



PEAKS[®] CMD

PROTEOMIC COMMAND LINE SOLUTION

Microsoft Windows User Guide
November, 2015



Bioinformatics Solutions Inc.
www.bioinfor.com

1. Introduction

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) based proteomics has become a routine approach in life science since the past decade. In a typical proteomics lab, vast amounts of experimental data are generated daily requiring batch analysis without manual intervention. Further, a growing number of institutions utilize analytical pipelines to perform fundamental proteomic identification tasks. To address each of these cases PEAKS CMD offers a command line-based solution for protein/peptide identification.

The following describes how to set up and configure PEAKS CMD on a Microsoft Windows computer ensuring prompt sample analysis either independently or as part of a customized pipeline.


2. Configuration

To meet the demands of analyzing large data sets, PEAKS CMD is designed for 64 bit Windows Operating Systems (XP, Vista, and 7). The amount of disk space required depends on the size of the user datasets. The two main factors affecting PEAKS performance are CPU and RAM.

- A recommended configuration can be an Intel Core processor, 8GB RAM.
- An ideal configuration may be closer to Intel Core i7 or Xeon processors, 16GB RAM or more (or 2GB per core).

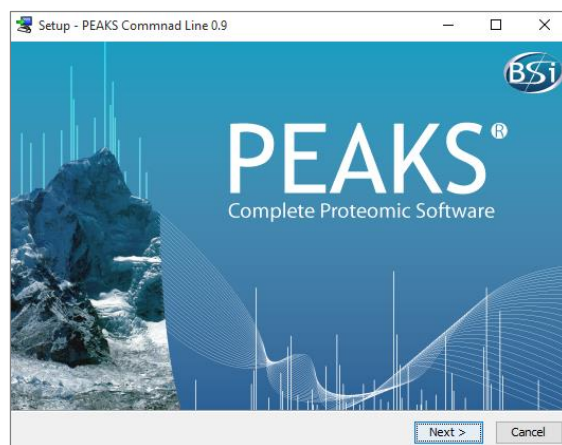
3. Installation

Step 0: Download the PEAKS CMD setup file to the intended computational resource (pipeline core or computer). In the event this file cannot be found, please contact sales@bioinform.com. They can also assist with registration key issues.

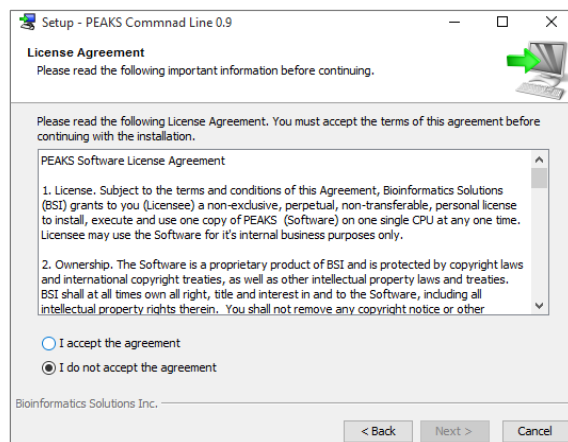
Step 1: Locate to the PEAKS CMD software package directory and double-click the setup file to launch the installation. The setup file can be easily recognized by its icon .

NOTE: If PEAKS CMD is intended to be installed on a Microsoft Windows Server, right-click the setup file and select “Run it as administrator”.

Step 2: Click the “Next” button to start the installation process.

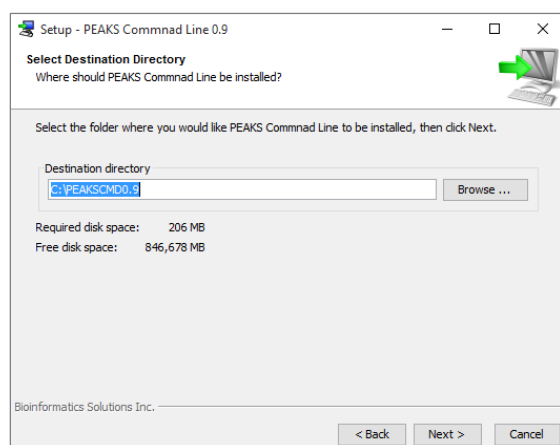


Step 3: Check the radio button “I accept the agreement” and click “Next”.

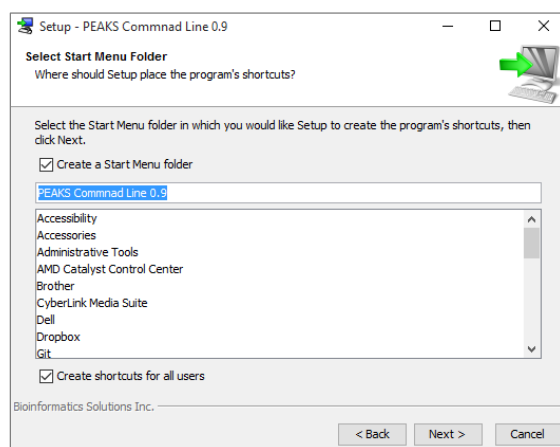


Step 4: Select the desired installation directory for PEAKS CMD using the “Browse ...” button and click “Next >”.

