**Exercises – Spectral Library Building using SpectraST**

*These exercises take you through the process of building a spectral library from sequence search results from the yeast SILAC dataset we have been using.*

* Create RAW spectral library
* Create consensus spectral library
* Generate quality-filtered SRM assay library
* Explore library building results

**Task 1. Create a RAW spectral library from confidently identified spectra**

* Open Petunia interface, Select **Files**, then navigate to *C:\TPP\data\class\Yeast\iprophet\*. Select interact.ipro.pep.xml, then **Copy**
* Navigate to *C:\TPP\data\*, select both mzML files, then **Copy**.

[only if necessary]

* Navigate to *C:\TPP\data\*class*\Yeast\spectrast\* and create a new directory called ‘speclibs’ Click link and **Set as Working Directory**
* Click **Paste** to place the input files in your working directory
* Select **TPPTools->SpectraST Tools->SpectraST Library Import** link.
	+ Specify ‘.pepXML’ as **File Format**
	+ Under **Choose Files**, click Add Files and select the pep.xml file
	+ Under **General Options**, enter Yeast\_raw as name of output file, and Course as dataset identifier
	+ Click on **Import Library Files**

 As the job starts, you will be redirected to the job output page, where you can see the command. Building the raw library should take 5-10 minutes. SpectraST scans through all the identifications in the pepXML file, selects only those above the probability threshold, then extracts the corresponding query spectrum from the source mzML file(s), and saves each one to the library.

SpectraST started at Fri Sep 02 11:24:50 2016.

Processing "c:/interact.ipro.pep.xml"...500...1000... ..13000...DONE!

Failed to open input file 'c:/TPP/data/class/Yeast/spectrast/speclibs/OR20080317\_S\_SILAC-LH\_1-1\_01.mzXML'.

10%...20%...30%...40%...50%...60%...70%...80%...90%...DONE!

Total number of spectra in library: **13481**

Total number of distinct peptide ions in library: **5589**

Total number of distinct stripped peptides in library: **2797**

CHARGE +1: 0 ; +2: 7144 ; +3: 5431 ; +4: 848 ; +5: 58 ; >+5: 0 ; Unk: 0

TERMINI Tryptic: 12952 ; Semi-tryptic: 529 ; Non-tryptic: 0

PROBABILITY >0.9999: 12218 ; 0.999-0.9999: 553 ; 0.99-0.999: 393 ; 0.9-0.99: 317 ; <0.9: 0

NREPS 20+: 0 ; 10-19: 0 ; 4-9: 0 ; 2-3: 0 ; 1: 13481

MODIFICATIONS C,Carbamidomethyl: 1019 ; C,Pyro-carbamidomethyl: 23

 E,Glu->pyro-Glu: 7

 K,Label:13C(6)15N(2): 4424

 M,Oxidation: 162

 Q,Gln->pyro-Glu: 187

**Task 2. Create a CONSENSUS spectral library from the RAW library**

* Select **TPPTools->SpectraST Tools->SpectraST Library Import** link.
	+ Specify ‘.splib’ as **File Format**
	+ Navigate to the speclibs directory [if necessary]
	+ Under **Choose Files**, select the checkbox next to the pep.xml file and **Remove.** Click Add Files, select Yeast\_raw.splib, and **Select**.
	+ Choose consensus as build action
	+ Under **General Options**, enter Yeast\_consensus as name of output file.
	+ Click on **Import Library Files**

SpectaST is now combining spectra identified for each peptide ion and will create a “consensus” spectrum for each peptide ion that has multiple replicates in the raw library raw.splib. For peptide ions with only a single replicate, SpectraST will also include them in the final library after some spectrum processing.

. There are far fewer ions, and no need to pull data from the mzML files, so this step is much quicker.

SpectraST started at Fri Sep 02 11:47:33 2016.

Creating CONSENSUS library from "c:/TPP/data/class/Yeast/spectrast/speclibs/Yeast\_raw.splib"

Importing ions...500...1000...1500...2000...2500...3000...3500...4000...4500...5000...5500...DONE!

Library file (BINARY) "Yeast\_consensus.splib" created.

Library file (TEXT) "Yeast\_consensus.sptxt" created.

