

Peptide Validation

2020.06

김현우

Database searching approach

- SEQUEST(Comet), Mascot

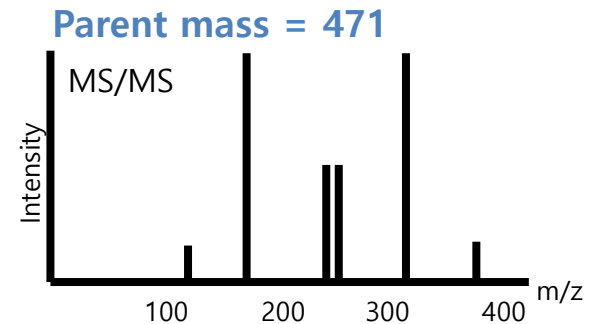
Candidate peptides

READ
DVGAE
DEAR
GAGVDA
EGDVA
...

Comparison



Experimental MS/MS spectrum

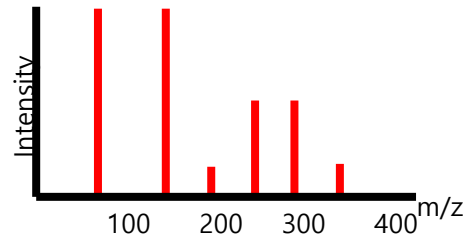


Peptide Assignment by MS/MS

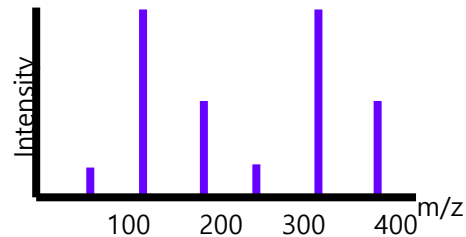
Candidate peptides

Theoretical MS/MS spectrum

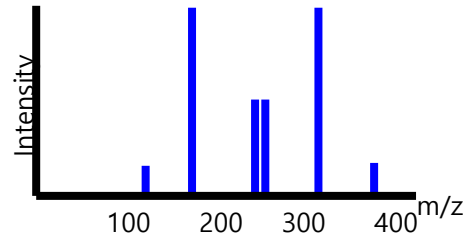
READ



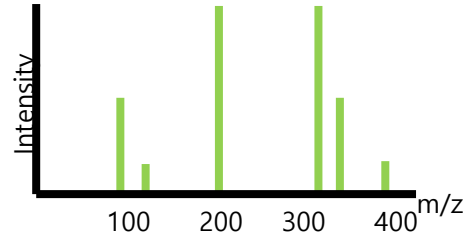
DVGAE



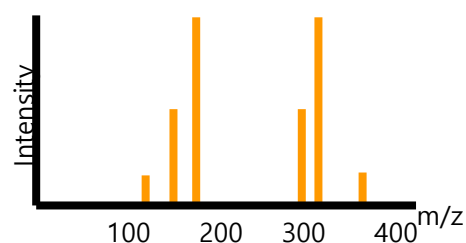
DEAR



GAGVDA

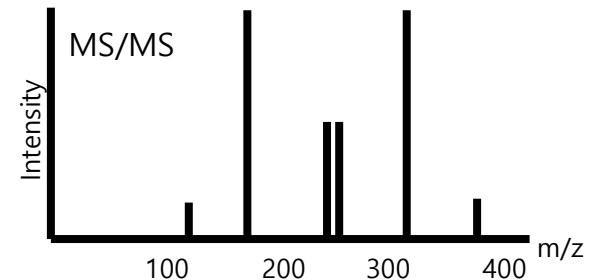


EGDVA

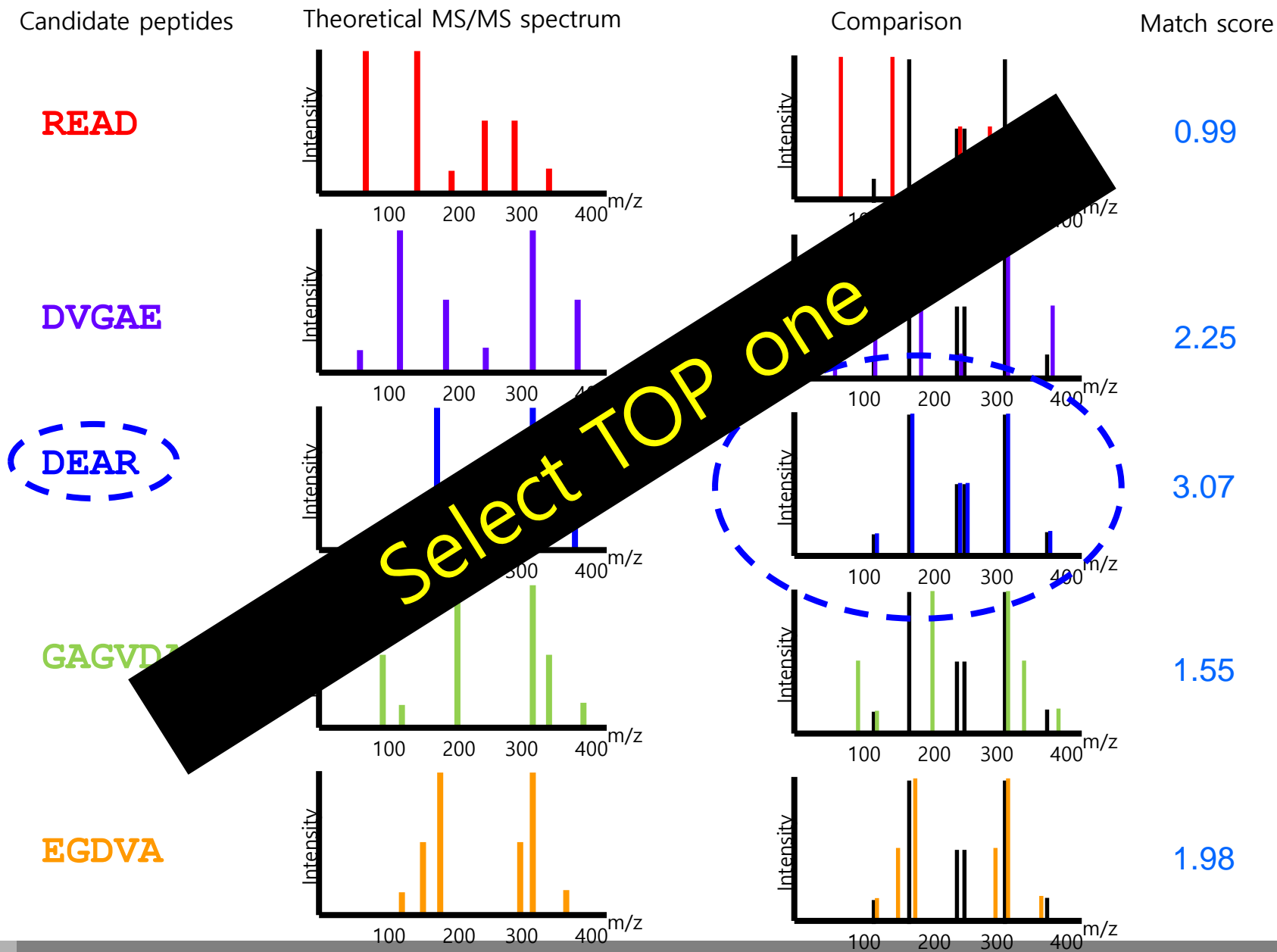


