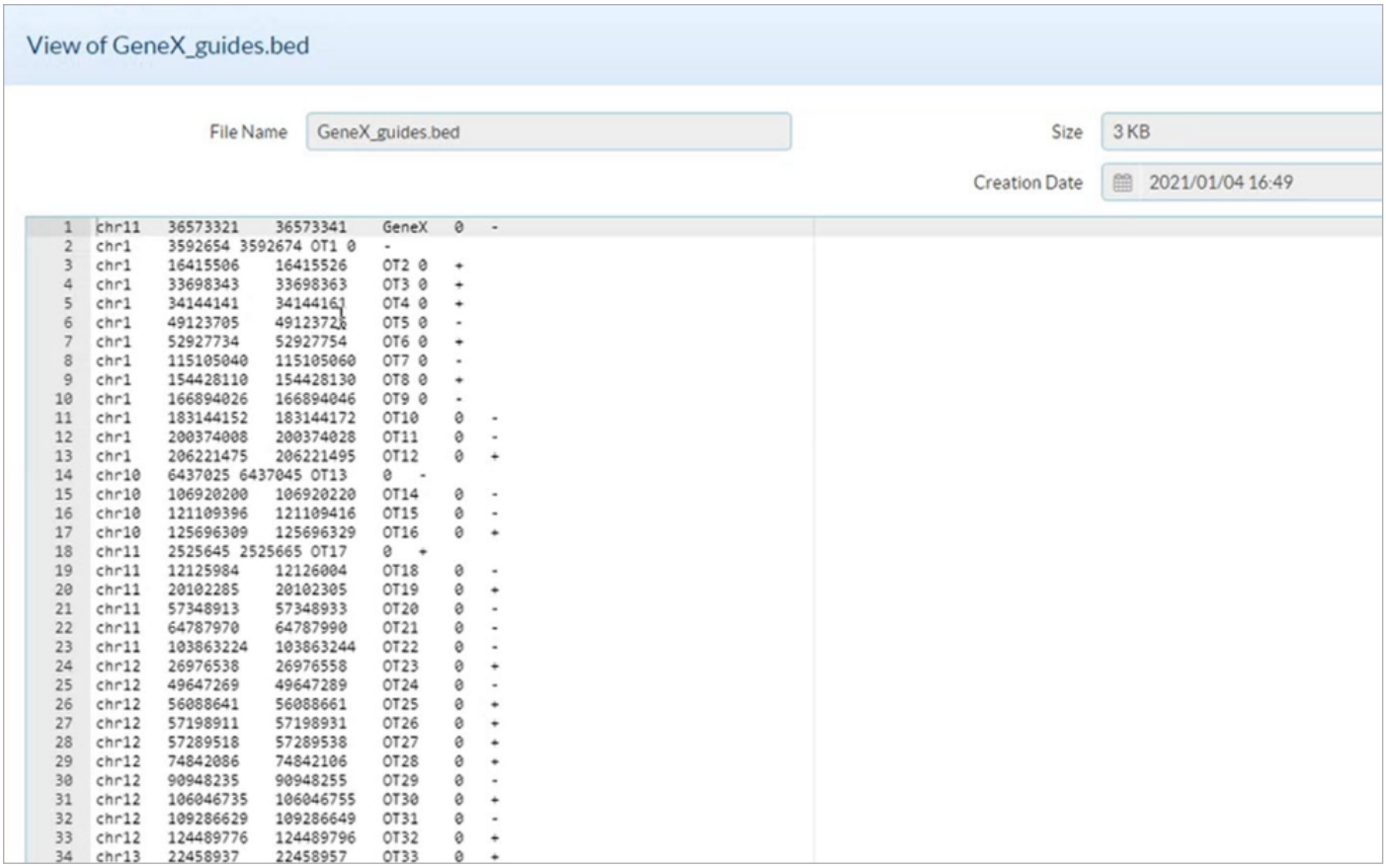
Data with multiple amplicons

If you are analyzing data with multiple amplicons in each sample, you will need to upload two 6-column BED-formatted files specifying amplicon and guide coordinates of the genome of interest. These BED files contain the following for all targets being analyzed:

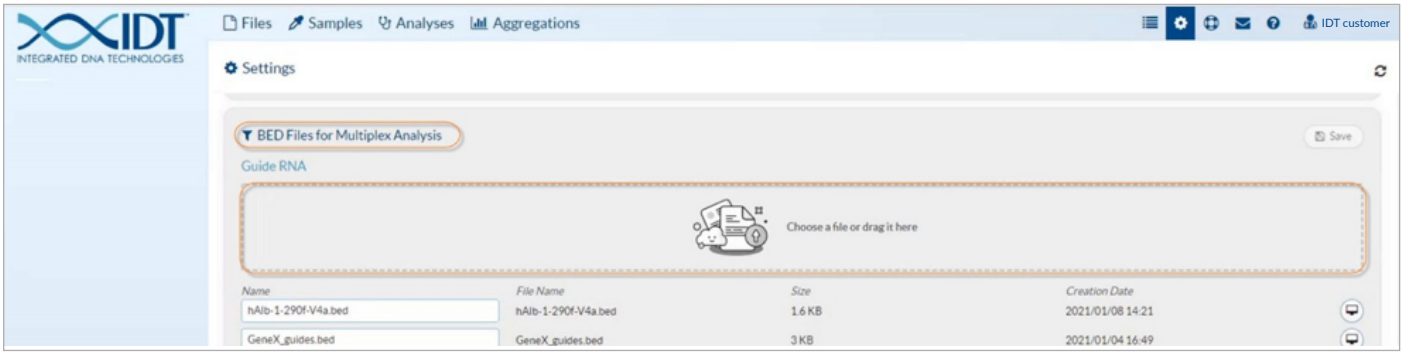
- chromosome
- start location
- stop location
- name
- strand

Important! Guide coordinates should not include the protospacer adjacent motif (PAM) but should include the correct strand information (+ or -) so that the guide is 5′ to 3′.

Note: If the rhAmpSeq Design Tool was used to generate the panel, then the amplicon BED file can be found in the Design Tool dashboard, as shown below. The file "Assay_panel.bed" will work for this.

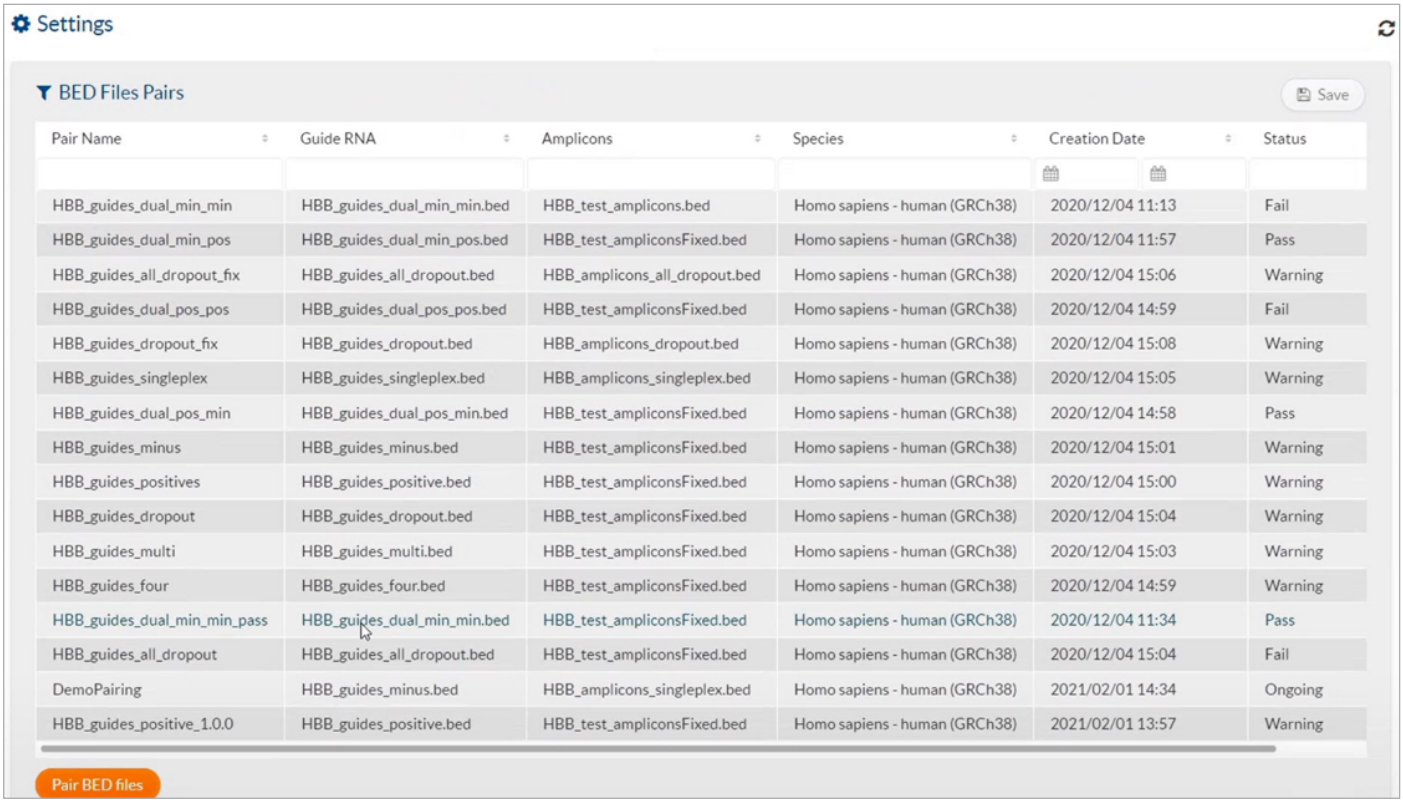


To upload these files, use the gear icon to open the **Settings** page, then scroll down to find the section called "BED Files for Multiplex Analysis." Upload your guide RNA and amplicon BED files into the specified sections by using the drag-and-drop upload method.



