DeepNovo documentation

For the latest version, please refer to GitHub: https://github.com/nh2tran/DeepNovo-DIA

First, you need to run the installer downloaded from GitHub to install DeepNovo.

Please ensure that you have internet connection in order to activate DeepNovo.
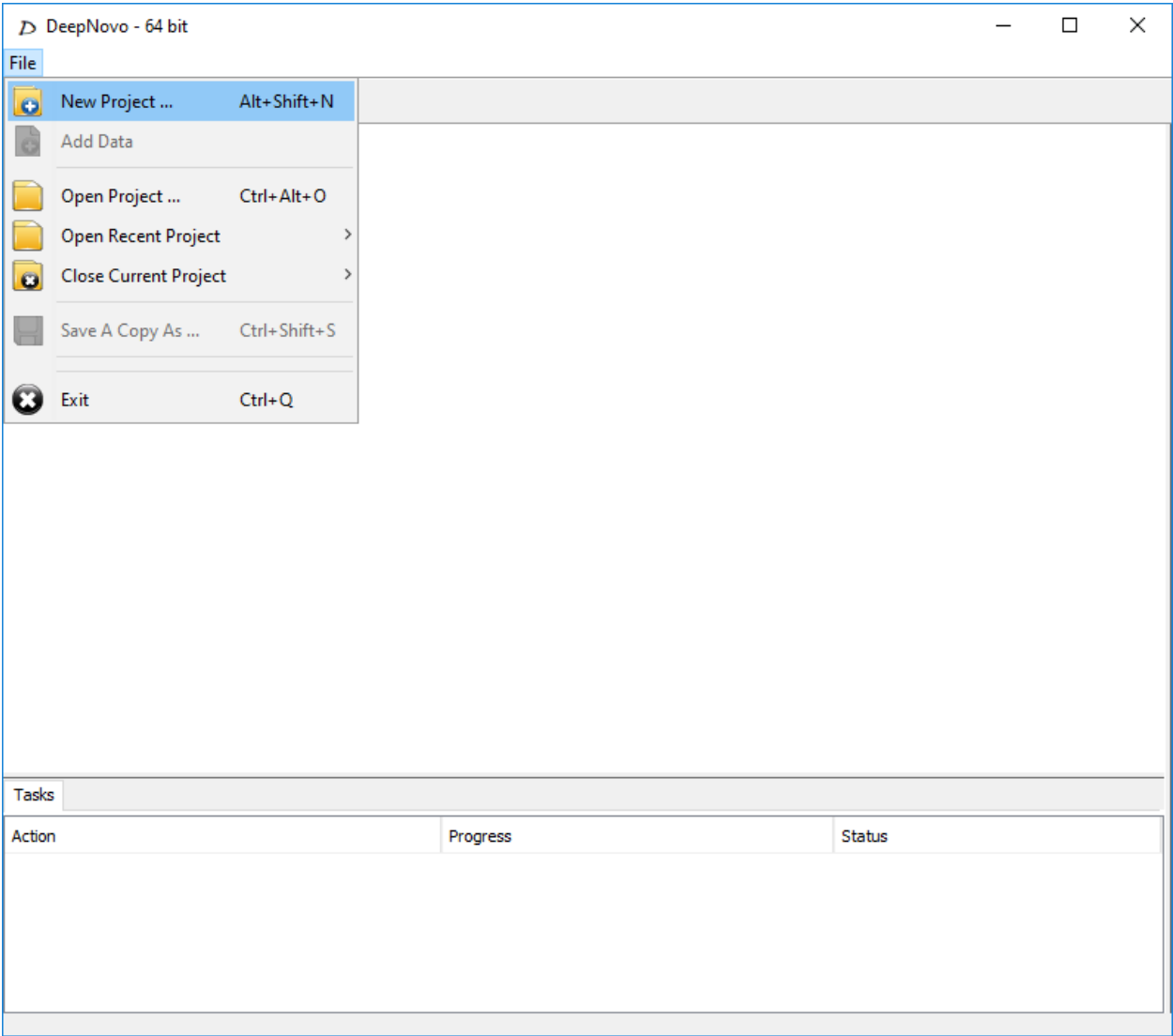
If you don't have Thermo MSFile Reader installed, please check the box below to install it.
