Protein Metrics Preview and Byonic Nodes in Proteome Discoverer 3.1

Fall 2023

Protein Metrics Inc., Cupertino, California, USA

Contents

Protein Metrics Preview and Byonic Nodes in Proteome Discoverer 3.1

Introduction

This document describes the basic usage of **Proteome Discoverer 3.1** nodes for Protein Metrics **Preview** and **Byonic** search engines, with focus given to the features that are specific to Proteome Discoverer 3.1

For more details on Preview and Byonic (e.g., setting parameters or interpreting results), please refer to the **Preview** and **Byonic** user manuals.

Installation

Prerequisite:

1. You must have a licensed and installed copy of Proteome Discoverer 3.1. Make sure that it is not running when performing additional installations.

2. If not already installed, install Java Runtime, either Oracle JRE (version 8 or above) *or* OpenJDK (version 11.0.02 or above).

3. If not installed already, install the current version of the Protein Metrics software suite (currently version 5.3), which includes both Preview and Byonic. The most recently version of the Protein Metrics software suite is available from http://proteinmetrics.com/. Make sure the license is activated.

4. Run the installer for the Preview node and the Byonic node.

Working with other nodes

Preview is a very quick and simple search engine. Preview:

- Automatically generates suggested parameters for a subsequent Byonic search – see Using Preview and Byonic together.

- Produces text output (formatted as HTML) rather than lists of proteins and peptides that can be used as input to other nodes.

There is generally no utility in connecting Preview with nodes other than those shown in the workflow diagram under Running Preview.

A **Byonic** workflow generally requires fewer nodes compared to other database search engines:

- Though possible, we do *not* recommend using a PSM Validation node (e.g., Percolator) in a Byonic workflow because Byonic already has its own machine learning optimization and FDR calculation.

- In our experience, spectrum pre-processing generally does not help Byonic and often makes results worse.

- For ETD data, we do not recommend using the Non-Fragment Filter node because Byonic can intelligently handle ETD spectra with large precursor peaks.

- Nodes such as the annotation and quantification nodes are compatible with the Byonic node.

- See the example workflow diagram under Running Byonic. We recommend using the analysis template "PMI-Byonic Template.pdAnalysis" as a starting point.

Running Preview

We recommend running the analysis template "PMI-Preview Template.pdAnalysis" as a starting point.

To load this template:

1. Create/open a study by clicking **File>New Study/Analysis** and providing a name and folder destination (Study Root Directory).
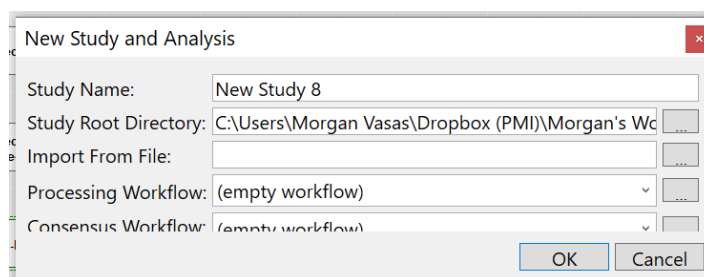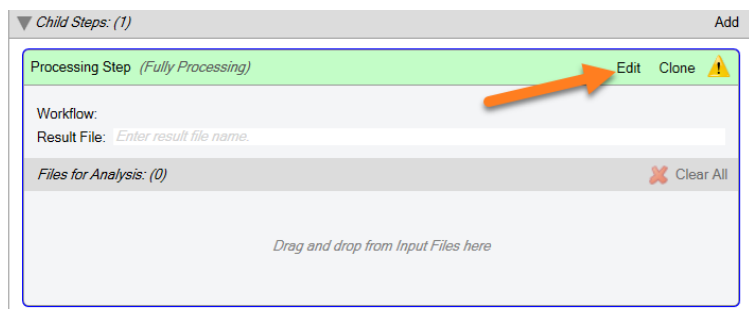
2. Perform either of the following:

    a. Click the **Open Analysis Template** button and browse to find "PMI-Preview Template.pdfAnalysis" in the Proteome Discoverer 3.1 standard workflows/templates folder (typically "C:\Users\Public\Documents\Thermo\Proteome Discoverer 3.1\Common Templates"
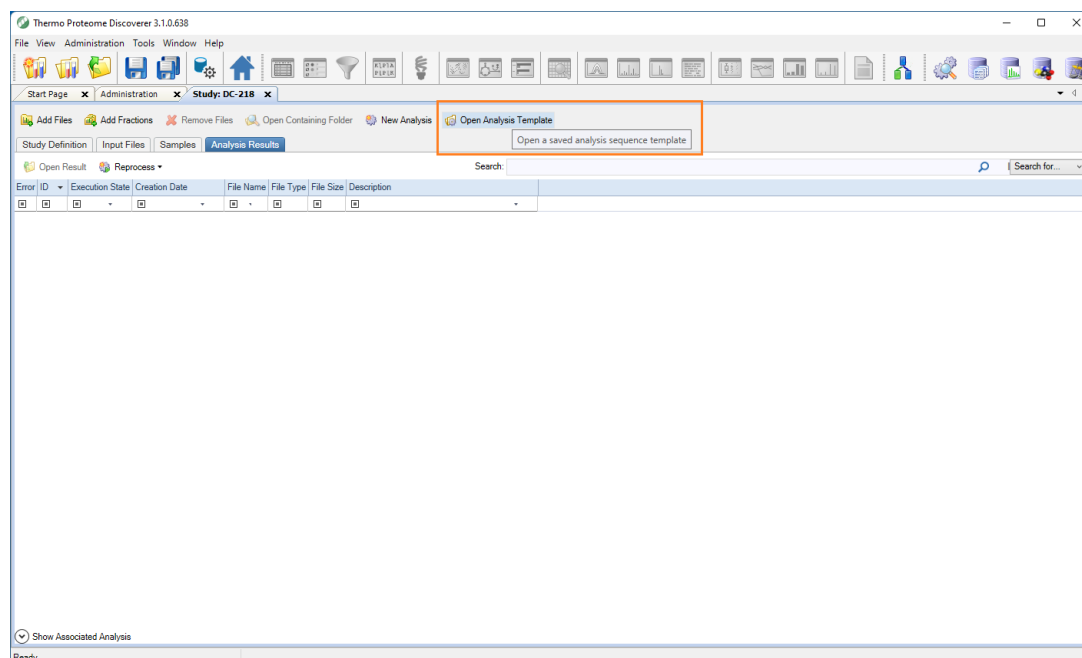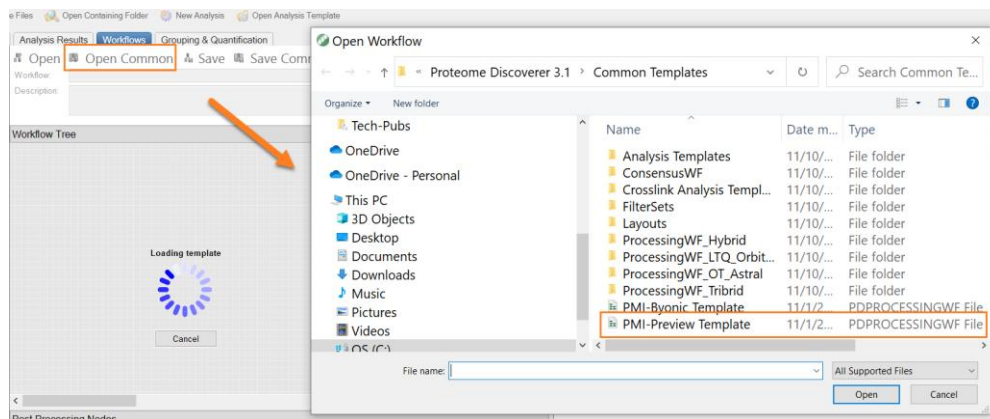

    b. **Or**, perform the following steps:

        i. Click "New Analysis"
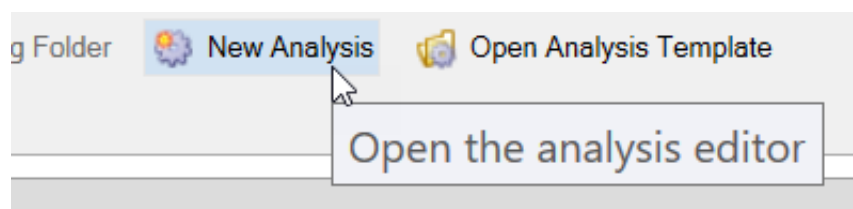
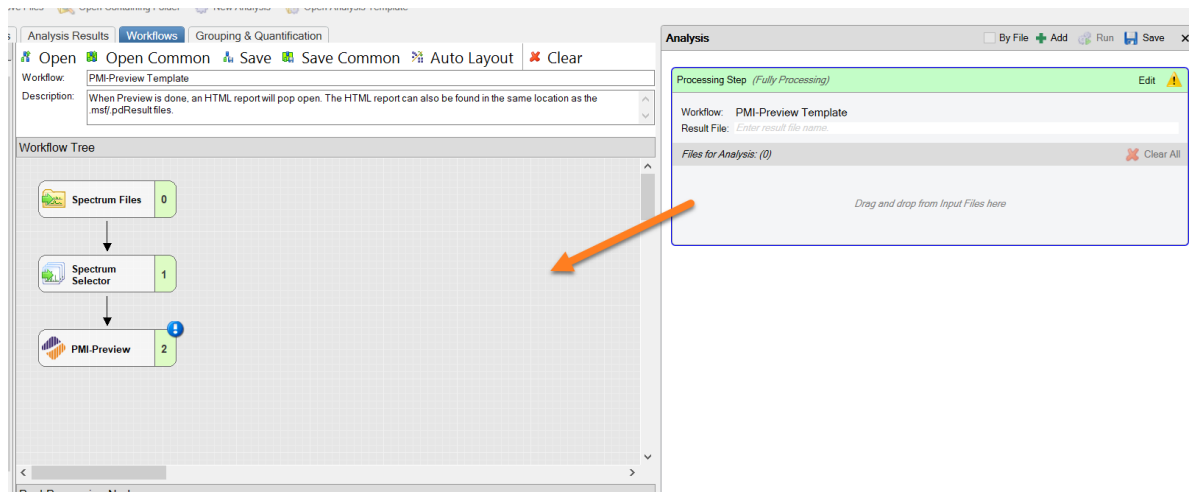   ii. Click "Edit" in the **Processing Step** window
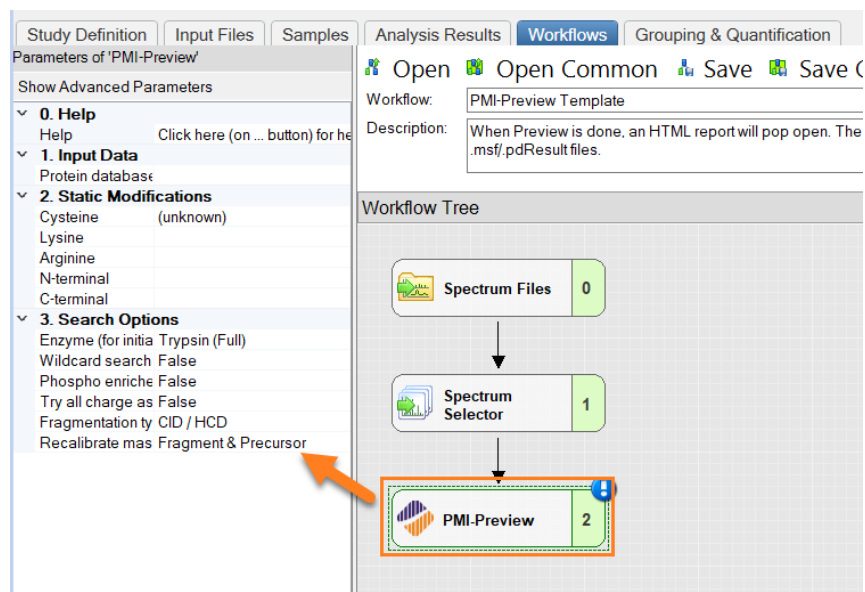
iii. Click "Open Common" and load the Processing workflow:

iv. Click on the x to delete the Consensus Workflow window as it is not needed.

3. Clicking anywhere on the **Processing Step** window will open the associated workflow:

4. Click on a node to see its parameters. Shown below are the parameters for the **PMI-Preview** node:

a. To see the advanced parameters, click **Show Advanced Parameters** above the list of parameters.

Note that the **FASTA/protein database** (1. Input Data) and modifications (2. Static Modifications) can be added or modified within these parameters.

Running Byonic

We recommend using the analysis template "PMI-Byonic-Template.pdAnalysis" as a starting point.