Quality control checking BED inputs

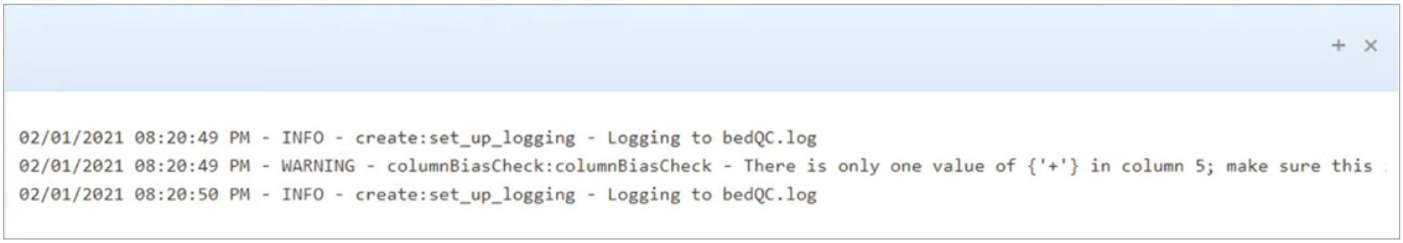
In order to use your BED files for multiplex analysis, the uploaded BED files need to be checked for errors. Navigate to the **Settings** page, then look for the "BED Files Pairs" section. This is where you QC-check your inputs to make sure they can successfully be used for analysis. To begin, select the orange **Pair BED files** button, as shown below.



You will enter:

- a unique name to reference the files being paired
- the pipeline version intended for analysis
- guide and amplicon BED files
- species these coordinates come from

These entries will initiate the QC check on the files, which will take approximately 5–10 minutes to complete. A status message of PASS, WARNING, or FAIL will appear when the pairing is complete.



Handling QC statuses

If a **PASS** status is shown, you are ready to start your analysis.

If a **WARNING** status is shown, double-click the pairing to understand why the paired BED files were flagged. These files will still be available for analysis; however, they should be reviewed to make sure there are no persistent problems with your experimental inputs.

If a **FAIL** status is shown, there is something wrong with the BED inputs that needs correction. Check the log and make the appropriate corrections in the files to reverse the failure.

Important! Once paired, BED files cannot be deleted from the analysis tool. This is to ensure they are available for previous analyses that used them. Try to provide unique and memorable names for these BED file pairs.

Analyzing singleplex experiments

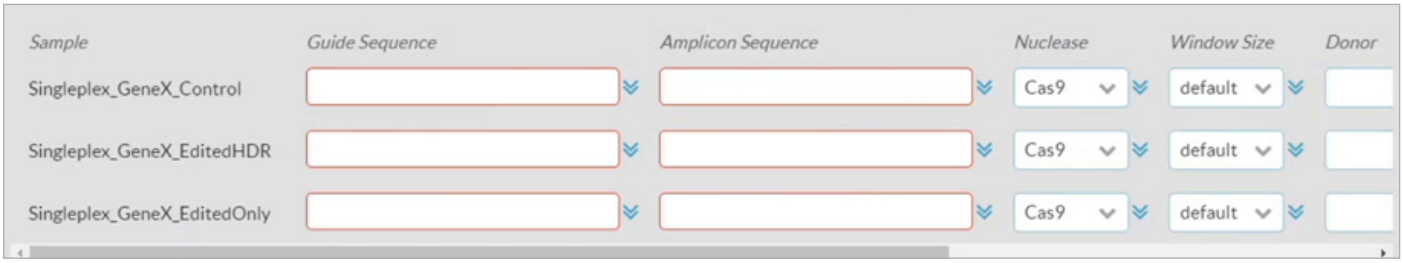
If you want to screen on-target guides and conditions, you will need to know how to analyze your singleplex experiment.

Important! Make sure that data has been completely uploaded and put into samples before proceeding.

First, select your samples. Navigate to the **Samples** page where you will be able to see all the samples that have FASTQ files associated with them. Select the groups of samples with a single amplicon that you want to analyze. Begin the analysis by clicking **Analyze Singleplex**.

To start the run, you will need to associate information and parameters with the run and individual samples. This is where you will also specify the version of the CRISPAltRations software to use for the analysis.

Before starting the run, provide the information highlighted by red boxes, namely the nucleotide sequences for the guide and the amplicon, 5′ to 3′, excluding the PAM.

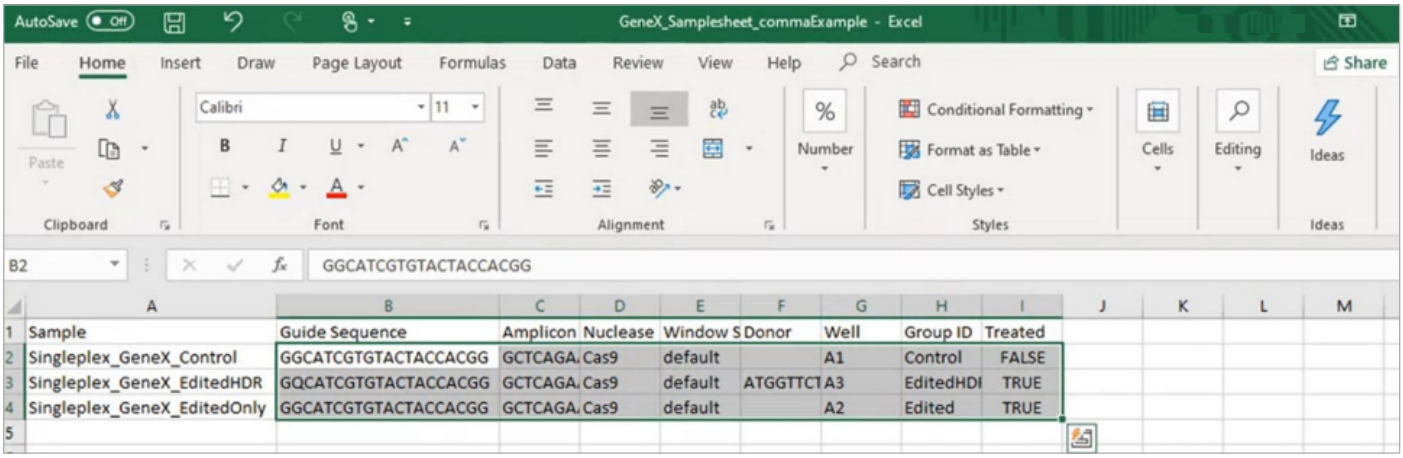


The additional parameters that can be configured on a run include:

- the applicable CRISPR nuclease
- window size for interrogating variants
- nucleotide sequences for HDR donors
- well position (for advanced visualizations)
- sample name
- treated vs. untreated sample

Note: You can quickly populate the same entry down an entire column at once by entering the information in the first row, then clicking the down arrow icon to the right of the field.

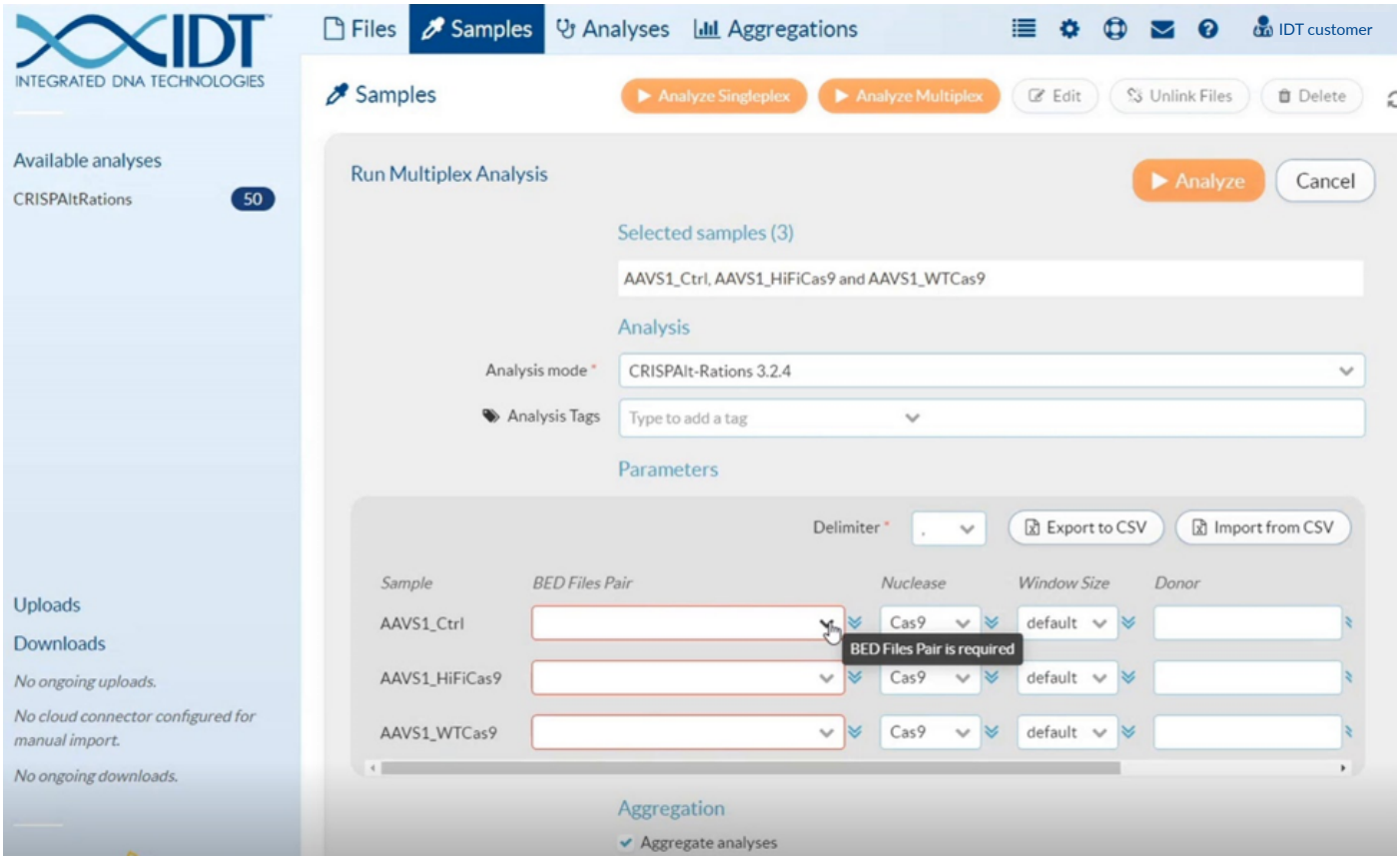
Note: For batch runs that have a large number of samples, you will need to upload a samplesheet to provide run information. To get a template of this samplesheet, click the **Export to CSV** button.