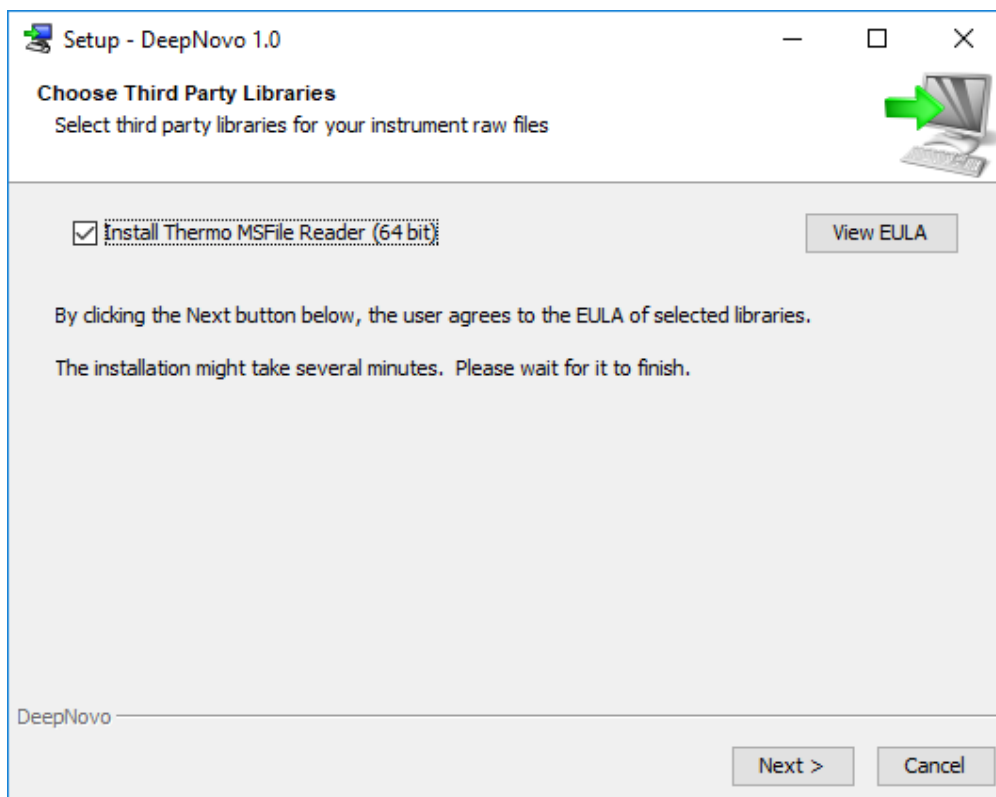
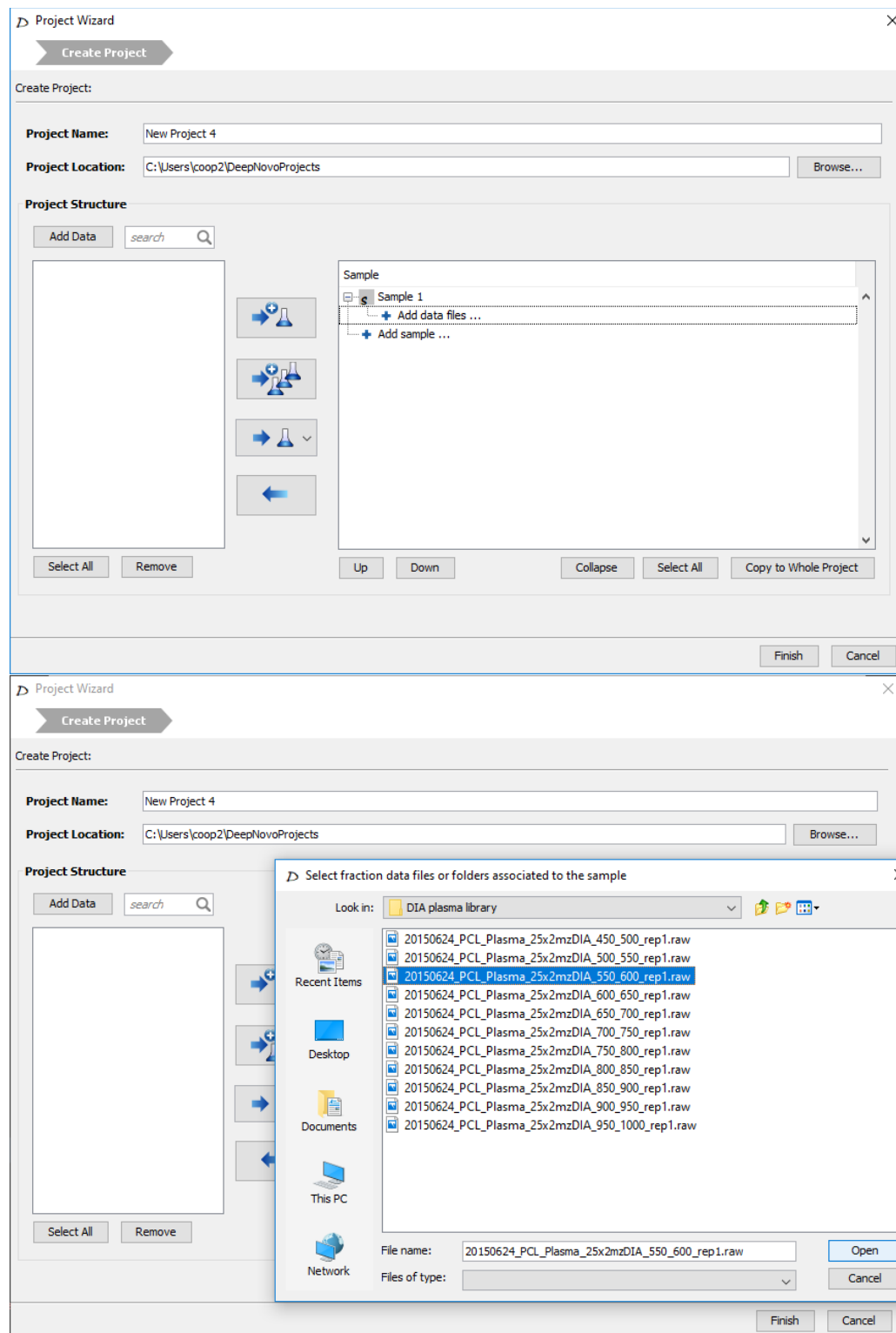Please ensure that you have internet connection in order to activate DeepNovo.
