

# PNNL Biodiversity Library

## Tutorial for the PNNL Biodiversity Library Skyline Plugin

---

Developed By: Michael Degan, Grant Fujimoto, Lillian Ryadinskiy, Samuel H Payne

Contact Information: Samuel Payne ([Samuel.Payne@pnnl.gov](mailto:Samuel.Payne@pnnl.gov))

Michael Degan ([Michael.Degan@pnnl.gov](mailto:Michael.Degan@pnnl.gov))

Grant Fujimoto ([Grant.Fujimoto@pnnl.gov](mailto:Grant.Fujimoto@pnnl.gov))

This tutorial is to help users install and use the PNNL Biodiversity Library plugin with Skyline.

Troubleshooting help will follow at the end of the tutorial. We note that this plugin does not work with Windows XP and must be used with Skyline version 3.1.1.7490 or greater. For information about the Skyline program, including how to download, please visit

<https://skyline.gs.washington.edu/labkey/project/home/begin.view?>

Further information about the Biodiversity Library available at:

<http://omics.pnl.gov/project-data/biodiversity-library>

### Program Goal

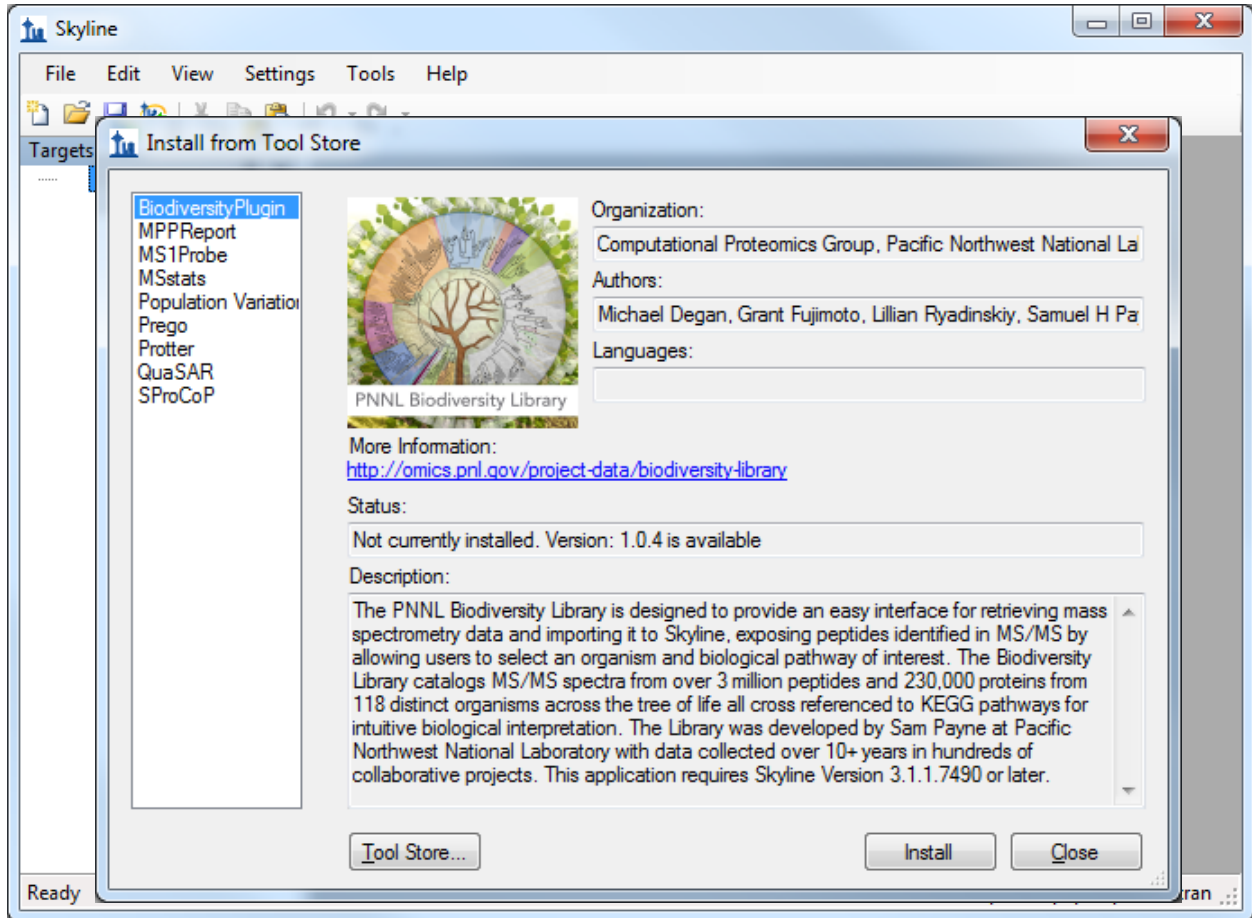
The goal of the plug-in is to help users browse the large spectral libraries from the PNNL Biodiversity Library and import relevant data into Skyline. Additional functionality allows users to import their own data into the plugin for easy navigation and loading into Skyline.

## Table of Contents

|  |           |
|--|-----------|
| <b>Program Goal</b> .....  | <b>1</b>  |
| <b>Installation</b> .....  | <b>3</b>  |
| <b>Running the Plugin</b> .....                                    | <b>4</b>  |
| <b>Select an Organism</b> .....                                    | <b>5</b>  |
| <b>Select Pathways of Interest</b> .....                           | <b>6</b>  |
| <b>Select proteins of interest through KEGG Pathway maps</b> ..... | <b>7</b>  |
| <b>Review and export</b> .....                                     | <b>8</b>  |
| <b>Export data to Skyline:</b> .....                               | <b>9</b>  |
| <b>View spectral data in Skyline</b> .....                         | <b>10</b> |
| <b>Customizing the plugin with your data</b> .....                 | <b>11</b> |
| Input Mass Spec Data .....   | 12        |
| Customizing Options .....  | 13        |
| Selecting an Organism .....  | 14        |
| Reviewing and Confirming .....                                     | 16        |
| <b>Settings in Skyline</b> .....                                   | <b>19</b> |
| <b>Troubleshooting</b> .....                                       | <b>20</b> |

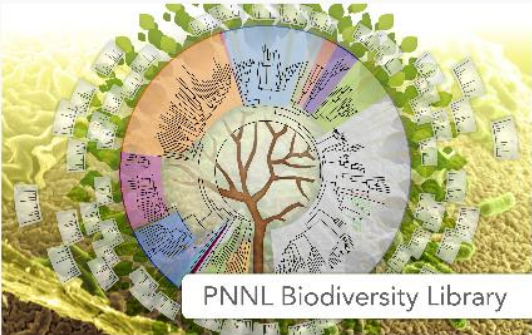
## Installation:

The PNNL Biodiversity Library Plugin for Skyline can be installed through the Skyline tool store interface. Tools -> Tool Store -> PNNL Biodiversity Library



## Running the Plugin:

PNNL Biodiversity Library



**Overview**

**Organism Chooser**

**Pathway Chooser**

**Pathway Visualization**

**Review and Export**

**Upload New Data**

The PNNL Biodiversity Library is designed to provide an easy interface for retrieving mass spectrometry data. This data can be exported into Skyline to assist in SRM assay design or DIA data analysis. The tool exposes peptides identified in MS/MS by allowing users to select an organism and biological pathway of interest. In total the Biodiversity Library catalogs MS/MS spectra from 3 million peptides and 230,000 proteins from 118 distinct organisms across the tree of life. All proteins are cross referenced to KEGG pathways for intuitive biological interpretation. The Library was developed by Sam Payne at [Pacific Northwest National Laboratory](#) with data collected over 10+ years in hundreds of collaborative projects.

The wizard helps users browse data using the following steps. Users can upload new data into the plugin by clicking "Upload New Data" on the left hand side at the bottom of the page.

- 1) Select an Organism.
- 2) Select pathways of interest.
- 3) View pathways and curate proteins.
- 4) Review and export data to Skyline.

PNNL Biodiversity Plugin Version: 1.0.4  
Database Source: [Biodiversity Library v2.1.230](#) Thursday, March 31, 2016  
Funding: US DOE, [Biological and Environmental Research](#)  
Contact: [Samuel Payne](#)

U.S. DEPARTMENT OF  
**ENERGY**

**Next >**

The plug-in is a wizard that walks users through the process of importing data into Skyline.

## Select an Organism:

You are able to select your desired organisms in one of two ways, either through the Taxonomy Explorer or by directly searching for the organism by name.

The Taxonomy Explorer organizes organisms through scientific naming convention, i.e. Kingdom, Phylum, Class and then individual strain of the organism. Human data is additionally broken down by tissue.

The Organism search (right panel) allows for users to type in a name and auto-filters the dataset to matching organisms.

The screenshot displays the PNNL Biodiversity Library interface. On the left is a vertical navigation menu with buttons for Overview, Organism Chooser, Pathway Chooser, Pathway Visualization, Review and Export, and Upload New Data. The main content area is titled "1) Select an Organism." and contains two panels: "Taxonomy Explorer" and "Organism search".

