Comet UI

User Guide

Comet UI Version 1.0.0.0

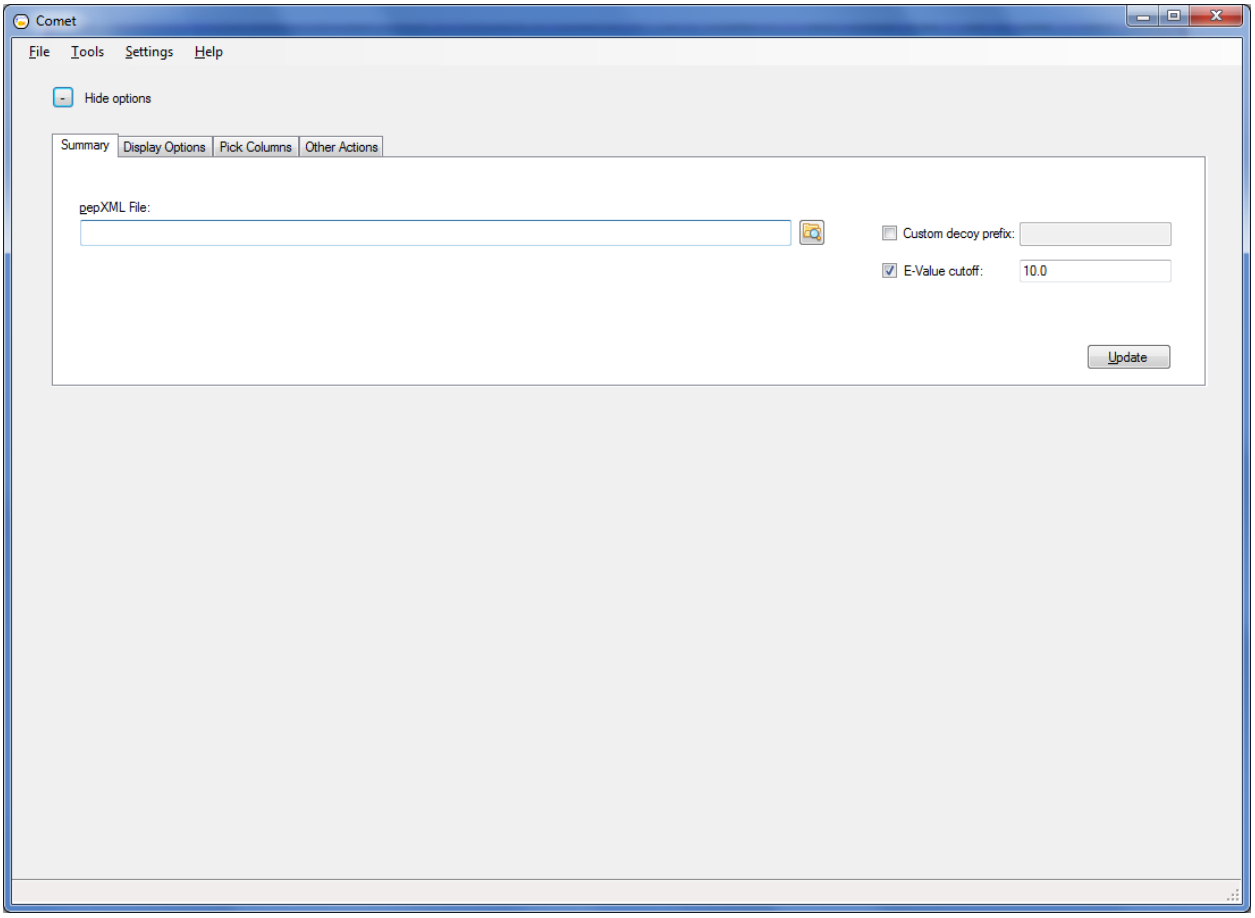
Comet Search Engine Version 2015.02 rev. 0

Contents

Introduction

This document describes features of the CometUI application, a Windows-based graphical user interface used to run the Comet search engine and interrogate results. Users can also use CometUI to edit Comet search parameters, and import/export them from/to a ".params" file used by the command line version of the Comet search engine.

Users can download CometUI from the Comet website. Launching the application will bring up the following window:



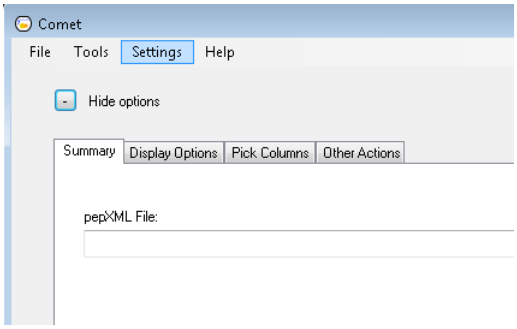
This document will describe how to navigate this window to perform the following tasks:

- View and Edit Search Settings
- Save Search Settings
- Import/Export Search Settings
- Run Search
- View Search Results

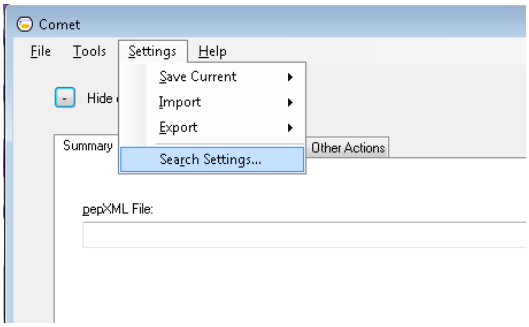
View and Edit Search Settings

All the search parameters and their default values in the Search Settings dialog correspond to Comet search parameters listed on the Comet website. To view/edit the search parameters, open the Search Settings dialog:

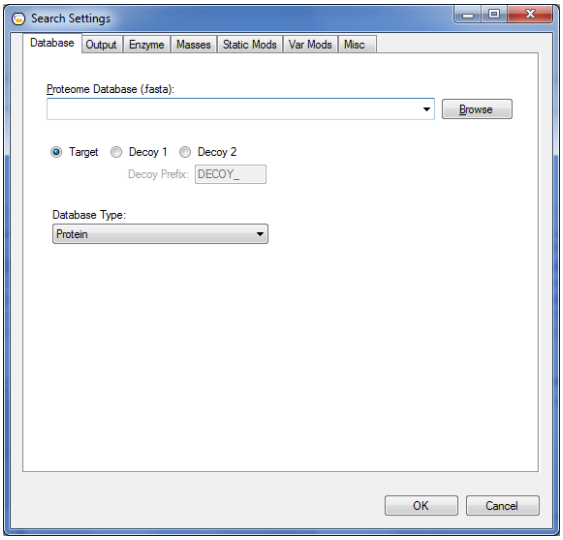
1. Click on the "**Settings**" menu:



2. Choose the "**Search Settings…**" menu item:

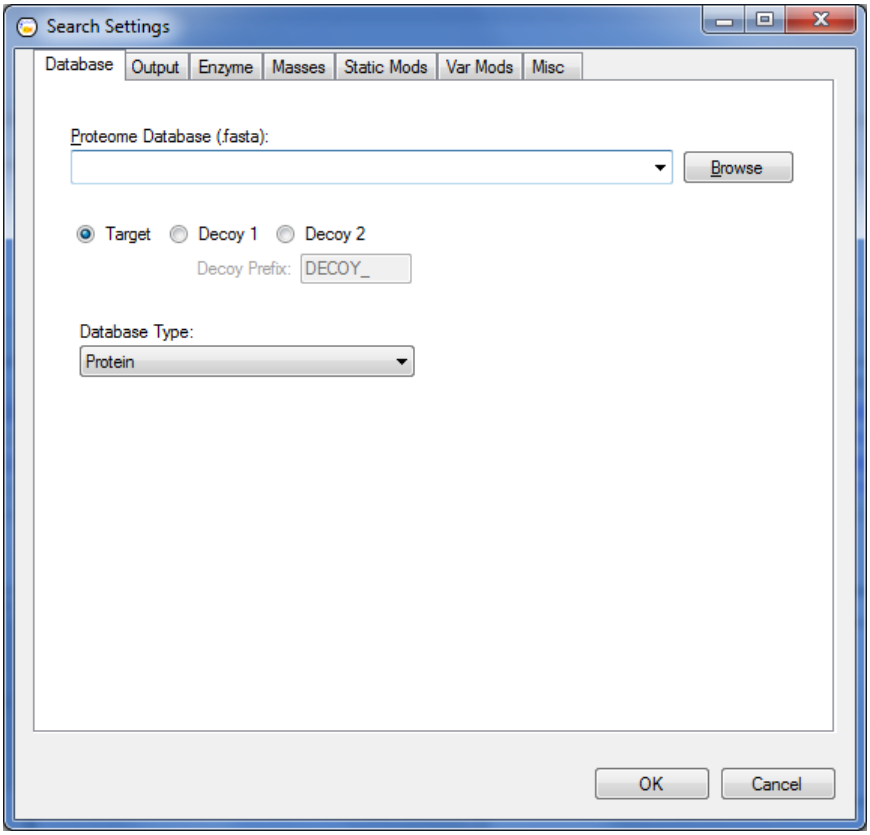


3. The following dialog should open up:



Database Search Settings Tab

The "**Database**" tab on the Search Settings dialog contains the following search parameters related to the database to be used for the Comet search:



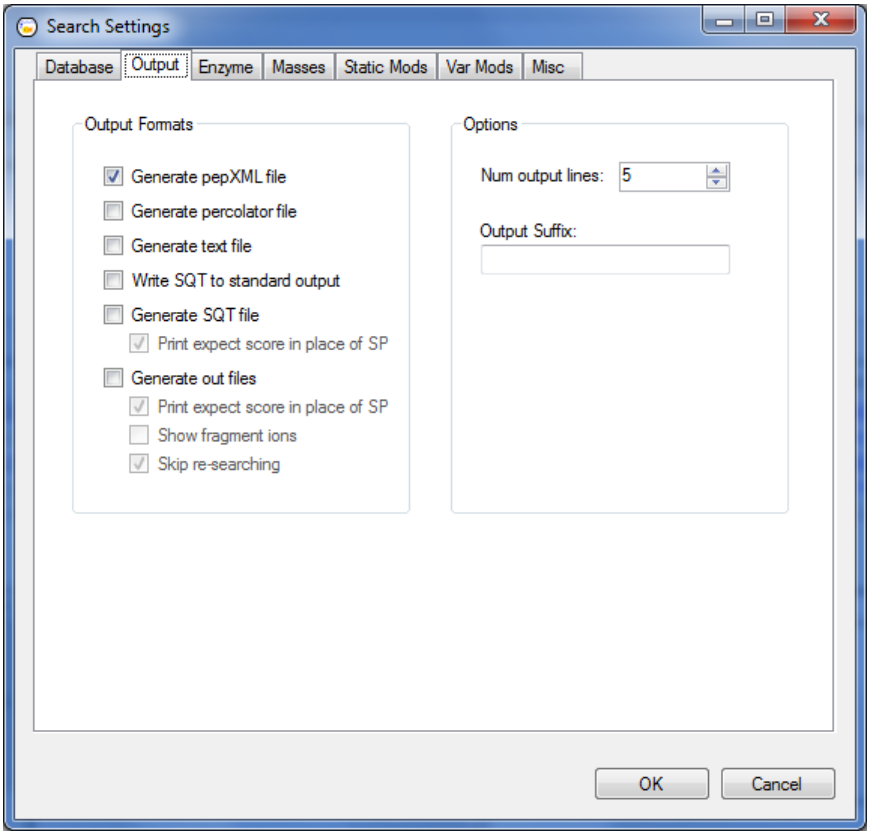
- "**Proteome Database**" edit box corresponds to the "*database_name*" Comet search parameter.

- "**Target/Decoy/Decoy 1**" radio buttons correspond to the "*decoy_search*" parameter, and the "**Decoy Prefix**" edit box corresponding to the "*decoy_prefix*" parameter.

- "**Database Type**" drop-down corresponding to the "*nucleotide_reading_frame*" parameter.

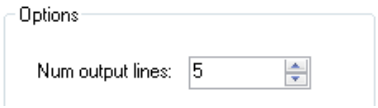
Output Settings Tab

The "**Output**" tab on the Search Settings dialog contains the following search parameters related to the type of output generated by the Comet search:



- The following output format checkboxes correspond to the Comet search parameters indicated:
  - "**Generate pepXML file**" → "*output_pepxmlfile*"
  - "**Generate percolator file**" → "*output_percolatorfile*"
  - "**Generate text file**" → "*output_txtfile*"
  - "**Write SQT to standard output**" → "*output_sqtstream*"
  - "**Generate SQT file**" → "*output_sqtfile*"
  - "**Print expect score in place of SP**" → "*print_expect_score*"
  - "**Generate out files**" → "*output_outfiles*"
  - "**Print expect score in place of SP**" → "*print_expect_score*"
  - "**Show fragment ions**" → "*show_fragment_ions*"
  - "**Skip researching**" →"*skip_researching*"

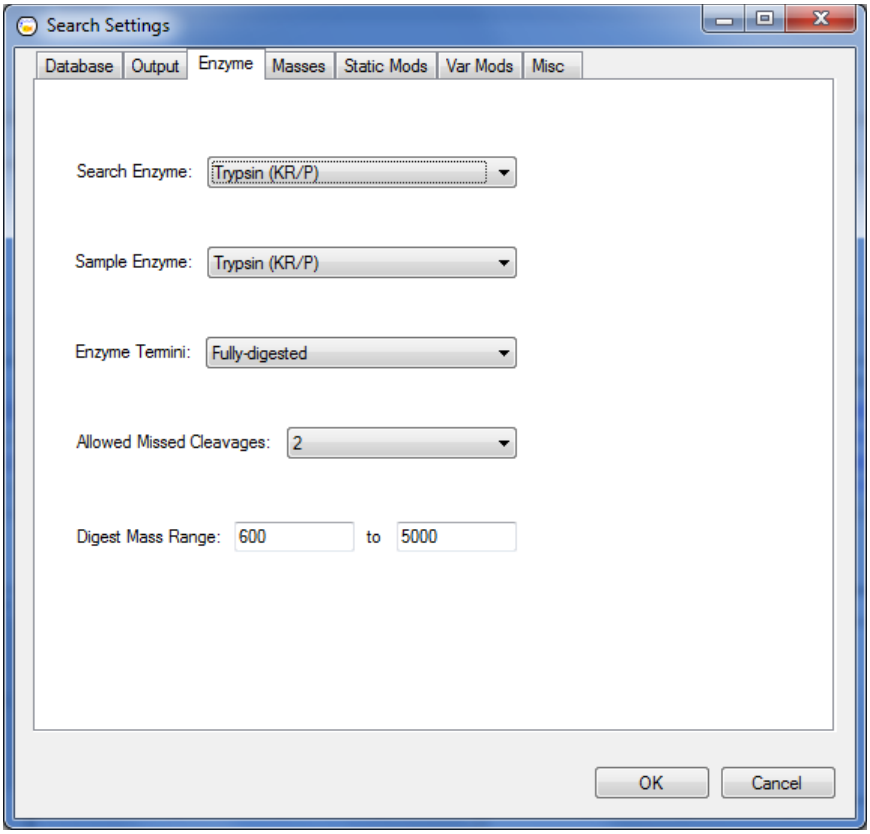
- The "**Num of output lines**" text box corresponds to the "*num_output_lines*" Comet search parameter.



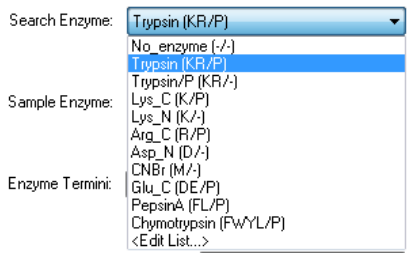
- The "**Output Suffix**" text box corresponds to the "*output_suffix*" Comet search parameter.