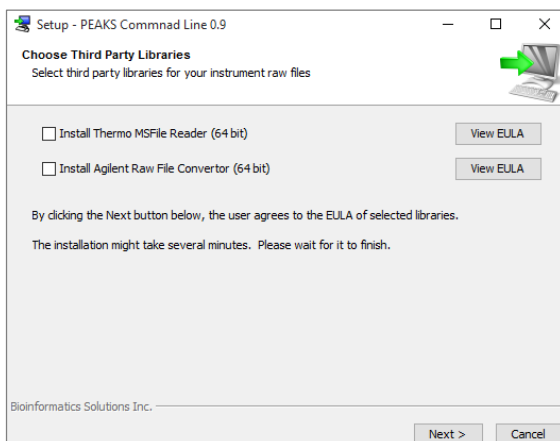
NOTE: From this step users can simply click the “< Back” button to navigate to the previous step for possible changes.



Step 5: Create the Start Menu Folder, and click “Next >”.

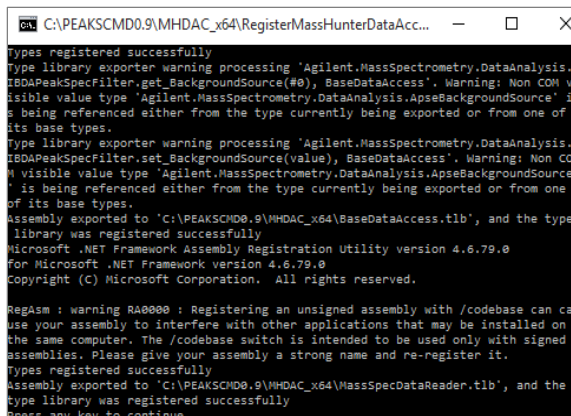


Step 6: Select the boxes of any required third party libraries, such as Thermo or Agilent, for direct loading instrument raw files. Click “Next >”. This will prompt the installation of third party software onto this resource.



NOTE: A DOS window may appear requesting “press any key to continue”, press any key to process to the next step.

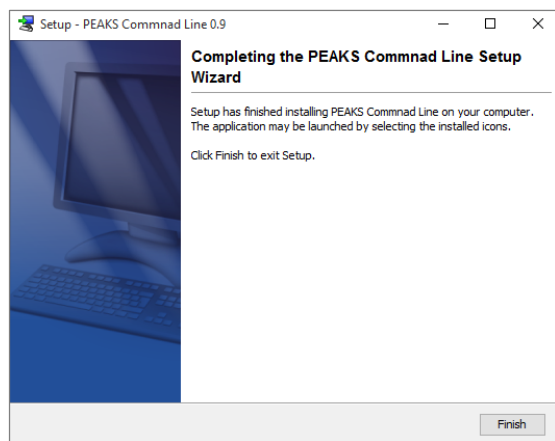
NOTE: If third party software installations are not required, click “Next >” directly.





```
C:\PEAKSCMD0.9\MHDAC_x64\RegisterMassHunterDataAcc...
Types registered successfully
Type library exporter warning processing 'Agilent.MassSpectrometry.DataAnalysis.I80APeakSpecFilter.get_BackgroundSource(#0), BaseDataAccess'. Warning: Non COM visible value type 'Agilent.MassSpectrometry.DataAnalysis.ApseBackgroundSource' is being referenced either from the type currently being exported or from one of its base types.
Type library exporter warning processing 'Agilent.MassSpectrometry.DataAnalysis.I80APeakSpecFilter.set_BackgroundSource(value), BaseDataAccess'. Warning: Non COM visible value type 'Agilent.MassSpectrometry.DataAnalysis.ApseBackgroundSource' is being referenced either from the type currently being exported or from one of its base types.
Assembly exported to 'C:\PEAKSCMD0.9\MHDAC_x64\BaseDataAccess.tlb', and the type library was registered successfully
Microsoft .NET Framework Assembly Registration Utility version 4.6.79.0
for Microsoft .NET Framework version 4.6.79.0
Copyright (C) Microsoft Corporation. All rights reserved.

RegAsm : warning RA0000 : Registering an unsigned assembly with /codebase can cause your assembly to interfere with other applications that may be installed on the same computer. The /codebase switch is intended to be used only with signed assemblies. Please give your assembly a strong name and re-register it.
Types registered successfully
Assembly exported to 'C:\PEAKSCMD0.9\MHDAC_x64\MassSpecDataReader.tlb', and the type library was registered successfully
Press any key to continue . . .
```

