Sigma-Aldrich®
Lab & Production Materials

Protocol

DEL NGS Analysis Site: Usage Instructions for DyNAbind® 10 Million Compound DNA-Encoded Library (DYNA002)

Summary

This document provides basic instructions for operating the MilliporeSigma DEL NGS analysis web portal to analyze data from using product DYNA002.

Analysis

To analyze the results of your next-generation sequencing, you can securely upload your data to the MilliporeSigma DEL NGS Analysis Portal (and perform an automated analysis to find the most likely candidates for drug development.

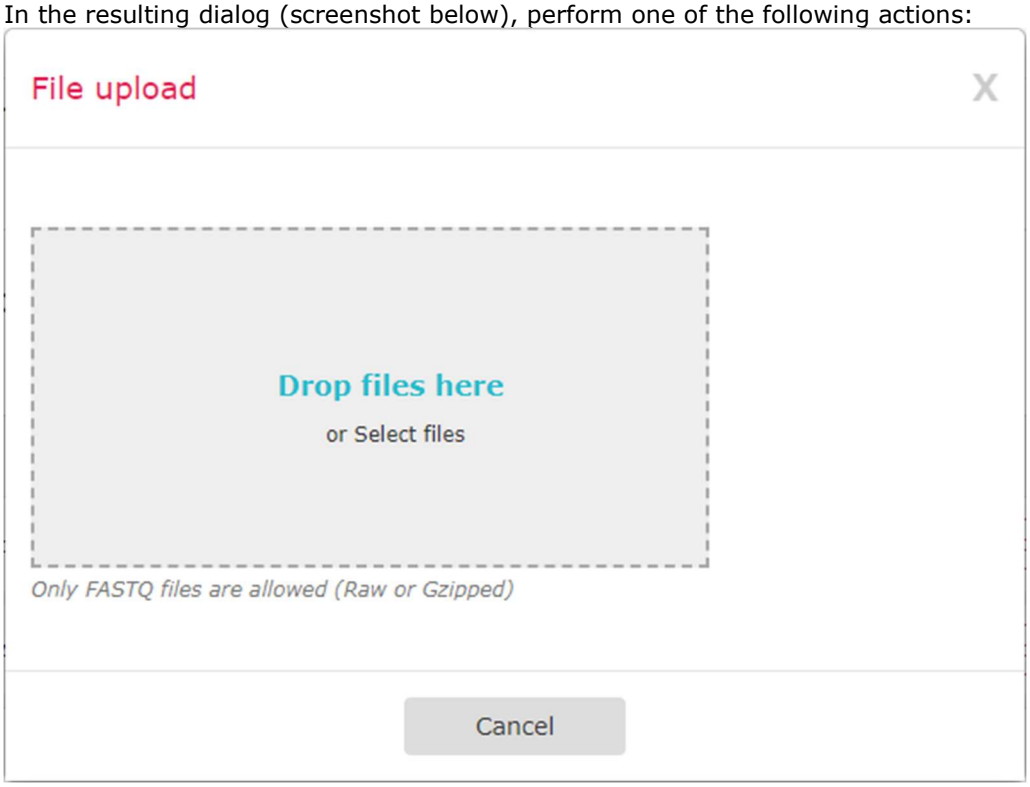
Registration

Before uploading any data to the server, you will need to register on the site. Go to the DEL NGS Analysis Portal and click the **Register Now** link. For organizational registration, simply use a single user's account for the organization.
Note: Because of the high storage requirements for NGS data, each user is limited to **30 GB** of FASTQ storage.

Uploading Data

When uploading files, the maximum size per file is 5 GB. Larger files can be split and loaded as multiple files. Make sure that any of the split points are between reads. To upload your data in the Analysis Portal:

1. Click **Upload FASTQ Files**.
2. In the resulting dialog (screenshot below), perform one of the following actions:



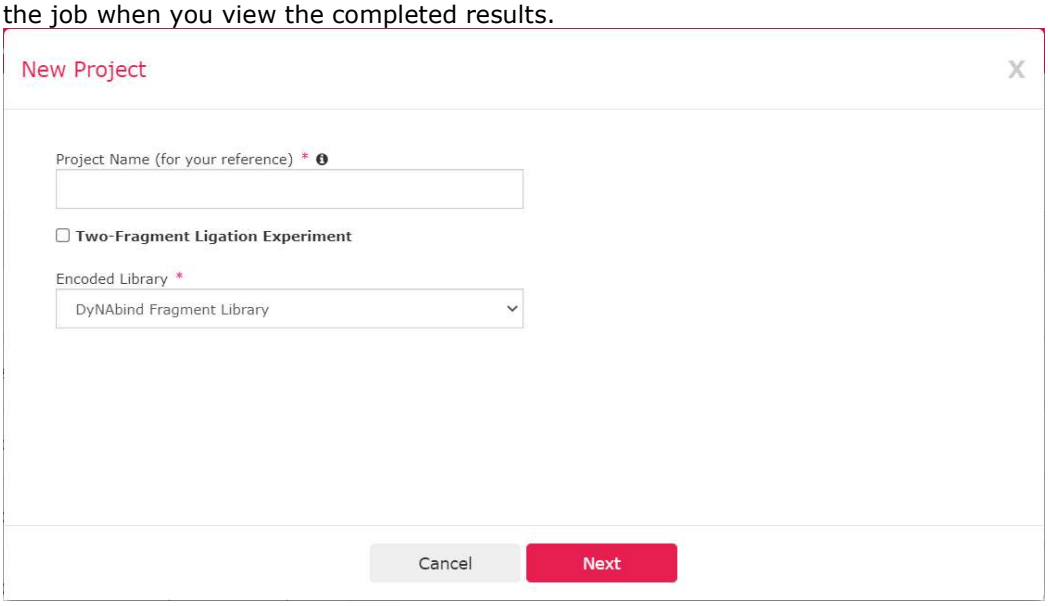
   a. Drop your file(s) onto the **Drop files here** section.
   b. Click **Drop files here** and select your file(s) from the file browser dialog that appears.
3. Wait for a message indicating that the file was successfully uploaded.

Analyzing Data

To analyze data, a new Job must be created on the site.
To create a job:

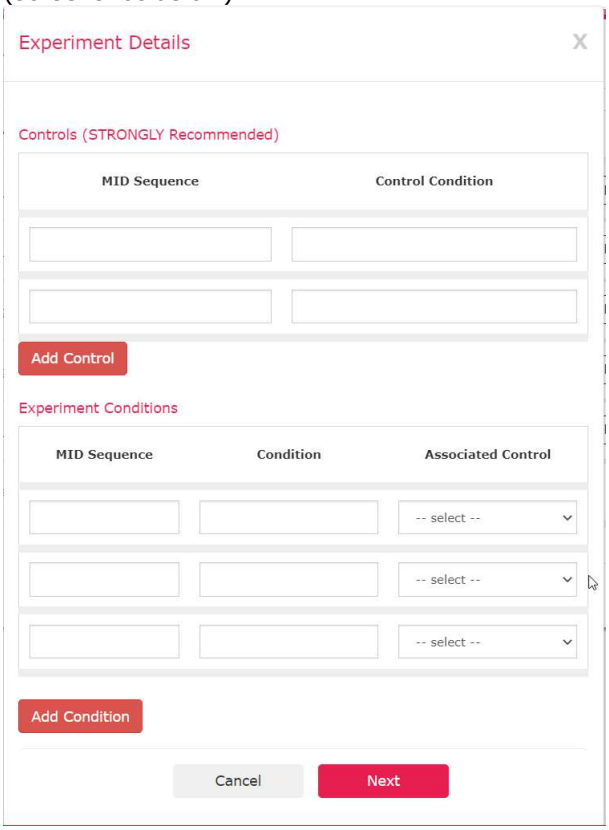
1. Click **Create Job For Analysis**.

2. In the resulting **New Project** dialog (screenshot below), enter a name for the job. This will be used to label the job when you view the completed results.



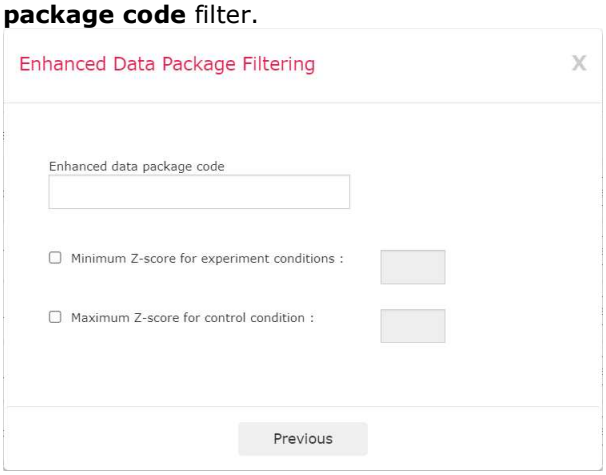
3. Select "DyNAbind 10 M Library" from the **Encoded Library** dropdown.
4. Select the **Enhanced Data Package** checkbox and click **Next**.
5. Use the **Select** checkboxes to pick the file(s) that will be analyzed.
    a. If one or more files are not listed, click **Upload New File** to open the **File Upload** dialog and upload the necessary file(s).