Enzyme Settings Tab

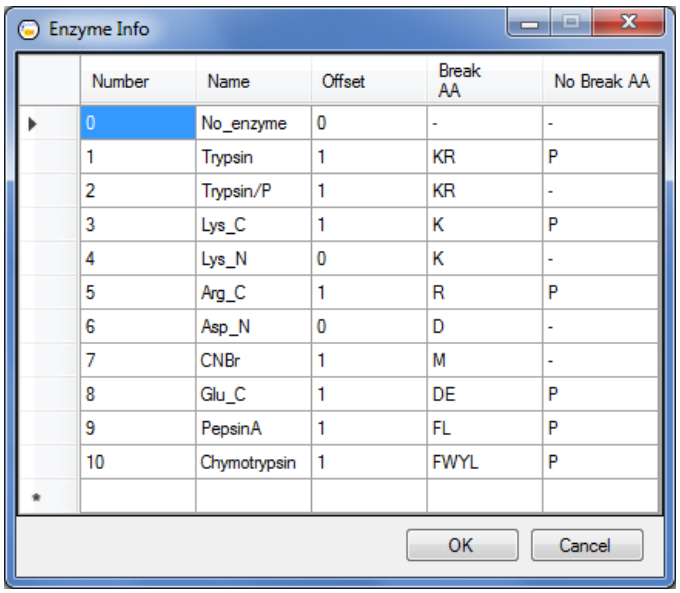
The "**Enzyme**" tab on the Search Settings dialog contains the following search parameters used in the Comet search:



- The "**Search Enzyme**" drop-down corresponds to the "*search_enzyme_number*" Comet search parameter:



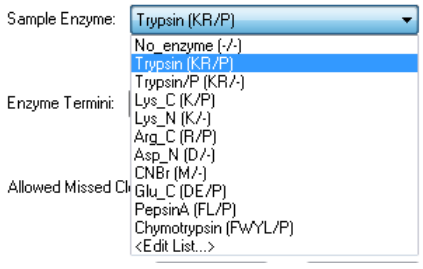
Clicking on the "**<Edit List…>**" item at the bottom of the drop-down list opens the following "**Enzyme Info**" dialog to allow users to edit the enzyme information listed in the drop down:



	Number	Name	Offset	Break AA	No Break AA
▶	0	No_enzyme	0	-	-
	1	Trypsin	1	KR	P
	2	Trypsin/P	1	KR	-
	3	Lys_C	1	K	P
	4	Lys_N	0	K	-
	5	Arg_C	1	R	P
	6	Asp_N	0	D	-
	7	CNBr	1	M	-
	8	Glu_C	1	DE	P
	9	PepsinA	1	FL	P
	10	Chymotrypsin	1	FWYL	P
*					

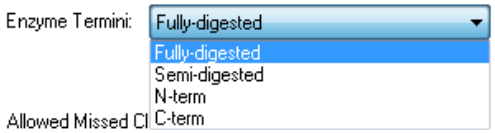
This dialog corresponds to the *"**[COMET_ENZYME_INFO]**"* list at the bottom of Comet ".params" files.

- The "**Sample Enzyme**" drop-down corresponds to the "*sample_enzyme_number*" Comet search parameter:

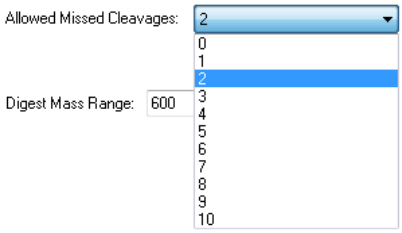


The **"<Edit List…>**" item opens up the same "**Enzyme Info**" dialog as above.

- The "**Enzyme Termini**" drop-down corresponds to the "*num_enzyme_termini*" Comet search parameter:



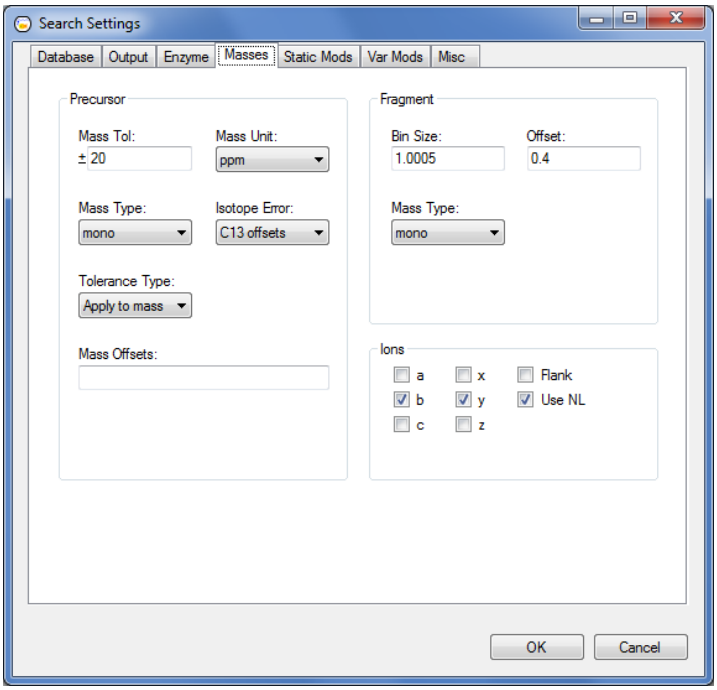
- The "**Allowed Missed Cleavages**" drop-down corresponds to the "*allowed_missed_cleavage*" Comet search parameter:



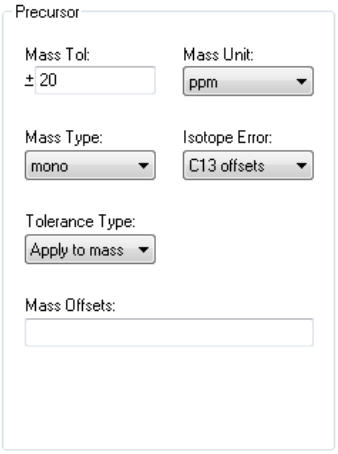
- The "**Digest Mass Range**" edit boxes correspond to the "*digest_mass_range*" Comet search parameter:

Mass Settings Tab

The "**Masses**" tab on the Search Settings dialog contains precursor and fragment mass options used in the Comet search, such as mass units and tolerances:

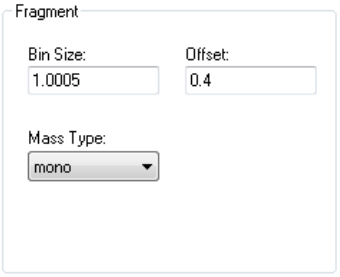


- The following "**Precursor**" mass options correspond to the Comet search parameters indicated:



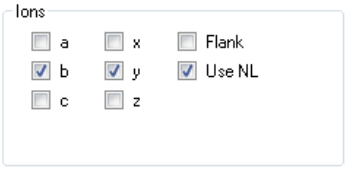
  - "**Mass Tol**" edit box → "*peptide_mass_tolerance*"
  - "**Mass Unit**" drop-down → "*peptide_mass_units*"
  - "**Mass Type**" drop-down → "*mass_type_parent*"
  - "**Isotope Error**" drop-down → "*isotope_error*"
  - "**Tolerance Type**" drop-down → "*precusor_tolerance_type*"
  - "**Mass Offsets**" edit box → "*mass_offsets*"

- The following "**Fragment**" mass options correspond to the Comet search parameters indicated below:



  - "**Offset**" edit box → "*fragment_bin_offset*"
  - "**Mass Type**" drop-down → "*mass_type_fragment*"
  - "**Bin Size**" edit box → "*fragment_bin_tol*"