**Taxonomy Explorer:** A hierarchical tree view of organisms. The "Mycobacterium tuberculosis H37Rv" entry is highlighted with a red box. The tree structure is as follows:

- Archaea
- Bacteria
  - Acidobacteria
  - Actinobacteria
    - Actinosynnema
    - Arthrobacter
    - Brachybacterium
    - Cellulomonas
    - Cryptobacterium
    - Gardnerella
    - Kineococcus
    - Mycobacterium
      - Mycobacterium tuberculosis H37Rv**
    - Nakamurella
    - Nocardiopsis
    - Pseudonocardia
    - Saccharomonospora
    - Sanguibacter
    - Slackia
    - Stackebrandtia
    - Thermobispora
    - Xylanimonas
  - Alphaproteobacteria
  - Bacteroidetes
  - Betaproteobacteria
  - Chlorobi

**Organism search:** A list of organisms matching the search criteria. The list includes:

- Acidiphilium cryptum
- Actinosynnema mirum
- Anabaena variabilis
- Anaeromyxobacter dehalogenans 2CP-C
- Anaplasma phagocytophilum HZ
- Arthrobacter sp. FB24
- Bacillus anthracis Ames
- Bacillus anthracis Sterne
- Bacillus subtilis subsp. subtilis 168
- Bacteroides fragilis 638R
- Bacteroides thetaiotaomicron
- Bartonella henselae Houston-1
- Borrelia burgdorferi B31
- Brachybacterium faecium
- Burkholderia mallei ATCC 23344
- Candidatus Pelagibacter ubique
- Caulobacter crescentus CB15
- Cellulomonas flavigena
- Cenarchaeum symbiosum
- Chloracidobacterium thermophilum
- Chlorobium tepidum
- Chloroflexus aurantiacus
- Cryptobacterium curtum
- Cyanothece sp. ATCC 51142
- Cyanothece sp. PCC 7424
- Cyanothece sp. PCC 7425
- Cyanothece sp. PCC 7822
- Cyanothece sp. PCC 8801
- Dehalococcoides mccartyi 195

Navigation buttons for "< Previous" and "Next >" are located at the bottom of the search panel.

## Select Pathways of Interest:

After you have selected an organism, you then choose the KEGG pathways you are interested in investigating. The coverage metric is dynamically calculated for each Organism/Pathway combination and shows the percentage of proteins in a pathway for which the Library contains proteomic data. You can select as many pathways as you desire.

**2) Select pathways of interest.**  
Pathway definitions and membership are defined by [KEGG](#). Navigate to pathways and click the checkbox. Multiple pathways can be selected. If you decide to use your own data, you can do so by clicking "Upload New Data" on the bottom left.

### Kegg Pathways

**Metabolism**

- Carbohydrate metabolism
  - Glycolysis / Gluconeogenesis 66.67% covered in MSMS
  - Citrate cycle (TCA cycle) 75.00% covered in MSMS
  - Pentose phosphate pathway 55.56% covered in MSMS
  - Fructose and mannose metabolism 60.00% covered in MSMS**
  - Galactose metabolism 75.00% covered in MSMS
  - Ascorbate and aldarate metabolism 75.00% covered in MSMS
  - Starch and sucrose metabolism 68.00% covered in MSMS
  - Amino sugar and nucleotide sugar metabolism 65.38% covered in MSMS
  - Pyruvate metabolism 53.85% covered in MSMS
  - Glyoxylate and dicarboxylate metabolism 68.75% covered in MSMS
  - Propanoate metabolism 56.25% covered in MSMS
  - Butanoate metabolism 85.71% covered in MSMS
  - C5-Branched dibasic acid metabolism 60.00% covered in MSMS
  - Inositol phosphate metabolism 42.86% covered in MSMS
- Energy metabolism
- Lipid metabolism
- Nucleotide metabolism
- Amino acid metabolism
- Metabolism of other amino acids
- Glycan biosynthesis and metabolism
- Metabolism of cofactors and vitamins
- Metabolism of terpenoids and polyketides
- Biosynthesis of other secondary metabolites

### Pathways Selected

Fructose and mannose metabolism

< Previous      Next >

## Select proteins of interest through KEGG Pathway maps:

The images for the pathways you selected will be created dynamically, highlighting all the KEGG orthologs for the organism in the pathway. If that protein has been observed in the spectral library, it will be highlighted initially as red; unobserved proteins are highlighted in blue. From here, you can select the proteins which are of interest to you in the pathway in two ways, either by clicking on individual boxes or by clicking the select/deselect all buttons on the right hand side of the screen. If an ortholog does become deselected it will then be highlighted as grey to show that the information will not be pushed to the next step of the process.

Additionally, you can save the pathway image by clicking the button on the right-hand side of the screen, below the Deselect All button. This will save the image for the pathway along with your current selection of orthologs as a .png file with a resolution of 300 dpi.

**3) Curate protein list for Organism: *Mycobacterium tuberculosis* H37Rv**

Pathways specified in Step 2 are overlaid with proteomics data for the organism selected in step 1. Proteins identified in the Biodiversity Library are shown in red. Proteins which are annotated in the genome but are not present in the MS data are shown in blue. Users can deselect any protein they wish to exclude by clicking on the corresponding red box on the pathway image. (Deselected proteins appear as grey boxes.) If you wish to use your own data instead for this organism, click "Upload New Data" on the bottom left.

**Fructose and mannose metabolism**

**Legend:**

- Red box: Protein annotated in *Mycobacterium tuberculosis* H37Rv and observed in MS/MS data
- Blue box: Protein annotated in *Mycobacterium tuberculosis* H37Rv and not observed in MS/MS data

**FRUCTOSE AND MANNOSE METABOLISM**

The pathway diagram shows the conversion of L-Sorbose to D-Fructose and D-Mannose, and their subsequent metabolism. Key enzymes and intermediates are labeled with EC numbers and names. Proteins are color-coded according to the legend: red for observed in MS/MS data and blue for not observed. For example, enzyme 5.4.2.8 (D-Mannose-1-phosphatase) is highlighted in red, while enzyme 2.4.1.1 (Mannan 6-phosphatase) is highlighted in blue.

**All Proteins Selected**

- Select All
- Deselect All
- Save Pathway Image

**Navigation:**

- < Previous
- Next >

## Review and export:

On the Review and Export tab, the application will display the total number of genes selected from all of your Organism/Pathway combinations as well as an individual breakdown of the proteins. If there are Organism/Pathway combinations you are no longer interested in, or individual proteins which you do not wish to export to Skyline, you can uncheck the box for the row(s).

**4) Review and export data to Skyline.**  
From this final tab, a table of genes for each organism and pathway are shown for review. By hitting the export button, all relevant data will be transferred to Skyline. This includes a FASTA file for all selected proteins, as well as the identified peptides in a bibliospec library.

| Organism                        | Pathway                         | # of Genes | Export?                             |
|---------------------------------|---------------------------------|------------|-------------------------------------|
| Mycobacterium tuberculosis H37R | Fructose and mannose metabolism | 10         | <input checked="" type="checkbox"/> |

**Organism: Mycobacterium tuberculosis H37Rv**  
**Pathway: Fructose and mannose metabolism**  
**Genes (10)**

| Export?                             | Accession | Protein name   |
|-------------------------------------|-----------|--|
| <input checked="" type="checkbox"/> | O53634    | gca: GDP-mannose 4,6-dehydratase                             |
| <input checked="" type="checkbox"/> | P9WN21    | glpX: fructose 1,6-bisphosphatase                            |
| <input checked="" type="checkbox"/> | P9WG43    | tpi: triosephosphate isomerase                               |
| <input checked="" type="checkbox"/> | P71790    | gmdA: GDP-D-mannose dehydratase GmdA                         |
| <input checked="" type="checkbox"/> | P71791    | epiA: nucleotide-sugar epimerase EpiA                        |
| <input checked="" type="checkbox"/> | P9WKD7    | rpiB: ribose-5-phosphate isomerase B                         |
| <input checked="" type="checkbox"/> | O05898    | manA: mannose-6-phosphate isomerase                          |
| <input checked="" type="checkbox"/> | O86374    | pmmA: phosphomannomutase PmmA                                |
| <input checked="" type="checkbox"/> | L7N6A5    | manB: D-alpha-D-mannose-1-phosphate guanylyltransferase ManB |
| <input checked="" type="checkbox"/> | O53360    | pmmB: phosphomannomutase PmmB                                |



## Export data to Skyline:

Once you are satisfied with your selection, you can then either select another organism and repeat the wizard steps, or export the data to Skyline.

Once you confirm the export to skyline, the application will create a FASTA for the proteins you indicated and download the Spectral Library files for the organisms of interest from the Biodiversity Library's Repository on MassIVE (<ftp://massive.ucsd.edu/MSV000079053/library/>). Both of these files (FASTA and Bibliospec library) are then imported to Skyline.

The screenshot shows a web application window titled "PNNL Biodiversity Library". On the left is a vertical navigation menu with buttons for "Overview", "Organism Chooser", "Pathway Chooser", "Pathway Visualization", "Review and Export" (which is highlighted in dark blue), and "Upload New Data".

