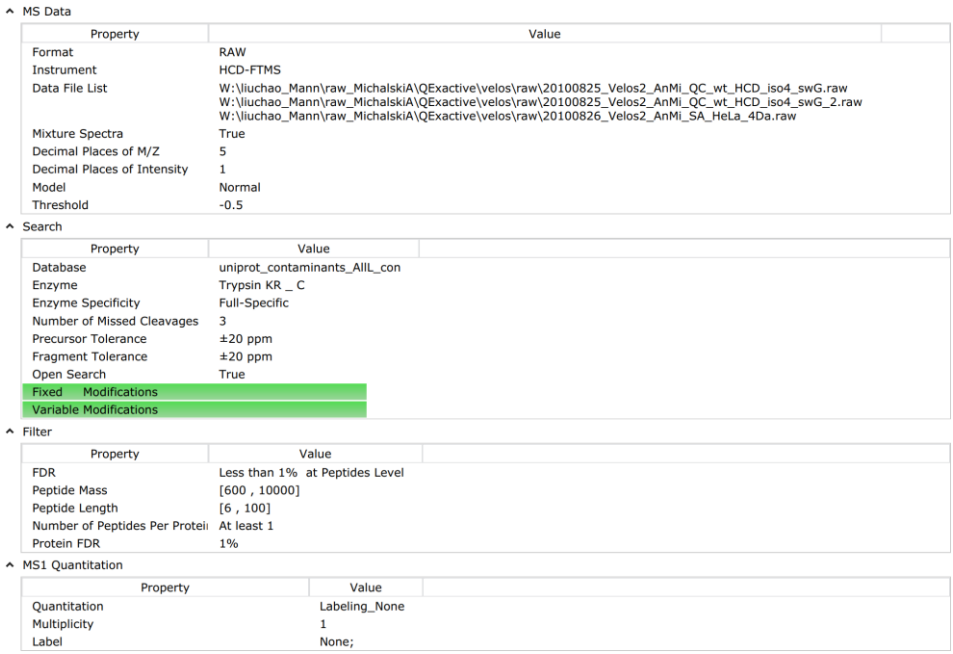
Supplementary Note

1. Dataset annotation

In this section, we provide the detailed parameter settings for dataset pre-processing.

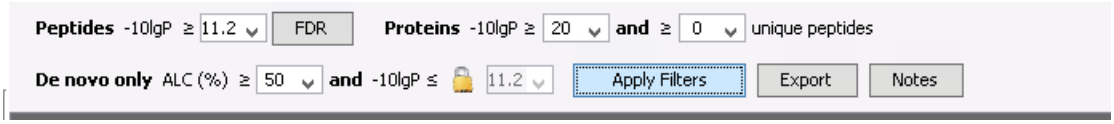
Open-pFind:

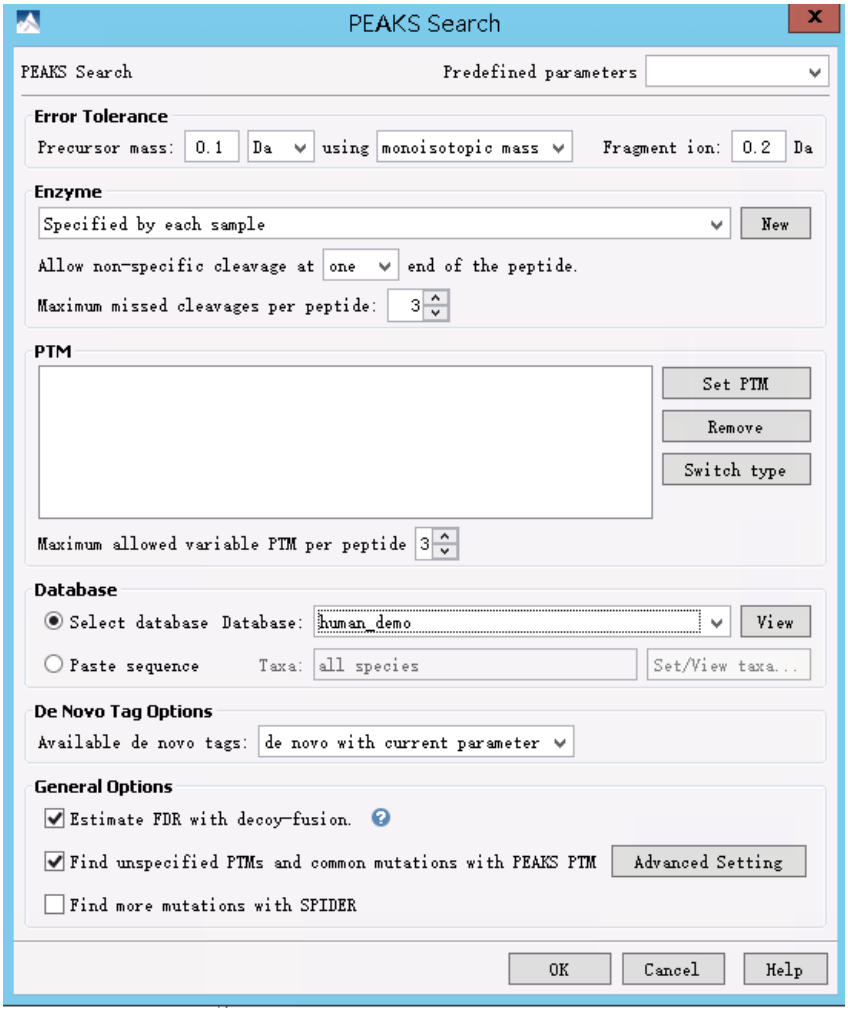
1) Tool download: http://pfind.ict.ac.cn/software/pFind3/index.html
2) Parameter setting interface:
   (Please note that no screenshots are all set by default)



PEAKS:

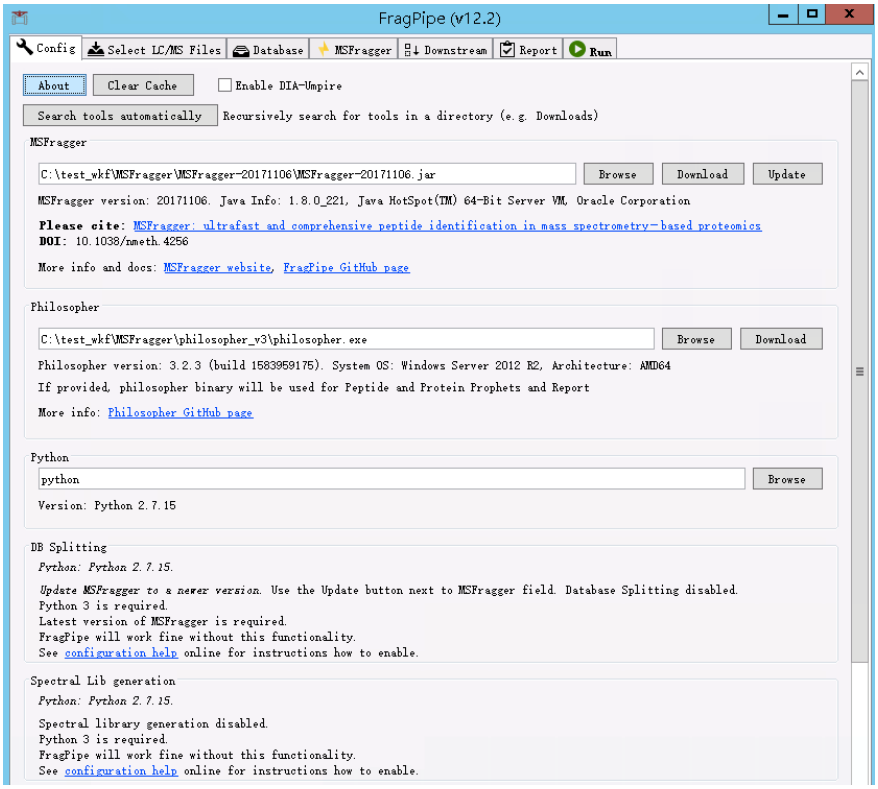
1) Tool download: https://www.bioinfor.com/download-peaks-studio/
2) Parameter setting interface:

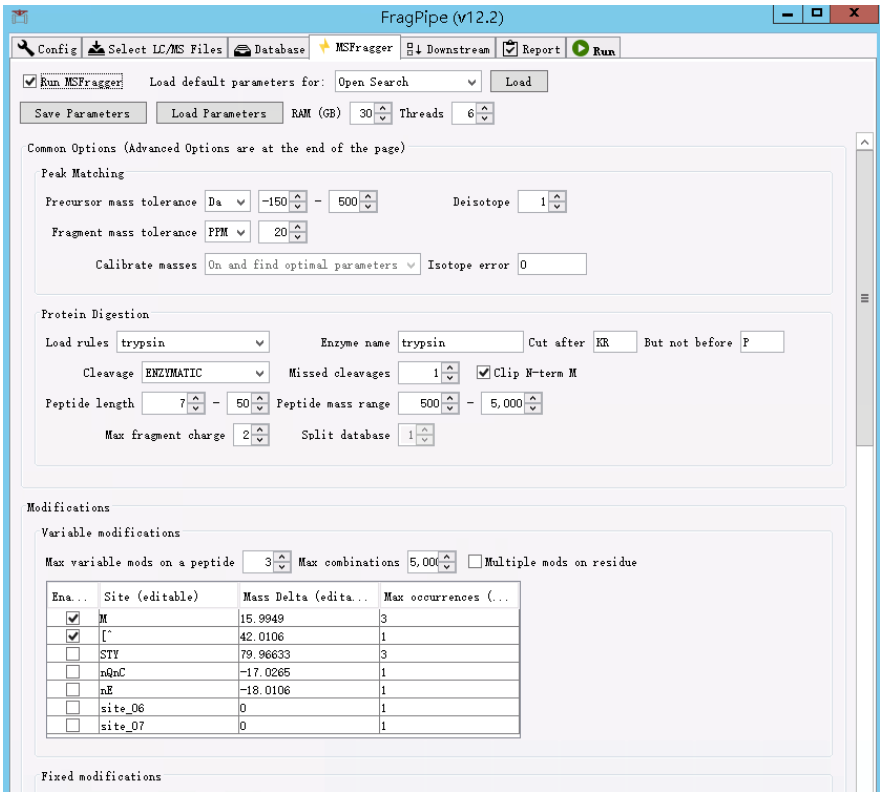




MSFragger:

1) Tool download: http://msfragger.nesvilab.org/
2) Parameter setting interface:


FragPipe (v12.2)
Config
Select LC/MS Files
Database
MSFragger
Downstream
Report
Run
About
Clear Cache
Enable DIA-Umpire
Search tools automatically
Recursively search for tools in a directory (e.g. Downloads)
MSFragger
C:\test_wkf\MSFragger\MSFragger-20171106\MSFragger-20171106.jar
Browse
Download
Update
MSFragger version: 20171106. Java Info: 1.8.0_221, Java HotSpot(TM) 64-Bit Server VM, Oracle Corporation
Please cite: MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics
DOI: 10.1038/nmeth.4256
More info and docs: MSFragger website, FragPipe GitHub page
Philosopher
C:\test_wkf\MSFragger\philosopher_v3\philosopher.exe
Browse
Download
Philosopher version: 3.2.3 (build 1583959175). System OS: Windows Server 2012 R2, Architecture: AMD64
If provided, philosopher binary will be used for Peptide and Protein Prophets and Report
More info: Philosopher GitHub page
Python
python
Browse
Version: Python 2.7.15
DB Splitting
Python: Python 2.7.15.
Update MSFragger to a newer version. Use the Update button next to MSFragger field. Database Splitting disabled.
Python 3 is required.
Latest version of MSFragger is required.
FragPipe will work fine without this functionality.
See configuration help online for instructions how to enable.
Spectral Lib generation
Python: Python 2.7.15.
Spectral library generation disabled.
Python 3 is required.
FragPipe will work fine without this functionality.
See configuration help online for instructions how to enable.