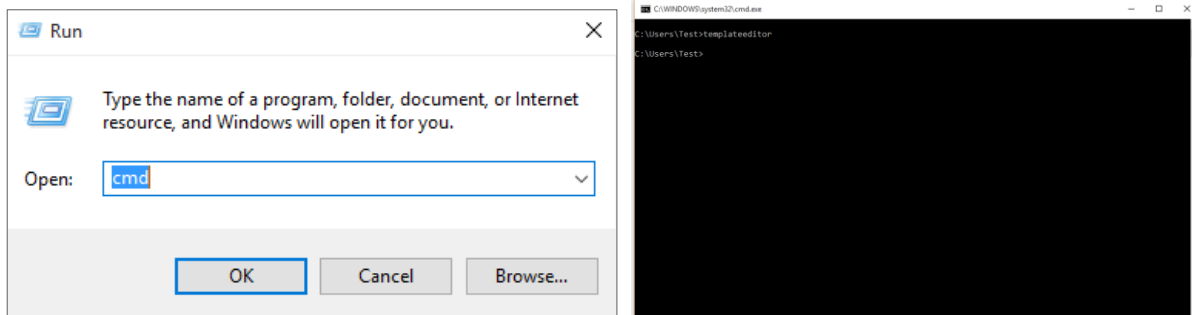
Step 7: The following dialog indicates the successful installation of PEAKS CMD. Click “Finish” to complete the installation.



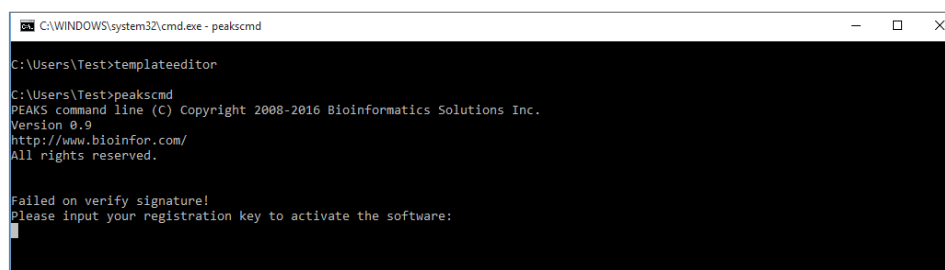
4. Registering PEAKS CMD

Microsoft Windows command line must be launched to run PEAKS CMD. To launch a command line, click Windows “Start” button or “ + S”, type “cmd” in the Search or Run line and press . Typically, a Windows command line will start to run as shown below:

NOTE: If PEAKS CMD is installed on a Microsoft Windows Server, you need to run Windows command line as administrator.



Type “peakscmd” in the Windows command line to run PEAKS CMD. The first time PEAKS CMD is run, it will ask for a registration key to activate the software.

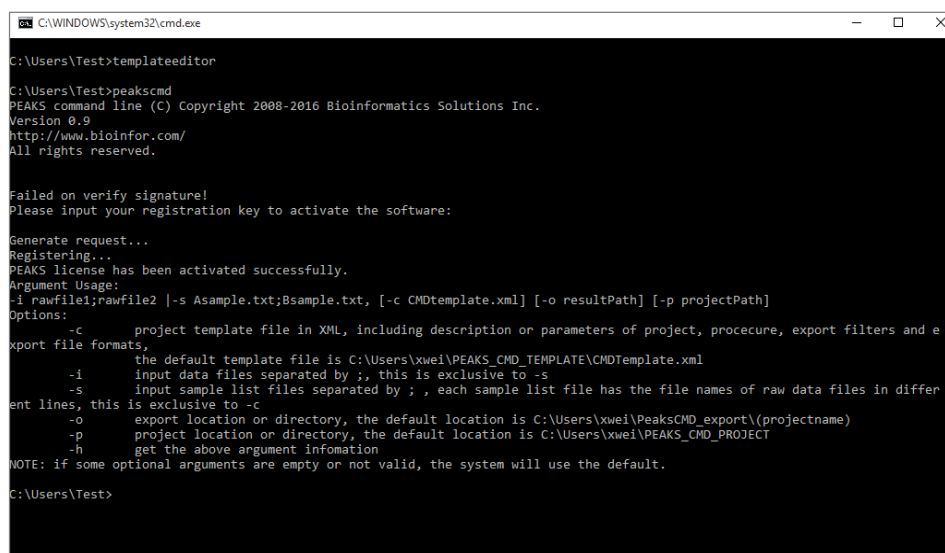


```
C:\WINDOWS\system32\cmd.exe - peakscmd
C:\Users\Test>templateeditor
C:\Users\Test>peakscmd
PEAKS command line (C) Copyright 2008-2016 Bioinformatics Solutions Inc.
Version 0.9
http://www.bioinform.com/
All rights reserved.

Failed on verify signature!
Please input your registration key to activate the software:
```

Type or paste in the registration key acquired from a BSI sales representative.

