Run PepQuery (version >=1.6) on immunopeptidomics data

1. Convert MS/MS raw data to MGF format using the following command line on Windows:

```
msconvert --filter "peakPicking true 1-2" --mgf *.raw
```

The tool msconvert can be downloaded from http://proteowizard.sourceforge.net/.

If there are multiple MGF files from multiple fractions (or runs) for one sample, they must be combined into one MGF file. In Linux, this can be done using "cat" command line as shown below. Assume all the mgf files for the sample (sample1) are in folder "sample1".

```
cat sample1/*.mgf >test.mgf
```

2. Generate tag file for PepQuery using Open-pFind:

Open-pFind can be downloaded from http://pfind.ict.ac.cn/software/pFind3/. It can only be run on Windows system.

Step 1: Load MS/MS data in MGF format as shown below. If there are multiple MGF files from multiple fractions (or runs) for one sample, they must be combined into one MGF file and then load the combined MGF file into Open-pFind.
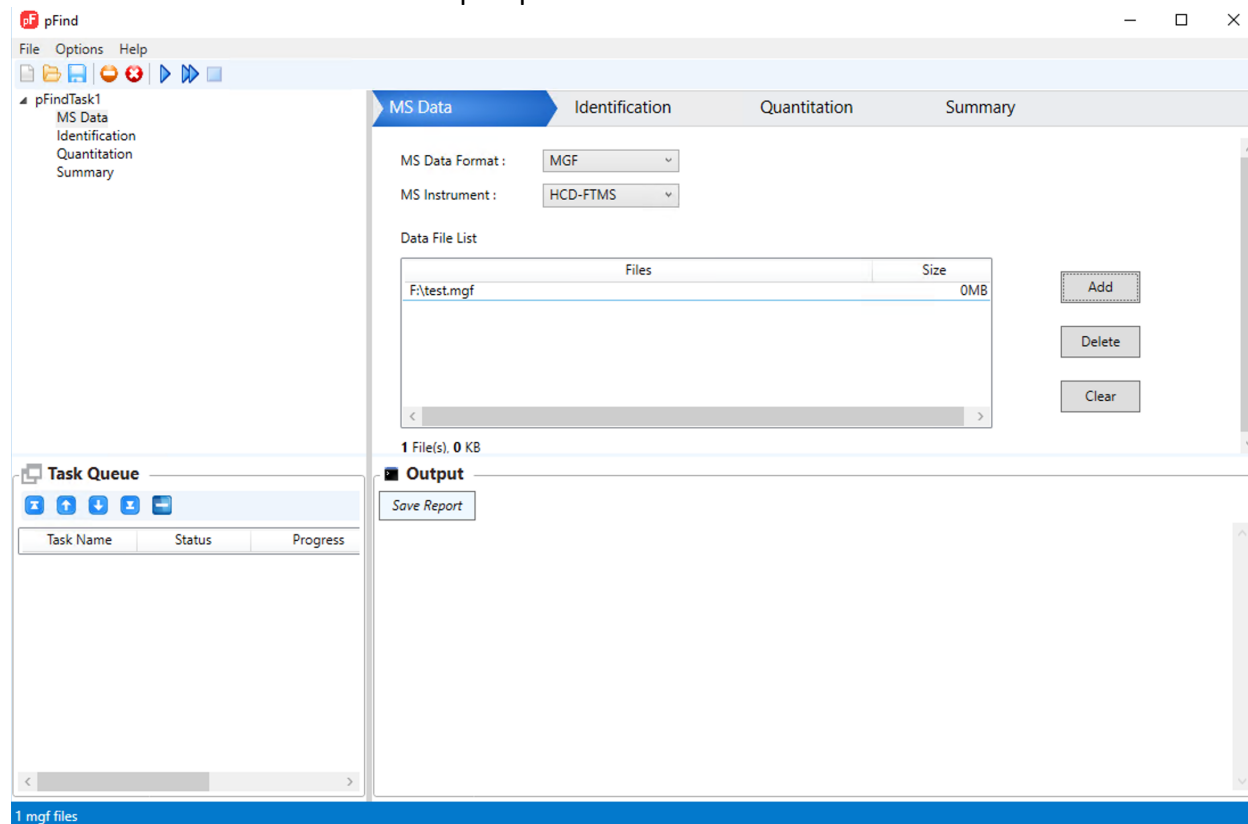
Step 2: Set peptide identification parameters.