6. Once all of the necessary files have been selected, click **Next** to bring up the **Experiment Details** dialog (screenshot below).



7. Enter each of the negative controls (beads without target) in the **Controls** section of the **Experiment Details** dialog as follows:
    a. Enter the MID (Molecular Identifier) associated with the control in the **MID Sequence** column.
    b. Enter a unique name in the corresponding **Control Condition** column.
8. Enter each of the treated conditions in the **Experiment Conditions** section as follows:
    a. Enter the condition's MID in the **MID Sequence** column.
    b. Enter a unique name for the condition in the **Condition** column.
    c. **IF** a control is associated with this treated condition, select the control from the **Associated Control** dropdown menu.
        Note: Controls are *STRONGLY* recommended, but the algorithm does not require them. You can remove the control to save on reagents, but your results will not be as reliable without it.
9. Click **Next** to navigate to the **Enhanced Data Package Filtering** dialog.

10. In the **Enhanced Data Package Filtering** dialog, enter the EDP code you purchased in the **Enhanced data package code** filter.



11. **IF** you wish to filter the output based on the Z-scores, select the appropriate checkbox(es) and enter your desired criteria. This is entirely optional and is only needed if you want to remove extra records that you know are not of interest.
    a. The **Minimum Z-score for experiment conditions** option will, for each sample, filter out all compounds where the treated/experimental condition's Z-score is *lower than* the given value. Please note that this applies separately for each sample, so a compound might be filtered out of one sample but be retained in another, depending on its Z-scores in those samples.
    b. The **Maximum Z-score for control condition** option will filter out all compounds in *all* samples associated with a control where the Z-score was *greater than* the given value.
12. Click **Create & Run**. When the analysis is complete, it will produce a report if it completed successfully.

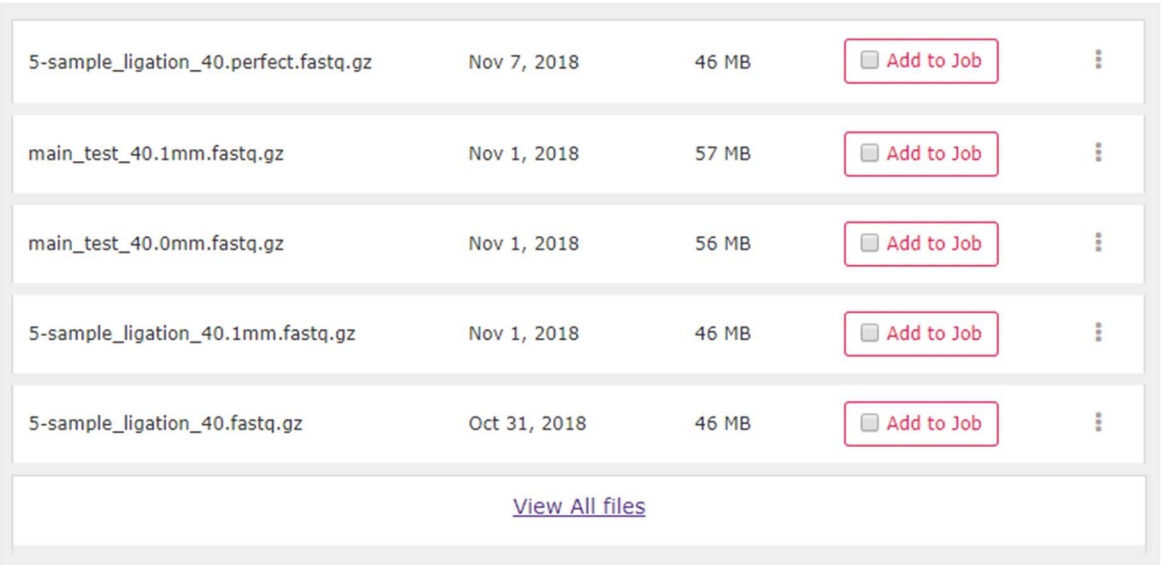
Viewing the Report

To view a report in the Analysis Portal:

- If the job you wish to view is in the **Recent Jobs** section, click the **View Report** button to show the report.
- Otherwise, click **View All Jobs** and scroll to the job of interest, then click **View Report**.