Usage information will be shown if the registration is successful.



```
C:\WINDOWS\system32\cmd.exe
C:\Users\Test>templateeditor
C:\Users\Test>peakscmd
PEAKS command line (C) Copyright 2008-2016 Bioinformatics Solutions Inc.
Version 0.9
http://www.bioinform.com/
All rights reserved.

Failed on verify signature!
Please input your registration key to activate the software:


Generate request...
Registering...
PEAKS license has been activated successfully.
Argument Usage:
-i rawfile1;rawfile2 -s Asample.txt;Bsample.txt, [-c CMDtemplate.xml] [-o resultPath] [-p projectPath]
Options:
-c      project template file in XML, including description or parameters of project, procedure, export filters and e
export file formats,
the default template file is C:\Users\xwei\PEAKS_CMD_TEMPLATE\CMDTemplate.xml
-i      input data files separated by ;, this is exclusive to -s
-s      input sample list files separated by ;, each sample list file has the file names of raw data files in differ
ent lines, this is exclusive to -c
-o      export location or directory, the default location is C:\Users\xwei\PeaksCMD_export\projectname
-p      project location or directory, the default location is C:\Users\xwei\PEAKS_CMD_PROJECT
-h      get the above argument information
NOTE: if some optional arguments are empty or not valid, the system will use the default.

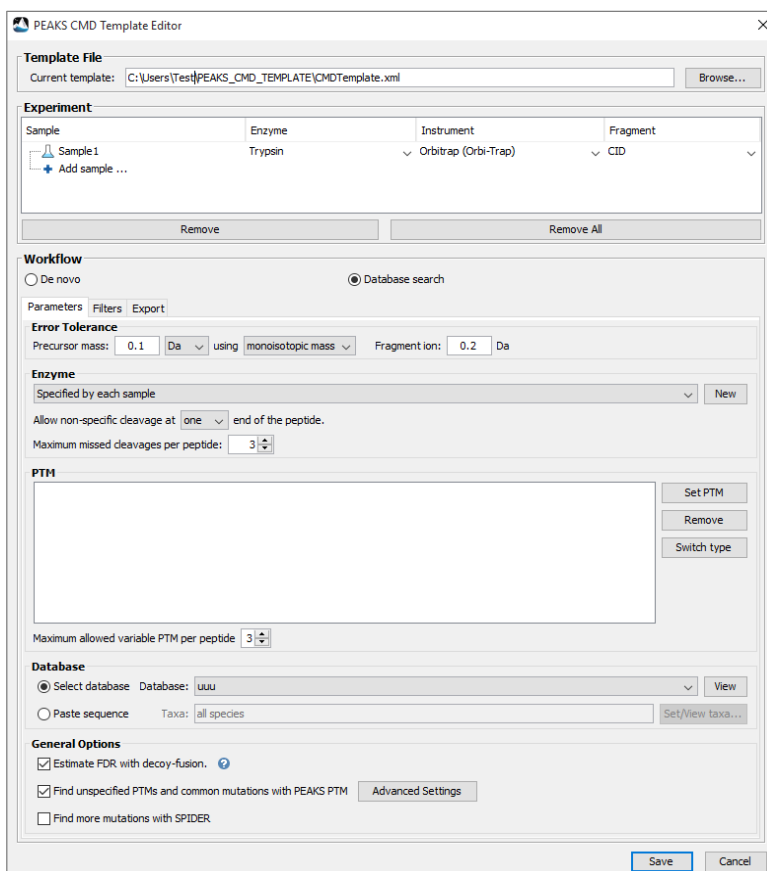
C:\Users\Test>
```

5. Using PEAKS CMD

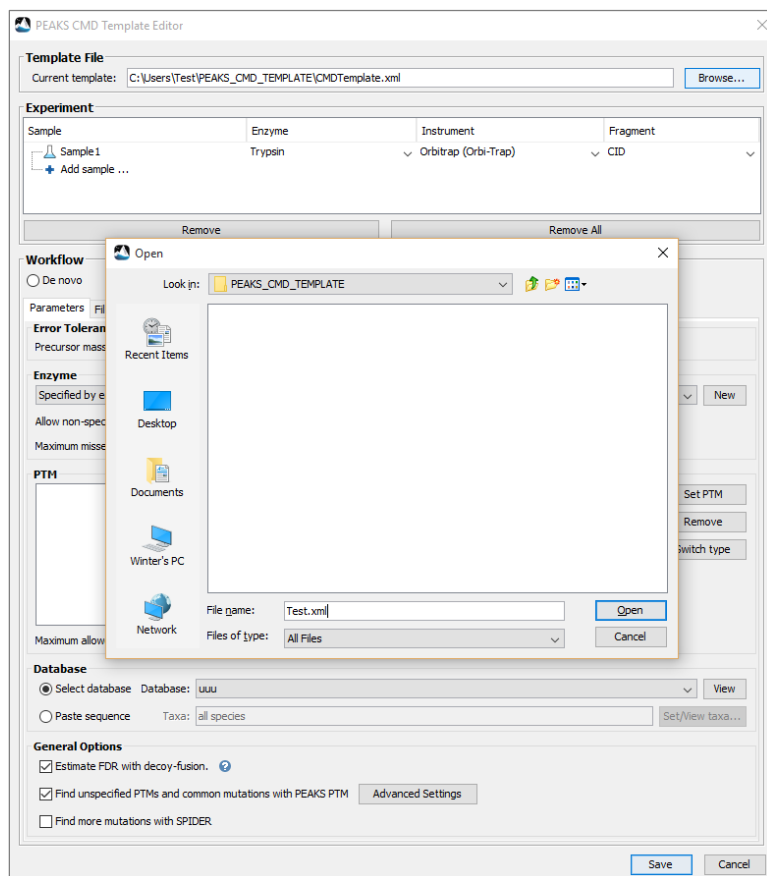
Performing protein/peptide identification using PEAKS CMD requires two steps: 1) generating a template file and 2) running PEAKS CMD with a template file.

