User Guide

XY-Meta is an efficiency and convenient tool for large-scale metabolome identification. By using XY-Meta, you can implement different search mode such as narrow search, open search and iterative search. Most of all, this tool can estimate the FDR of identification results with Target-Decoy strategy. This document can help you learn the usage of XY-Meta quickly, please read it carefully. First and foremost, checking the running memory of your computer must have more than 4G. Secondly, preparing the input data for running program. Finally, adjusting the parameter settings of database search.

**Parameter setting**

All the necessary parameters and its explanation are filled in the **parameter.default** in the **config** folder. When you start to run XY-Meta, you can place the parameter.default in any folder.

**How to run XY-Meta**

You can run XY-Meta by using command line easily. The following is the explain about parameters in the command line:

-S The pathway of parameter.default

-D The pathway of query spectra

-R The pathway of reference spectra

You can start to run XY-Meta just like this:

XY-Meta.exe -S \XY-Meta-Win\config\parameter.default -D \XY-Meta-Win\query\GNPS\_NIST.mgf -R \XY-Meta-Win\database\MoNA-export-MassBank.mgf

After the program is done, a table of the identification results will be generated under the folder where there is to contain query spectra.

You can modify the parameter.default to change the search mode and make this tool to suit for your task.

**Narrow search mode**

The narrow search mode is the most commonly used to implement metabolome identification. Before we start to run the search engine, we always set the precursor mass tolerance between 5-100ppm. Limiting the mass tolerance of precursor is strictly able to increase the accuracy of identification results.

search\_pattern=1/2/4

precur=1

tolerance\_precur=5-100ppm

**Open search mode**

The open search mode is error-tolerance search mode with a large precursor mass tolerance. When using open search, the precursor mass tolerance can be set to more than 100Da. Open search mode can be used to find out some molecules which are modified by unexpected chemical group. This search mode has been proved that it could implement the FDR control by Target-Decoy strategy in proteome identification. Surely, the open search mode is also suited for metabolome identification in finding “dark matter”.

search\_pattern=1/2/4

precur=0/1

tolerance\_precur=100ppm-500Da

**Iterated search mode**

The reference database is always too large so that the estimated FDR is higher than the actual FDR when we are identifying proteins in meta-proteomics research. In this case, iterated search mode can concentrate the reference database in first search and implements the FDR control final step. When the reference database is too large or identifying the metabolites in meta-metabolomics research, Iterated search mode can increase calculated efficiency significantly.

search\_pattern=3

precur=1

tolerance\_precur=5-100ppm