The main content area is titled "4) Review and export data to Skyline." and contains the following text: "From this final tab, a table of genes for each organism and pathway are shown for review. By hitting the export button, all relevant data will be transferred to Skyline. This includes a FASTA file for all selected proteins, as well as the identified peptides in a bibliospec library."

Below this text is a summary table:

| Organism                         | Pathway                         | # of Genes | Export?                             |
|----------------------------------|---------------------------------|------------|-------------------------------------|
| Mycobacterium tuberculosis H37Rv | Fructose and mannose metabolism | 10         | <input checked="" type="checkbox"/> |

Below the summary table, the selected organism and pathway are listed: "Organism: Mycobacterium tuberculosis H37Rv" and "Pathway: Fructose and mannose metabolism".

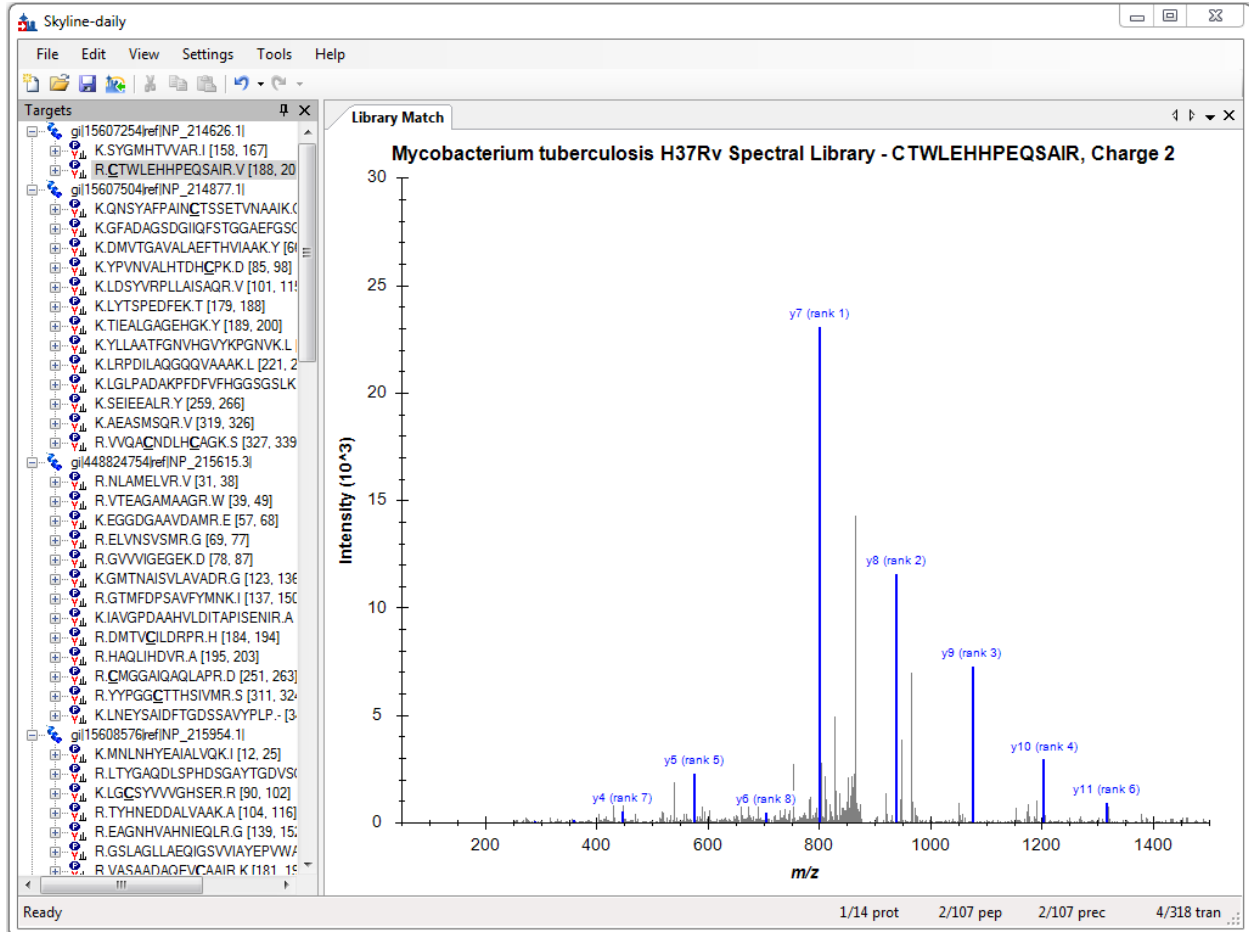
Underneath, it says "Genes (10)" and displays a table of 10 genes:

| Export?                             | Accession | Protein name   |
|-------------------------------------|-----------|--|
| <input checked="" type="checkbox"/> | O53634    | gca; GDP-mannose 4,6-dehydratase                             |
| <input checked="" type="checkbox"/> | P9WN21    | glpX; fructose 1,6-bisphosphatase                            |
| <input checked="" type="checkbox"/> | P9WG43    | tpi; triosephosphate isomerase                               |
| <input checked="" type="checkbox"/> | P71790    | gmdA; GDP-D-mannose dehydratase GmdA                         |
| <input checked="" type="checkbox"/> | P71791    | epiA; nucleotide-sugar epimerase EpiA                        |
| <input checked="" type="checkbox"/> | P9WKD7    | rpiB; ribose-5-phosphate isomerase B                         |
| <input checked="" type="checkbox"/> | O05898    | manA; mannose-6-phosphate isomerase                          |
| <input checked="" type="checkbox"/> | O86374    | pmmA; phosphomannomutase PmmA                                |
| <input checked="" type="checkbox"/> | L7N6A5    | manB; D-alpha-D-mannose-1-phosphate guanylyltransferase ManB |
| <input checked="" type="checkbox"/> | O53360    | pmmB; phosphomannomutase PmmB                                |

At the bottom of the interface are three buttons: "Select another Organism" (highlighted with a red box), "< Previous", and "Confirm to Export" (highlighted with a red box).

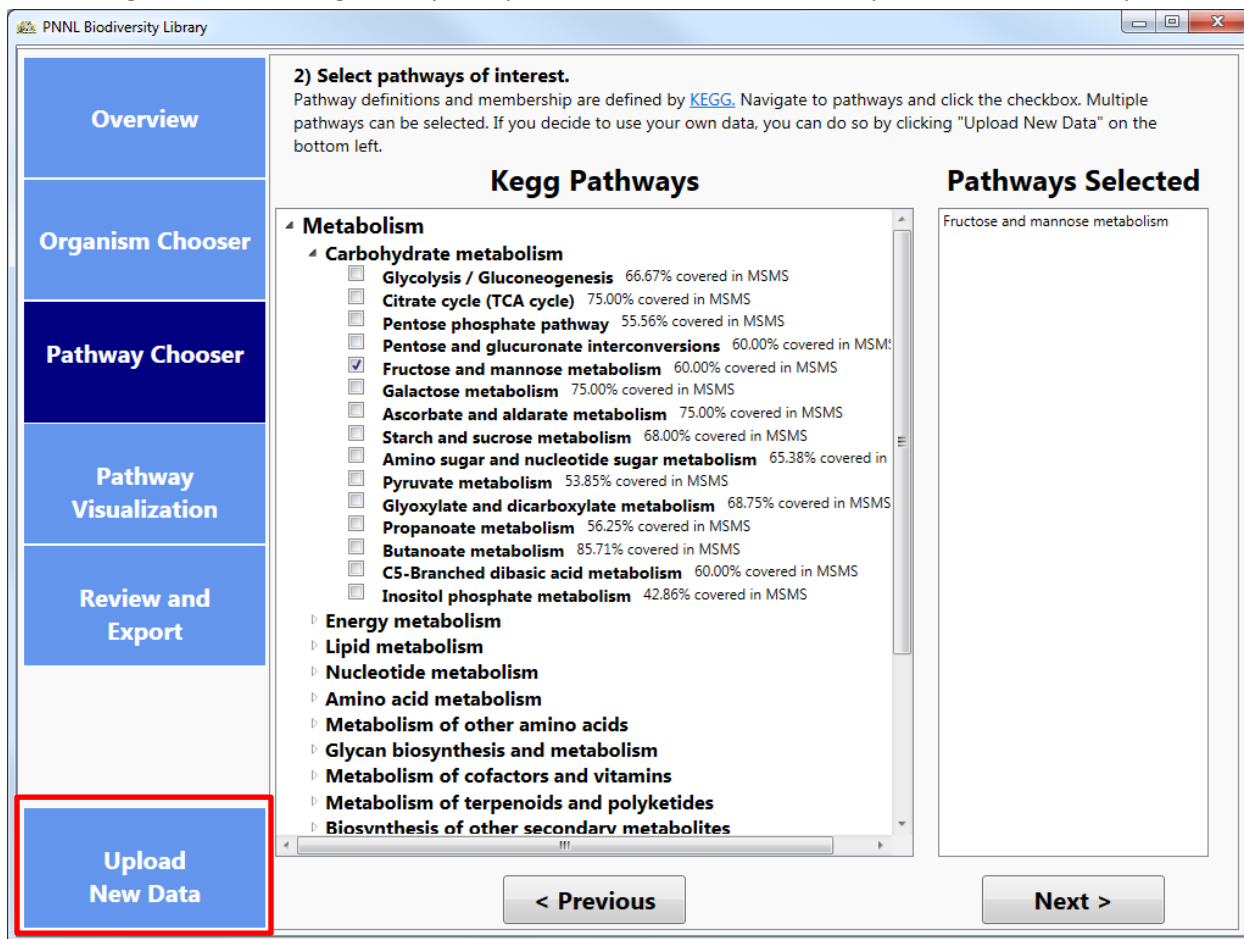
## View spectral data in Skyline:

The plugin will automatically propagate the information downloaded into Skyline, where it uses your peptide settings to create peptides for proteins in the FASTA and uses the Spectral Library downloaded through the plugin to perform library matching.