Experimental MS/MS spectrum

Parent mass = 471



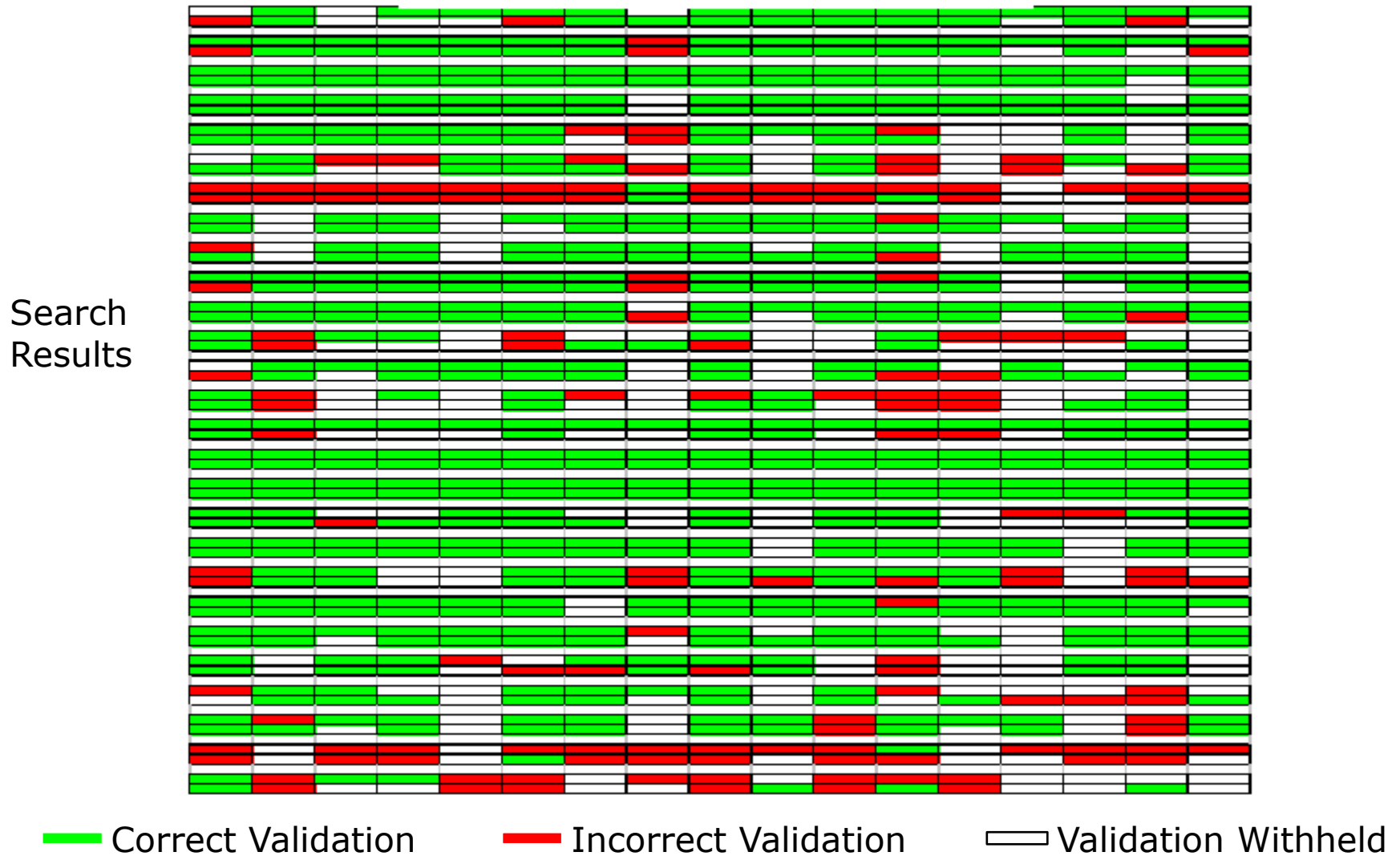
Peptide Assignment by MS/MS



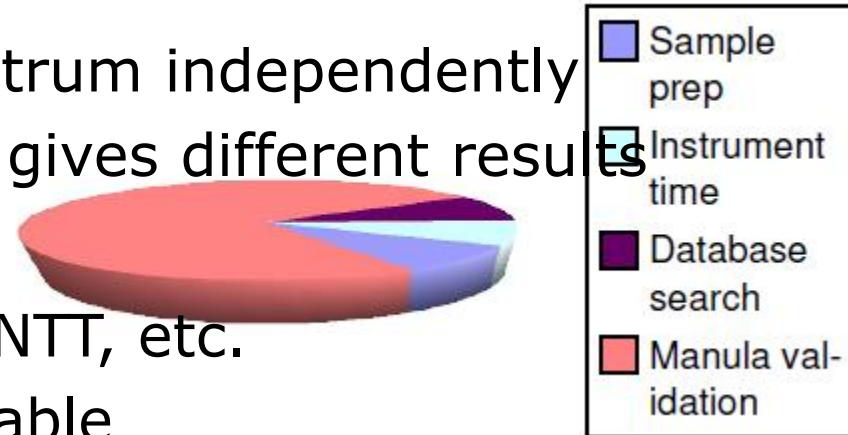
- Peptide assignment
 - interpret each MS/MS spectrum independently
 - different analysis software gives different results
- Manual validation
 - filtering by search scores, NTT, etc.
 - subjective opinion is inevitable
 - hard to estimate the error rate
 - when dataset gets large?
- Statistical validation
 - model-based validation: probabilistic model for score distribution
 - target-decoy: estimate false discovery rate based on the match to the “decoy” database

