Pseudocode for FDR & Precision test

Ziyang Chen

May 19, 2021

Algorithm 1: Annotation

Data: *rawdata* is the target peak list, *lib* is the target-decoy list used by `NetID`

Result: $Ann_{mz}, Ann_{node}, Ann_{edge}, Ann_{ni}$ are annotation lists by each method

1 $SCORE_{mz}, SCORE_{node}, SCORE_{edge}, Ann_{ni}$ are generated using `NetID` with $rawdata, lib$ as input

`//` $SCORE$`s are score matrix of the peak list using mz, node, edge score`

2 $Ann_{mz}, Ann_{node}, Ann_{edge}$ are decided using $SCORE_{mz}, SCORE_{node}, SCORE_{edge}$ by obtaining a similar amount of $Unknown$ in $Ann_{mz}, Ann_{node}, Ann_{edge}, Ann_{ni}$

Algorithm 2: Target-Decoy library Generator

Data: target library lib_{in} that is used by `NetID`

Result: a 1 : 1 target-decoy library lib_{out}

1 decoys D is generated by substitute a H of every formula in lib_{in} by an element randomly chosed from imp

`//` imp `is a list of implausible elements which removed all elements used in target library`

2 $lib_{out} = lib_{in} \bigcup D$

Algorithm 3: FDR calculator

Data: Ann is a annotation list of length M

Result: FDR is the false discovery rate

1 $FP =$ number of annotations that have implausible element

2 $FDR = \frac{FP}{M-FP}$

Algorithm 4: Precision calculator

Data: Ann is a annotation list of length M, GT is a anootation list of 314 peaks that are mannually annotated

Result: P is the percentage of correct annotated peaks of Ann with respect to GT

$P = \frac{N_{correct}}{314}$

Algorithm 5: FDR&Precision test

Data: $rawdata$ is the target peak list, lib_{in} is the target library, GT is a anootation list of 314 peaks that are mannually annotated, REP is the time of repeatation

Output: FDR is a $4 \times REP$ matrix of false discovery rate for 4 methods, P is a $4 \times REP$ matrix of Precision for 4 methods

$FDR, P \longleftarrow [\,]_{4\times REP}$

for i *in* $1 : REP$ **do**

  $lib_{out} =$ `Target-Decoy library Generator` (lib_{in})

  $Ann_{mz}, Ann_{node}, Ann_{edge}, Ann_{ni} =$ `Annotation` $(rawdata, lib_{out})$

  $FDR[, i] =$ [`FDR calculator` (Ann_{mz}), `FDR calculator` (Ann_{node}), `FDR calculator` (Ann_{edge}), `FDR calculator` (Ann_{ni})]

  $P[, i] =$ [`Precision calculator` (Ann_{mz}), `Precision calculator` (Ann_{node}), `Precision calculator` (Ann_{edge}), `Precision calculator` (Ann_{ni})]