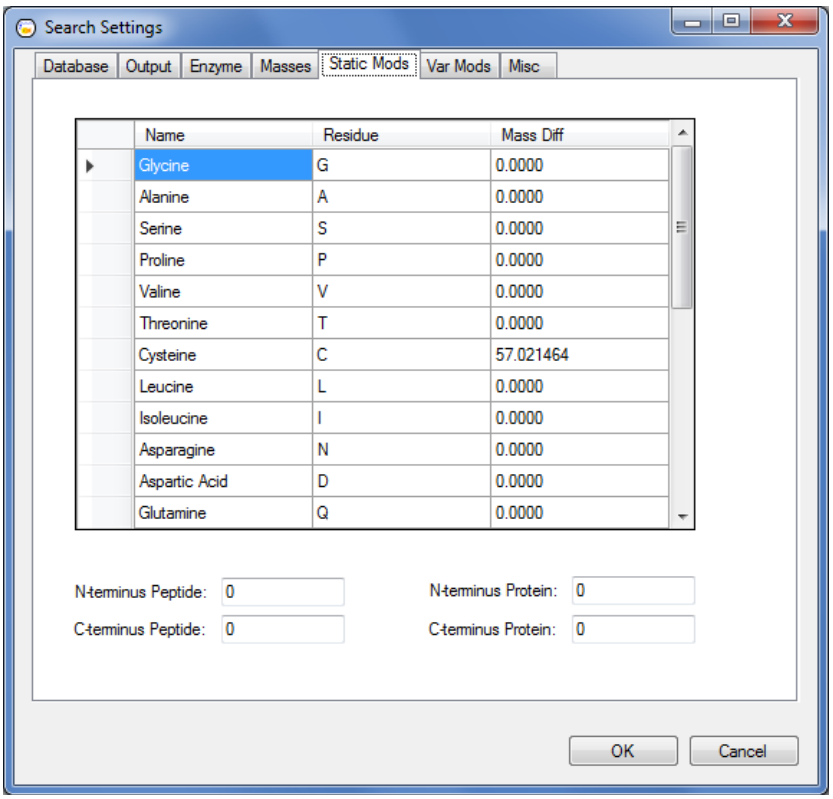
- The following "**Ions**" mass options correspond to the Comet search parameters indicated below:



  - "**a**" check box → "*use_A_ions*"
  - "**b**" check box → "*use_B_ions*"
  - "**c**" check box → "*use_C_ions*"
  - "**x**" check box → "*use_X_ions*"
  - "**y**" check box → "*use_Y_ions*"
  - "**z**" check box → "*use_Z_ions*"
  - "**Use NL**" check box → "*use_NL_ions*"

Static Mods Tab

The "**Static Mods**" tab on the Search Settings dialog contains static modification parameters used by the Comet search:



- The following table items correspond to the Comet search parameters indicated:

Name	Residue	Mass Diff
Glycine	G	0.0000
Alanine	A	0.0000
Serine	S	0.0000
Proline	P	0.0000
Valine	V	0.0000
Threonine	T	0.0000
Cysteine	C	57.021464
Leucine	L	0.0000
Isoleucine	I	0.0000
Asparagine	N	0.0000
Aspartic Acid	D	0.0000
Glutamine	Q	0.0000
Lysine	K	0.0000
Glutamic Acid	E	0.0000
Methionine	M	0.0000
Ornithine	O	0.0000
Histidine	H	0.0000
Phenylalanine	F	0.0000

- "**Glycine**" → "*add_G_glycine*"
- "**Alanine**" → "*add_A_alanine*"
- "**Serine**" → "*add_S_serine*"
- "**Proline**" → "*add_P_proline*"
- "**Valine**" → "*add_V_valine*"
- "**Threonine**" → "*add_T_threonine*"
- "**Cysteine**" → "*add_C_cysteine*"
- "**Leucine**" → "*add_L_leucine*"
- "**Isoleucine**" → "*add_I_isoleucine*"
- "**Asparagine**" → "*add_N_asparagine*"
- "**Aspartic Acid**" → "*add_D_aspartic_acid*"
- "**Glutamine**" → "*add_Q_glutamine*"
- "**Lysine**" → "*add_K_lysine*"
- "**Glutamic Acid**" → "*add_E_glutamic_acid*"
- "**Methionine**" → "*add_M_methionine*"
- "**Ornithine**" → "*add_O_ornithine*"
- "**Histidine**" → "*add_H_histidine*"
- "**Phenylalanine**" → "*add_F_phenylalanine*"

Arginine	R	0.0000	
Tyrosine	Y	0.0000	
Tryptophan	W	0.0000	
User Amino Acid	B	0.0000	
User Amino Acid	J	0.0000	
User Amino Acid	U	0.0000	
User Amino Acid	X	0.0000	
User Amino Acid	Z	0.0000	

- "**Arginine**" → "*add_R_arginine*"
- "**Tyrosine**" → "*add_Y_tyrosine*"
- "**Tryptophan**" → "*add_W_tryptophan*"
- "**User Amino Acid (B)**" → "*add_B_user_amino_acid*"
- "**User Amino Acid (J)**" → "*add_J_user_amino_acid*"
- "**User Amino Acid (U)**" → "*add_U_user_amino_acid*"
- "**User Amino Acid (X)**" → "*add_X_user_amino_acid*"
- "**User Amino Acid (Z)**" → "*add_Z_user_amino_acid*"

- The "**N-terminus Peptide**" text box corresponds to the "*add_Nterm_peptide*" parameter:

  N-terminus Peptide: 0

- The "**C-terminus Peptide**" text box corresponds to the "*add_Cterm_peptide*" parameter:

  C-terminus Peptide: 0

- The "**N-terminus Protein**" text box corresponds to the "*add_Nterm_protein*" parameter:

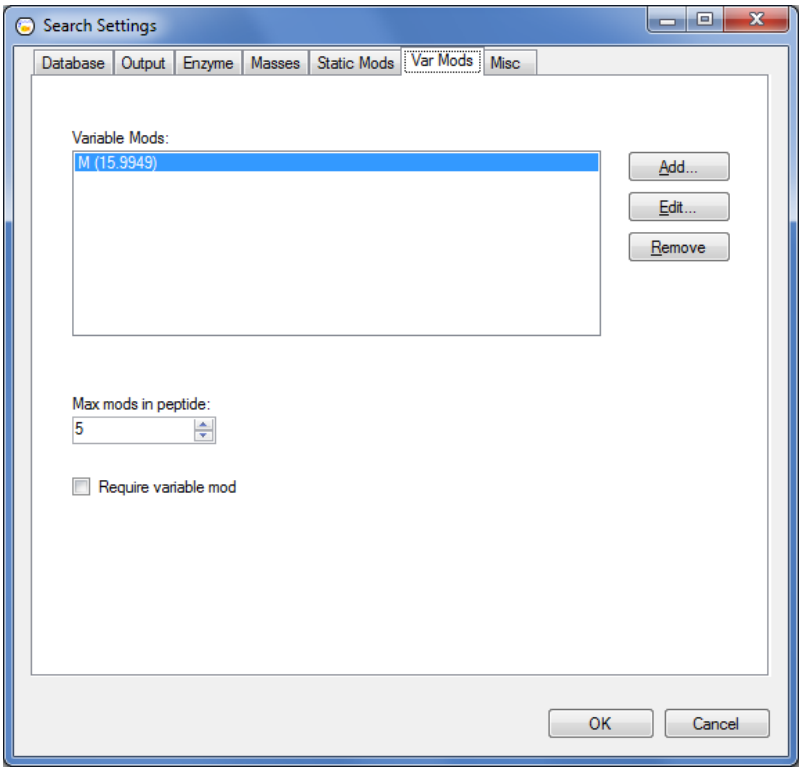
  N-terminus Protein: 0

- The "**C-terminus Protein**" text box corresponds to the "*add_Cterm_protein*" parameter:

  C-terminus Protein: 0

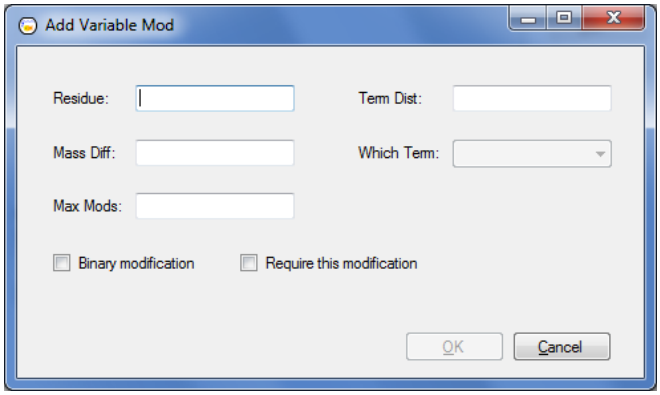
Var Mods Tab

The "**Var Mods**" tab on the Search Settings dialog contains variable modification parameters used by the Comet search:



- The items in the "**Variable Mods**" list below correspond to the Comet variable mod parameters "*variable_mod0N*", where N ranges from 0-9.