(Un)reliability of manual validation

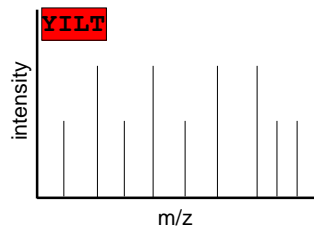
Manual Authenticators



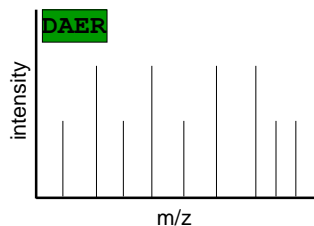
- Peptide assignment
 - interpret each MS/MS spectrum independently
 - different analysis software gives different results
- Manual validation
 - filtering by search scores, NTT, etc.
 - subjective opinion is inevitable
 - hard to estimate the error rate
 - when dataset gets large?
- Statistical validation
 - model-based validation: probabilistic model for score distribution
 - target-decoy: estimate false discovery rate based on the match to the “decoy” database



Target/Decoy method



3.15



2.47

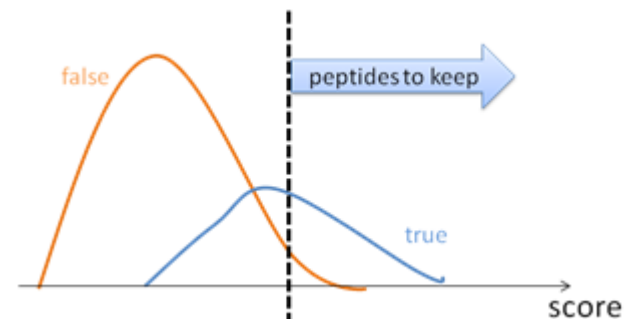
>Protein A (Target Sequence)

MEMEKEFEQIDKSGSWAAIYQDIDVGAEDFPCRVAKLPK
NKNRNRVYRDVSPFDHSRKREADDNDYINASLIKMEEAQR
SYILTTQQIDKSGSWAAIYQDIRHEASDFHEASDFPCRVA
KLPKNKDEARYMEKEFEQIDKGAGVDADIRHEMEKEFEQ
IDKSGSWAAIYQDIRHE

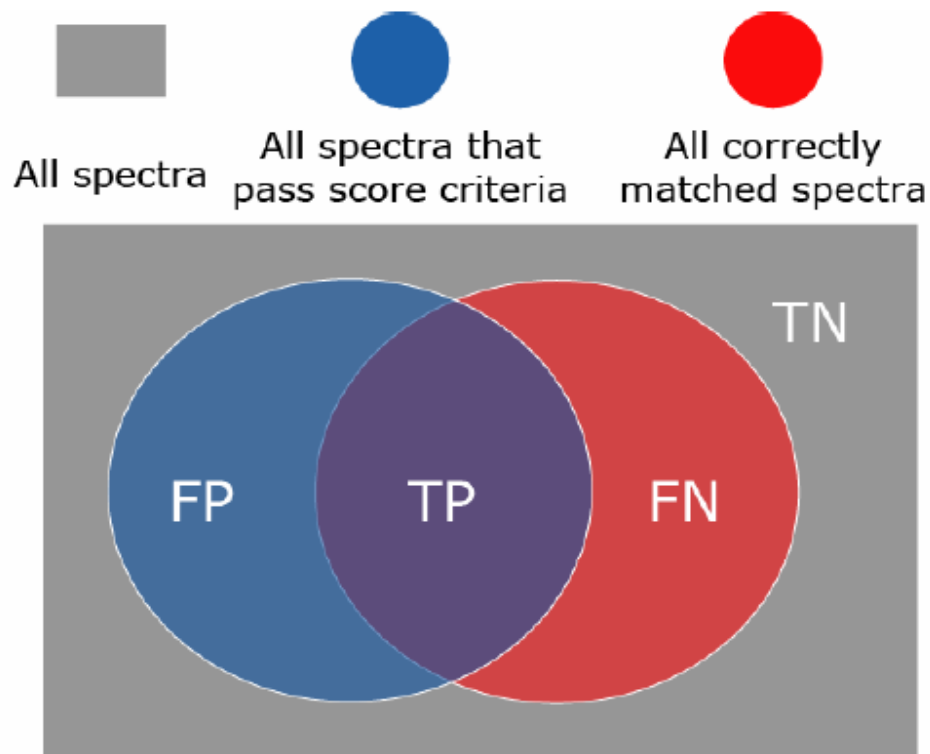
>Reversed Protein A (Decoy Sequence)

EHRIDQYIAAWSGSKDIQEFEEKEMEHRIDADVGAGKDIQ
EFEKEMYRAEDKNKPLKAVRCPFDSAETHFDSAETHRIDQY
IAAWSGSKDIQQTLIYSRQAEEMKILSANIYDNDDAERK
RSHDFPSVDYRNRNKNKPLKAVRCPFDEAGVDIDQYIA
AWSGSKDIQEFEEKEMEM

- $T = 1000$ # of matches to the target sequence (above score threshold)
- $D = 20$ # of matches to the decoy sequence (above score threshold)
- False Discovery Rate = ?



Target/Decoy method



TP = True Positive

TN = True Negative

FP = False Positive

FN = False Negative

$$\text{FP rate} = \frac{\text{FP}}{\text{TP} + \text{FP}}$$

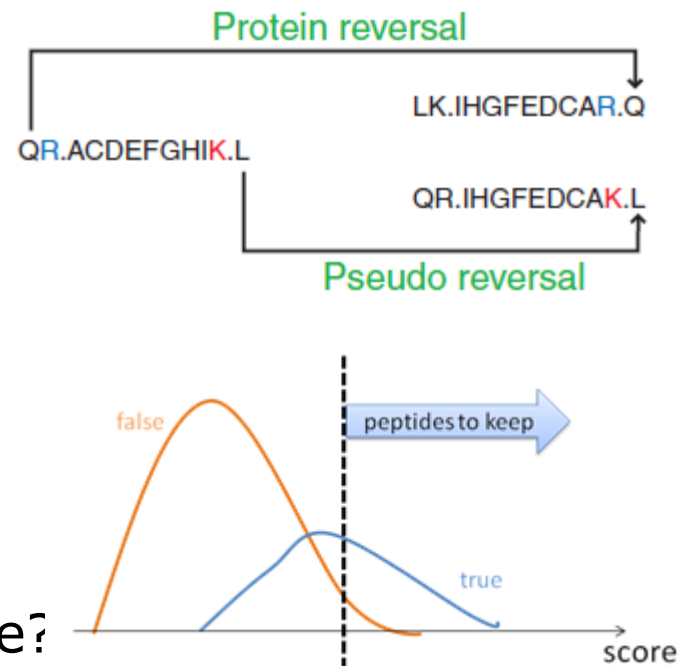
$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