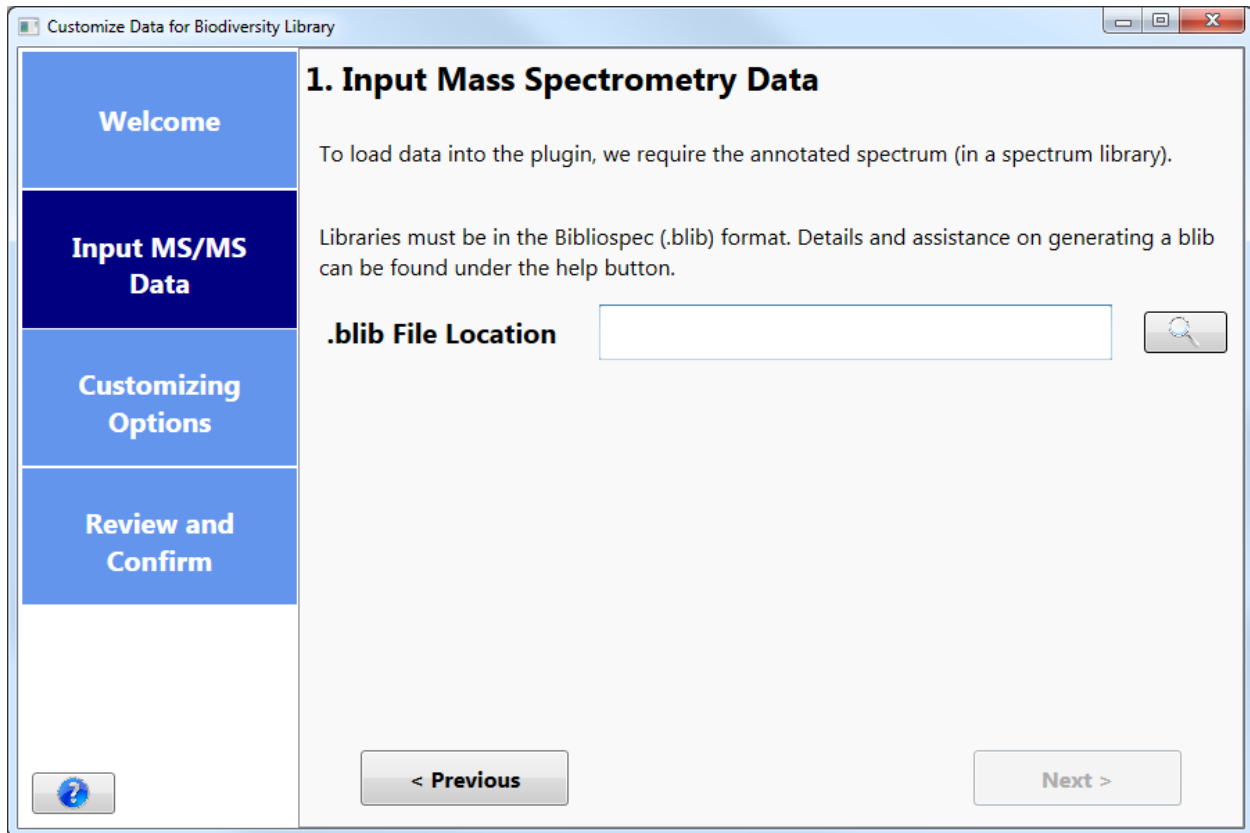


## Customizing the plugin with your data:

The Biodiversity Plugin enables users to update the MS/MS data to include personal data and use the same pathway oriented interface to load their data directly into Skyline. Users start customizing their data by selecting the **Upload New Data** button at the bottom left of the window. This starts a new wizard to guide users through the upload process. Users can choose to upload new data at any time.



## Input Mass Spec Data:



Only one file is required, a [Bibliospec formatted](#) spectral library. Please consult the [online documentation](#) for details about building a Bibliospec library. To build a library from mzIdentML results, and example command line is:

```
~> BlibBuild.exe -c 0.9999 E:\path\to\PSM_results.mzid E:\path\to\Library.blib
```

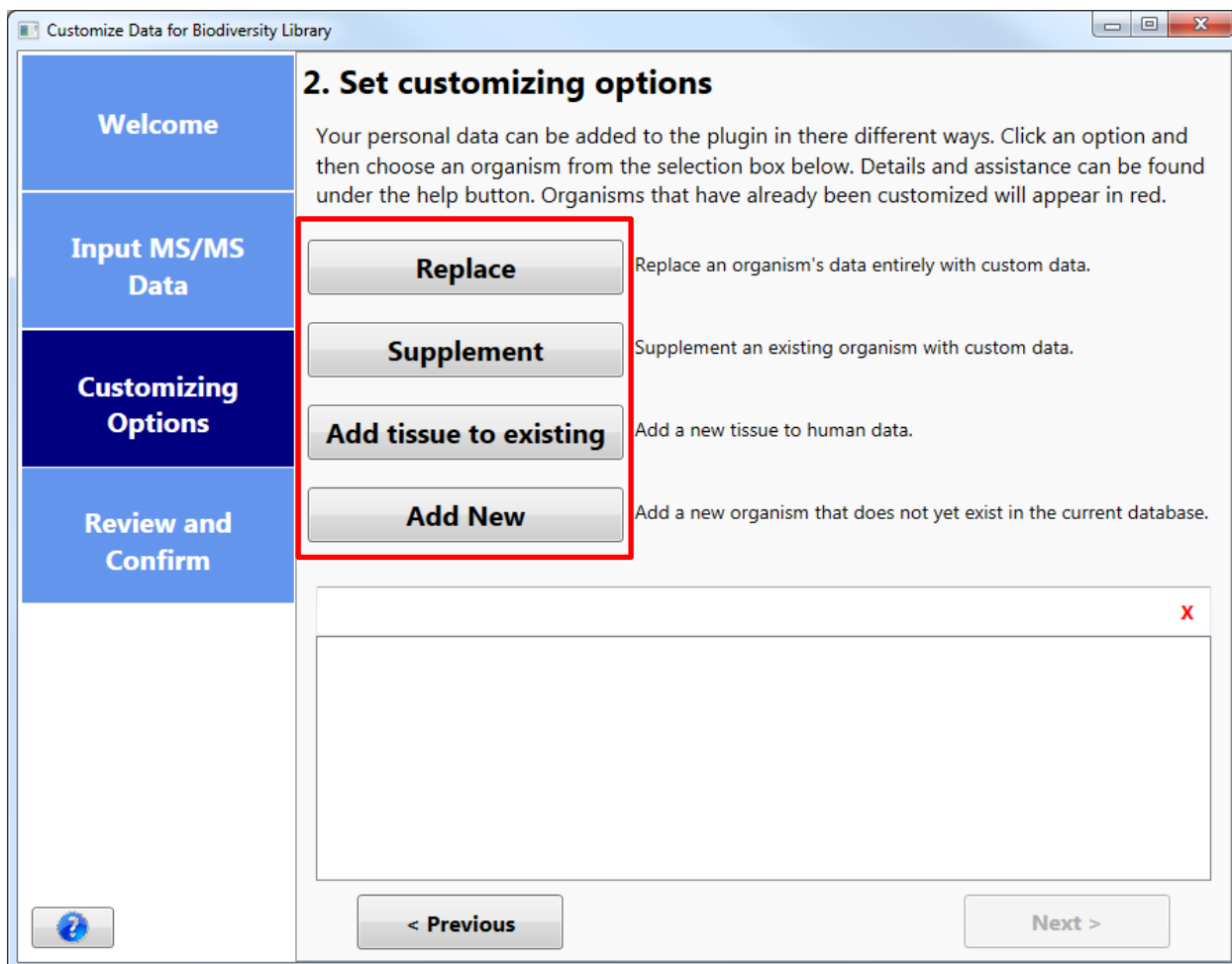
It is then necessary to filter the Blib file to contain ONE and only one PSM per peptide. Therefore, you must filter the above library file, for example:

```
~> BlibFilter.exe -b 1 E:\path\to\Library.blib E:\path\to\FilteredLibrary.blib
```