After specifying your sample information in Excel, save and upload this file directly into the interface using the **Import from CSV** button.

Before starting your run, designate an "Aggregation" name, which creates a single location to interrogate all samples in the run. You can add additional metadata and unique descriptions to help label samples and experimental details for tracking results before starting the run.

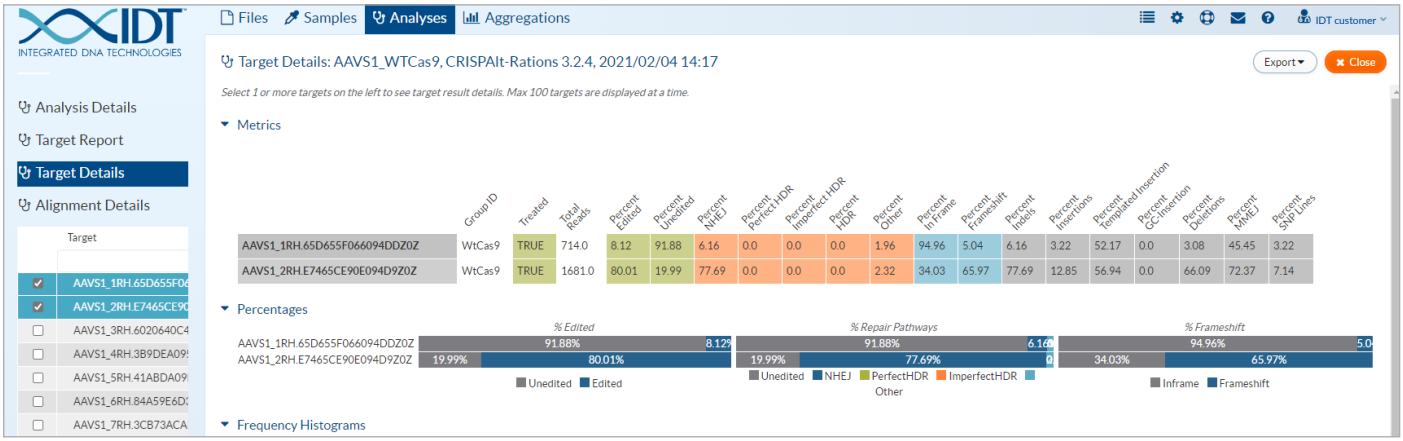
Caution: When a red box highlights any of the sample information fields, you will not be able to start the run. Check the hover text on any parameter with a red box to find acceptable inputs and make corrections accordingly. When all inputs are satisfied and no red boxes appear, then you are ready to begin the run.



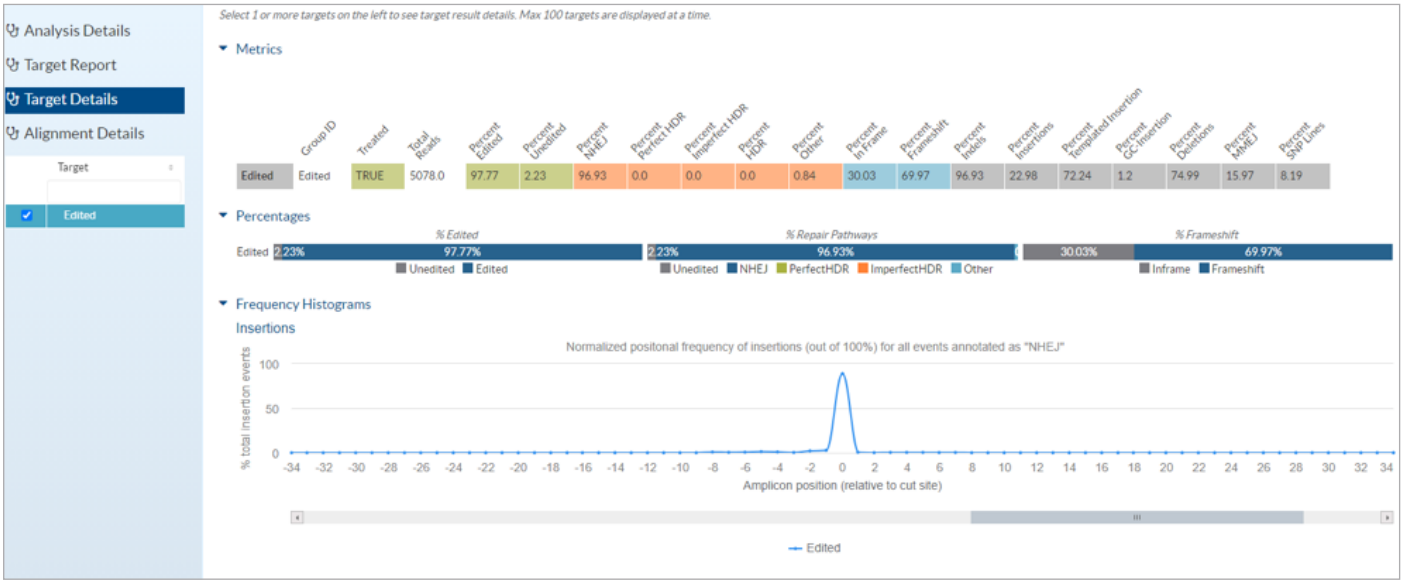
To check the status or results of a run, navigate to the **Analyses** page. As soon as cloud resources begin processing, a progress bar should appear. Analyses can take between 20 mins–2 hrs to complete.

Interpreting singleplex experiments

View your results by double-clicking the sample on the **Analyses** page. This opens the landing page where you will interpret the results of your experiment. View the graphics for all targets amplified in the "Target Details" section. You can also easily export the graphics in the interface for selected targets using the **Export** button in the top right corner of the page.