- For the Byonic node, this template uses the Byonic node's default parameters.

- For other nodes in the workflow, this template sets some parameters that are optimized for Byonic and differ from the Proteome Discoverer default values. For more details, see the notes at the end of this section.

Steps to load this template:

1. Create/open a study by clicking **File>New Study/Analysis** and providing a name and folder destination (Study Root Directory).
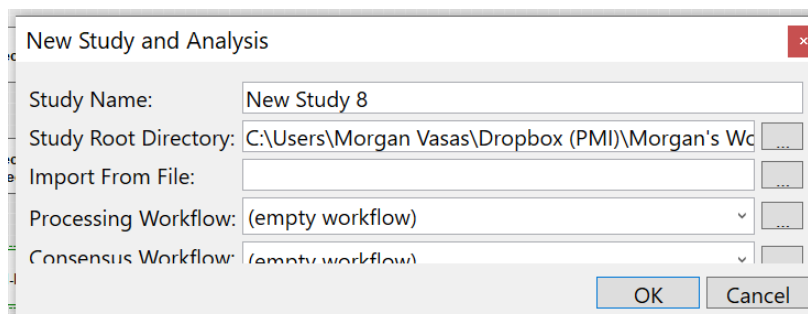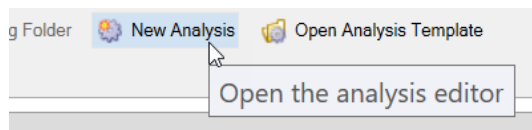