## Customizing Options:

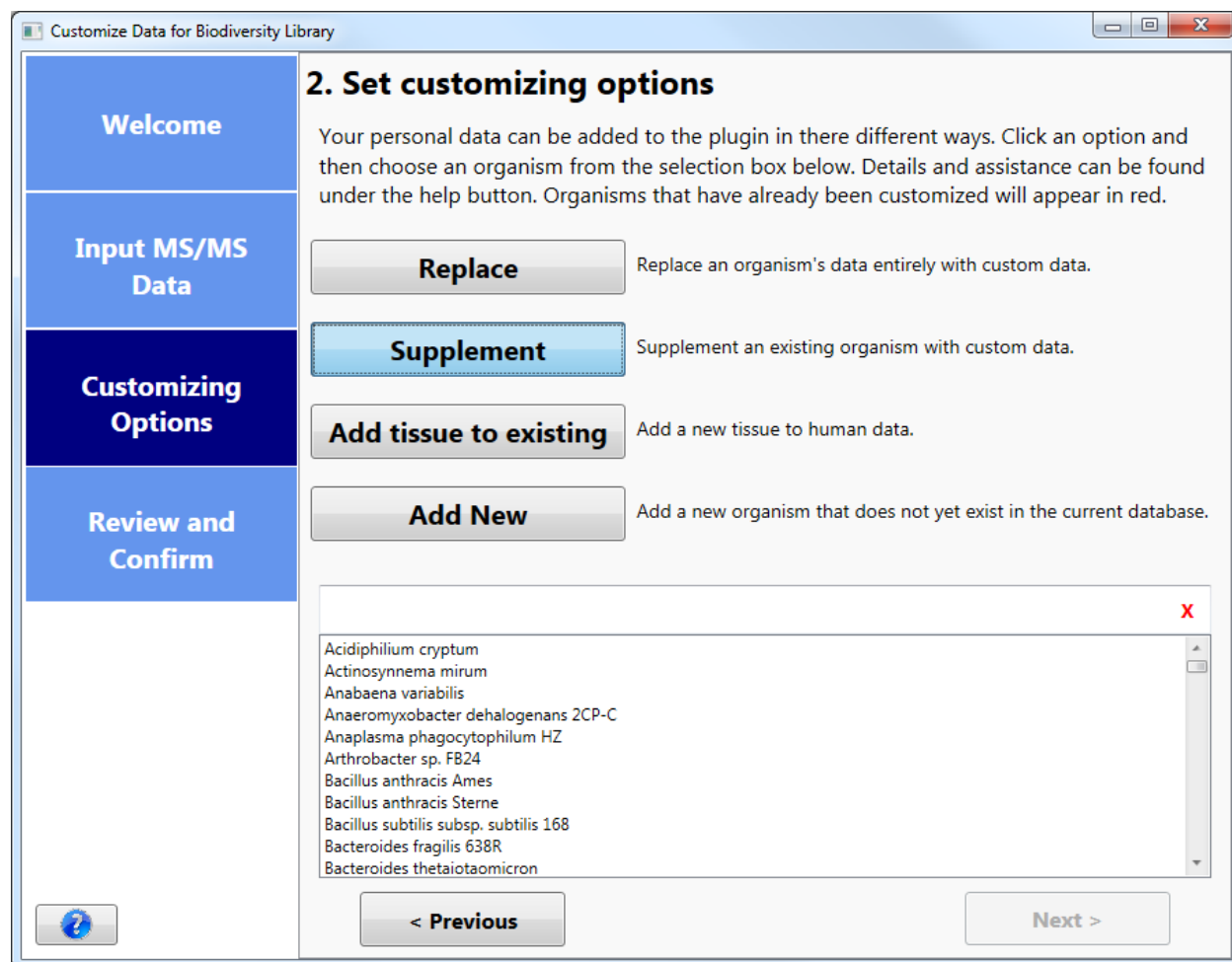
There are four options for updating organism data:

- Replace the existing organism entirely with custom data.
- Supplement an existing organism with additional data.
- Add a new tissue to an organism that is already in the database.
- Add a new organism that does not yet exist in the current database.



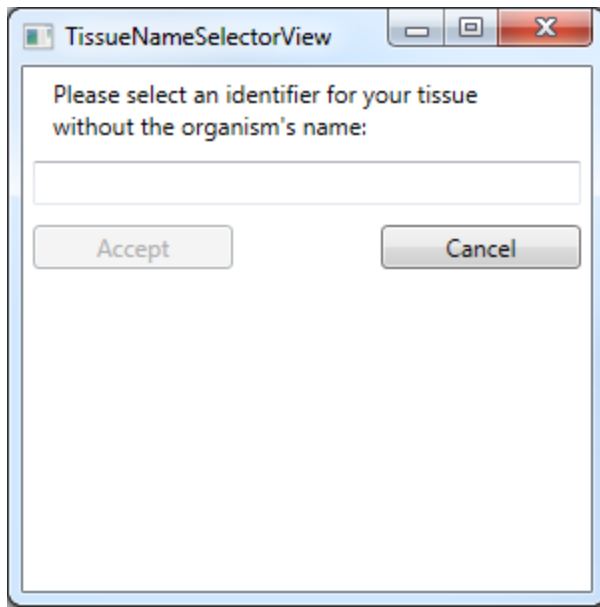
## Selecting an Organism:

Once the customizing option has been selected, the list box at the bottom of the window will populate with organisms to search and choose from.



Once you have selected an organism, the next button will enable. Once you click the **Next** button, the wizard will begin collecting the results. Users will be navigated to the Review and Confirm tab where they will wait for their results to appear.

Users also have the option of adding a new tissue to the human data. To add new tissue, click the **Add tissue to existing** button and then select the organism that you would like to add new tissue to. Currently, the plugin has the capability of adding tissue only to Home Sapiens. After you have selected the organism and click **Next**, the following window will appear.

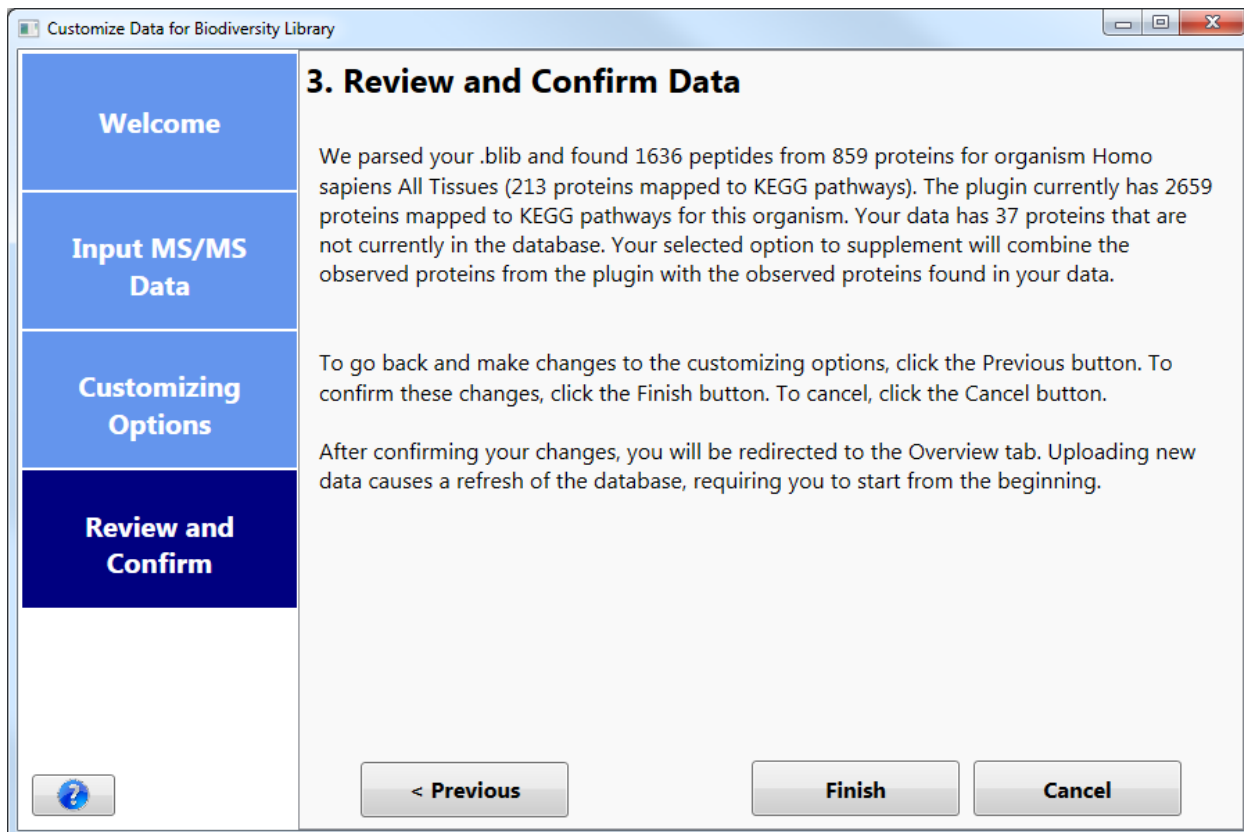


Here, users will be prompted to enter a name for the tissue. The named you enter will automatically be appended to the name of the organism you are adding tissue for. For example, if you entered "Hippocampus" as the tissue name, the full name will be Home Sapiens Hippocampus.