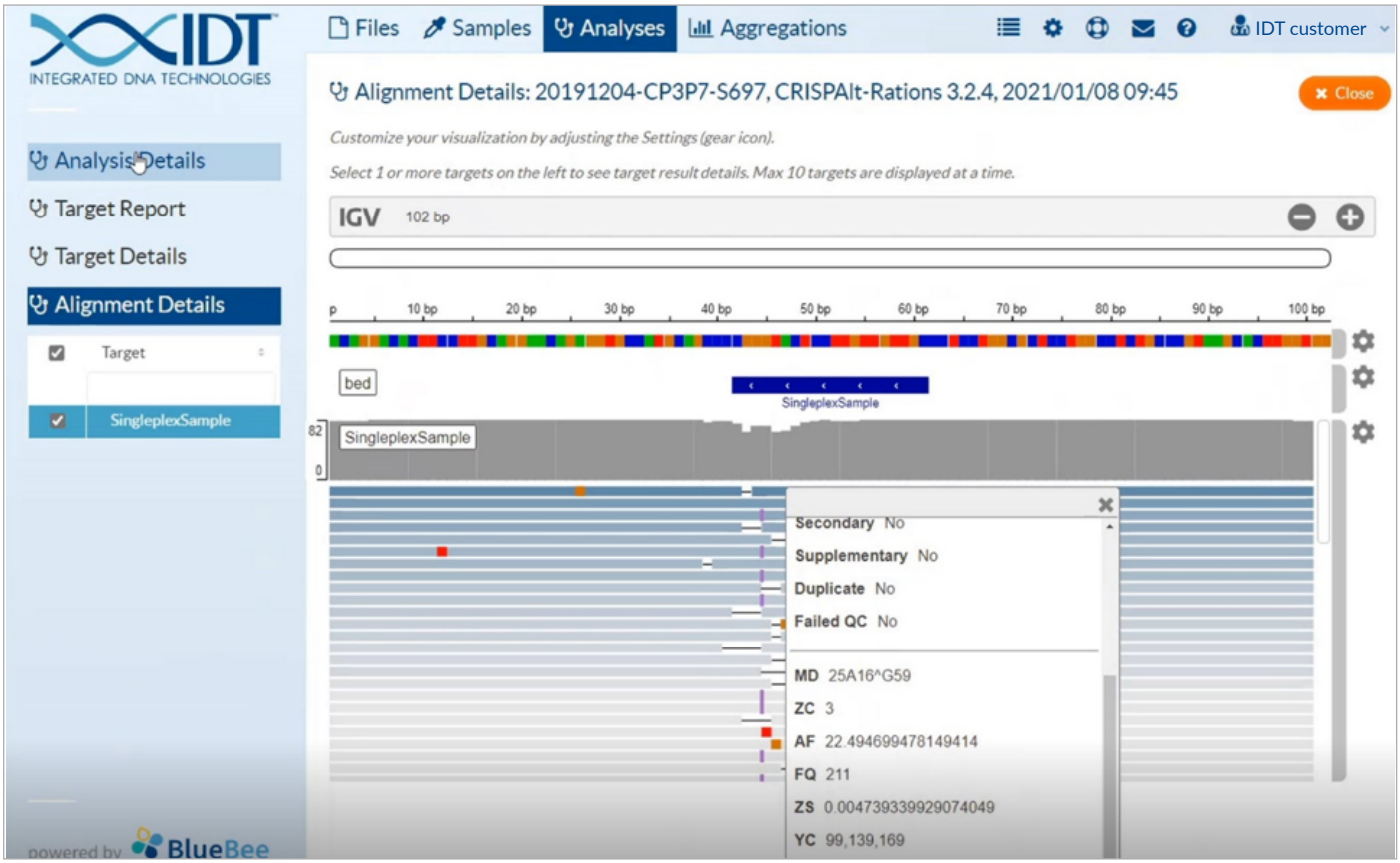


Also, on this page you will be able to find the frequency of the editing, the specific repair pathways used, and any frameshifting events that may have occurred. Dive deeper to explore the exact position of insertion and deletion events in the sequenced sample, even the size of these events. We include a frequency histogram showing the normalized positional frequency of insertions, as shown below.

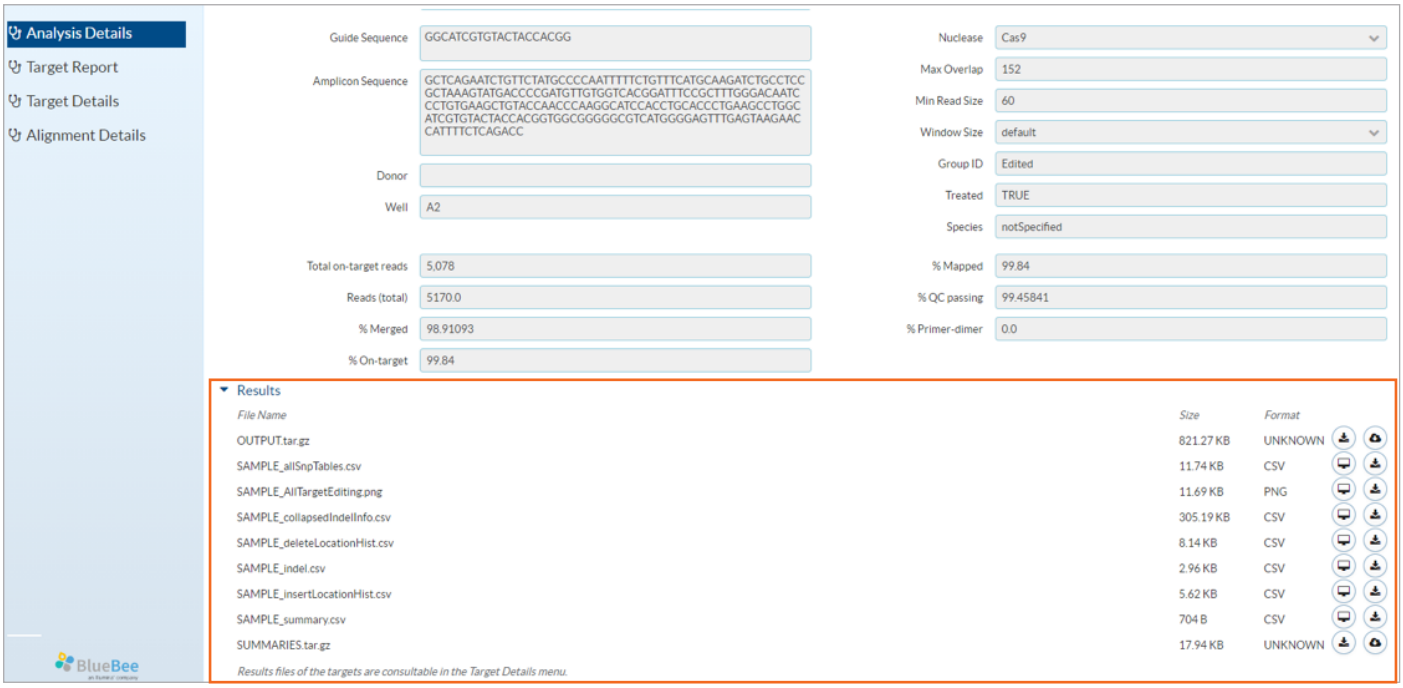


Viewing allele details

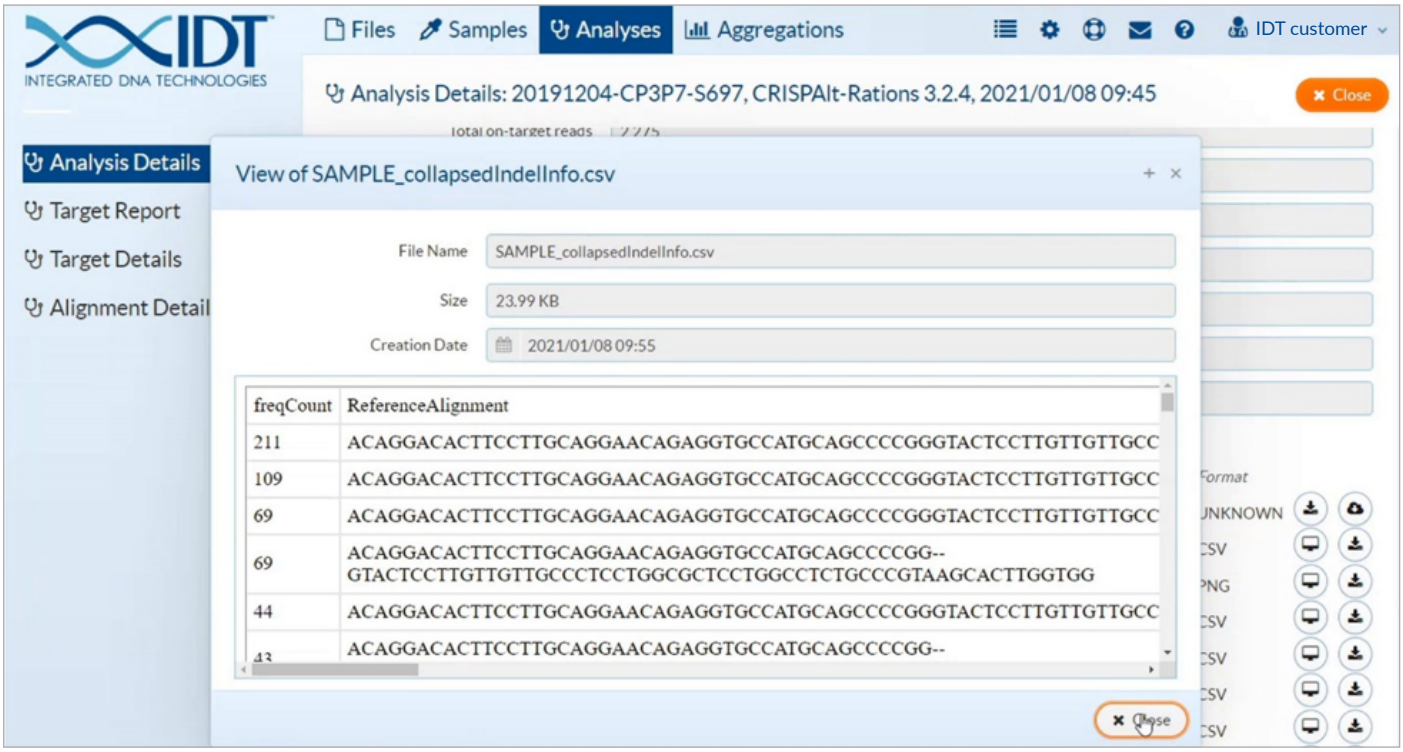
Go to the **Alignment Details** page to view the distribution of alleles in the sequenced population using the Integrative Genomics Viewer (IGV).
Click on the event of interest to see both the frequency of the event and the read count.



For those with advanced graphical considerations, the raw data we use to render the graphics can be downloaded from the **Analysis Details** page.



On this page, you can see general sample information, as well as an assortment of other files for either viewing or downloading.