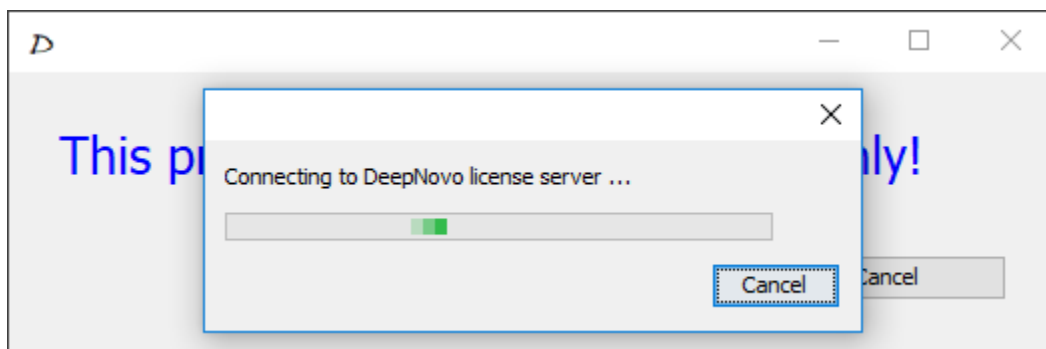
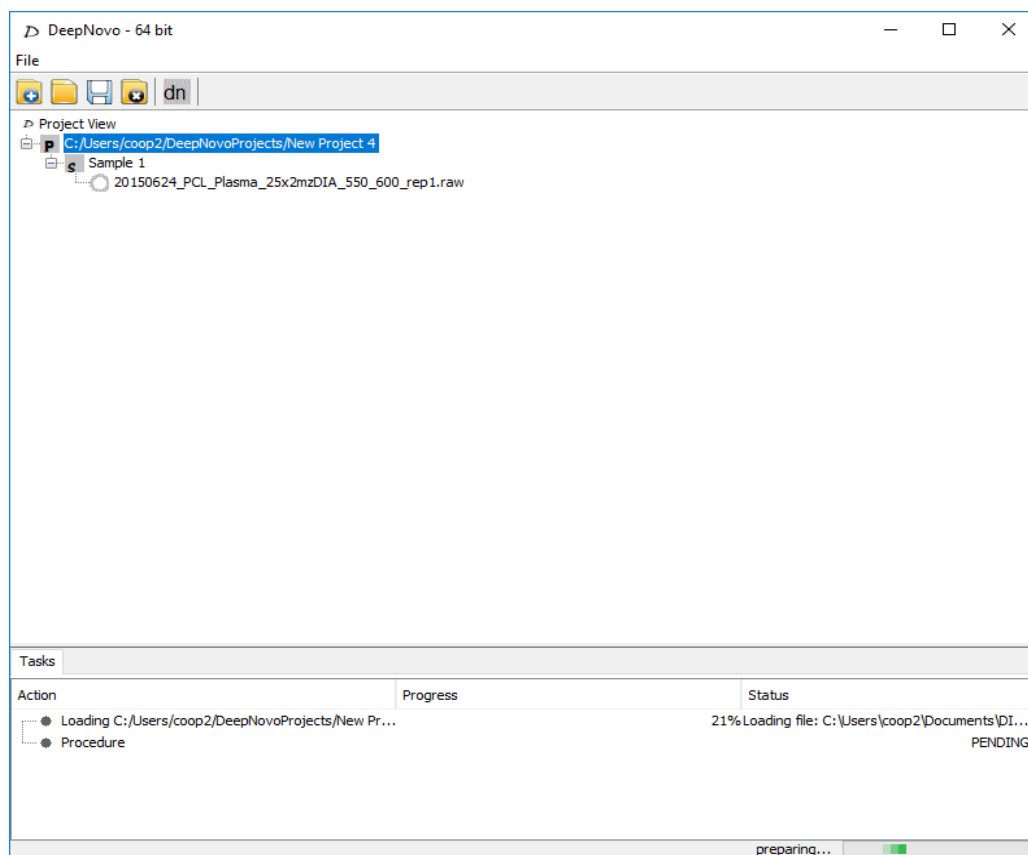To start a new project, click the "File" menu and select "New Project".
