

T-61.6070 Special course in bioinformatics 1

Exercise 30.4

1. Why protein identification is a challenging task?
2. What kind of information does pepHMM model?
3. Identify protein, if possible, for the following synthetic MS/MS data using your favourite MS/MS search tool. comment on results and do report search parameters

data:

BEGIN IONS

TITLE=Spectrum 1

PEPMASS=3000.6

1365.6477	411.1400
1383.6386	349.4500
1425.616	2193.2000
1434.7539	751.9100
1475.7353	1934.5000
1491.7199	907.8700
1493.724	1564.7000
1707.7676	280.7800
1791.7448	267.8400
1838.9766	79.8830
1894.0492	91.6580
1901.9144	297.3200
1920.9151	236.7500

END IONS

4. Compute emission probability using pepHMM framework for the following data?

T_b=0.124, T_y=0.235,

I_b=300.5, I_y=400.3

$$\lambda_b = 4.2$$

$$\lambda_y = 6.2$$

$$\mu = -0.038$$

$$\sigma = 0.119$$