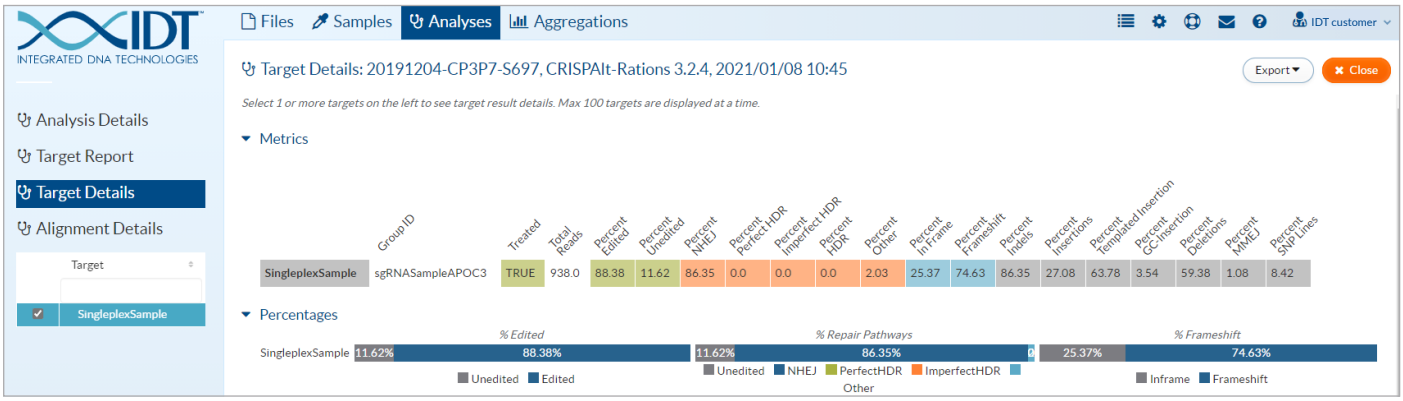


Viewing the target report

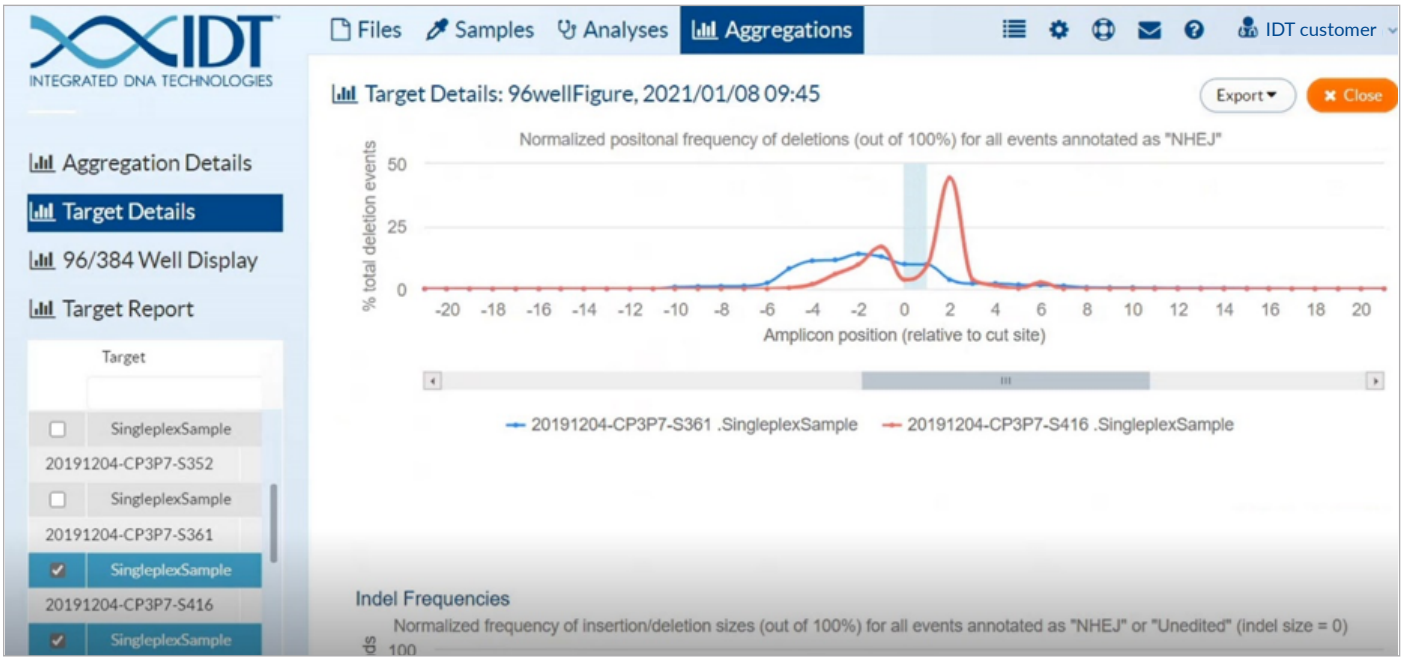
To interpret and compare results of different samples within an experiment, go to the **Aggregations** tab.

On the **Target Report** landing page, view all the samples that were aggregated for summarization. Any report that you view can easily be exported to Microsoft Excel. From the **Target Report** page, sort or filter the experimental results as needed, and create additional graphics in the software of your choice.

To compare a sample with another, click **Target Details**. On this page you will find the same graphics that were available for individual samples; however, here you can also overlay results from other relevant samples within the whole experiment.



In the graphic below, the run contained two samples with the same guide, treated under different conditions. Notice that although the editing efficiency is different, the overall indel profile of these two samples is similar, as expected.

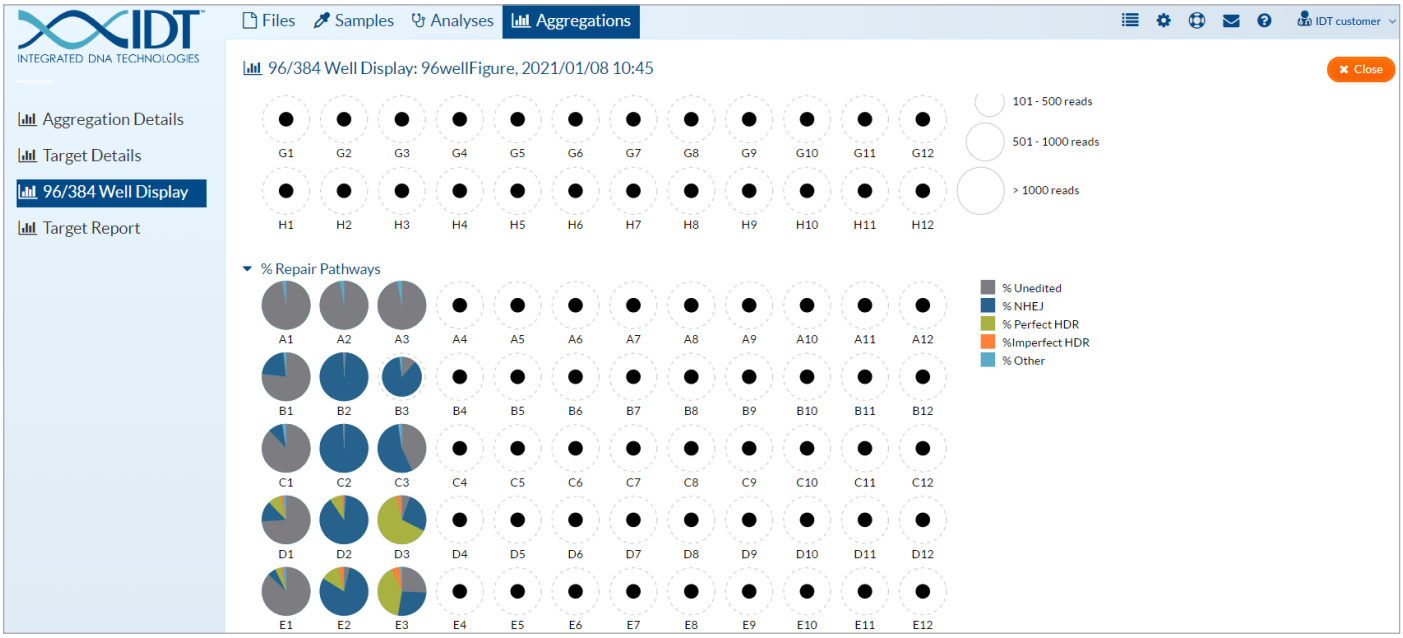


Screening guides

When screening a large number of guides under varying conditions and treatments, you can generate graphics in 96- or 384-well plate formats to find conditions that are optimal for your experimental goals. The graphics in 96- or 384-well plate formats are only available for singleplex analyses. To create these graphics, simply specify the "Well" information for each singleplex sample prior to sample analysis.