5.1 Generating a Template File

The PEAKS CMD Workflow, including configurations of MS file loading, search types, parameters, protein/peptide filtration, and result exporting, is customized by users and stored in a template file. To run PEAKS CMD tasks, users must first create a new (or edit an existing) template file. A GUI-based editor, TemplateEditor, is provided to create or edit template files. Users can launch the editor by typing “TemplateEditor” in Windows command line and press . The editor will appear as presented right:



Step 1: Create or edit the template file in the “Template File” section. The path shown in the “Current Template” text field is the default path to save a template file. Users can also save template files in a preferred directory and with different files name by clicking “Browse...”.



Step 2: Add or remove samples in the “Experiment” section. PEAKS CMD supports analyzing an experiment consisting of multiple samples; add a sample by clicking the “Add sample” or remove selected samples by clicking the “Remove” button. Clicking the “Remove All” button removes all the samples listed in the experiment.

NOTE: Adding sample files in this step is not necessary. Users only need to set the sample numbers and names for the experiment.

Step 3: Select a task for PEAKS CMD. Currently two types of PEAKS CMD tasks are provided in the “Workflow” section: 1) “De novo” for *de novo* sequencing, and 2) “Database Search” for PEAKS database search.

Step 4: Configure PEAKS CMD task parameters. Typically, *de novo* sequencing and database searching require slightly different parameter configurations. The following steps demonstrate how to configure the parameters required in *de novo* sequencing and database searching, respectively.

Option 4.1: Check the “De novo” radio button.

Step 4.1.1: Configure the *de novo* sequencing parameters under the “Parameters” tab.

The screenshot shows the PEAKS CMD Template Editor window. The 'Template File' section shows the current template path. The 'Experiment' section contains a table with columns: Sample, Enzyme, Instrument, and Fragment. It lists 'SampleA' and 'SampleB' with 'Trypsin' as the enzyme and 'Orbitrap (Orbi-Trap)' as the instrument. The 'Workflow' section has 'Database search' selected. The 'Parameters' tab is active, showing 'Error Tolerance' (Precursor mass: 10 ppm, Fragment ion: 0.5 Da), 'Enzyme' (Specified by each sample), and 'PTM' (Maximum allowed variable PTM per peptide: 3). The 'Database' section shows 'Select database' with 'swissprot' selected. The 'General Options' section has 'Estimate FDR with decoy-fusion' checked.

The screenshot shows the PEAKS CMD Template Editor window with the 'De novo' workflow selected. The 'Experiment' section contains a table with columns: Sample, Enzyme, Instrument, and Fragment. It lists 'Sample 1' with 'Trypsin' as the enzyme and 'Orbitrap (Orbi-Trap)' as the instrument. The 'Workflow' section has 'De novo' selected. The 'Parameters' tab is active, showing 'Error Tolerance' (Precursor mass: 0.1 Da, Fragment ion: 0.2 Da), 'Enzyme' (Specified by each sample), and 'PTM' (Maximum allowed variable PTM per peptide: 3). The 'General Options' section has 'Report up to 5 candidates per spectrum'.

Step 4.1.2: Set the *de novo* sequencing result filter under the “Filters” tab.

The screenshot shows the PEAKS CMD Template Editor window. The 'Template File' section at the top shows the current template path as 'C:\Users\Test\PEAKS_CMD_TEMPLATE\Test.xml'. Below this is the 'Experiment' section, which contains a table with columns for Sample, Enzyme, Instrument, and Fragment. The table has one row with 'Sample 1', 'Trypsin', 'Orbitrap (Orbi-Trap)', and 'CID'. There are 'Remove' and 'Remove All' buttons below the table. The 'Workflow' section is below the experiment table, with 'De novo' selected as the workflow type. The 'Parameters' tab is active, showing a filter for 'De novo' with the condition 'ALC (%) ≥ 50'. The 'Filters' tab is also visible. The 'Save' and 'Cancel' buttons are at the bottom right.

Sample	Enzyme	Instrument	Fragment
Sample 1	Trypsin	Orbitrap (Orbi-Trap)	CID

Workflow: ☒ De novo ☐ Database search

Parameters | Filters | Export

De novo ALC (%) ≥ 50

Step 4.1.3: Select the files required for export under the “Export” tab.

The screenshot shows the PEAKS CMD Template Editor window. It has a title bar with a close button. The main area is divided into several sections:

- Template File:** A text field showing the current template path: `C:\Users\Test\PEAKS_CMD_TEMPLATE\Test.xml`, with a **Browse...** button to its right.
- Experiment:** A table with four columns: **Sample**, **Enzyme**, **Instrument**, and **Fragment**.

Sample	Enzyme	Instrument	Fragment
Sample 1	Trypsin	Orbitrap (Orbi-Trap)	CID
+ Add sample ...			