Once you click **Accept**, the wizard will begin collecting the results. Users will be navigated to the Review and Confirm tab where they will wait for their results to appear.

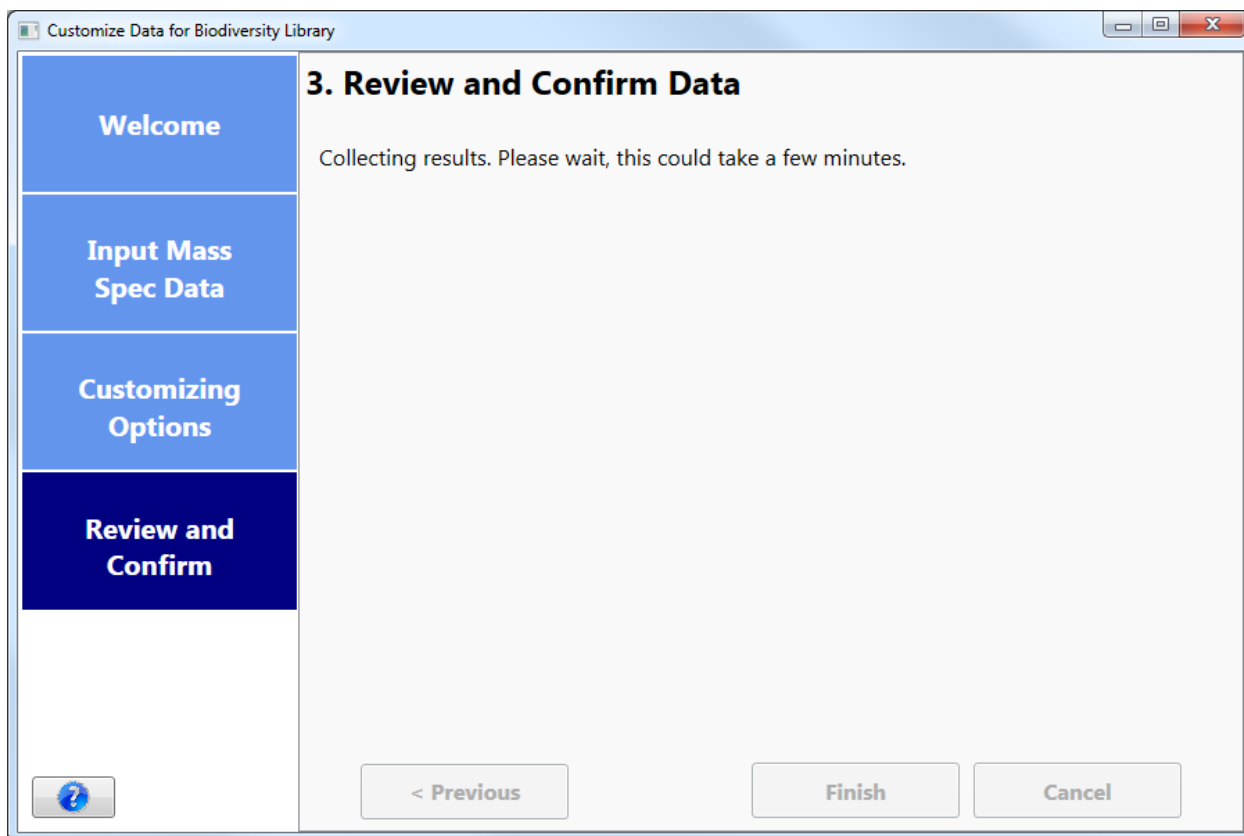
## Reviewing and Confirming:

Once the results have been collected, a message will appear prompting you to confirm the results and finish the process by clicking the **Finish** button, cancel the process by clicking the **Cancel** button, or go back and modify your customizing options by clicking the **Previous** button.





**Please note:** Adding a new organism can take longer than Replacing or Supplementing. If the process takes longer than expected, a message (shown below) will appear, allowing you to wait there until the results are collected.



For your convenience, any organism that was customized by you will appear in red. Unmodified organisms will remain black.

**1) Select an Organism.**  
Data in the Biodiversity Library is organized by the organism. To retrieve data, please select an organism of interest, either through the Phylogeny explorer or through the Organism search box. If you want to analyze data from multiple organisms, please export the data one organism at a time. If you would like to add or customize an organism, click "Upload New Data" on the left hand side at the bottom of the page. Organisms that have been customized are shown in red.

### Phylogeny Explorer

- ▲ Eukaryota
  - ▲ Animalia
    - ▲ Homo sapiens
      - Homo sapiens Adult Adreanal Gland
      - Homo sapiens Adult B cell
      - Homo sapiens Adult CD4 T Cell
      - Homo sapiens Adult CD8 T Cell
      - Homo sapiens Adult Colon
      - Homo sapiens Adult Esophagus
      - Homo sapiens Adult Frontal Cortex
      - Homo sapiens Adult Gallbladder
      - Homo sapiens Adult Heart
      - Homo sapiens Adult Kidney
      - Homo sapiens Adult Liver
      - Homo sapiens Adult Lung
      - Homo sapiens Adult Monocytes
      - Homo sapiens Adult NK Cells
      - Homo sapiens Adult Ovary
      - Homo sapiens Adult Pancreas
      - Homo sapiens Adult Platelets
      - Homo sapiens Adult Prostate
      - Homo sapiens Adult Rectum
      - Homo sapiens Adult Retina
      - Homo sapiens Adult Spinal Cord
      - Homo sapiens Adult testis
      - Homo sapiens Adult Urinary Bladder
      - Homo sapiens All Tissues**
      - Homo sapiens Fetal Gut
      - Homo sapiens Fetal Heart
      - Homo sapiens Fetal Liver