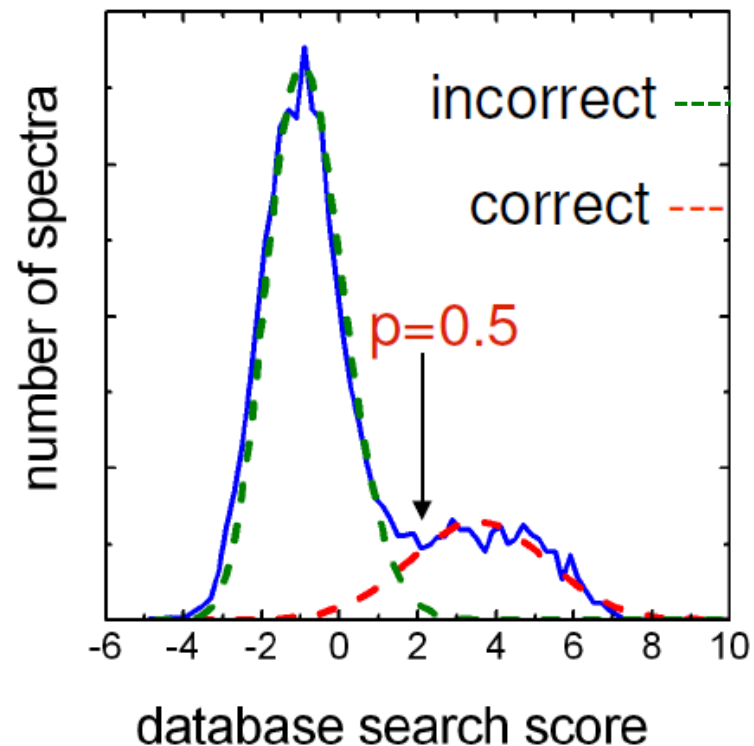
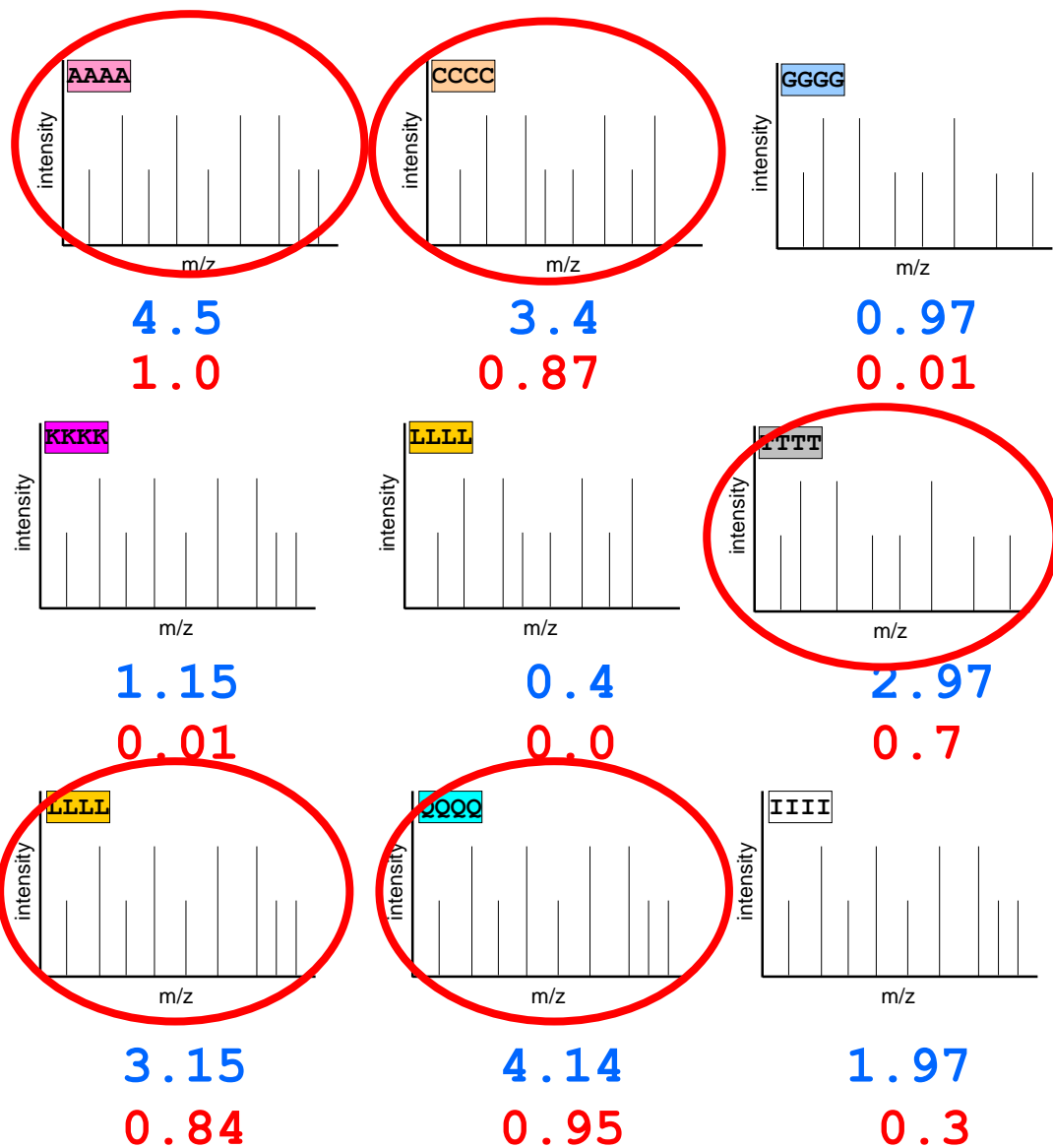
$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

Target/Decoy method

- Target/Decoy method is meaningful only for large datasets.
 - Decoy database? (database size, amino acid composition, peptide length distribution, precursor mass distribution)
 - Reversed sequence
 - Pseudo-reverse sequence
 - Random (Shuffled) sequence
 - Pseudo-shuffle
 - Separated or concatenated?
 - Threshold score 30
 - Match to the target score 50
 - Match to the decoy score 40
 - Is this counted as a false positive?
 - FDR calculated as: # of decoy hits / # of target hits above a certain threshold.
- The diagram illustrates the Target/Decoy method for peptide identification. It shows a protein sequence 'QR.ACDEFGHIK.L' being reversed to 'LK.IHGFEDCAR.Q' (Protein reversal) and a pseudo-reversed sequence 'QR.IHGFEDCAK.L' (Pseudo reversal). Below, a score distribution graph shows two overlapping curves: an orange curve for 'false' (decoy) and a blue curve for 'true' (target). A vertical dashed line marks a threshold score. The area under the orange curve to the right of the threshold is labeled 'false', and the area under the blue curve to the right of the threshold is labeled 'true'. A blue arrow labeled 'peptides to keep' points to the right, indicating the selection of peptides with scores above the threshold.



