

Deliberator

Table of contents

1 Overview.....	2
2 Algorithm.....	2
3 Usage.....	3
4 Releases.....	4
4.1 Rel0.19.....	4
4.2 Rel0.18.....	4
4.3 Rel0.17.....	4
4.4 Rel0.16.....	5
4.5 Rel0.15.....	5

1. Overview

This application generates a decoy library from a given annotated MS library.

Please cite:

Ahrné E., Ohta Y., Nikitin F., Scherl A., Lisacek F., Müller M., An improved method for the construction of decoy peptide MS/MS spectra suitable for the accurate estimation of false discovery rates, *Proteomics.*, 2011, 11 (7), pp 4085-4095

2. Algorithm

The method is based on the shuffling of each MS precursor peptide and the generation of decoy spectra obtained after the theoretical fragmentation.

Each generated spectrum fragment then follows a different fate:

1. the mz is recalculated for any Annotated fragments
2. the Non-Annotated (NA) fragment is kept or sampled from the overall NA distributions

Note:

The decision to sample a NA peak depends on its deviation from the overall population (of closed mz fragments sharing the same precursor charge in the precursor mz neighborhood) - The less it deviates the more probable it will be sampled.

PSEUDO CODE

```
BEGIN
1. parse the annotated spectra library file
   (with --etd option: precursor neighbors fragments removed)
2. make n distributions of all NA fragments by precursor charge (n charges)
   for future sampling
   (example at\_charge +3)
   FOR each original_spectrum DO
3. [make decoy spectrum]
   count <- 0
3.1 shuffle precursor peptide sequence (do not shuffle N/C terminal
   amino-acids (peptidase footprint))
3.2 create the fragmentation spectrum from the shuffled precursor
   /// TODO: (specific kind of fragmentation ??)
   FOR all peaks DO
     IF annotated peak THEN
       recalculate the new mzs
     ELSE IF mz closed to the baseline THEN
       pick a peak mz in the NA sample (see 2)
     ELSE
       keep the peak
     FI
   ROF
4. [compute spectrum match (score [0, 1[]]
```

```

    score <- original_spectrum versus decoy_spectrum (with only annotated
peaks)
    IF score > dot_product_threshold (decoy ~ original) and count < 10
THEN
    count <- count + 1
    GOTO 3.1
    ELSE
5.    [write decoy spectrum]
    flush the most different (lowest score) decoy spectrum in output
file
    FI
    ROF
END

```

3. Usage

```

usage: Deliberator <mslib> [-a] [-c] [-d <arg>] [--decoy-tag]
    [--default-decoy-tag] [--default-pep-decoy-tag] [-h] [-i <arg>]
[--log
    <arg>] [-o <arg>] [-p <arg>] [--pep-decoy-tag <arg>] [-q] [-r <arg>]
[-s
    <arg>] [-t <arg>] [-v] [-w <arg>]
-a,--average                set the average mass mode for
                             peptide mass calculation
                             by default: MONOISOTOPIC.
-c,--concat-libs            concatenate libs
                             by default: false.
-d,--dp-threshold <arg>    define the dot-product threshold
                             ([0-1[) for spectrum shuffling
                             by default: 0.7.
--decoy-tag                 set this decoy tag in 'Comments' of
                             decoy spectra
                             by default: No tag.
--default-decoy-tag        set this default decoy tag in
                             'Comments' of decoy spectra
                             by default: 'DECOY_'.
--default-pep-decoy-tag    set this default decoy tag in
                             peptide 'Name' of decoy spectra
                             by default: 'decoy_'.
-h,--help                  print this message.
-i,--setting-file <arg>    give a property file with all input
                             settings.
--log <arg>                define the log file.
-o,--output <arg>          set the output filename (.msp or
                             .sptxt file only).
-p,--precision <arg>      define the number of fractional
                             digits for output
                             by default: 2.
--pep-decoy-tag <arg>     set this decoy tag in peptide 'Name'
                             of decoy spectra

```

```

-q,--quiet                by default: No tag.
                           quiet mode (verbose off)
-r,--render-dir <arg>    by default: false.
                           render NA peak histograms
                           (render-dir/hist) and original +
decoy spectra (render-dir/scan).
-s,--sampling-prob <arg> warning: execution time x10.
                           define the probability of sampling
                           non-annotated peaks ([0-1[)
                           by default: -1.0.
-t,--tol <arg>           define the tolerance for mz fragment
                           peak comparison
                           by default: 0.1.
-v,--version             print the version info.
-w,--sampling-interval-width <arg> define the bin width of
for sampling              non-annotated (NA) peaks histograms
                           by default: 100.

```

4. Releases

The latest version is v0.19 - [Download app](#) with the [default properties file](#).

4.1. Rel0.19

Type	Change
Fixed Bug	Deliberator was generating the decoy library in sptxt format only. It now produces sptxt or msp file given the file name extension.
Fixed Bug	The MW is now evaluated to the "exact molar mass of the peptide ion" and not to the neutral molar mass.

4.2. Rel0.18

Type	Change
Fixed Bug	While shuffling peak sequence, some peak mz were becoming negative after recomputation given their fragment type. We now ignore them.

4.3. Rel0.17

Type	Change
------	--------

Fixed Bug	MS1/MS2 dist computation sometimes crashed
-----------	--------------------------------------------

4.4. Rel0.16

Type	Change
Runtime behavior Change	Parsed spectra (needed in MS1/MS2 distributions process) are not kept in memory anymore. They are now serialized in hadoop file reducing the memory overhead.

4.5. Rel0.15

Type	Change
Decoy process progression Change	The progression is now visible in a progress bar.