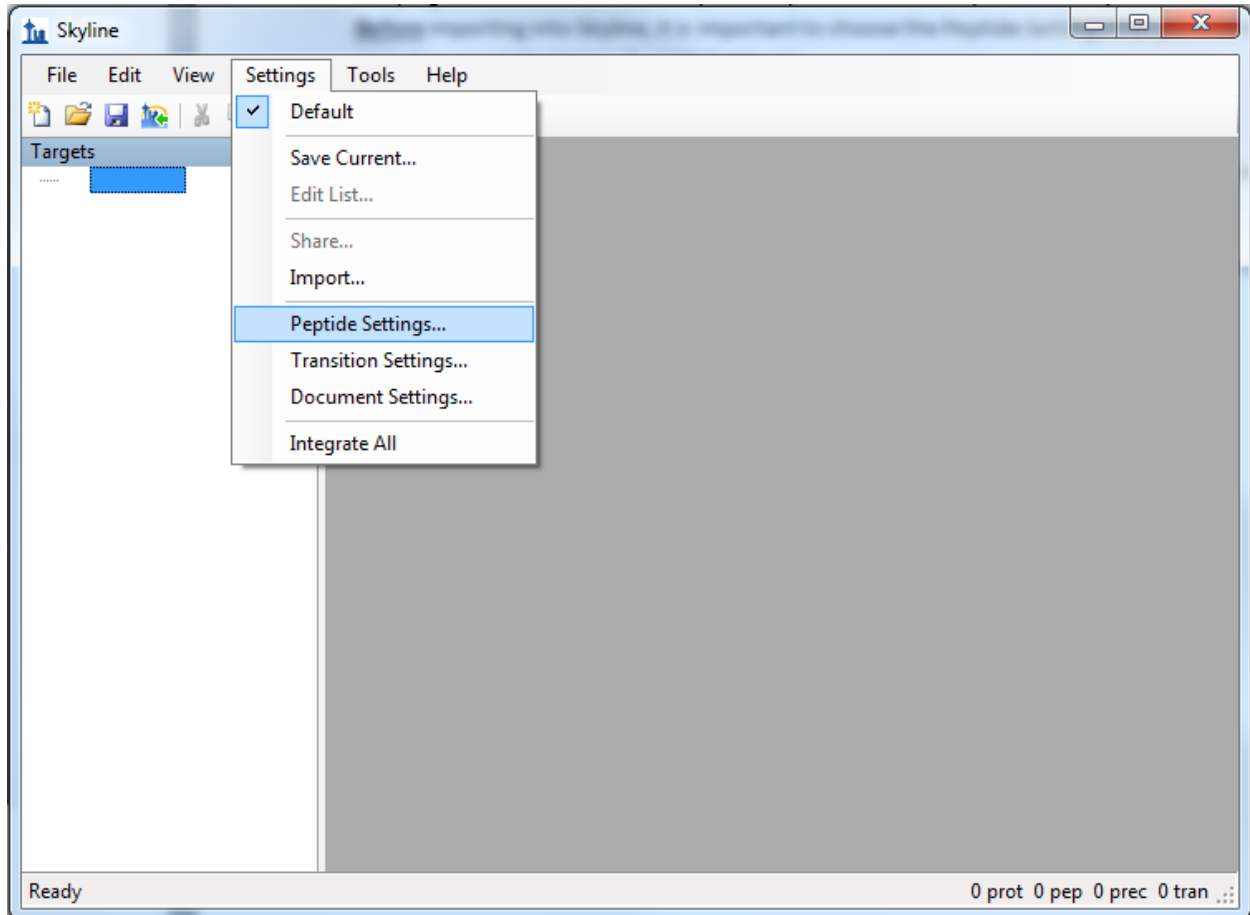
### Organism search

- Homo sapiens Adult Ovary
- Homo sapiens Adult Pancreas
- Homo sapiens Adult Platelets
- Homo sapiens Adult Prostate
- Homo sapiens Adult Rectum
- Homo sapiens Adult Retina
- Homo sapiens Adult Spinal Cord
- Homo sapiens Adult testis
- Homo sapiens Adult Urinary Bladder
- Homo sapiens All Tissues**
- Homo sapiens Fetal brain
- Homo sapiens Fetal Gut
- Homo sapiens Fetal Heart
- Homo sapiens Fetal Liver
- Homo sapiens Fetal Ovary
- Homo sapiens Fetal Placenta
- Homo sapiens Fetal Testis
- Kineococcus radiotolerans
- Kosmotoga olearia
- Lactobacillus crispatus
- Lactobacillus gasseri ATCC 33323
- Methanosarcina barkeri Fusaro
- Methanospirillum hungatei
- Mycobacterium tuberculosis H37Rv
- Nakamurella multipartita
- Nocardiopsis dassonvillei
- Novosphingobium aromaticivorans
- Pelobacter carbinolicus
- Prevotella ruminicola

< Previous      Next >

## Settings in Skyline

The plug-in creates a FASTA of the proteins you choose and imports that sequence file into Skyline. **Before** importing into Skyline, it is important to choose the Peptide Settings that you wish to use for your document (Settings > Peptide Settings).



After the FASTA has been imported, users will not be able to make changes in Digestion, Filter or Modifications tabs.

Additionally, it is recommended to set Transition Settings before importing the FASTA. Be sure to select precursor charge states (in the Filter Tab).

## Troubleshooting

- The application is taking an unusually long time while generating a FASTA from my selection of proteins.
  - Depending on your Skyline settings, Skyline could be seeing that there are proteins incoming from this FASTA which do not contain peptides according to the current filter settings. While this window is open from Skyline, our application is waiting for a response. To allow the information to be fully exported, select “Keep” so that the proteins will export and the .blib file(s) which download afterwards will provide a library match and populate the peptides for these new proteins.
- No proteins appeared as red in the KEGG pathway maps.
  - This could be due to two possibilities:
    - The organism contains proteins which should be observed in MS/MS space on this pathway, but were not observed in experimental results. These ortholog boxes will still be highlighted as blue.
    - The organism does not contain proteins which should be observed in MS/MS space on this pathway. In this case, no ortholog boxes will be highlighted on the image at all.
- Spectral Library doesn't download.
  - Due to the server based location of the Spectral libraries, our application does require a stable internet connection to acquire these files. If you have confirmed that your internet connection is stable, the issue is most likely due to the repository on MassIVE being inaccessible. Please contact either [ccms-web@proteomics.ucsd.edu](mailto:ccms-web@proteomics.ucsd.edu) or [ccms-web@cs.ucsd.edu](mailto:ccms-web@cs.ucsd.edu) for further assistance with their FTP connection.
- No proteins were downloaded from NCBI.
  - Due to the application creating FASTA information dynamically based on protein selection, our application does require a stable internet connection to acquire these files from NCBI at runtime. If you have confirmed that your internet connection is stable, the issue is most likely due to an issue with NCBI's FTP connection. Please send an email directly to [info@ncbi.nlm.nih.gov](mailto:info@ncbi.nlm.nih.gov) for further assistance with their FTP connection.
- My favorite protein was not found or exported to Skyline.
  - There are a number of reasons why a specific protein was not pushed into skyline.
    - Skyline peptide settings (missed cleavages)
    - Protein not observed in spectral library
      - To add custom observed data, see the *Customizing the plugin with your data* section of the tutorial
    - Protein is not functionally classified by KEGG
  - To manually insert new protein:
    - File -> Insert -> FASTA and copy your FASTA snippet in (Skyline will automatically populate the protein with peptides based on your digestion settings)
    - Edit -> Insert -> Proteins and fill out the table.
  - To manually insert a peptide go to Edit -> Insert -> Peptides and fill out the table.
- No proteins were imported into Skyline
  - This could be due to two possibilities:
    - The plugin was unable to obtain FASTA information for all proteins selected for export. NCBI or your local computer may be experiencing connectivity issues.

- No proteins were selected for export from the plugin. The Spectral Library for the organism(s) examined still should export, but a message will display informing you that no proteins were selected and thus no FASTA was created
- Error Messages:
  - When the Biodiversity Plugin encounters an error, the user is prompted with the option to send an error report. We highly encourage all users who encounter problems to send the error report along with a description of what lead up to the failure. These messages allow our developers to emulate the failure state so the problem can be diagnosed and fixed. We will try to address issues as soon as possible.
  - My version of Skyline is out of date.
    - Our application requires Skyline to be version 3.1.1.7490 or later. To update to the most recent version, please go to [Skyline Daily's webpage](#) to download the latest version of Skyline.
  - Unable to establish connection to NCBI to acquire FASTA for your organisms.
    - The plugin queries NCBI in real time and requires a stable internet connection. NCBI or your local computer may be experiencing connectivity issues.
  - MassIVE Server Unreachable
    - The plugin downloads the spectral library files from the UCSD MassIVE FTP server and requires a stable internet connection. MassIVE or your local computer may be experiencing connectivity issues.
- How do I add my data to the Biodiversity Library?
  - The Biodiversity Library is curated by Samuel Payne at Pacific Northwest National Laboratory. For more information about your data being visible in the Library, please contact [Samuel.Payne@pnnl.gov](mailto:Samuel.Payne@pnnl.gov).
- Observed protein count is zero when customizing data.
  - This could be due to two possibilities:
    - The organism you selected to update doesn't match the organism that was used in your results
    - The identifiers in your data are not Uniprot identifiers (e.g. RefSeq, Genbank, GI numbers)
    - The organism you selected does not use Uniprot as identifiers in KEGG

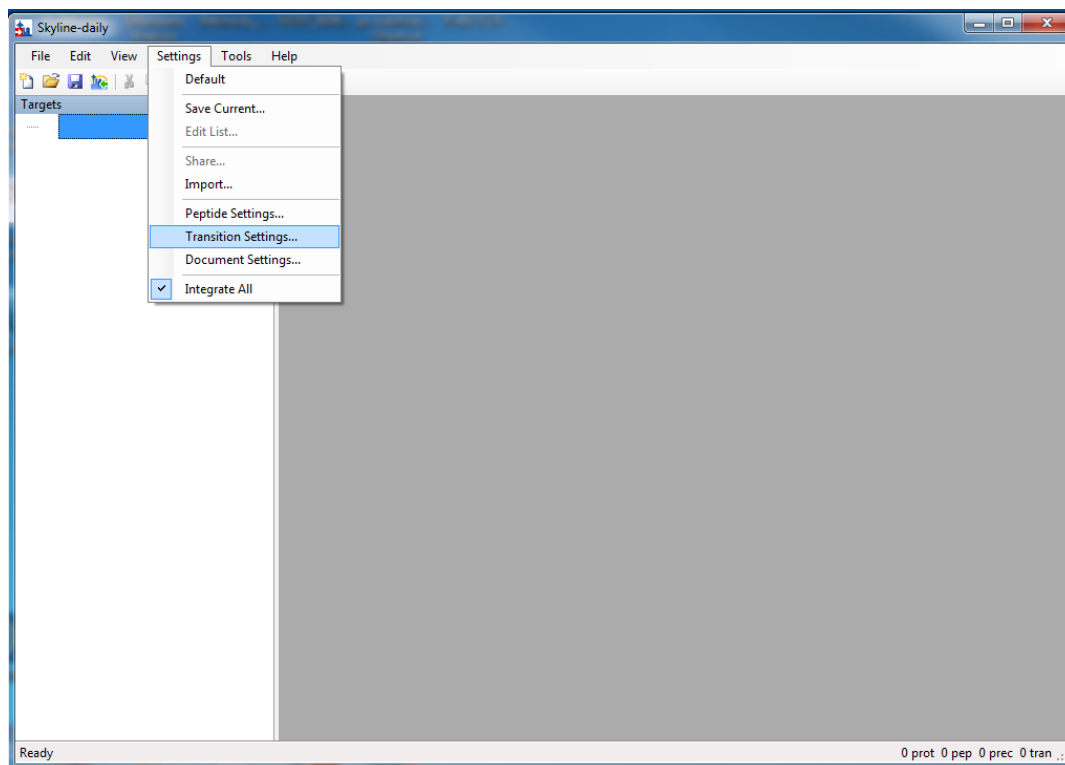
## Biodiversity Plugin for DIA Data Analysis