Please note that the database used in this search should be a reference database with all the sequences are known proteins. Any novel peptide or protein sequences should not be included in this database. This database should be the same database used for PepQuery parameter "-db". Open-pFind search requires to search a target-decoy database. Please follow the instruction of Open-pFind to generate such database by taking a target only database as input.
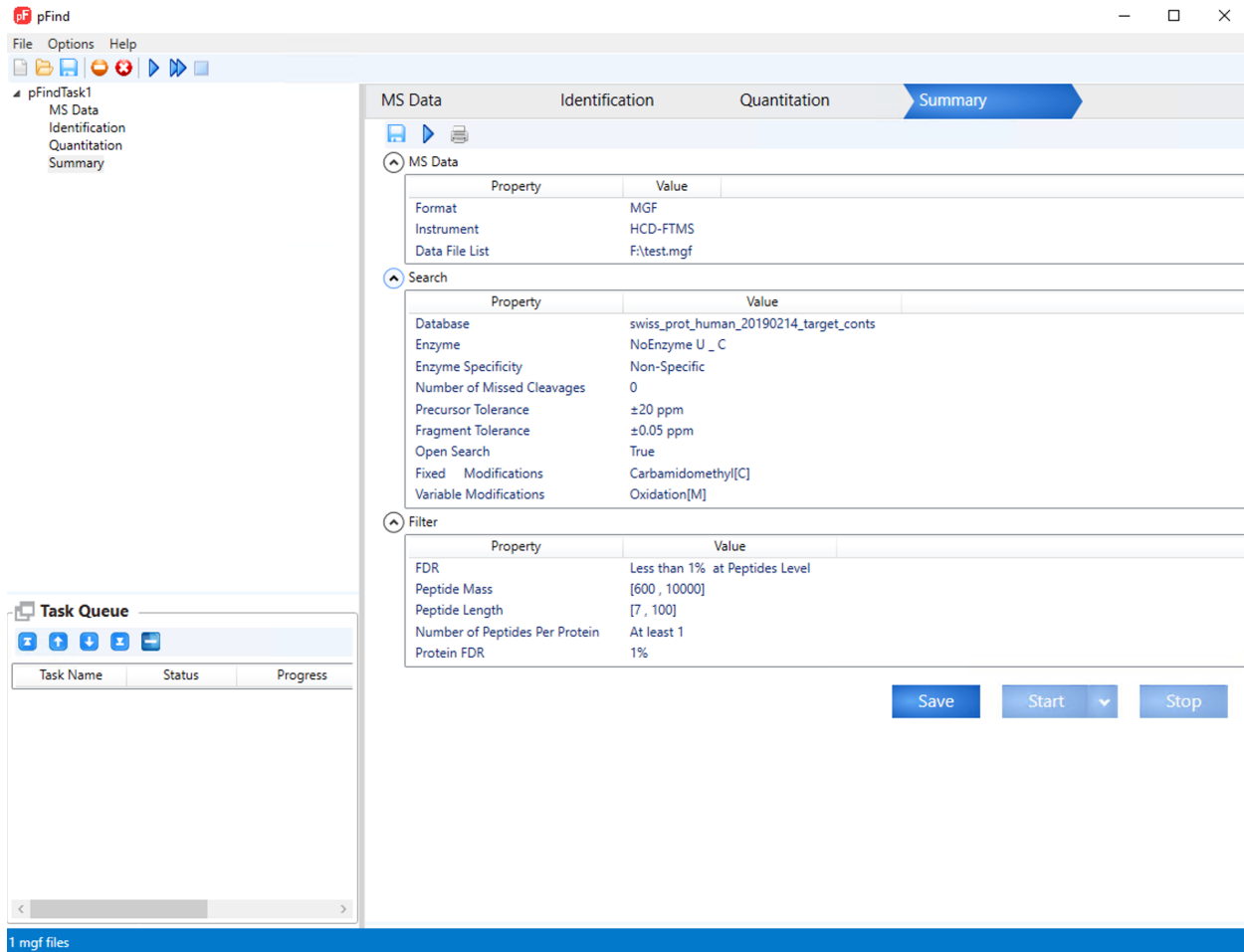
Other parameters such as enzyme, precursor tolerance, fragment tolerance, fixed modification, variable modification should be consistent with the setting for PepQuery.

Step 3: Start search.

After the search is finished, users can find the result files in folder "result".

3. PepQuery search:

Step 1: Prepare tag file for PepQuery search. File "test.qry.res" is generated by Open-pFind.

```
java -cp pepquery.jar main.java.util.generateTags test.qry.res >test.tags
```

Step 2: Run PepQuery:

```
java -jar pepquery.jar -varMod 117 -fixMod 6 -db ./protein.pro-ref.fasta -um -tag test.tags -pep input_pep_for_vali
dation.txt -ms test.mgf -minScore 12 -itol 0.05 -e 0 -tol 20 -indexType 2 -o test
```

Please note if there is no fixed modification in the experiment, then -fixMod should be set as "0". In default, Open-pFind uses the spectrum title (string in row with "TITLE=") in the MGF file as index for each spectrum, so -indexType should be set as 2.