

# Proteomics

WS 2014/15

## Exercises 7

### 1. Theoretical spectrum

Consider the peptide ACLHVR. Construct a theoretical spectrum when the fragments have charge 1. So What is the mass-to-charge of following ions  $a_3, b_2, b_4, b_5, c_4, y_3, y_4, y_5$ ? (Use nominal residual mass: A: 71, C: 103, L 113, H: 137, V 99, R 156, N-terminal group 1, C-terminal group 17; residue 1 is A.)

### 2. Theoretical spectrum 2

Assume that two peaks with nominal  $m/z$  values 172 and 343 are in a spectrum. Investigate if these can correspond to the same peptide.

What type of ion is the peak 172 Th? (b/y)

### 3. Mascot-Mowse score

Given a peptide sequence  $d$  of mass  $m=1000$ , it is placed in peptide intervals  $x, y, z$ . For one protein mass interval  $j$ , we found the number of theoretical peptides belonging to class  $(x, j)$   $(y, j)$   $(z, j)$  is 20, 40, 50, respectively. Then what is the scoring for peptide  $d$ , if  $\max_i n_{i,j} = 100$ ?

### 4. X!Tandem scoring

Given:

- Experimental spectrum:  $(m/z, \text{intensity})$ : (227.3, 50), (321.3, 30), (330.5, 20), (374.7, 90), (418.7, 100), (544.7, 60), (593.7, 30), (685.7, 20), (839.0, 70)
- Theoretical spectrum of charge 1(b+y):  $m/z$ : 227, 276, 330, 375, 490, 544, 593, 650.

Given the fragment mass tolerance = 0.5Da, calculate the X!Tandem dot product for the experiment spectrum and the given theoretical spectrum.

### 5. X!Tandem E-Value

Assume X!Tandem found a set of PSMS with dot product  $x \in \{20, 22, 25, 34, 34, 34, 35, 45, 46, 47, 53, 57, 59, 62, 66, 80, 92, 95, 99, 119\}$ . If we divide these scores into 10 bins and get the frequency  $f(x)$  of each score  $x$ . What is the value of the survival function  $s(x)$  at score 80? What is the corresponding E-value of score  $x=80$ ?

## 6. Modification

Given a peptide sequence of length 7 with mass 706 Da, we calculate mass difference between b4(+1) and b5(+1) is 101 Da. This is the residue mass of T. If allowing fixed modification phosphorylation(80 Da shift), where could we find the b7 ion of charge 2? (the m/z value of the peak)

## 7. PeptideProphet

Assume Sequest found a PSM with  $s_1(Xcorr) = 0.4$ ,  $s_2(Sp) = 80$ ,  $s_3(DeltaCn) = 0.08$ ,  $s_4(LnrSp) = 3$ . Using discriminant analysis, we found the weights:  $c_1 = 0.5$ ,  $c_2 = 0.05$ ,  $c_3 = 0.875$ ,  $c_4 = 0.05$ ,  $c_0 = -0.42$ . After measuring all the PSMs, the discriminant score positive distribution was observed as  $N(4,4)$ . What is  $p(F|+)$  of this PSM?

## 8. PeptideProphet 2

Assume a search engine found two PSMs of same charge: (K)GGASPK and (K)CKSYLEDTI. PeptideProphet did discriminant analysis on these two PSMs and both were assigned the same discriminant score =5.5. Which of the two PSMs could have a higher PeptideProphet-computed probability of being correct?

## 9. ConsensusID

Assume for a spectrum, peptide QRESTATDILQK was found by X!Tandem, Mascot and OMSSA. After comparing all the hits found by the three engines, this peptide-spectrum match was assigned the similarity score: 0.9,0.5,0.4, of each engine. What is the consensus score for this match?

## 10. Experiment

Download `velos005614.mzML` from [http://svn.code.sf.net/p/open-ms/code/Tutorials/UM\\_2014/Example\\_Data/OpenMS/small/](http://svn.code.sf.net/p/open-ms/code/Tutorials/UM_2014/Example_Data/OpenMS/small/) and the human SwissProt database.

- Use XTandemAdapter in Knime/OpenMS to identify the peptides. Try fragment mass tolerance = 10ppm and 100ppm. You can use FileInfo to see the number of identified peptides.
- Use XTandemAdapter again and try fixed modification phosphorylation STY.
- Construct decoy sequences for the protein database and attach decoy to the end of the original fasta file. try again to search with XTandem. Use PeptideIndexer to index each target-decoy PSMs and calculate the FDR by FalseDiscoveryRate. See how many identified PSMs with FDR lower than 0.1 by apply IDFilter.
- Search with XTandemAdapter and OMSSAAdapter against target-decoy database, then try ConsensusID node. This time see if the number of PSMs with  $FDR \leq 0.1$  is changed.