Usage of the Biodiversity Plugin for hypothesis-driven DIA data analysis proceeds in a manner very similar to that which is described for SRM/PRM method development. However, several additional steps, common in the analysis of DIA data, are recommended. If you are unfamiliar with the analysis of DIA data in Skyline, you are encouraged to consult the webinar and tutorial for Data Independent Analysis, which can be found at Please consult the Data Independent Acquisition tutorial on the Skyline website. The tutorial and webinar can be found at

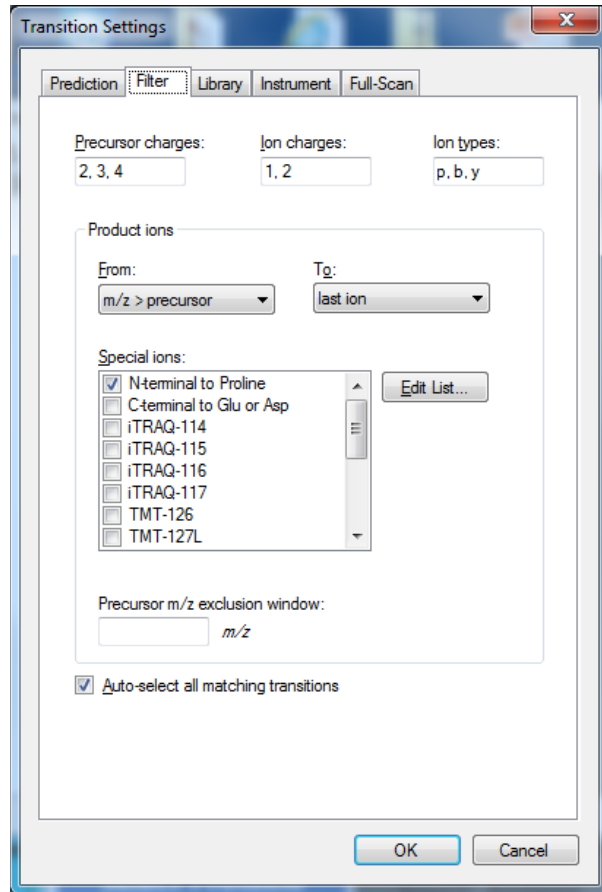
[https://skyline.gs.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=tutorial\\_dia](https://skyline.gs.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=tutorial_dia).

To analyze DIA data using the Biodiversity Plugin:

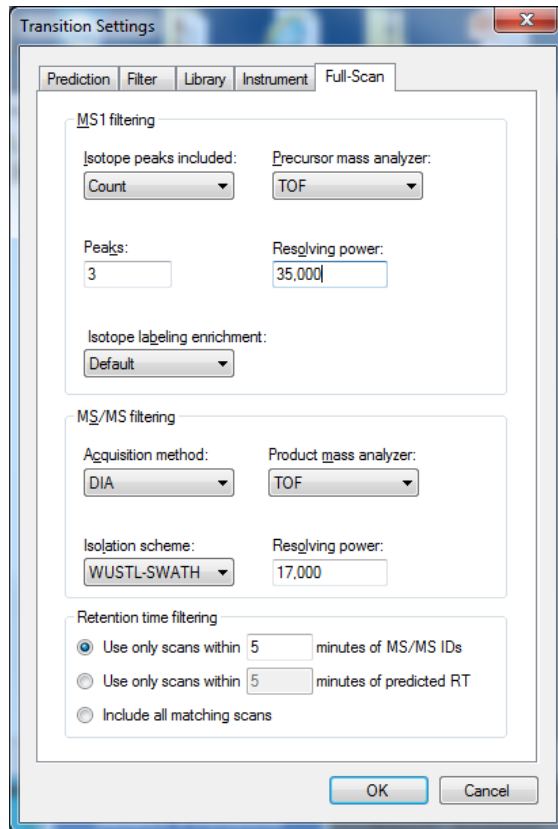
1. Start Skyline. If you have not already done so, install the PNNL Biodiversity Library Plugin for Skyline through the Skyline tools store interface (Tools --> Tool Store --> Biodiversity Plugin. See p. 3 of the Biodiversity Plugin Tutorial).
2. Within Skyline, choose the Peptide Settings you wish to use for your document (Settings --> Peptide Settings). After the FASTA file has been imported, you will not be able to make changes in Digestion, Filter, or Modifications tabs (see p. 18 of this Tutorial). On the Library tab, ensure that no spectral libraries are selected. A new library, designed for your experiment, will be generated by the Biodiversity Plugin.
3. Choose the appropriate Transition Settings for your DIA experiment (Settings --> Transition Settings).



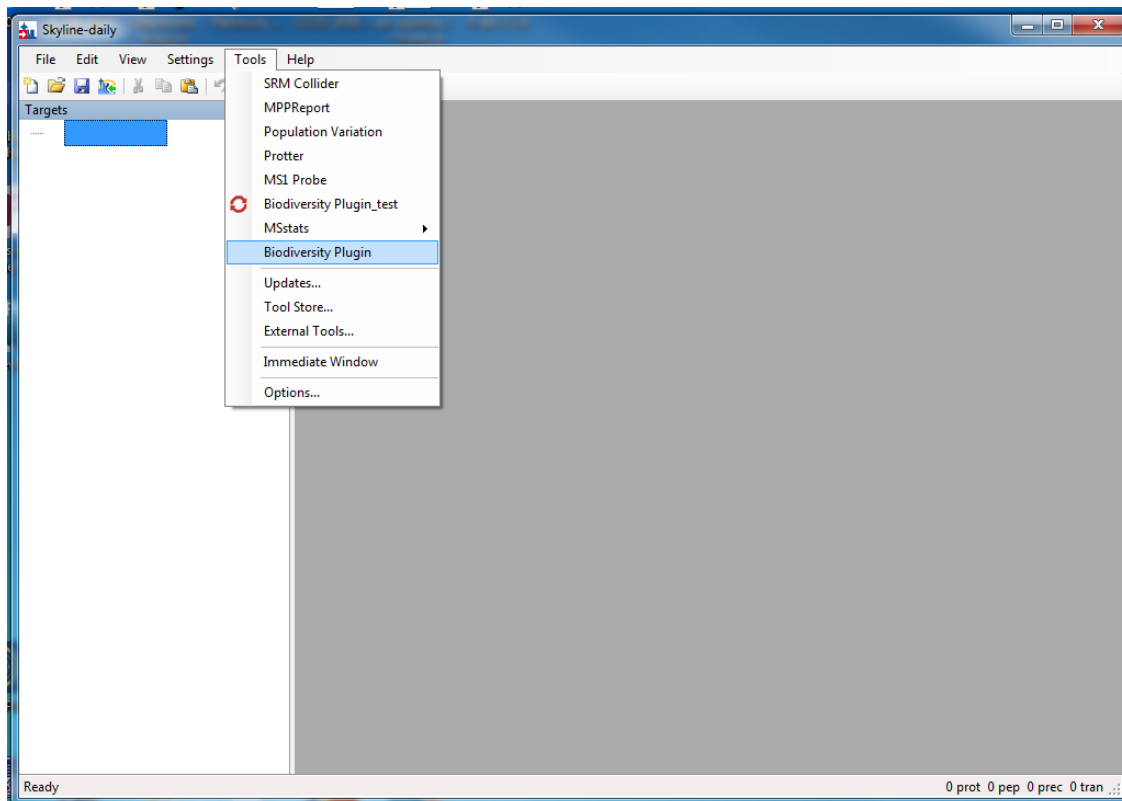
- a. In the Filter tab, set desired precursor and ion charges, and also select ion types (p, b, y for precursor, b ions, y ions) as desired. Select desired mass range and rules for in designated area for Product ions.



- b. In the Library tab, make sure that ion match tolerance is set at a value that is appropriate for your data files.
- c. In the Instrument tab, make sure that the mass range (Min m/z and Max m/z) match the data acquisition range for your data files and that the Method match tolerance is appropriate for your instrument and acquisition settings.
- d. In the Full-Scan tab, in the MS1 filtering area, set Isotope peaks included to Count with 3 peaks. Settings for Precursor mass analyzer and Resolving power should be appropriate for the instrument used for data acquisition (either Orbitrap or TOF).
- e. In the Full-Scan tab, in the MS/MS filtering area, set the Acquisition method to DIA and select the appropriate product mass analyzer (either Orbitrap or TOF). Generate an isolation scheme that matches what was used in data acquisition. This step is extremely important; data will not import correctly if the Isolation Scheme is not correct. Once these steps are completed, the Full-Scan tab should look something like this:



4. Start the Biodiversity Plugin (Tools --> Biodiversity Plugin).





5. For processing of DIA experiments, it is highly recommended that you upload a local spectral library in order to have accurate retention time information for your particular LC-MS system. See pp. 11-17 of this Tutorial for instructions.
6. Choose desired organism (created with data upload) and KEGG pathway, export .fasta file to Skyline as described in pp. 5-10 of this Tutorial.
7. Import DIA data files into Skyline and proceed with analysis. If additional details are needed, please see the Skyline Data Independent Acquisition tutorial and webinar.