Below the table are **Remove** and **Remove All** buttons.
- Workflow:** A section with two radio buttons: **De novo** (selected) and **Database search**. Below this are three tabs: **Parameters**, **Filters**, and **Export**. The **Export** tab is active, showing a list of items to be exported:

Item	File
<input checked="" type="checkbox"/> De novo peptides	(de novo peptides.csv)
<input checked="" type="checkbox"/> De novo peptides - pepxml	(de novo peptides.xml)
<input checked="" type="checkbox"/> All de novo candidates	(all de novo candidates.csv)
<input checked="" type="checkbox"/> All peptide features	(peptide features.csv)

At the bottom right of the window are **Save** and **Cancel** buttons.

Option 4.2: Check the “Database search” radio button.

Step 4.2.1: Configure the database search parameters under the “Parameters” tab. Users can also choose to expand their search by utilizing PEAKS PTM and SPIDER. To do this check the “Find unspecified PTMs and common mutations with PEAKS PTM” and “Find more mutations with SPIDER” buttons at the bottom of the window.

The screenshot shows the PEAKS CMD Template Editor window. The 'Template File' section at the top shows the current template path. The 'Experiment' section contains a table with columns for Sample, Enzyme, Instrument, and Fragment. The 'Workflow' section has radio buttons for 'De novo' and 'Database search' (selected). Below this are tabs for 'Parameters', 'Filters', and 'Export'. The 'Parameters' tab is active, showing 'Error Tolerance' settings for precursor mass and fragment ion, 'Enzyme' settings for cleavage and missed cleavages, 'PTM' settings for allowed variable PTMs, 'Database' settings for selecting a database or pasting a sequence, and 'General Options' with checkboxes for FDR estimation, PTM/mutation finding, and SPIDER. The 'Find more mutations with SPIDER' checkbox is highlighted with a red box. The bottom of the window has 'Save' and 'Cancel' buttons.

Sample	Enzyme	Instrument	Fragment
Sample 1	Trypsin	Orbitrap (Orbi-Trap)	CID

Workflow

☐ De novo ☒ Database search

Parameters | Filters | Export

Error Tolerance

Precursor mass: 0.1 Da using monoisotopic mass Fragment ion: 0.2 Da

Enzyme

Specified by each sample [New]

Allow non-specific cleavage at one end of the peptide.

Maximum missed cleavages per peptide: 3

PTM

Maximum allowed variable PTM per peptide: 3

Database

☒ Select database Database: uu View

☐ Paste sequence Taxa: all species Set/View taxa...

General Options

☒ Estimate FDR with decoy-fusion. ?

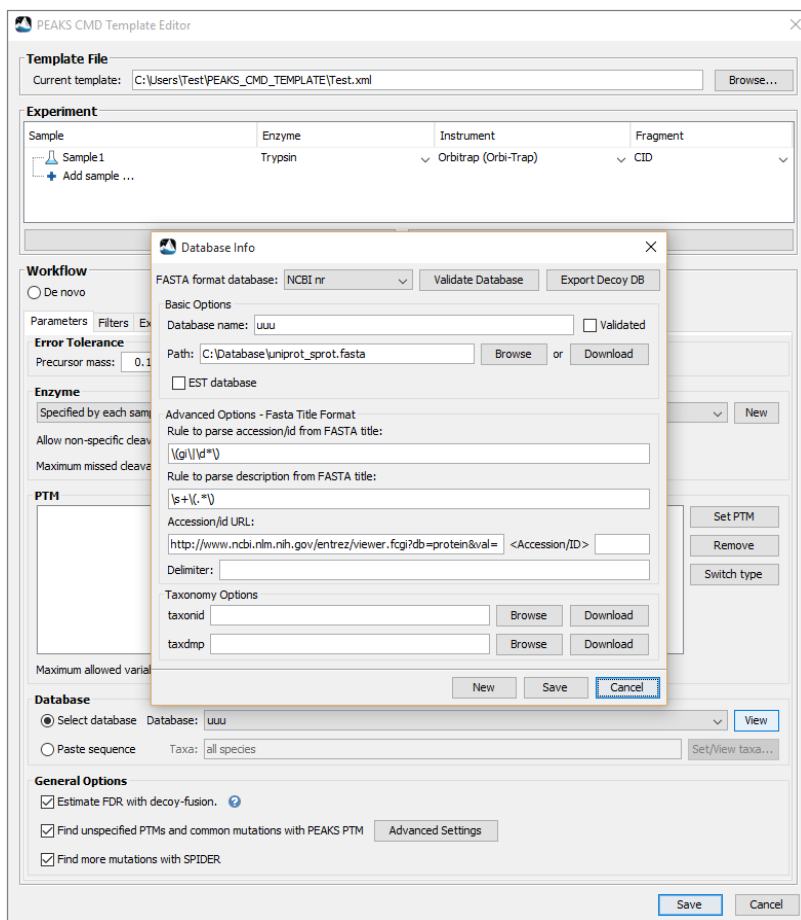
☒ Find unspecified PTMs and common mutations with PEAKS PTM Advanced Settings

☒ Find more mutations with SPIDER

Save Cancel

Configure the protein sequence database in the “Database” section under the “Parameters” tab. A protein sequence database is necessary for performing the PEAKS database search. Users can either select one of the existing protein databases created in a previous search or create a new database entry from a FASTA file.