FragPipe (v12.2)
Config
Select LC/MS Files
Database
MSFragger
Downstream
Report
Run
Run MSFragger
Load default parameters for:
Open Search
Load
Save Parameters
Load Parameters
RAM (GB)
30
Threads
6
Common Options (Advanced Options are at the end of the page)
Peak Matching
Precursor mass tolerance
Da
-150
500
Deisotope
1
Fragment mass tolerance
PPM
20
Calibrate masses
On and find optimal parameters
Isotope error
0
Protein Digestion
Load rules
trypsin
Enzyme name
trypsin
Cut after
KR
But not before
P
Cleavage
ENZYMATIC
Missed cleavages
1
Clip N-term M
Peptide length
7
50
Peptide mass range
500
5,000
Max fragment charge
2
Split database
1
Modifications
Variable modifications
Max variable mods on a peptide
3
Max combinations
Multiple mods on residue
Ena...
Site (editable)
Mass Delta (edita...
Max occurrences (...
M
15.9949
3
[^
42.0106
1
STY
79.96633
3
nQnC
-17.0265
1
nE
-18.0106
1
site_06
0
1
site_07
0
1
Fixed modifications

2. Baseline Setting

In this section, we give the parameter setting for the baselines employed in the experiments.

InsPecT:

1) Tool download: http://proteomics.ucsd.edu/Software/Inspect/
2) Dataset format conversion: http://tools.proteomecenter.org/wiki/index.php?title=Main_Page
3) Database Setup: run CMD as belows

```
> python PrepDB.py FASTA [myDB.fasta]
```

4) Parameter InputFile:

```
speactra, [FILENAME.mgf]
protease, Trypsin
mod, 57, C, fix
TagCount, 100
TagLength, 5
```

5) Run the InsPect:

```
> InsPecT.exe -i InputFile.text -o OutputFile.txt
```

(In order to output the extracted tags, we use the "DEBUG" model, more detail can be refer to document: http://proteomics.ucsd.edu/Software/Inspect/InspectDocs/)

PepNovo+:

1) Tool download: https://github.com/jmchilton/pepnovo
2) Run the pepnovo+:

```
>PepNovo.exe -file [mgfPath] -model CID_IT_TRYP
C+57:M+16 -digest NON_SPECIFIC -tag_length 5 -
num_solutions 100 -fragment_tolerance 0.01 >
[OutputFilePath]
```

(Note that we utilize the high-resolution data version for tag extraction. In practice, If the MS/MS spectra come from high-resolution instruments, the sequencing performance can be improved by manipulating the tolerances. For instance if the spectra have fragment tolerances of 0.01, this can be set with the flag: -fragment_tolerance 0.01 .)

SVM:

1) Package download: https://www.csie.ntu.edu.tw/~cjlin/libsvm/
2) Parameter setting:

Input features	Peak intensity Edge mass error

	Node relevance degree (same as tag discriminator)
Feature Normalization	Yes
Kernel	RBF
C	512
Gamma	0.03125
Validation	5-fold cross-validation

3) Train the SVM model:
```
>svm-train.exe -t 2 -g 0.03125 -c 512 -v 5 -s 1
[DatasetName] [ModelName]
```
4) Predict the model using trained SVM model:
```
>svm-predict.exe [DatasetName] [ModelName] [OutputFile]
```

3. Experimental Results Supplementary

Due to the limitation of paper space, we will present more results here.
(To be continued)