Once your run is complete, go to the **Aggregations** tab where you can find graphics displaying the following:

- frequency of editing
- repair pathway utilization
- frameshift frequency



Analyzing multiplex experiments

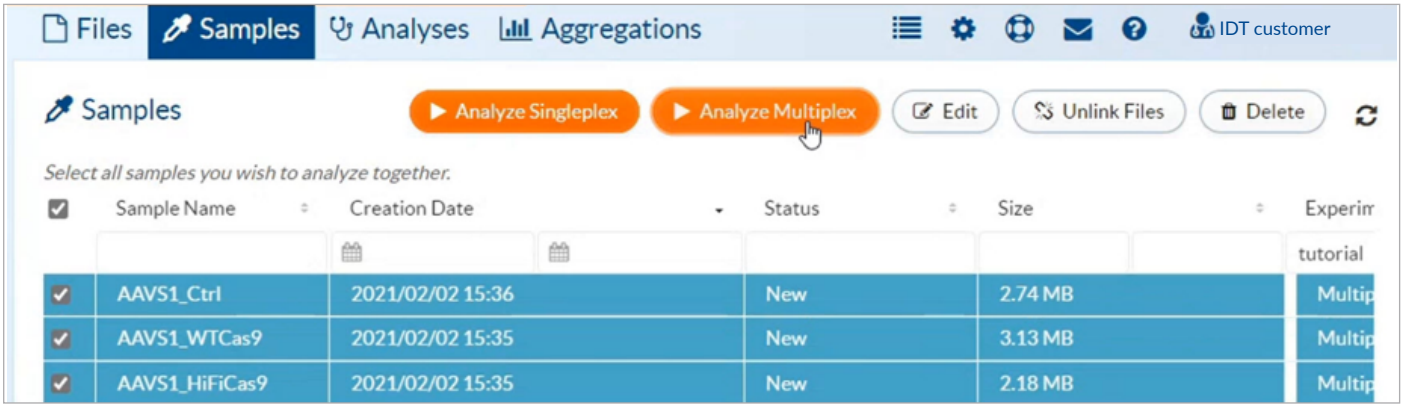
The most common application for multiplex CRISPR NGS is to quantify on- and off-target editing effects of a guide. In our example, we show how to analyze samples that have multiple amplicons per sample, known as multiplexed samples.

Important: Make sure the following have been completed before proceeding.

- Data have been uploaded and created into samples
- Guide and amplicon BED files have been uploaded and paired

Tip: For a deep dive into pairing your guide and amplicon BED files, watch the section of the rhAmpSeq CRISPR Analysis Tool video How to quality control check multiplex file inputs.

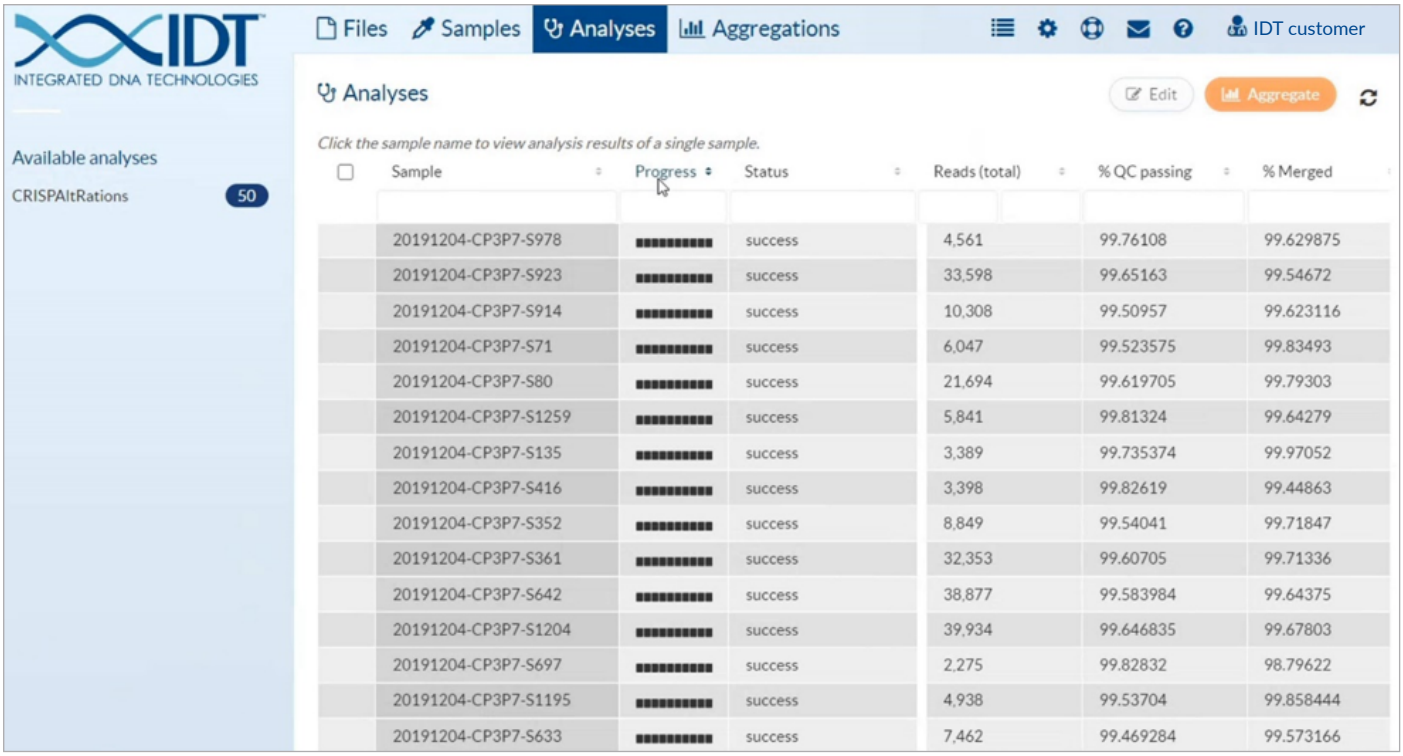
To begin, select the samples you want to analyze, then click **Analyze Multiplex**.



Once your data has successfully been checked for quality control, select your named guide/amplicon pairs and other relevant run information, including:

- the appropriate CRISPR nuclease
- window size for interrogating variants
- nucleotide sequences for HDR donors
- sample name information
- treated vs. untreated

Navigate to the **Analyses** page to check the status or results of your analysis. As soon as cloud resources begin processing, you will see the progress bar. Analyses can take anywhere from 20 mins–2 hrs to complete. Your turnaround time will depend on your read depth plus the number of amplicon targets being quantified per sample.



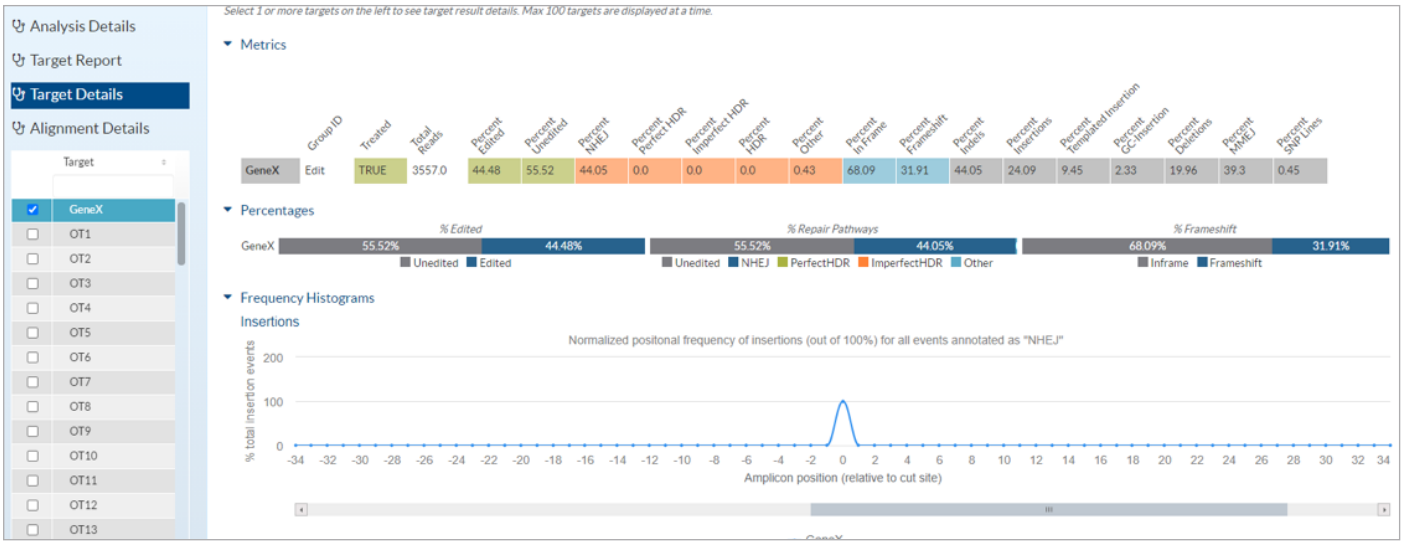
Interpreting multiplex experiments

When you have completed your analysis, the **Target Details** landing page displays. This page contains graphics for all amplified targets. Each target can be toggled on or off by checking or unchecking the checkboxes to the left of the targets in the table. Also, the active display on the page can easily be exported by using the **Export** button in the top right corner.