If there is no database entry in the list, users must first click the “View” button and create a database through the prompts in a pop-up “Database Info” dialog by then clicking the “New” button. Otherwise, users can simply select an existing database entry for the subsequent data analysis. Note that users can always click the “View” button to modify an existing database entry in PEAKS CMD.



Step 4.2.1: Set the database search result filter under the “Filters” tab.

The screenshot shows the PEAKS CMD Template Editor window. The 'Template File' section at the top shows the current template path as 'C:\Users\Test\PEAKS_CMD_TEMPLATE\Test.xml'. Below this is the 'Experiment' section, which contains a table with columns for Sample, Enzyme, Instrument, and Fragment. The table has one row with 'Sample 1', 'Trypsin', 'Orbitrap (Orbi-Trap)', and 'CID'. There are 'Remove' and 'Remove All' buttons below the table. The 'Workflow' section is below the experiment table, with radio buttons for 'De novo' and 'Database search'. The 'Database search' option is selected. Below the workflow section are three tabs: 'Parameters', 'Filters', and 'Export'. The 'Filters' tab is active, showing settings for 'Peptides' and 'Proteins'. The 'Peptides' section has a radio button for 'FDR' (selected) and a value of '1.0' with a '%' symbol. The 'Proteins' section has a radio button for '-10lgP' (selected) and a value of '20.0'. There is also a section for 'unique peptides' with a value of '0'. At the bottom right of the window are 'Save' and 'Cancel' buttons.

Sample	Enzyme	Instrument	Fragment
Sample 1	Trypsin	Orbitrap (Orbi-Trap)	CID

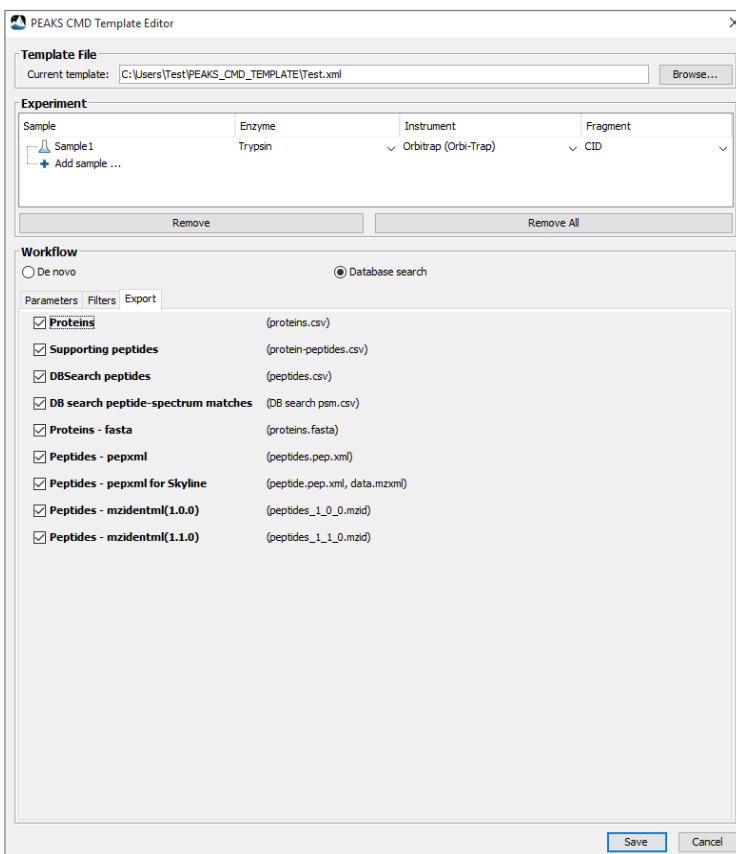
Workflow: ☐ De novo ☒ Database search

Parameters Filters Export

Peptides ☒ FDR % ☐ -10lgP

Proteins ☐ -10lgP and unique peptides

Step 4.2.2: Select the files required for export under the “Export” tab.



Step 5: Click the “Save” button to save the template file and close the TemplateEditor.

