# Batman\_usage

**experimental procedures**

1. Create "batman" folder. Under the "batman" folder, create a folder for each expression template.
2. If you use the expression template provided by DECEPTICON, the folder needs to be named with a fixed name (such as ciber\_base, TRef\_base, BRef\_base, quan\_base and immu\_base).
3. If you use custom expression template(s), you need to name the folder as "sig \_ 1", "sig \_ 2", etc.



1. Run BATMAN

Examples:

library(batman)

bm = batman (runBATMANDir = "./ciber \_ base", showPlot = FALSE)

## This will create the folder "runBATMAN" in current working directory,

## within the folder "runBATMAN", a subfolder "BatmanInput" contains all the

## input files batman uses. Users can modify "metabolitesList.csv",

## "batmanOptions.txt" and so on to change the settings of batman.

## Please check "BatmanInput" for details on how to adjust input parameters.

########################################################################

## The following is an example of what will be displayed in R

## and what value the user could input:

########################################################################

## batman...

## Number of burn-in iterations: 4000

## Number of post-burn-in iterations: 100

##

## The template file used is

## 1: The default template of multiplets in multi\_data.csv file.

##

## Loading multi\_data.csv...

## Percentage completed...

## | | 0%

## Size of each spectrum is 393.

## Size of metabolite list is 22.

## Constructing chain data structure...

## time used is 0 seconds.

## Running MCMC...

## |=================================================== | 80%

## time used for burnin is 76 seconds.

## |==================================================================| 100%

## time used is 95 seconds.

## saving posteriors...

##

## time elapsed

## 95.61

## second.

## Reading in saved data in folder

## .../user\_specified\_dir/runBATMAN/BatmanOutput/07\_Dec\_17\_19\_18

## Completed.

########################################################################

## Alternatively if more than 1 spectrum are included without using fixed effect

## (in batmanOptions.txt file, set

## "Same concentration for all spectra (fixed effect) (1/0): 0"),

## user will be asked to input the following parameter:

########################################################################

## How many parallel processes (multicores) do you want to run

## the multi-spectra analysis?

## (Enter 1 for running them sequentially.)

##

## Parallel processing of multi spectra currently cannot display

## progress bar (or any words), if you input is > 1, please be patient

## for the results :)

##

## 1: 2 ## user input

## time elapsed

## 78.79

## second.

## Reading in saved data in folder

## .../user\_specified\_dir/runBATMAN/BatmanOutput/07\_Dec\_17\_35\_53

## Completed.

1. adjust input parameters

“BatmanInput” folder:

The user can modify the parameter values in the following input files (do not change the name of these files): batmanOptions.txt, metabolitesList.csv, multi\_data\_user.csv, NMRdata.txt.

3.1 batmanOptions.txt:

ppmRange - ppm ranges for analysis: (1.2, 1.6) (2.1, 2.8)

* Change to ppmRange - ppm ranges for analysis: (1, xx)

xx indicates the number of common genes of the expression template and bulk data.

specNo - Ranges of spectra number to be included (e.g. 1,3-4 etc.): 1-3, 5

* Change to specNo - Ranges of spectra number to be included (e.g. 1,3-4 etc.): 1-xx

xx indicates the number of samples in bulk data.

paraProc - No of parallel processes (multicores) (only 1 core will be used for single spectrum): 1

* (optional) Multi-core parallel operation can be set to improve the running speed.

downSamp - Down sampling factor: 3

* Change to downSamp - Down sampling factor: 1

nItBurnin - Number of burn-in iterations: 200

* Integer, this is the number of burn-in iterations. The number of iterations after burn in will be asked when running batman. In case of unsatisfactory fitting effect, it can be appropriately reduced.

nItPostBurnin - Number of post-burn-in iterations: 100

* Integer, this is the number of post-burn-in iterations. The posterior samples will be saved in the frequency specifified by the next parameter. If the fitting effect is not ideal, this parameter can be appropriately reduced or increased, generally greater than 10000.

rdelta - Truncation of the prior on peak shift (ppm): 0.030

* Chang to rdelta - Truncation of the prior on peak shift (ppm): 0.002

Note: Other parameters do not need to be adjusted;

The fitting result can be observed by the pdf map in the "BatmanOutput" folder. The more the blue line and the green line overlap, the better the fitting effect will be.

3.2 metabolitesList.csv:

List of cell type names to be fitted. Put "%"in front of the cell type name to comment out any cell type for batman analysis.

3.2 multi\_data\_user.csv:

Metabolite: The name of cell type.

pos\_in\_ppm: The middle position of ppmRange.

couple\_code: -2

J\_constant: 1,xx (xx indicates the number of common genes of the expression template and bulk data.)

relative\_intensity:1

overwrite\_pos: n

overwrite\_truncation: n

Include\_multiplet: The default is "1" and all multiplets belong to the listed

cell type will be used. Set to "0" to exclude certain multiplet(s) from listed

cell type(s).

3.3 NMRdata.txt

The first column of the document is the common genes (with the numbers 1, 2, 3,. . . Means), and each subsequent column represents a sample.

3.4 “PureSpectraTemplate” folder

The txt file name is the name of each cell type. The first column in the file is genes (the same as the first column in NMRdata.txt), and the second column is gene expression values.

Batman\_data will provide the required data of five expression templates: ciber, TRef, BRef, quan and immu. Users need to modify the parameters according to their own bulk data.