Probabilistic model-based: PeptideProphet



- Combine search scores into a single discriminant score.
- The discriminant score, F , can be computed as

$$F(x_1, x_2, \dots, x_S) = c_0 + \sum_{i=1}^S c_i x_i$$

- with constant c_0 and weights c_i derived such that the ratio of between-class variation to within-class variation is maximized.
- The discriminant score can be substituted to enable tractable calculation.

$$p(+|F) = \frac{p(F|+)p(+)}{p(F|+)p(+)+p(F|-)p(-)}$$

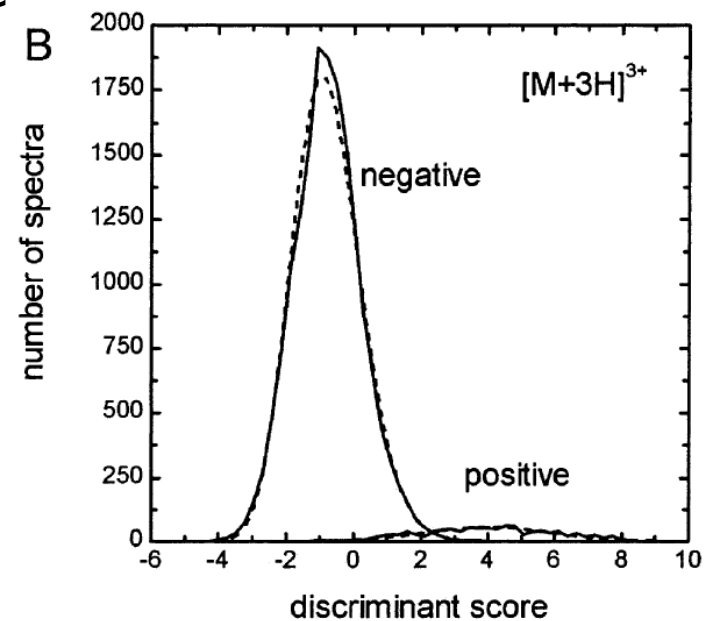
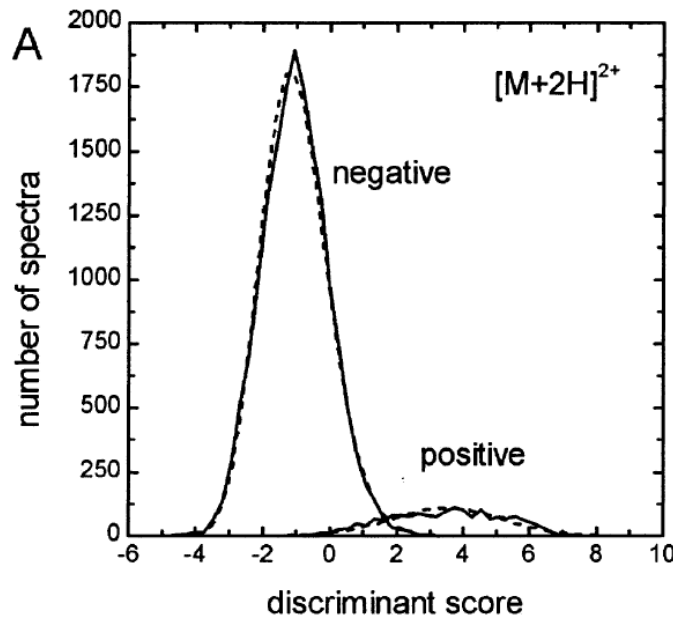
$$p(+|F, NTT) = \frac{p(F|+)p(NTT|+)p(+)}{p(F|+)p(NTT|+)p(+)+p(F|-)p(NTT|-)p(-)}$$

Table 1. Discriminant Functions Derived from Training Dataset Spectra of $[M + 2H]^{2+}$ and $[M + 3H]^{3+}$ Precursor Ions^a

variable	$[M + 2H]^{2+}$		$[M + 3H]^{3+}$	
	coefficient	correlation	coefficient	correlation
Xcorr'	8.362	0.798	9.933	0.698
ΔC_n	7.386	0.746	11.149	0.806
ln SpRank	-0.194	-0.510	-0.201	-0.491
d_M	-0.314	-0.306	-0.277	-0.251
constant	-0.959		-1.460	

– Among spectra of $[M + 2H]^{2+}$

- 84% of correct had discriminant scores of 1.7 or greater,
- whereas 99% of incorrect had scores below that