From the **Target Details** page, you can find a range of data, including:

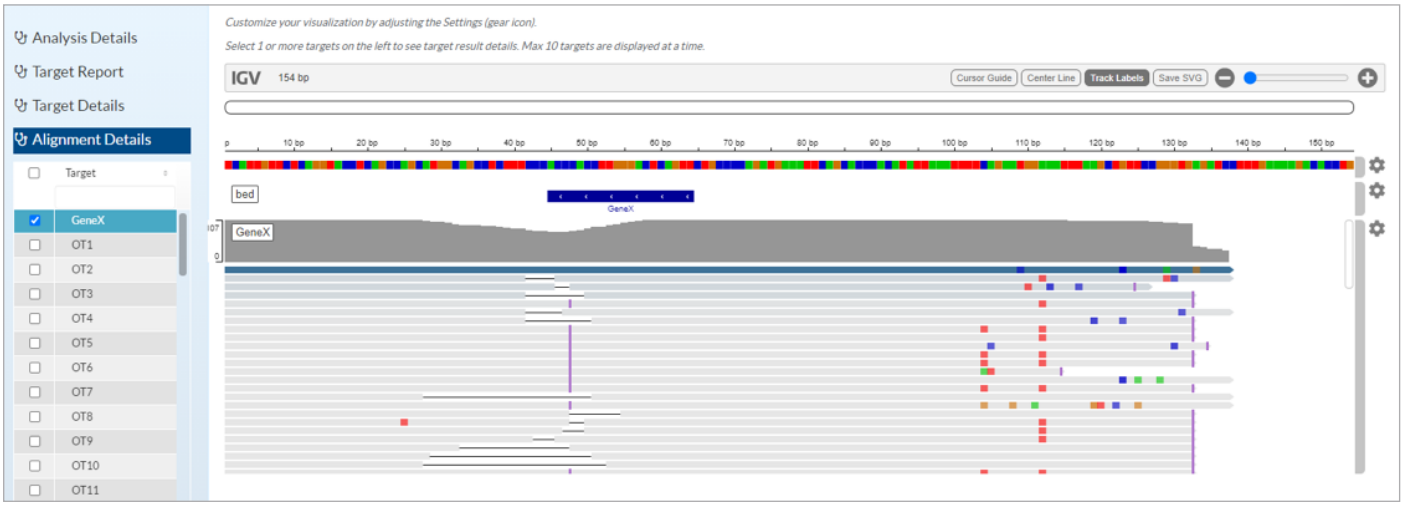
- frequency of editing
- repair pathways
- frameshifting events

Additionally, you can explore the position of any insertion and deletion event that occurred in the sequenced sample and the size of these events. We have included a frequency histogram that shows the normalized positional frequency of insertions below.

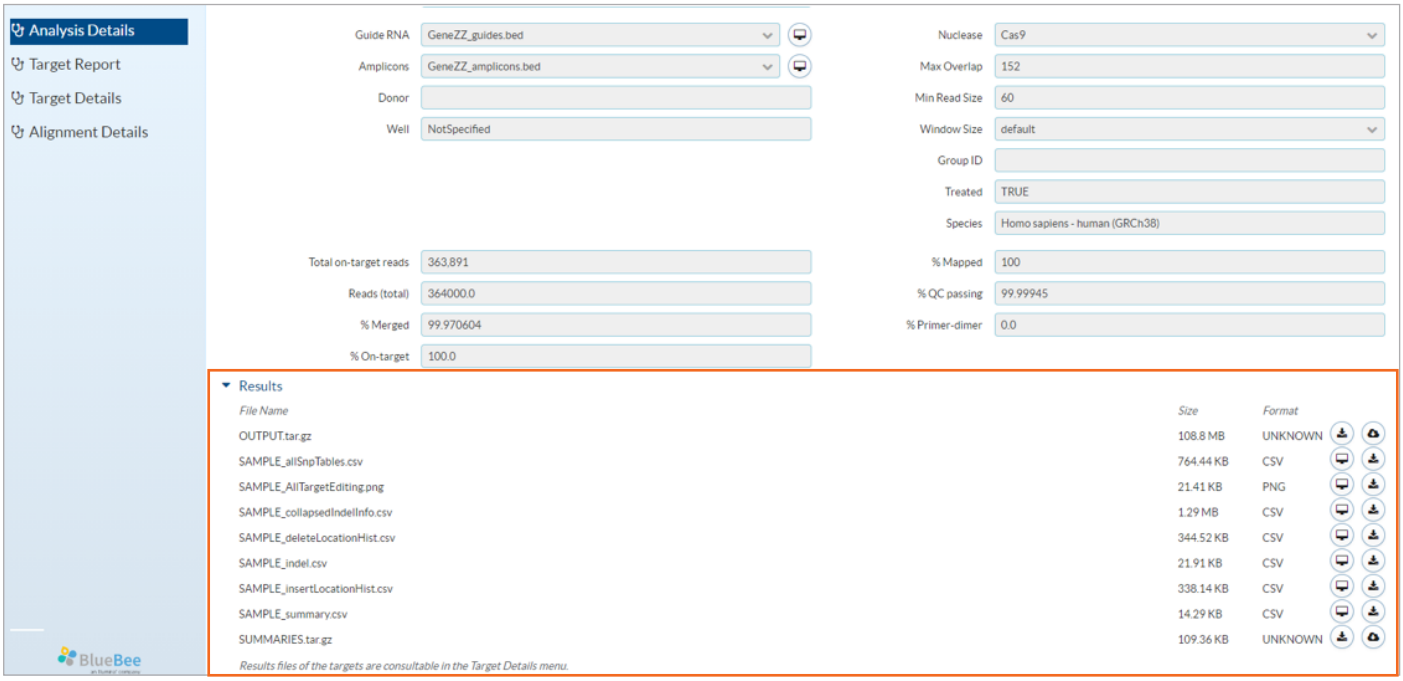


Viewing allele details

Go to the **Alignment Details** page to view the distribution of alleles in the sequenced population using IGV. Click on the event of interest to see both the frequency of the event and the read count.

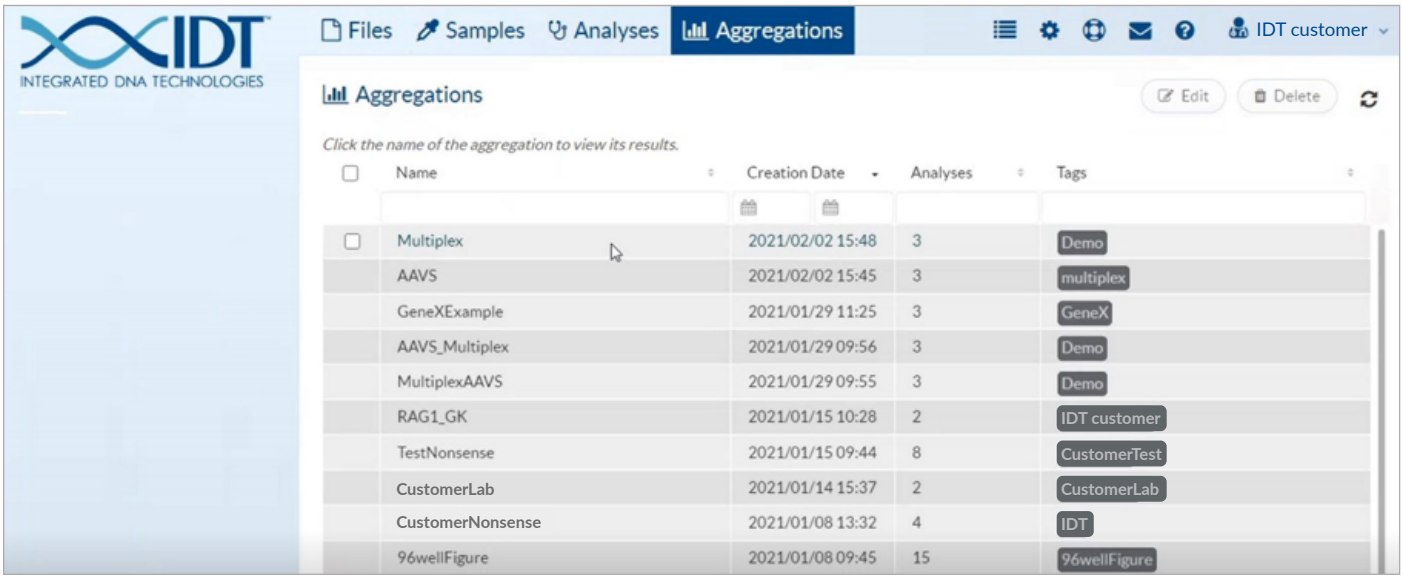


For those with advanced graphical considerations, the raw data we use to render these graphics can be downloaded from the **Analysis Details** page. On this page, you can see general sample information, as well as an assortment of other files for either viewing or downloading.