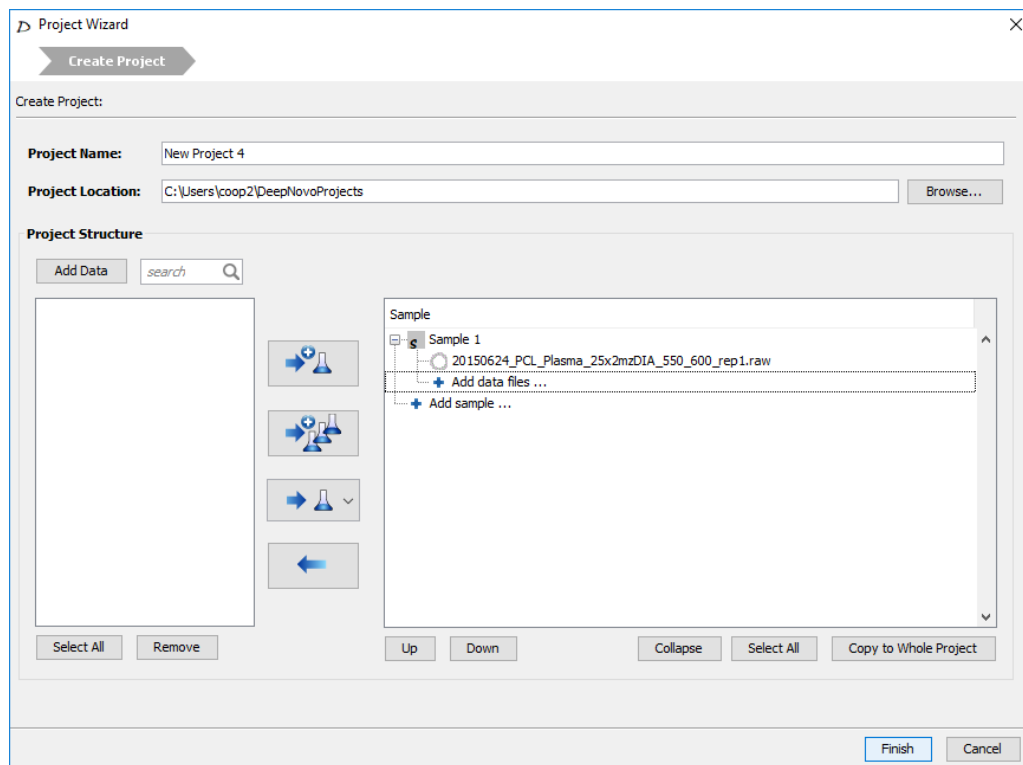
Next, enter the project name, select the desired location, and add one or multiple DIA .raw files.