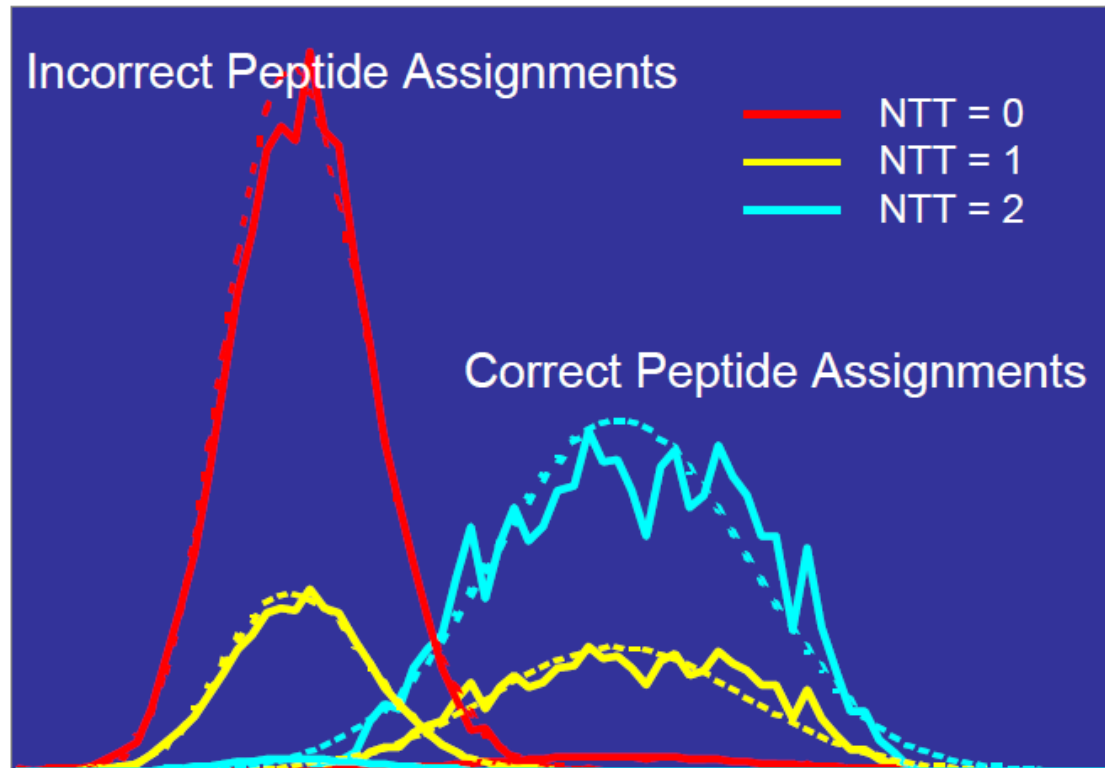


Training dataset(solid line) and Gaussian and gamma distributions(dashed line).

Gaussian: $p(F|+) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(F-\mu)^2/2\sigma^2}$

gamma: $p(F|-) = \frac{(F-\gamma)^{\alpha-1} e^{-(F-\gamma)/\beta}}{\beta^\alpha \Gamma(\alpha)}$

Semi-parametric PeptideProphet



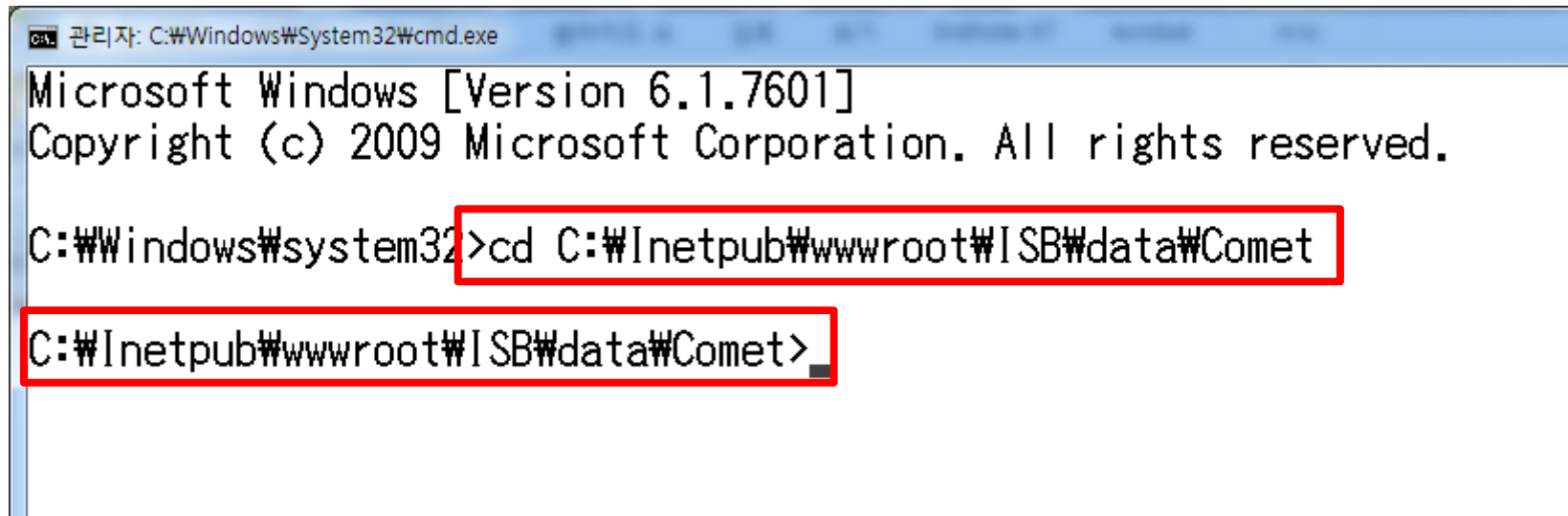
- Decoy search results => distribution for incorrect assignments
- EM algorithm to estimate distributions of correct assignments
 - NTT (number of tryptic termini)

Target/Decoy and PeptideProphet

- Cmd 창 실행
 - 시작메뉴
 - Cmd 입력



- Cmd 창
 - cd C:\Inetpub\wwwroot\ISB\data\Comet



```
관리자: C:\Windows\System32\cmd.exe
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.

C:\Windows\system32>cd C:\Inetpub\wwwroot\ISB\data\Comet

C:\Inetpub\wwwroot\ISB\data\Comet>
```

- Cmd 창

- java -jar CometTD.jar -i Sample.txt -fdr 0.01 -d XXX_

```
C:\Inetpub\wwwroot\ISB\data\Comet>java -jar CometTD.jar -i Sample.txt -fdr 0.01 -d XXX_
```

```
To Analyze      : Sample.txt  
Designated FDR  : 0.01  
Decoy Proteins starting with XXX_
```

```
_TargetDecoy for Charge State: 2, No. of Enzymatic Termini: 2  
Threshold=90.2000 | FDR= 0.96% | Target= 418 | Decoy= 4
```

```
_TargetDecoy for Charge State: 3, No. of Enzymatic Termini: 2  
Threshold=43.0000 | FDR= 0.00% | Target= 31 | Decoy= 0
```

```
_TargetDecoy for Charge State: equal to or more than 4, No. of Enzymatic Termini  
: 2  
Threshold=17.5000 | FDR= 0.00% | Target= 1 | Decoy= 0
```

```
_Analysis Report_  
At designated FDR 1.00%, Overall No. Identifications: 450  
Actual FDR= 0.89% | Target= 450 | Decoy= 4
```

TargetDecoy for Charge State: 2, No. of Enzymatic Termini: 2
Threshold=90.2000 | FDR= 0.96% | Target= 418 | Decoy= 4