To add a variable modification, click on the "**Add…**" button on the right to open up the "**Add Variable Mod**" dialog below, and use this dialog to specify the various fields of the variable modification (these variable mod fields are described in more detail on the Comet website):



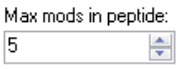
- The "**Residue**" text box corresponds to the second field on the variable mod parameter – the residue(s) that the modifications are possibly applied to.
- The "**Mass Diff**" text box corresponds to the first field on the variable mod parameter – specify the modification mass difference.
- The "**Max Mods**" text box corresponds to the fourth field on the variable mod parameter – specify the maximum number of modified residues possible in a peptide for this modification entry.
- The "**Binary modification**" check box corresponds to the third field on the variable mod parameter – specify whether the modification is a variable modification or a binary one.
- The "**Require this modification**" check box corresponds to the seventh field on the variable mod parameter – specify whether peptides must contain this modification.
- The "**Term Dist**" text box corresponds to the fifth field on the variable mod parameter – specify the distance the modification is applied to from the respective protein terminus.
- The "**Which Term**" drop-down corresponds to the sixth field on the variable mod parameter – specify which protein terminus the distance constraint is applied to.

To edit a variable mod, go back to the "**Variable Mods**" tab, select the variable mod item in the "**Variable Mods**" list and click on the "**Edit…**" button to open up the "**Edit Variable Mod**" dialog,

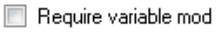
which is virtually identical to the "Add Variable Mod" dialog above, except with the fields populated to reflect the variable mod selected from the list.

To remove a variable mod from the "**Variable Mods**" list, just select the item and click on the "**Remove**" button.

- The "**Max Mods in Peptide**" text box corresponds to the "*max_variable_mods_in_peptide*" parameter:

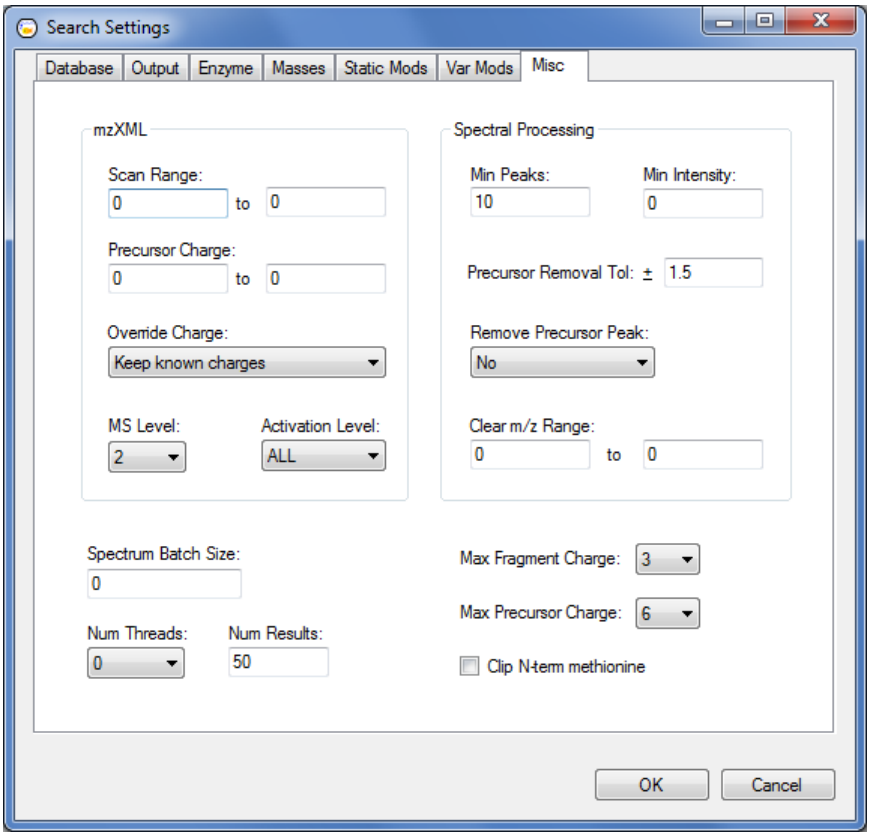


- The "**Require variable mod**" check box corresponds to the "*require_variable_mod*" parameter:

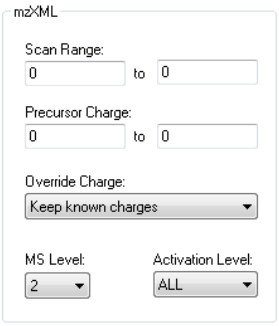


Misc Tab

The "**Misc**" tab on the Search Settings dialog contains miscellaneous Comet search parameters:

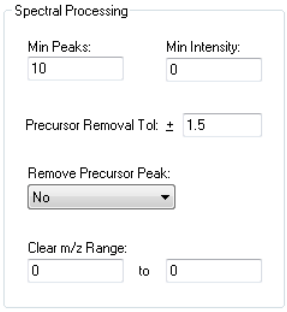


- The following "**mzXML**" fields correspond to the following Comet search parameters:



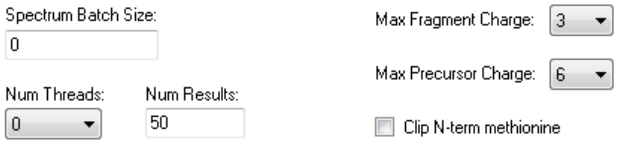
  - "**Scan Range**" text boxes → "*scan_range*"
  - "**Precursor Charge**" text boxes → "*precursor_charge*"
  - "**Override Charge**" drop-down → "*override_charge*"
  - "**MS Level**" drop-down → "*ms_level*"
  - "**Activation Level**" text boxes → "*activation_method*"

- The following "**Spectral Processing**" fields correspond to the following Comet search parameters:



  - "**Min Peaks**" text box → "*minimum_peaks*"
  - "**Min Intensity**" text box → "*minimum_intensity*"
  - "**Precursor Removal Tol**" text box → "*remove_precursor_tolerance*"
  - "**Remove Precursor Peak**" drop-down → "*remove_precursor_peak*"
  - "**Clear m/z Range**" text boxes → "*clear_mz_range*"

- The rest of the miscellaneous parameters correspond to the following Comet search parameters:

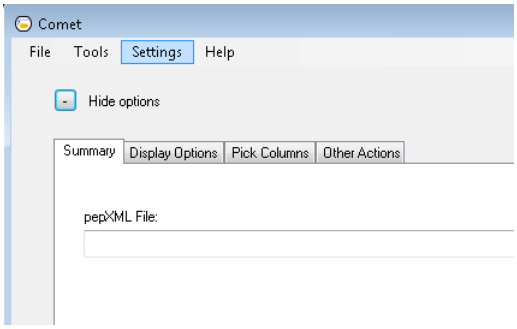


  - The "**Spectrum Batch Size**" text box corresponds to the "*spectrum_batch_size*" parameter.
  - The "**Num Threads**" drop-down corresponds to the "*num_threads*" parameter.
  - The "**Num Results**" text box corresponds to the "*num_results*" parameter.
  - The "**Max Fragment Charge**" drop-down corresponds to the "*max_fragment_charge*" parameter.
  - The "**Max Precursor Charge**" drop-down corresponds to the "*max_precursor_charge*" parameter.
  - The "**Clip N-term methionine**" checkbox corresponds to the "*clip_nterm_methionine*" parameter.