5.2 Running PEAKS CMD with a Template File

PEAKS CMD supports six arguments. Running PEAKS CMD with the argument “-h” or “-H” generates the list of six valid arguments as shown below:

```

C:\WINDOWS\system32\cmd.exe
C:\Users\Test>peaks cmd -h
PEAKS command line (C) Copyright 2008-2016 Bioinformatics Solutions Inc.
Version 0.9
http://www.bioinform.com/
All rights reserved.

Argument Usage:
-i rawFile1;rawFile2 [-s Asample.txt;Bsample.txt, [-c CMDtemplate.xml] [-o resultPath] [-p projectPath]
Options:
-c      project template file in XML, including description or parameters of project, procedure, export filters and e
export file formats,
        the default template file is C:\Users\xwei\PEAKS_CMD_TEMPLATE\CMDTemplate.xml
-i      input data files separated by ;, this is exclusive to -s
-s      input sample list files separated by ;, each sample list file has the file names of raw data files in differ
ent lines, this is exclusive to -c
-o      export location or directory, the default location is C:\Users\xwei\PeaksCMD_export\projectname
-p      project location or directory, the default location is C:\Users\xwei\PEAKS_CMD_PROJECT
-h      get the above argument information
NOTE: if some optional arguments are empty or not valid, the system will use the default.

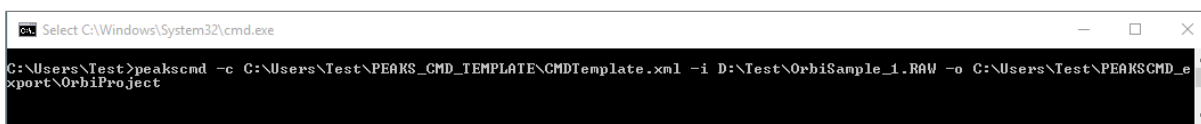
C:\Users\Test>

```

Among the six arguments, -i or -s are mandatory for adding the sample spectra data and are mutually exclusive. A user will then choose one of them and specify the required files or file lists. All other arguments are optional. If the input values for some of the optional arguments are not valid, PEAKS CMD will use the pre-specified default values.

Two examples of using the -i arguments are shown below:

1) Using the -i argument for loading only one data file:

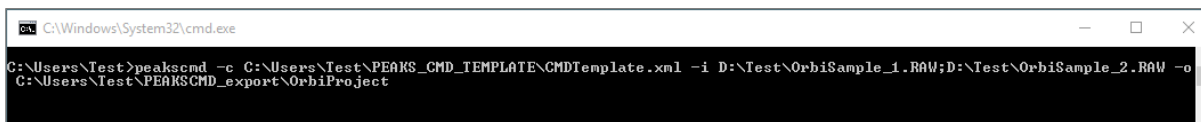


```

C:\Users\Test>peakscommand -c C:\Users\Test\PEAKS_CMD_TEMPLATE\CMDTemplate.xml -i D:\Test\OrbiSample_1.RAW -o C:\Users\Test\PEAKSCMD_export\OrbiProject

```

2) Using the -i argument for loading multiple data files:



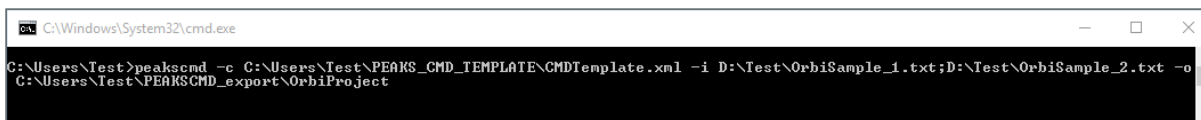
```

C:\Users\Test>peakscommand -c C:\Users\Test\PEAKS_CMD_TEMPLATE\CMDTemplate.xml -i D:\Test\OrbiSample_1.RAW;D:\Test\OrbiSample_2.RAW -o C:\Users\Test\PEAKSCMD_export\OrbiProject

```

The -s argument requires one or multiple text files in which each line lists an absolute path of a data file. When using the -s argument, the number and the names of the sample files must be identical to the number and names of the samples in template file generation using the TemplateEditor.

An example of using the -s argument with multiple sample files is shown below:



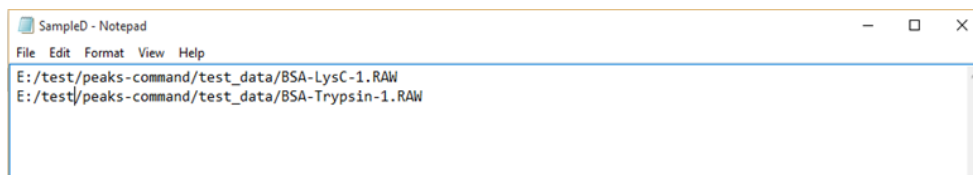
```

C:\Users\Test>peakscommand -c C:\Users\Test\PEAKS_CMD_TEMPLATE\CMDTemplate.xml -i D:\Test\OrbiSample_1.txt;D:\Test\OrbiSample_2.txt -o C:\Users\Test\PEAKSCMD_export\OrbiProject

```

An example of a sample file is shown below:

NOTE: The spectral data files in each sample file are listed line by line, or separated by “;” in a single line.



```

E:/test/peaks-command/test_data/BSA-LysC-1.RAW
E:/test/peaks-command/test_data/BSA-Trypsin-1.RAW

```

If values are not specified for the optional arguments the default template file is used automatically, the project is saved into a default folder, and the results is exported into a default folder. The absolute path of the default template file is C:\Users\user_name\PEAKS_CMD_TEMPLATE\CMDTemplate.xml, the default project location for is C:\Users\user_name\PEAKS_CMD_PROJECT, and the default directory of the exporting files is C:\Users\user_name\PeaksCMD_export/project_name.

PEAKS CMD supports the direct loading of RAW files from ThermoFisher, Waters, and Agilent instruments. Multiple types of text-based spectral data files, *eg.* PKL, MGF, mzXML, and mzML, are also supported.