introduction\_for\_si

## Data preparation

All the raw MS data should be transfered to mzXML format using ProteoWizard.

Then the mzXML data placed in tow folders named as “POS” and “NEG” according to their mode.

A “metabolite.info.csv” must be outside the “POS” and “NEG” folder which is used to give the information for metabolites.

The metabolite.info must have 11 columns, which can be got from the examples.

**Note: The Submitter column must same with names of mzXML datasets. For example, if metabolite is in dataset “MixA\_NCE25.mzXML”, then the “Submitter” of this metabolite is “MixA”. Please see the demo data.**

## Run databaseConstruction

###set work directory  
setwd(".")  
databaseConstruction(  
 path = ".",  
 version = "0.0.1",  
 mz.tol = 15,  
 rt.tol = 30,  
 threads = 3  
)

The MS2 database will be saved in the path.