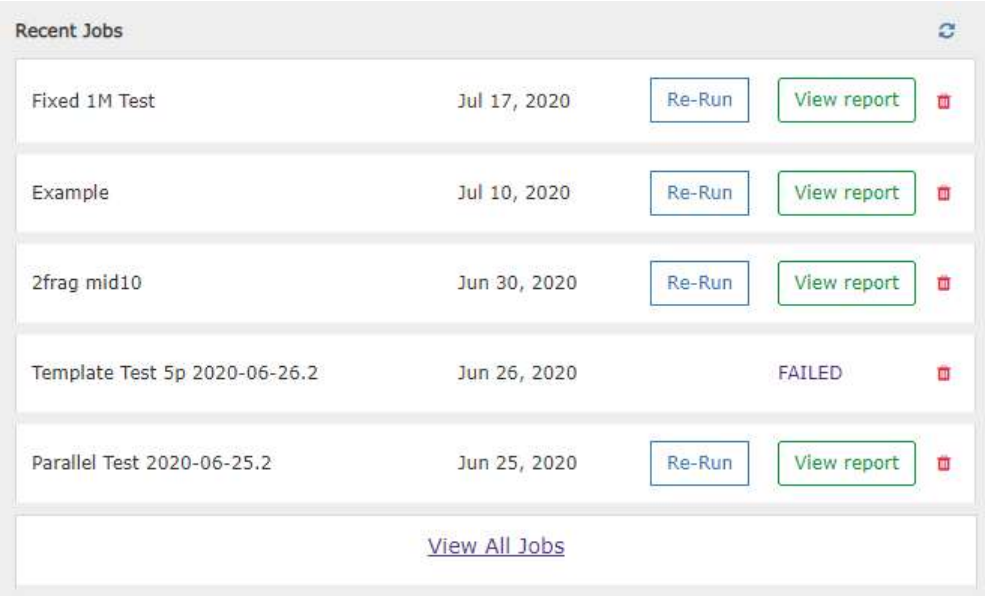
Running Files with Different Settings

If your files have already been uploaded, you can expedite the file selection step during job creation by using the **Add to Job** checkboxes for the desired files on the portal main page (see screenshot below).



Analysis Report

Once the analysis completes successfully, you can view the results by clicking **View Report** in the **Recent Jobs** section. Due to the size, the report itself must be downloaded—the report page itself only shows the top 50 hits for each sample. The output will be exported as a ZIP file, containing two files: `run_config.csv`, containing the run configuration, and `report.csv`, containing the Z-scores and SMILES strings, with the columns `Sample Name`, `Z-Score`, and `SMILES 10m`. ***NOTE:*** Due to the size of the library, the `report.csv` file can easily exceed the Microsoft Excel file limit of 1 048 576 rows. In such a case, you will have to use a different program to view it, appropriate to your own use case.



If the analysis fails, follow the steps in the **Troubleshooting Guide** section at the end of this protocol.

Z-score Calculations for Report

The analysis report will show the structures of the top compound hits, their SMILES strings, and the corresponding Z-scores. The Z-score for each compound is taken within the sample by subtracting the mean number of hits μ from the number of hits x for the compound and dividing the result by the standard deviation σ. In equation form, the Z-score Z is calculated as follows:

$$Z = \frac{x - \mu}{\sigma}$$

Only the top 50 results (up to 60 if ties would increase the number of top hits over 50) are included. The Z-scores are first filtered to remove any hits whose negative control (beads without target) Z-score is greater than the sample Z-score *or* whose negative control Z-score is ≥3.

Troubleshooting Guide

Observation	Possible Cause	Recommended Solution
File does not upload for several minutes	Network error	Refresh the page and try re-uploading the file
	Browser error	Refresh the page. If the file does not appear, or is not the right size, try re-uploading the file.
File upload is rejected	Not enough space remaining	Compress the FASTQ file with gzip. If the file upload is still rejected because there is not enough space, delete any unneeded FASTQ files from the account before trying again.
	Incorrect file extension	The file extension must be .fastq or .fastq.gz.
MID rejected by Experiment Details dialog	Incorrect MID	Ensure that the MID is a valid 12-bp DyNAbind MID, as listed in your kit's documentation.
	MID not unique	Each condition must have a unique MID. For example, you cannot use the same MID for two different experimental conditions or for an experimental condition and its control. If you have more conditions than unique MIDs (including any control conditions), you will need to run multiple assays and analyses.
Failed analysis	No MID matches for one or more samples	Ensure that the correct MIDs were entered.
		Ensure that the uploaded FASTQ file corresponds to the experimental setup.
	Invalid file format	Ensure that the contents of the uploaded sequencing file are in FASTQ format.

To Place an Order or Receive Technical Assistance

In the U.S. and Canada, call toll-free
1(800)-645-5476

For other countries across Europe, call
+44 (0) 115 943 0840

For other countries across Europe
and the world, please visit
SigmaAldrich.com/offices

For Technical Service, please visit
SigmaAldrich.com/techservice

MilliporeSigma
400 Summit Road
Burlington, MA 01803
SigmaAldrich.com