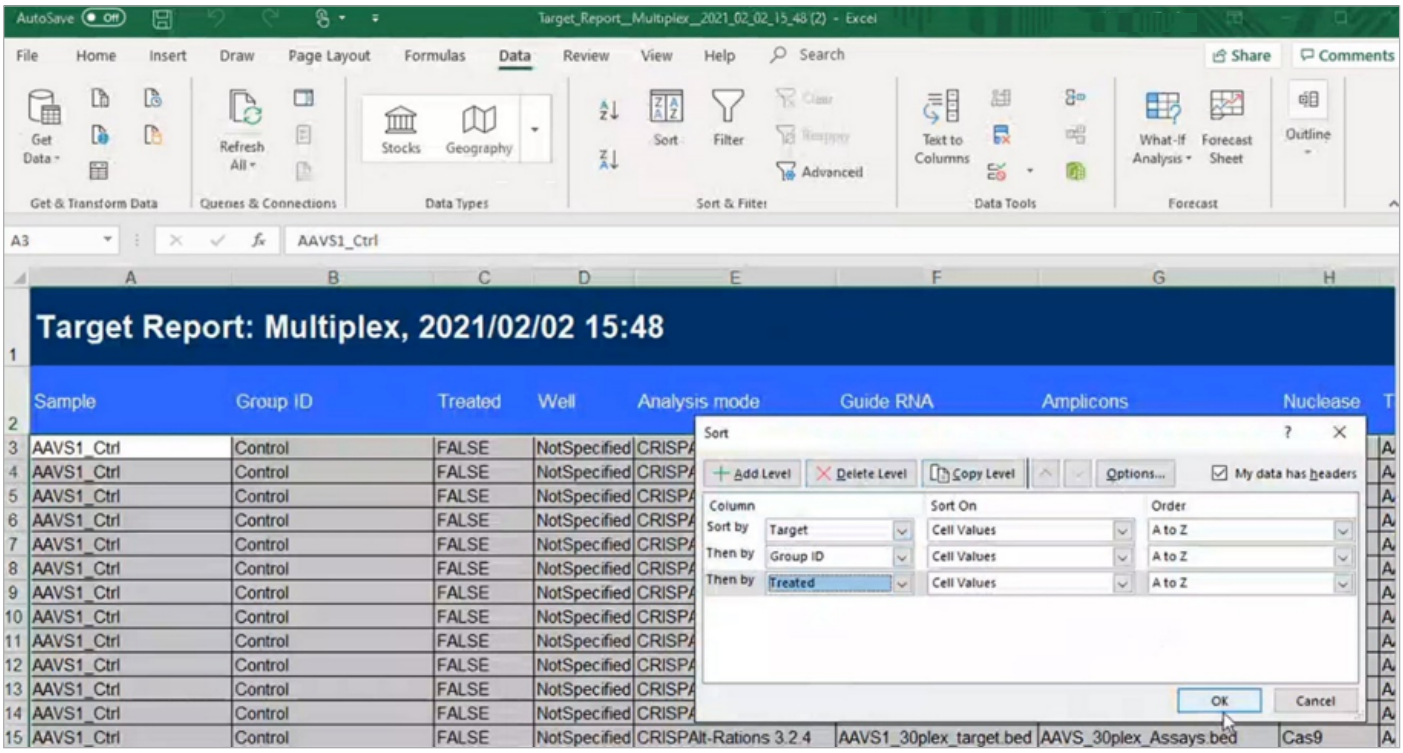


Viewing the aggregated results

Go to the **Aggregations** tab to interpret and compare results of different samples within an experiment.

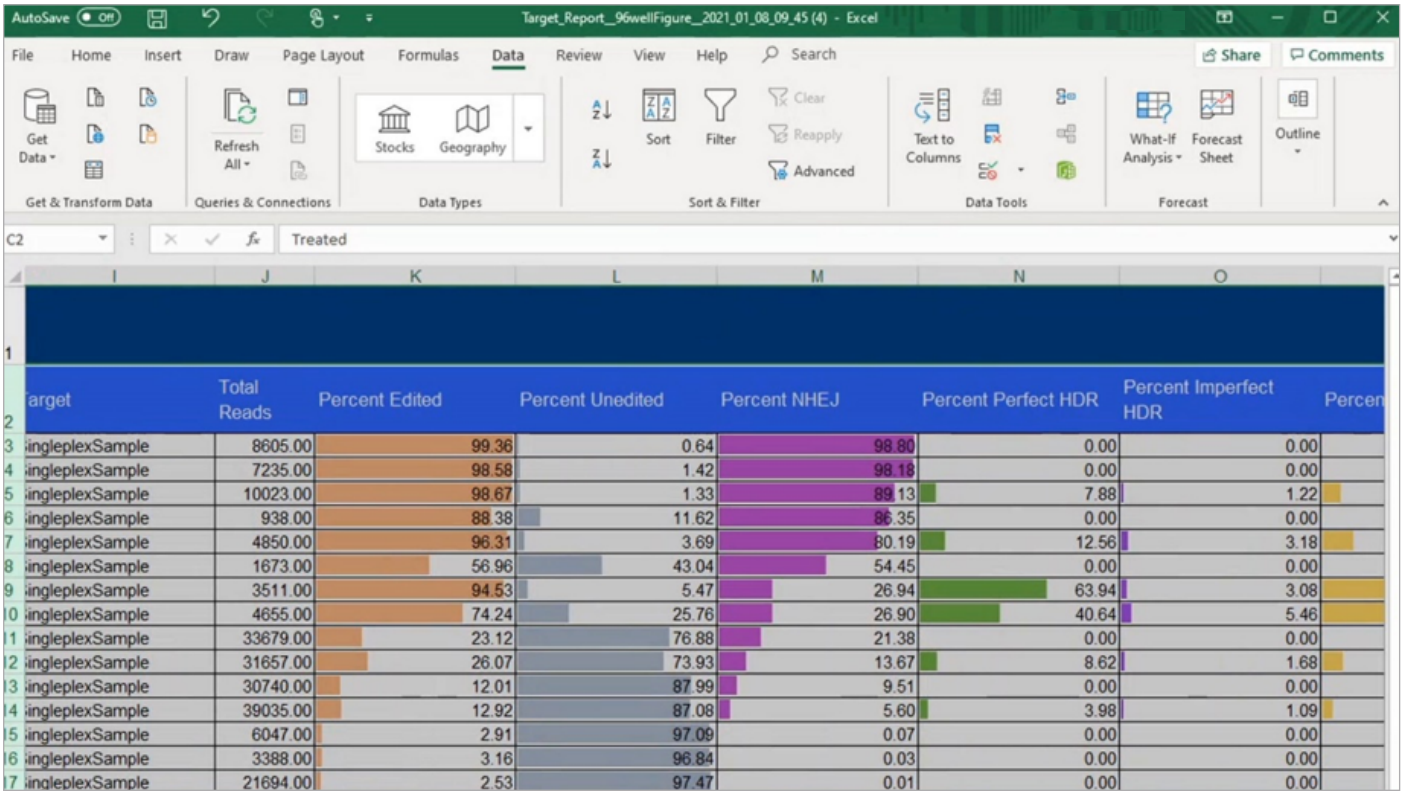


On the **Target Report** landing page, view all the samples that were aggregated for summarization, then export this information easily to Excel. With Excel, you can easily sort or filter experimental results as needed, then render additional graphics in the software of your choice.

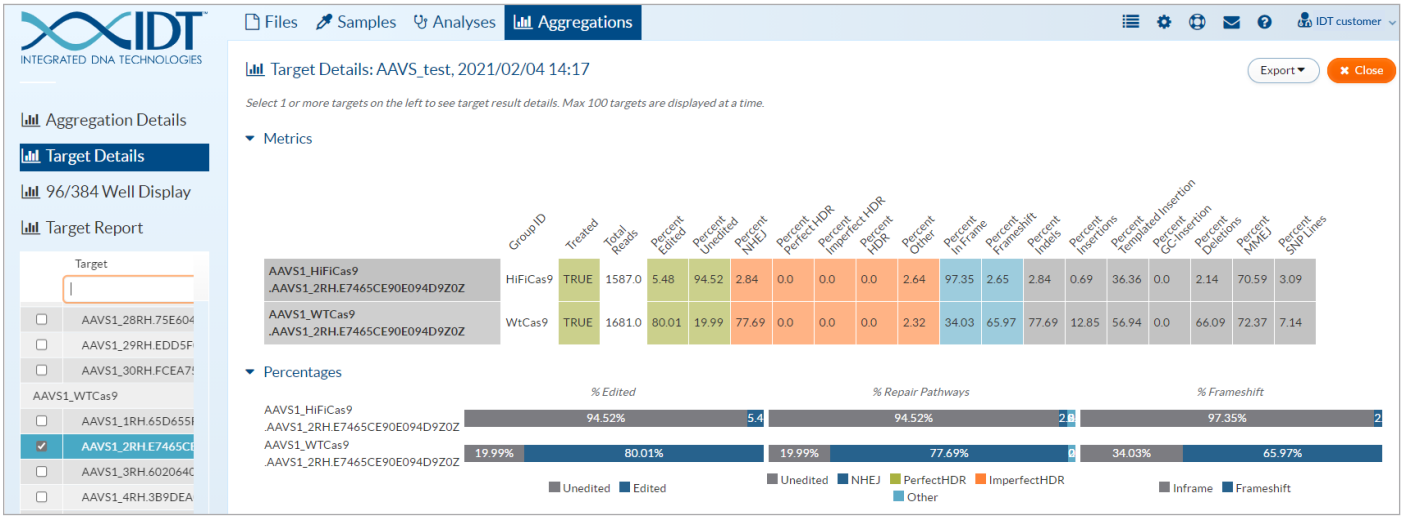


For on- and off-target editing experiments, this summary can be especially powerful for comparing treatment and control experiments. You can even use numbers in this table to implement a statistical test which can determine significant off-target effects in the experiment.

A quick sort by **Target**, **Group ID**, and **Treated** can quickly pair treatments with their corresponding controls to visualize editing differences using the conditionally formatted output.



To investigate differences between samples further, go to the **Target Details** page. Here, you can overlay the results of targets from different experimental samples, then export these graphics using the **Export** button in the top right corner. The graphics provided here are the same as when analyzing a single sample, but single samples cannot be overlaid. An example of this is shown below.



A complete video tutorial version of this guide can be viewed on the rhAmpSeq CRISPR Analysis Tool page.

Using the rhAmpSeq CRISPR Analysis Tool

Technical support: applicationsupport@idtdna.com

5655 0421