M/Z Index file "Yeast\_consensus.spidx" created.

Peptide Index file "Yeast\_consensus.pepidx" created.

Total number of spectra in library: 5589

Total number of distinct peptide ions in library: 5589

Total number of distinct stripped peptides in library: 2797

CHARGE +1: 0 ; +2: 2991 ; +3: 2157 ; +4: 403 ; +5: 38 ; >+5: 0 ; Unk: 0

TERMINI Tryptic: 5298 ; Semi-tryptic: 291 ; Non-tryptic: 0

PROBABILITY >0.9999: 5038 ; 0.999-0.9999: 240 ; 0.99-0.999: 167 ; 0.9-0.99: 144 ; <0.9: 0

NREPS 20+: 46 ; 10-19: 78 ; 4-9: 557 ; 2-3: 2800 ; 1: 2108

**Task 3. Apply quality filters and extract SRM assays from consensus library**

* Select **TPPTools->SpectraST Tools->SpectraST Library Import** link.
	+ Specify ‘.splib’ as **File Format**
	+ Navigate to the speclibs directory [if necessary]
	+ Under **Choose Files**, select the checkbox next to Yeast\_raw.splib and **Remove.** Click Add Files, select Yeast\_consensus.splib, and **Select**.
	+ Choose quality filter as build action
	+ Under **General Options**, enter Yeast\_Q2\_mrm as name of output file.
	+ Set Quality Level to Remove to 2:
	+ Add the following to additional parameters box: -cM -cQ16
	+ Click on **Import Library Files**

This command will also complete very quickly. SpectraST is evaluating each consensus, and discarding those which don’t meet the quality standard. It will also produce a tab-delimited SRM assay text file, which can be used to do targeted proteomics.

Level 1: Remove impure spectra

Level 2: Level 1 + spectra that have a spectrally similar counterpart in the library with a conflicting identification

Level 3: Level 2 + spectra whose peptide sequence has no shared sub-sequence with any other peptides in the library

Level 4: Level 3 + singleton spectra

Level 5: Level 4 + inquorate spectra (with a user-defined quorum)

SpectraST started at Fri Sep 02 13:09:33 2016.

Applying QUALITY FILTER to "c:/TPP/data/class/Yeast/spectrast/speclibs/Yeast\_consensus.splib".

Importing peptide ions...500...1000...1500...2000...2500...3000...3500...4000...4500...5000...5500...DONE!

Library file (BINARY) "Yeast\_consensus\_Q2.splib" created.

Library file (TEXT) "Yeast\_consensus\_Q2.sptxt" created.

M/Z Index file "Yeast\_consensus\_Q2.spidx" created.

Peptide Index file "Yeast\_consensus\_Q2.pepidx" created.

MRM Table file "Yeast\_consensus\_Q2.mrm" created.

Total number of spectra in library: 5438

Total number of distinct peptide ions in library: 5438

Total number of distinct stripped peptides in library: 2723

CHARGE +1: 0 ; +2: 2908 ; +3: 2095 ; +4: 397 ; +5: 38 ; >+5: 0 ; Unk: 0

TERMINI Tryptic: 5160 ; Semi-tryptic: 278 ; Non-tryptic: 0

PROBABILITY >0.9999: 5038 ; 0.999-0.9999: 240 ; 0.99-0.999: 91 ; 0.9-0.99: 69 ; <0.9: 0

NREPS 20+: 46 ; 10-19: 78 ; 4-9: 557 ; 2-3: 2795 ; 1: 1962

**Task 4. Investigate library building success**

* A. First we will generate HTML view of a spectral library
* Select **TPPTools->SpectraST Tools->Lib2HTML** link.
	+ Remove any existing libraries, select Yeast\_Q2\_mrm.splib
	+ Choose quality filter as build action
* B. While that is running, click on **Files** and navigate to speclibs
	+ Click on the **View** link by spectrast.log
	+ Search for the word ‘Start’ with your browser (cntl-F)
* C. Click on the **View** link next to Yeast\_Q2\_mrm.splib
	+ Note how few transitions there are per entry
	+ Read header info for n\_replicates
	+ Note that almost all peaks are y/b ions