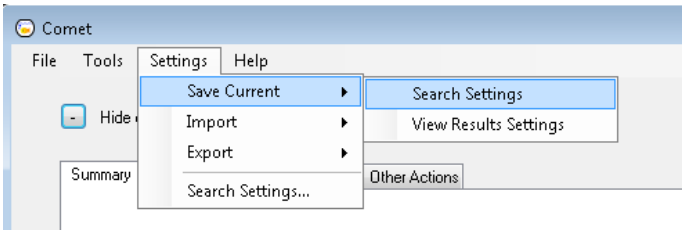
Save Search Settings

To save any changes made to the search settings:

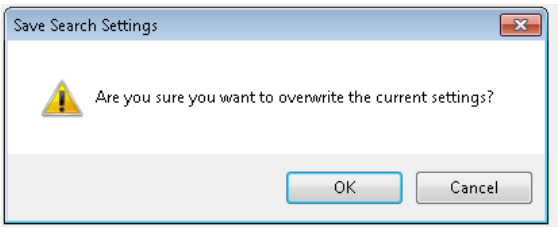
1. Click on the "**Settings**" menu:



2. Choose the "**Save Current… → Search Settings**" menu item:

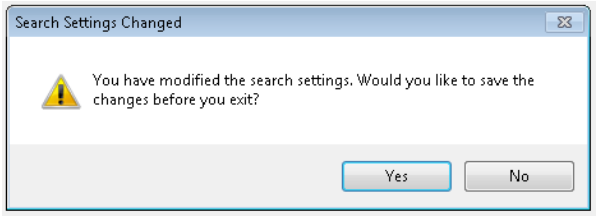


3. Click on the "**OK**" button on the confirmation dialog that pops up:



Now, if the CometUI application is rebooted, the changes made to the search settings will be preserved.

Alternatively, if changes were made to the search settings, and the user tries to close the CometUI application without saving these changes, the following dialog will appear giving the user a chance to save the changes by clicking on the "**Yes**" button.



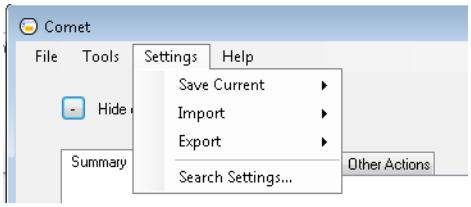
Import/Export Search Settings

The CometUI application allows users to import Comet search settings from a ".params" Comet parameter file, as well as export the settings in the UI to a ".params" file.

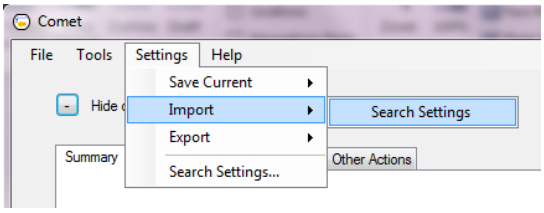
Import Search Settings

To import the search settings:

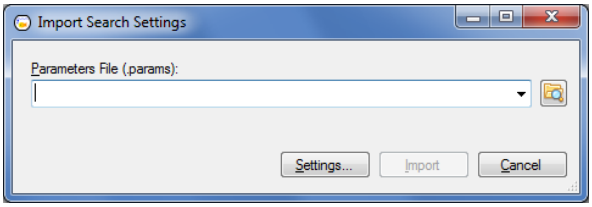
1. Click on the "**Settings**" menu:



2. Choose the "**Import → Search Settings**" menu item:

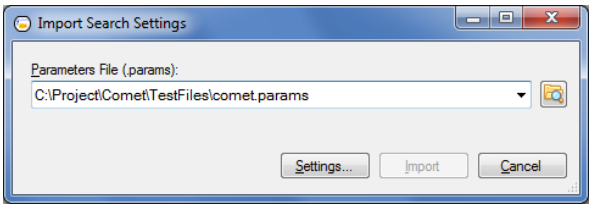


3. The following "**Import Search Settings**" dialog should open up:



Note: The "**Settings…**" button can be used to view the current settings in CometUI.

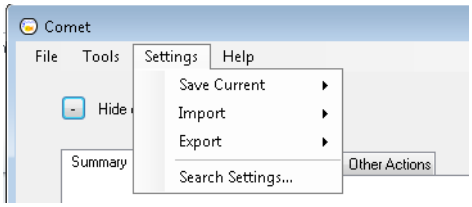
4. Use the folder browse button on this dialog to browse to the "*.params*" file to be imported, and click on the "**Import**" button:



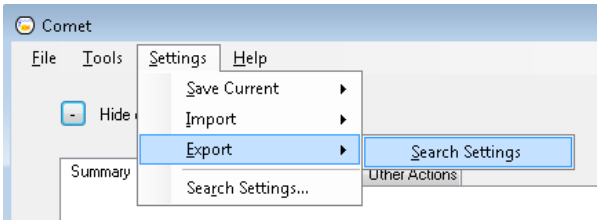
Export Search Settings

To export the search settings:

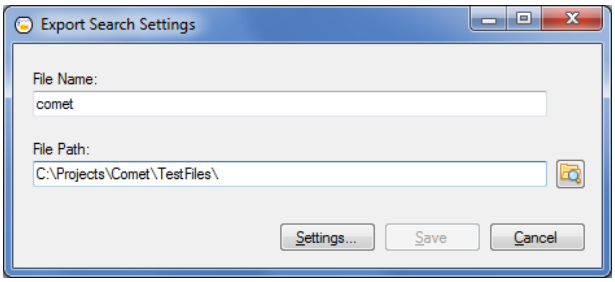
1. Click on the "**Settings**" menu:



2. Choose the "**Export → Search Settings**" menu item:

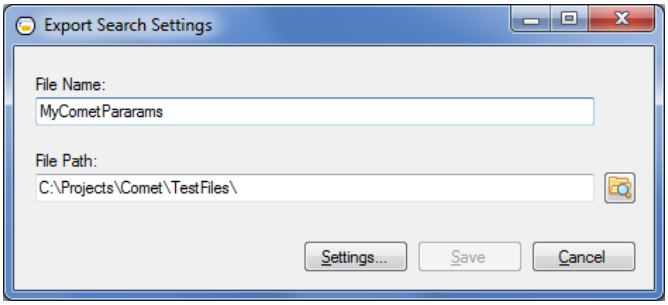


3. The following "**Export Search Settings**" dialog should open up:



Note: The "**Settings…**" button can be used to view/edit the current settings.

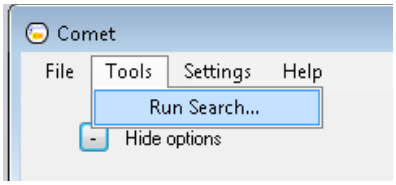
4. Choose a "**File Name**" (default is "comet"), a "**File Path**" (default is the current directory the CometUI.exe is running in) and click on the "**Save**" button, and the exported ".params" file with the specified name will get saved in specified location:



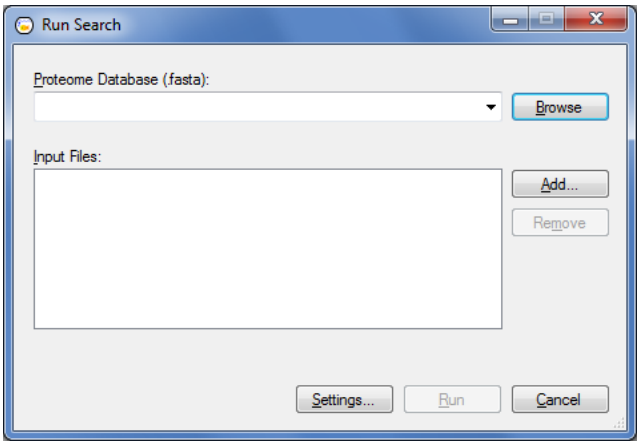
Run Search

CometUI provides the user with a graphical user interface to run the Comet search engine. To run a search:

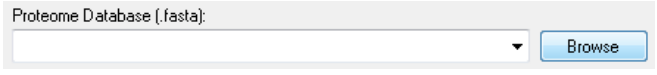
1. Click on the "Tools" menu and select the "Run Search…" menu item:



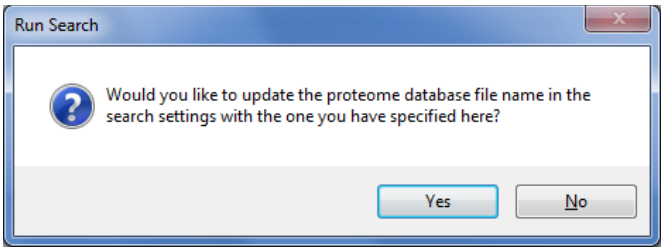
2. This will bring up the following window:



3. Specify a "**Proteome Database (.fasta)**" file using the "**Browse**" button:



If the file specified is different than the one in the search settings, the following message box will pop up:

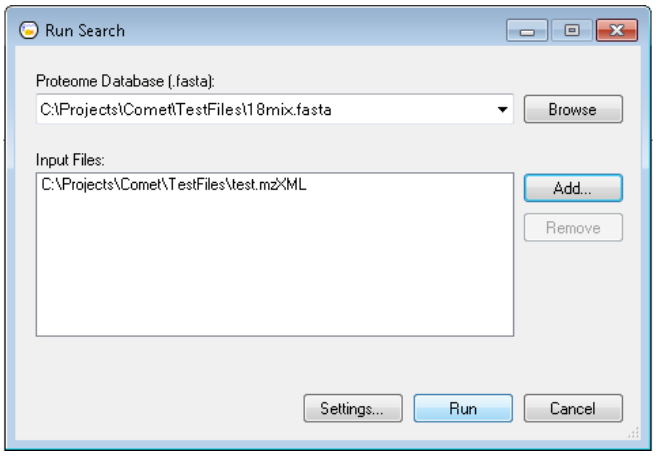


Click "**OK**" to update the proteome database file name in the search settings to this file, or "**Cancel**" to leave the search settings unchanged.