• Threshold 확인

- C:\Inetpub\wwwroot\ISB\data\Comet 폴더 이동
- Sample.txt 파일 excel에서 열기

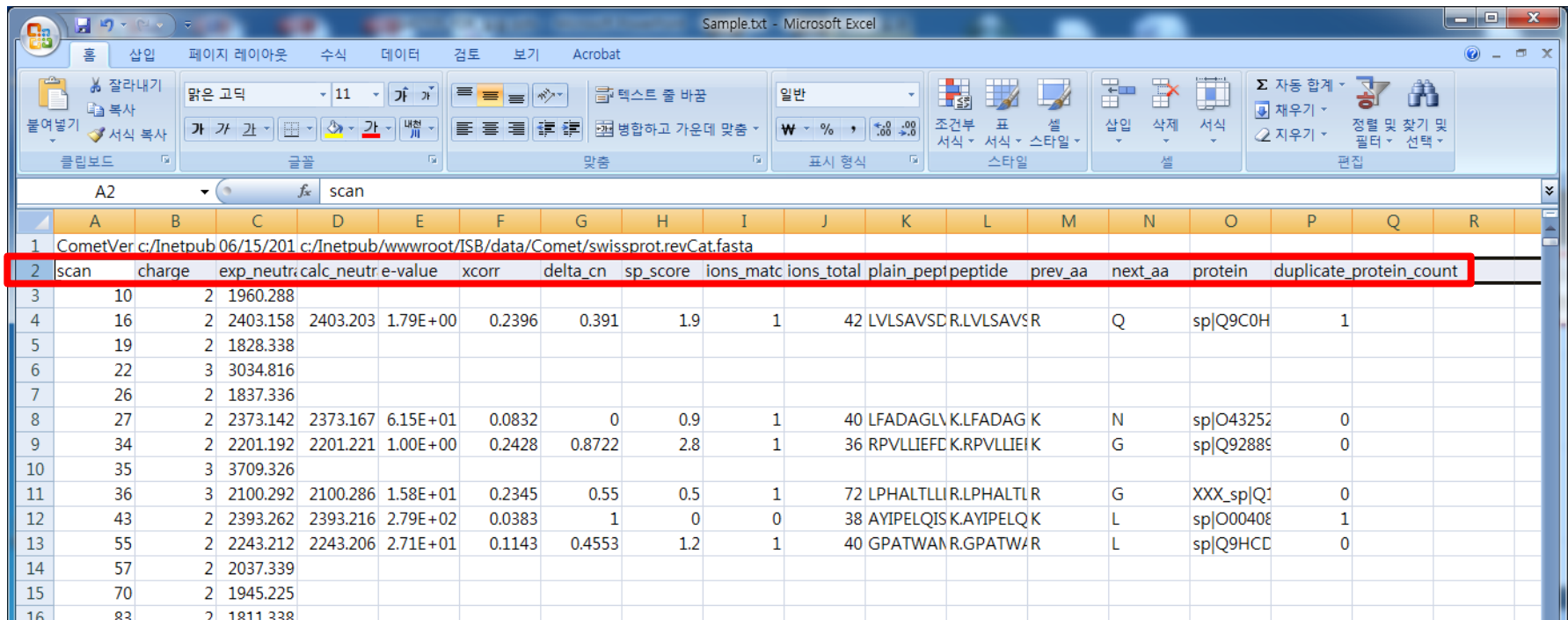
Sample.txt - Microsoft Excel																	
홈 삽입 페이지 레이아웃 수식 데이터 검토 보기 Acrobat																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View, Help]																	
[Ribbon: Font, Insert, Layout, References, Send to PDF, Review, View																	

Target/Decoy 방법

TargetDecoy for Charge State: 2, No. of Enzymatic Termini: 2
Threshold=90.2000 | FDR= 0.96% | Target= 418 | Decoy= 4

- Threshold 확인

- 두 번째 행 선택
- Shift + Ctrl + L



scan																
1	CometVer c:/Inetpub 06/15/201 c:/Inetpub/wwwroot/ISB/data/Comet/swissprot.revCat.fasta															
2	scan	charge	exp_neutr	calc_neutr	e-value	xcorr	delta_cn	sp_score	ions_mate	ions_total	plain_pept	peptide	prev_aa	next_aa	protein	duplicate_protein_count
3	10	2	1960.288													
4	16	2	2403.158	2403.203	1.79E+00	0.2396	0.391	1.9	1	42	LVLSAVS	LVLSAVS R	Q		sp Q9C0H	1
5	19	2	1828.338													
6	22	3	3034.816													
7	26	2	1837.336													
8	27	2	2373.142	2373.167	6.15E+01	0.0832	0	0.9	1	40	LFADAGL	K.LFADAG K	N		sp O43252	0
9	34	2	2201.192	2201.221	1.00E+00	0.2428	0.8722	2.8	1	36	RPVLLIEFC	K.RPVLLIE K	G		sp Q92885	0
10	35	3	3709.326													
11	36	3	2100.292	2100.286	1.58E+01	0.2345	0.55	0.5	1	72	LPHALTLLI	R.LPHALT L R	G		XXX_sp Q1	0
12	43	2	2393.262	2393.216	2.79E+02	0.0383	1	0	0	38	AYIPELQIS	K.AYIPEL Q K	L		sp O00408	1
13	55	2	2243.212	2243.206	2.71E+01	0.1143	0.4553	1.2	1	40	GPATWANR	.GPATWA R	L		sp Q9HCD	0
14	57	2	2037.339													
15	70	2	1945.225													
16	83	2	1811.338													

TargetDecoy for Charge State: 2, No. of Enzymatic Termini: 2
Threshold=90.2000 | FDR= 0.96% | Target= 418 | Decoy= 4