2. Perform either of the following:

a. Click **Open Analysis Template** and browse to find "PMI-Byonic Template.pdfAnalysis" in the Proteome Discoverer 3.1 standard workflows/templates folder (typically "C:\Users\Public\Documents\Thermo\Proteome Discoverer 3.1\Common Templates")
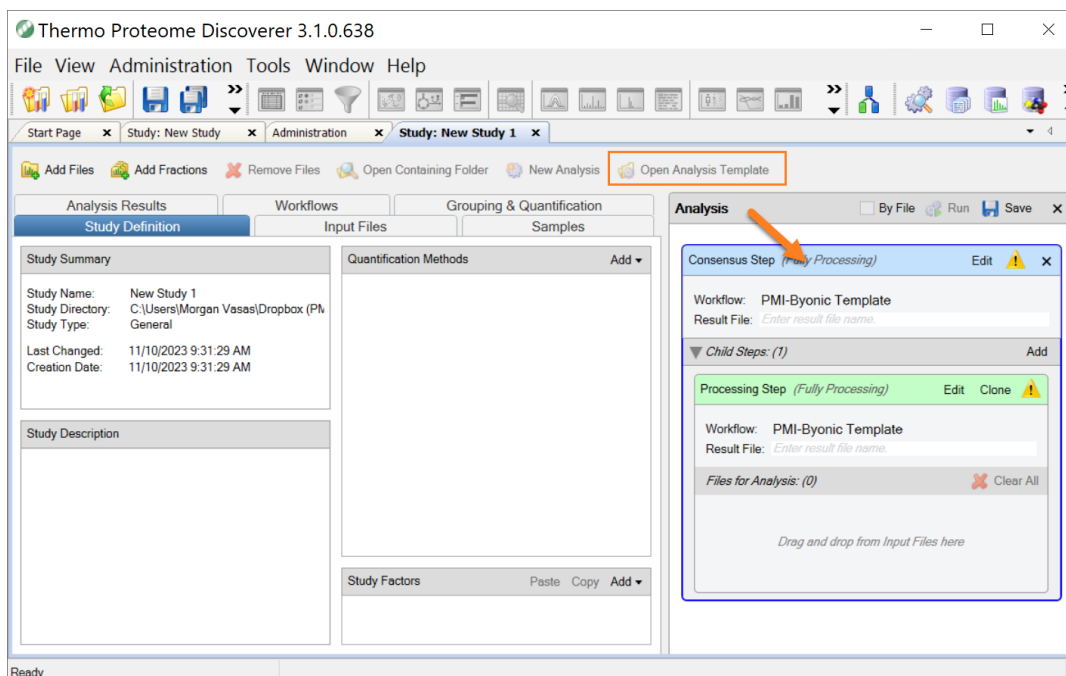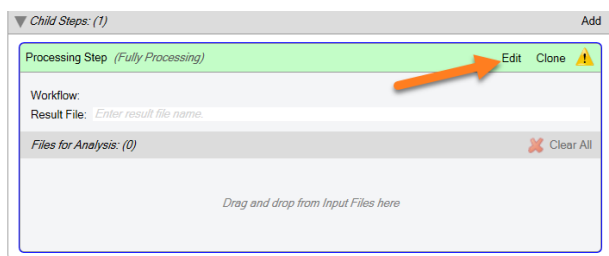
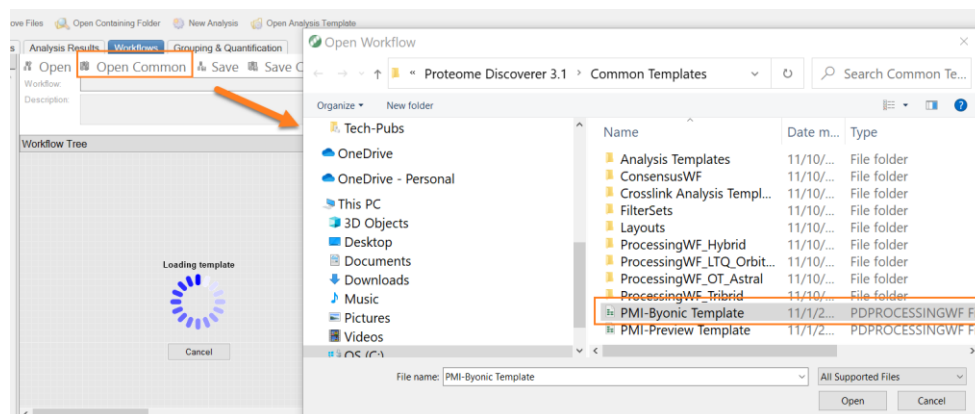b. **Or**, perform the following steps:
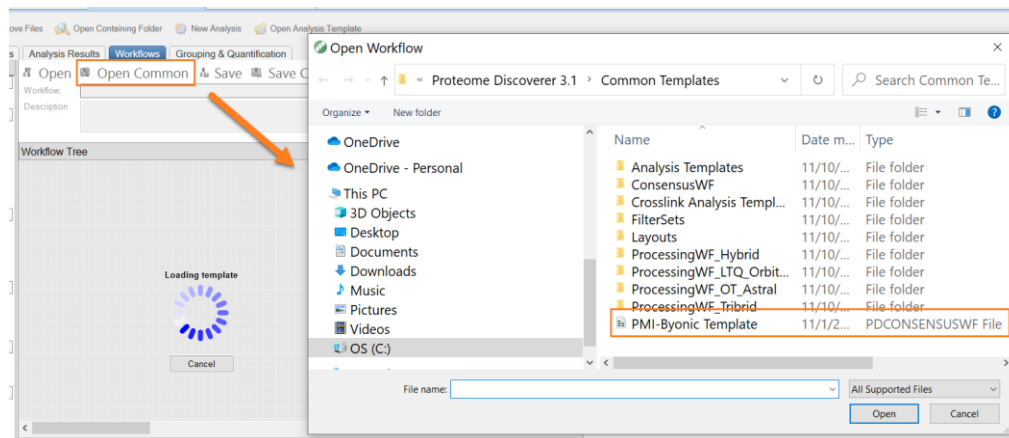
 i. Click **New Analysis**:

 ii. Click **Edit** in the **Processing Step** window:

 iii. Click "Open Common" and load the processing workflow template as shown below:

iv. Click anywhere on the **Consensus Step** window:

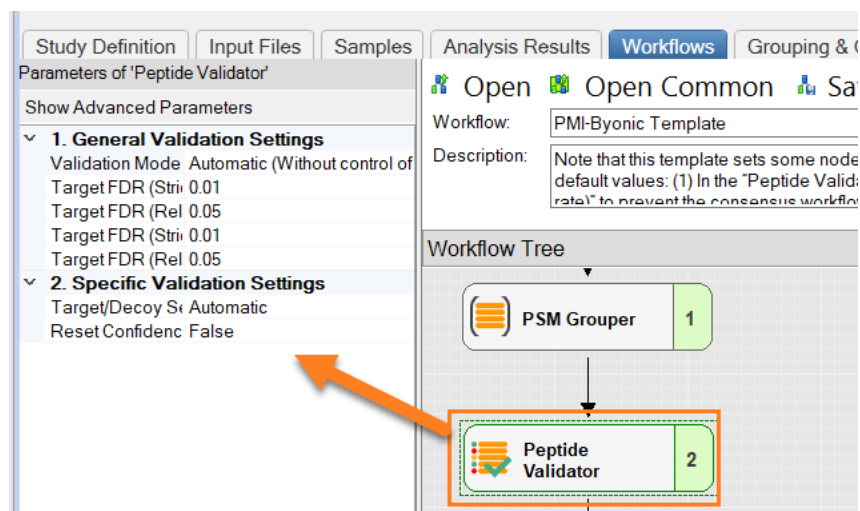v. Click Open Common and load the consensus workflow template as shown below:

3. Click on either window (Processing and Consensus) to open the respective workflow. Shown below is the Consensus workflow:
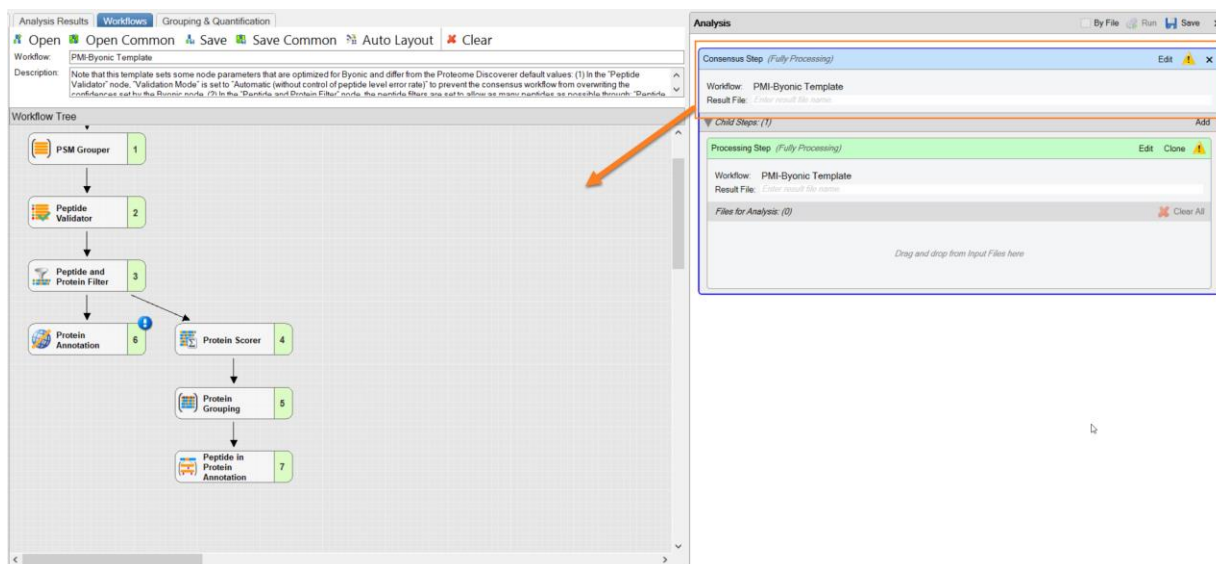
4. Click on a node to see its parameters. Shown below are the parameters of the **Peptide Validator** node:

    a. To see the advanced parameters, click **Show Advanced Parameters** above the list of parameters.
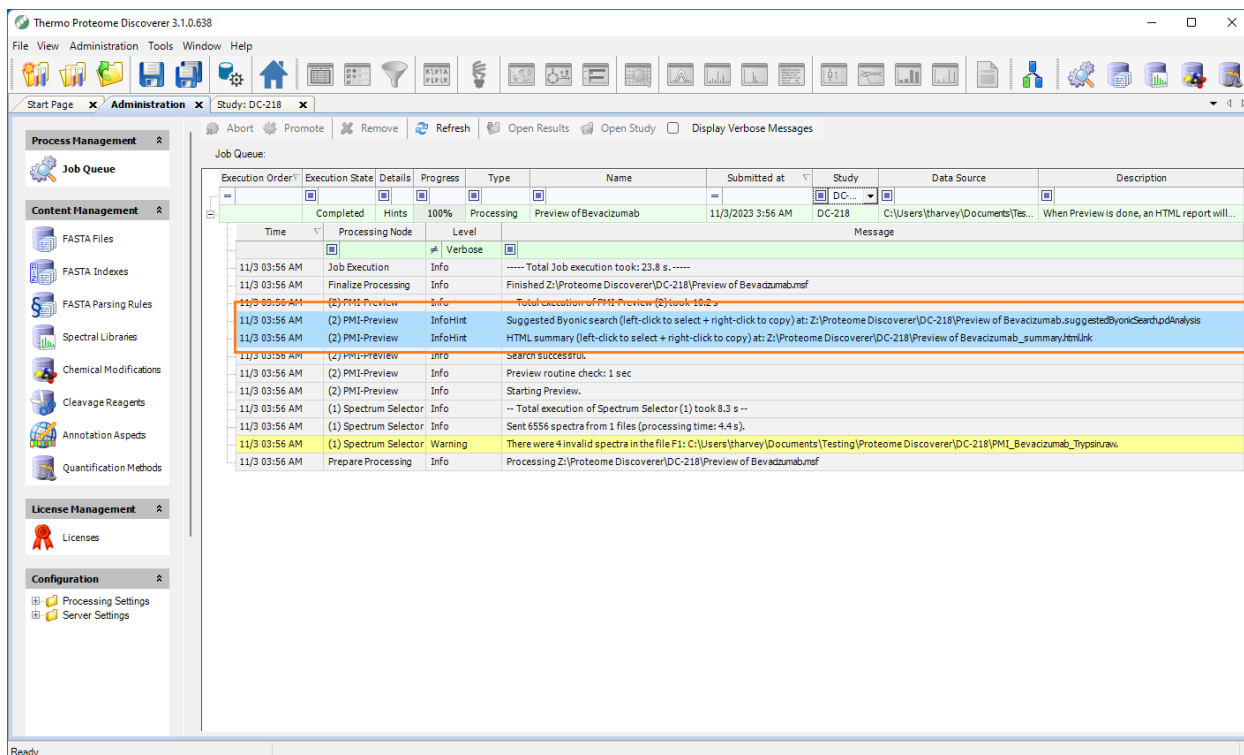
Notes on the **Processing** workflow:

- In the **Spectrum Selector** node, the "Max. Precursor Mass" is set to 0 Da, which means no limit – any precursor mass is allowed. This setting is critical when analyzing topdown data and can be important when analyzing glycopeptide data.

  - Additionally, the "MS Order" scan event filter is set to **Any**. This allows MS[1] spectra to pass through into the Byonic node. Unlike most other database search nodes, Byonic can reanalyze the MS[1] for precursor charge and m/z. This is particularly important for topdown.

Notes on the **Consensus** workflow:

- In the **Peptide Validator** node, the "Validation Mode" is set to "Automatic (Without control of peptide level error rate)" to prevent the consensus workflow from overwriting the confidences set by the Byonic node.

- In the **Peptide and Protein Filter** node, the peptide filters are set to allow through as many peptides as possible:

  - "Peptide Confidence At Least" is set to **low**.

  - "Keep Lower Confident PSMs" is set to **true**.

  - "Minimum Peptide Length" is set to **4** (the lowest possible setting).

- In the **Protein Grouping** node, parsimony is *disabled*. Byonic can identify multiple peptides from a single MS/MS spectrum but parsimony allows only one peptide identification per spectrum.

Using Preview and Byonic together

Preview generates, in addition to an HTML report, a suggested Byonic search if Preview has made a sufficient number of good identifications). Note that this is a *suggested* search. In general, the parameters should be reviewed and, if necessary, modified to fit the experiment. Preview is good at finding modifications that are consistently found sample-wide (for example, in-vitro modifications from sample handling and processing.). However, Preview may miss post-translational modifications that are found only on a few proteins, and Preview does *not* look for glypcopeptides. If you are interested in finding specific post-translational modifications or glycopeptides, you should carefully review and adjust the Byonic parameters suggested by Preview.

To load the suggested Byonic search, open the .suggestedbyByonicSearch.pdAnalysis file, which is named similarly to the Proteome Discoverer .msf/.pdfResult files.

Also note that the locations of the Preview HTML report and the suggested Byonic search (.suggestedByonicSearch.pdfAnalysis file) are shown in the search details in the job queue (see the above image).