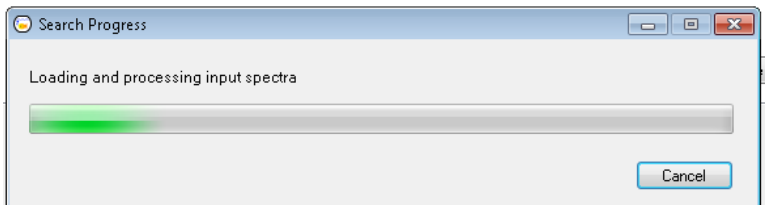
4. Next, click on the "**Add…**" button next to the "**Input Files**" list box to specify an input file (or multiple input files):

Formats currently supported are: ".mgf", ".mzxml", ".mzml", ".ms2", ".cms2" and ".raw".

5. Once at least one input file has been added, the "**Run**" button will become active:

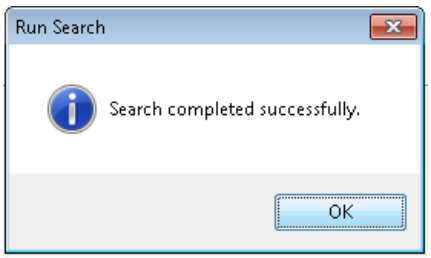


6. Clicking the "**Run**" button will launch the search, and a "**Search Progress**" dialog should appear:

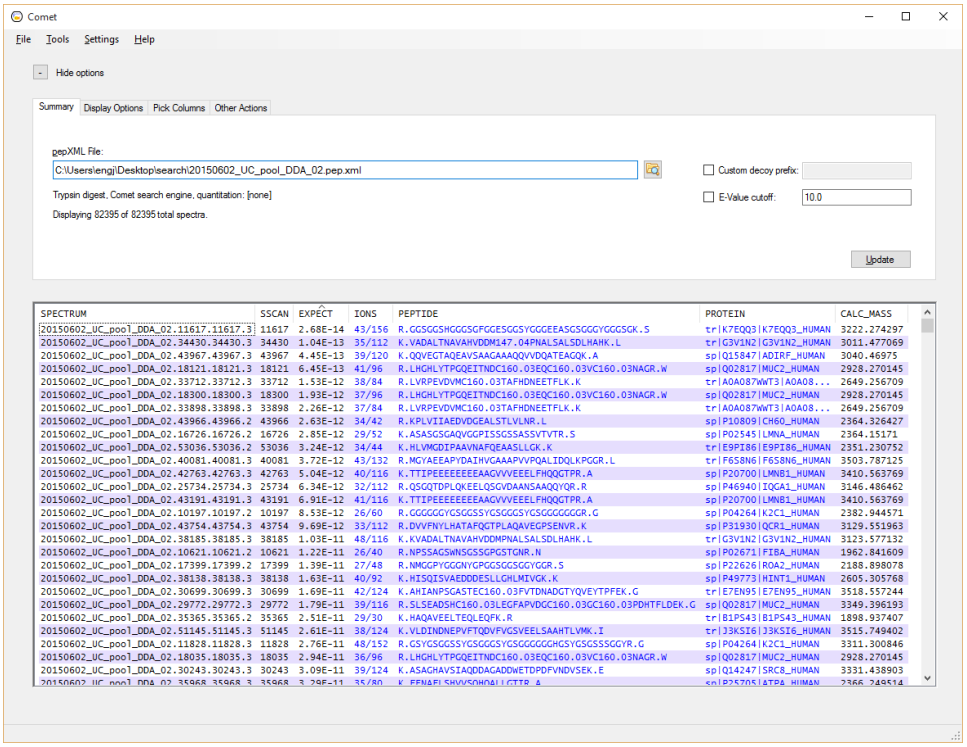


7. If the search completes successfully, the notification below should appear, and clicking "**OK**" on it should bring up the results automatically in the results viewer, which will be discussed next.

(Note that if multiple input files were specified, only the results of the FIRST file in the list will be displayed in the results viewer.)



8. The results in the results viewer may look something like the following:



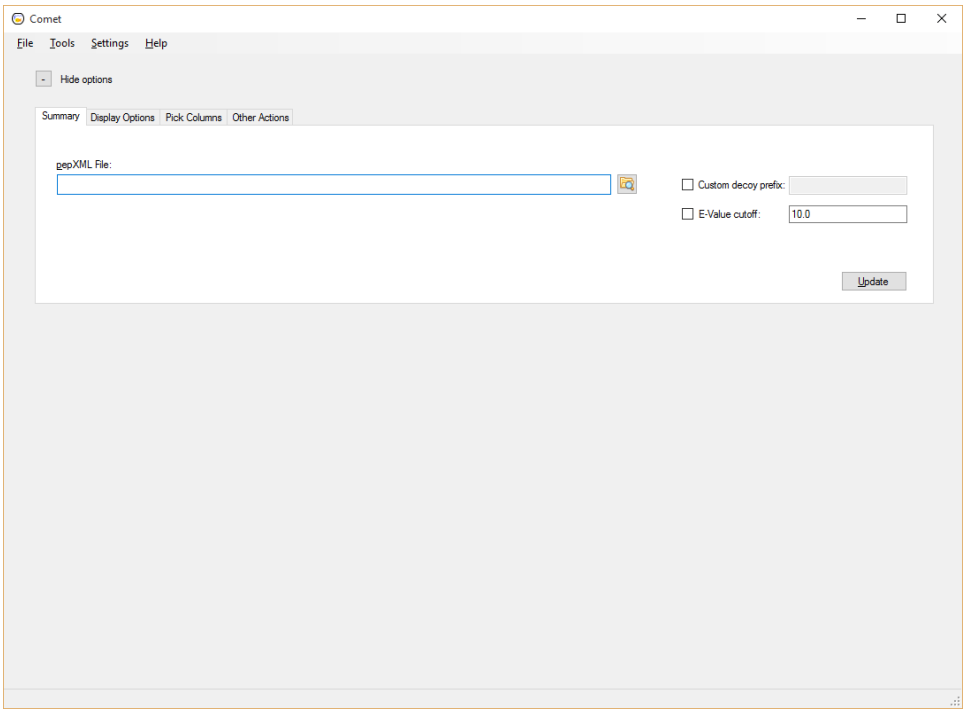
View Results

CometUI allows users to visualize Comet search results. Currently, only Comet searches generating a "*pep.xml*" file can be viewed in the results viewer.

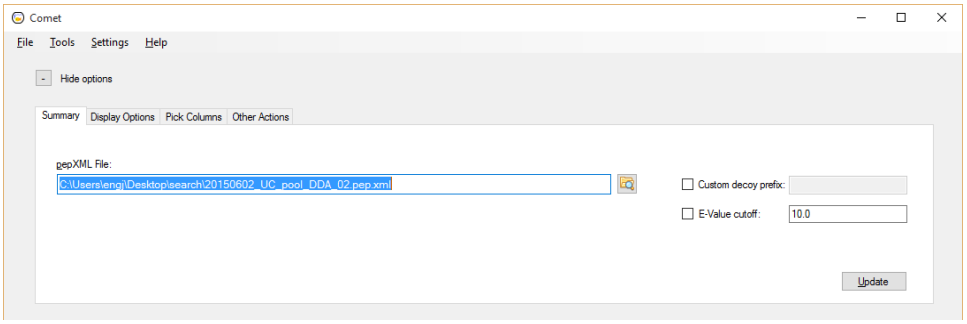
Note that results from a Comet search run via the CometUI will load automatically once the search completes successfully, provided a "*pep.xml*" file was specified as one of the output formats. (Please see the previous section, "Run Search", for details on running a Comet search from the UI.)