We suggest to start with one small sample file to test the software first, before doing real analysis.

Then, click "Finish" and wait for the data pre-processing to complete.

After the pre-processing has finished (node "PREPROCESS 1" has appeared), select the project name and click the button "dn" and then "OK" to run *de novo* sequencing. This process may take long time, please be patient. We suggest to start with one small sample file to test the software first, before doing real analysis. The example below (one raw file from the plasma dataset) took ~45 minutes on a moderate laptop (CPU i5, RAM 4 Gb).
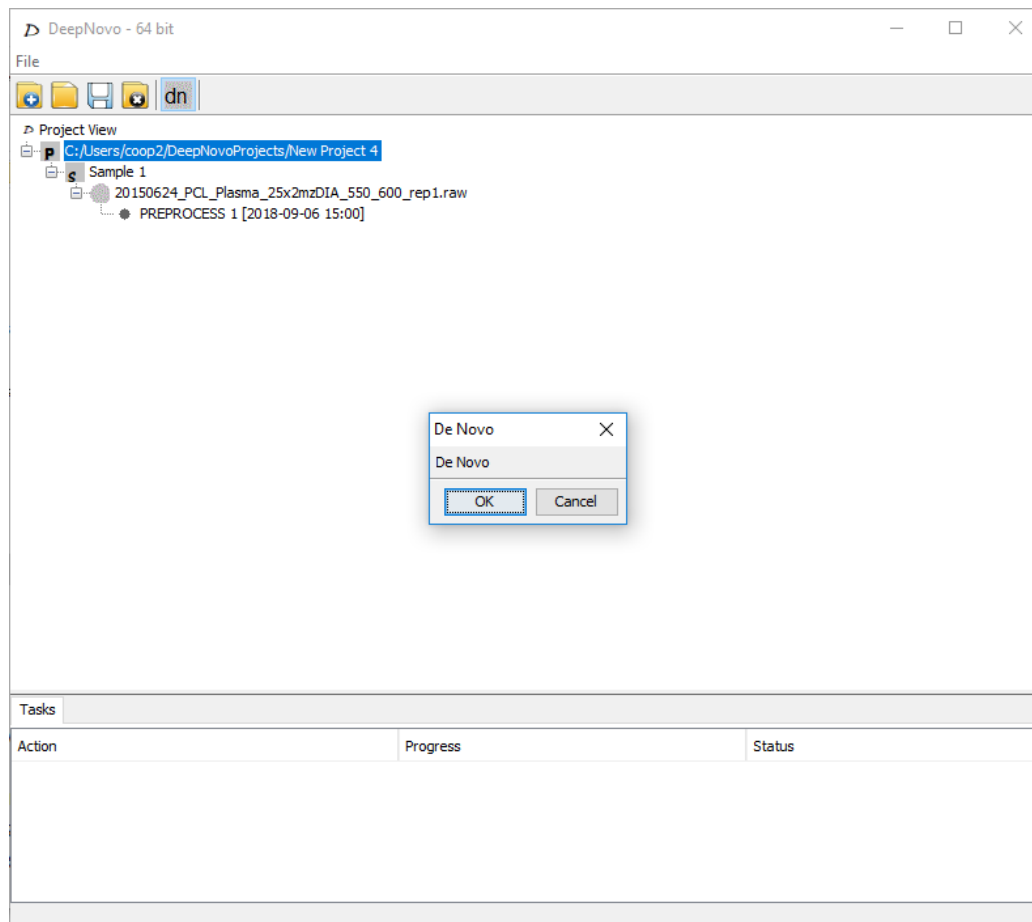
After the *de novo* sequencing has finished, right-click on the node "DENOVO 2" and check the project folder for the file with extension ".csv.deepnovo_denovo". This file includes all *de novo* results by DeepNovo. In addition, the two files with extensions ".csv" and ".mgf" contain the list of precursor features and DIA spectra, respectively. DeepNovo output includes the following columns:

- feature_id
- feature_area
- predicted_sequence
- predicted_score
- predicted_position_score: positional score for each amino acid in predicted_sequence
- precursor_mz
- precursor_charge
- protein_access_id: not relevant
- scan_list_original: list of scan ids of DIA spectra associated with this feature
- scan_list_middle: list of DIA spectra used for *de novo* sequencing
- predicted_score_max: same as predicted_score, not relevant