Users may view a previously generated "*pep.xml*" file by following the steps below:

1. Launch CometUI to open up the main window:

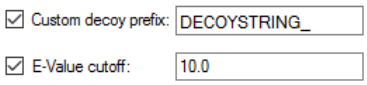


2. In the "**Summary**" tab, click the button with the folder icon next to the "**pepXML File**" edit box to browse to a "*pep.xml*" file you want to open (or simply type in the path to the file):

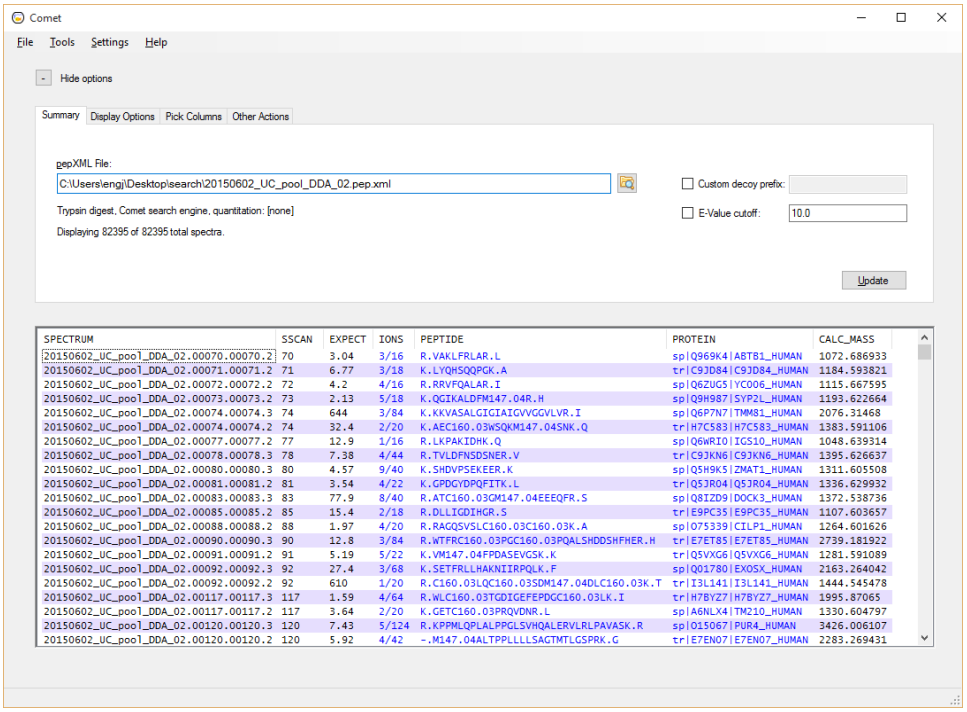


3. If the Comet search was run with a custom decoy prefix (other than what is specified in the "*decoy_prefix*" Comet parameter), check the "**Custom decoy prefix**" checkbox, and specify the decoy prefix string in the text box next to it.

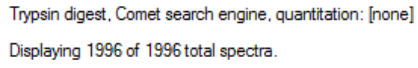
   Select the E-Value cutoff checkbox and enter a cutoff value to apply an expectation value filter to the results.



4. Click on the "**Update**" button , and the results should appear in a list right below the "**Summary**" tab as follows:

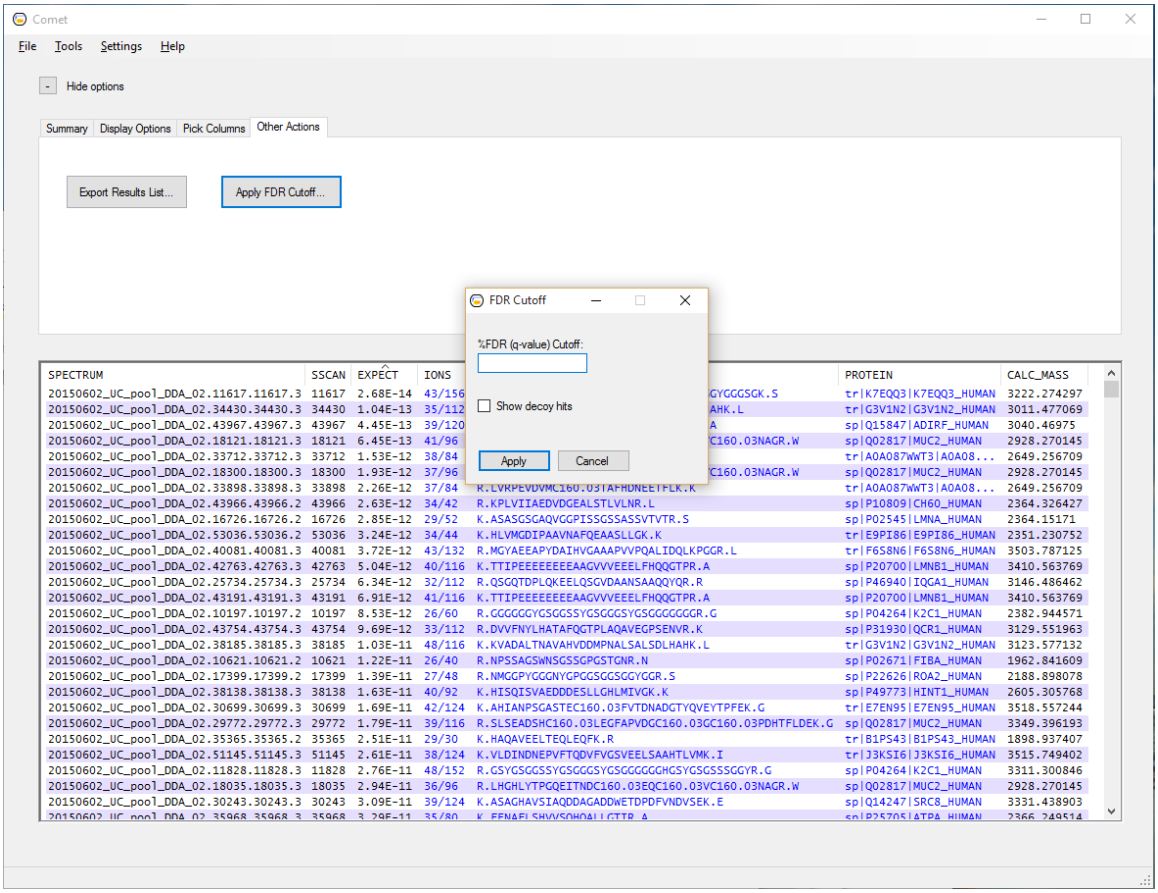


5. A summary of information about the results will appear right below the "pepXML File" edit box:



From here on, users can interact with the search results list in several ways, such as clicking on the column headers to sort the list by any of the columns, or clicking on a value in an interactive column with hyperlinks in it (e.g. **Ions**, **Peptide** and **Protein** columns). Users can also manipulate the list by using options available in the "**Display Options**", "**Pick Columns**" and "**Other Actions**" tabs.

If the search results were from a target-decoy search, a false discovery rate (FDR) analysis can be performed at the PSM level. Click on the “Other Actions” tab and “Apply FDR Cutoff” button to bring up the FDR cutoff dialogue. Specify the FDR cutoff, e.g. enter “5” for a 5% cutoff then hit the “Apply” button. This will filter the identifications at that FDR cutoff. To include the decoy hits in the list, select the “Show decoy hits” checkbox.



The links in the “IONS” column will display the MS/MS spectrum. The links in the “PEPTIDE” column will paste that peptide into an NCBI BLAST page in your browser. Finally, the links in the “PROTEIN” column will display the protein sequence with the identified peptide highlighted.