- Threshold 확인
 - sp_score에서 숫자 내림차순 정렬

Sample.txt - Microsoft Excel

홈삽입페이지 레이아웃수식데이터검토보기Acrobat

잘라내기
붙여넣기
클립보드

암은 고딕 11
가
가
가
글꼴

텍스트 줄 바꿈
병합하고 가운데 맞춤
맞춤

일반
W %
표시 형식

조건부 서식
표 서식
셀 스타일

삽입
삭제
서식
셀

자동 합계
채우기
지우기
편집

정렬 및 필터
정렬 및 찾기 및 필터
선택

H2

sp_score

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	CometVer c:\Inetpub 06/15/201 c:\Inetpub\wwwroot\ISB\data\Comet\swissprot.revCat.fasta																	
2	scan	charge	exp_neu	calc_neu	e-value	xcorr	delta_cr	sp_score	ions_ma	ions_tot	plain_p	peptide	prev_aa	next_aa	protein	duplica	protein_count	
3	10	2	1960.288															
4	16	2	2403.158	2403.203	1.79E+0			1.9	1	42	LVLSAVSD	R.LVLSAVSR	Q		sp Q9C0H	1		
5	19	2	1828.338															
6	22	3	3034.816															
7	26	2	1837.336															
8	27	2	2373.142	2373.167	6.15E+01			0.9	1	40	LFADAGLV	K.LFADAG K	N		sp O43252	0		
9	34	2	2201.192	2201.221	1.00E+00			2.8	1	36	RPVLLIEFC	K.RPVLLIE K	G		sp Q92889	0		
10	35	3	3709.326															
11	36	3	2100.292	2100.286	1.58E+01			0.5	1	72	LPHALTLRI	L.LPHALTL R	G		XXX_sp Q	0		
12	43	2	2393.262	2393.216	2.79E+02			0	0	38	AYIPELQIS	K.AYIPELQ K	L		sp O00408	1		
13	55	2	2243.212	2243.206	2.71E+01			1.2	1	40	GPATWANR	R.GPATWAN R	L		sp Q9HCD	0		
14	57	2	2037.339															
15	70	2	1945.225															
16	83	2	1811.338															
17	86	2	2345.009	2345.033	1.89E+01			0	0	38	GEEEDPEV	K.GEEEDPI K	V		sp Q9UNX	0		
18	88	2	841.6557															
19	96	3	3298.93	3298.87	2.71E+02			0	0	112	VHNLISIIG	R.VHNLISI R	S		XXX_sp O6	1		

숫자 오름차순 정렬(S)
숫자 내림차순 정렬(D)
생기주 정렬(M)
"sp_score"에서 필터 해제(C)
색 기준 필터(F)
숫자 필터(E)
(모두 선택)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
확인
취소

TargetDecoy for Charge State: 2, No. of Enzymatic Termini: 2
Threshold=90.2000 | FDR= 0.96% | Target= 418 | Decoy= 4

- Threshold 확인
 - Charge에서 2 선택, 확인 클릭

The screenshot shows a Microsoft Excel spreadsheet titled 'Sample.txt - Microsoft Excel'. The spreadsheet contains a table of protein search results. The columns are labeled as follows: A (CometVer), B (scan), C (charge), D (exp_neu), E (calc_neu), F (e-value), G (xcorr), H (delta_cr), I (sp_scor), J (ions_ma), K (ions_to), L (plain_p), M (peptide), N (prev_aa), O (next_aa), P (protein), Q (duplica), and R (protein_count). The table lists various protein entries with their corresponding scores and values. On the left side, a filter menu is open for the 'charge' column. The menu shows a list of values: 2, 3, 4, 5, and 6. The value '2' is selected and highlighted with a red box. Below the list, the '확인' (Confirm) button is also highlighted with a red box.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	CometVer	c:/inetpub/06/15/201	c:/inetpub/wwwroot/ISB/data/Comet/swissprot.revCat.fasta														
2	scan	charge	exp_neu	calc_neu	e-value	xcorr	delta_cr	sp_scor	ions_ma	ions_to	plain_p	peptide	prev_aa	next_aa	protein	duplica	protein_count
			989.5707	989.5698	5.19E-13	2.1866	0.5453	660.2	10	14	LGFQVWL	R.LGFQVWR	N	sp Q01995	0		
			1081.484	1081.483	5.70E-19	2.1754	0.6547	441.1	9	16	DAEAWFT	K.DAEAWfK	T	sp P08727	2		
			1720.926	1720.918	1.98E-09	2.872	0.5769	440	12	28	TNDQMVA	K.TNDQMfK	S	sp P51665	0		
			1110.64	1110.64	7.46E-16	2.4637	0.7422	432.8	11	22	AVLIAGQF	R.AVLIAGfR	T	sp Q9Y230	0		
			1511.83	1511.831	8.06E-10	2.5982	0.6209	421	12	26	IQGLTVEQ	R.IQGLTVER	L	sp Q14204	0		
			1764.838	1764.835	1.38E-12	3.0313	0.7497	418.1	14	30	LDILDTAG	R.LDILDfAR	E	sp P62070	2		
			1490.79	1490.788	3.76E-12	2.6289	0.7066	414.8	11	24	NQGNTWIR	N.QGNTfR	T	sp P01023	2		
			1510.867	1510.872	4.39E-08	2.5625	0.6352	412.2	11	26	VVLLEDLA	K.VVLLEDfK	T	sp Q96HY0	1		
			1121.573	1121.572	1.80E-12	2.6024	0.6281	398.4	10	18	IDYIAGLD	R.IDYIAGfL	G	sp P07741	1		
			1128.614	1128.614	1.32E-11	2.5147	0.4394	379.1	9	18	LSELEAAL	K.LSELEAAK	A	sp P05787	9		
			1041.582	1041.582	1.70E-07	1.8048	0.4318	337.5	8	16	ALQALEEL	R.ALQALEfR	L	sp Q15149	8		
			1280.714	1280.713	3.55E-08	2.1987	0.7485	334.6	8	18	LLELQYFIS	R.LLELQYfR	D	sp Q92896	2		
			1319.681	1319.672	1.81E-10	2.378	0.5486	333.3	9	22	YLIANATN	K.YLIANATK	V	sp P27348	0		
			1128.614	1128.614	2.81E-10	2.4038	0.2909	326	10	18	LSELEAAL	K.LSELEAAK	A	sp P05787	9		
			1074.613	1074.611	8.38E-09	1.938	0.5258	319.5	8	16	MLISILTER	K.MLISILfK	S	sp P12429	0		
			1108.577	1108.576	1.19E-09	1.9995	0.2727	318.8	10	16	LLEYDVTIR	LLEYDfVR	E	sp Q9HDC0	1		
			1708.858	1708.857	1.14E-09	2.2337	0.7466	318.6	9	26	NLTQYSW	K.NLTQYSfK	T	sp Q9UJ70	1		

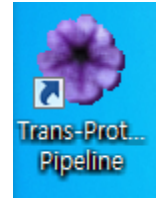
Target/Decoy 방법

TargetDecoy for Charge State: 2, No. of Enzymatic Termini: 2
Threshold=90.2000 | FDR= 0.96% | Target= 418 | Decoy= 4

- Threshold 확인
 - Threshold 확인 = 90.2

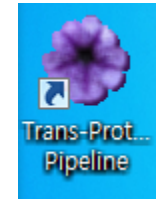
401	3179	2	948.4730	948.4099	3.93E-04	1.1771	0.2339	93.1	5	14	MINI DLSK K.MINI DLSK	I	sp P30578	U
462	7077	2	1444.721	1444.72	7.36E-05	1.3655	0.5741	92.8	6	24	WDDSGNIK.WDDSGIK	Q	sp P26232	7
463	10466	2	1385.77	1385.767	5.75E-05	1.7048	0.5607	92.8	6	22	VGWEQLL R.VGWEQIR	T	sp P12814	3
464	1412	2	685.3932	685.3911	3.71E-05	1.4022	0.0689	92.6	4	10	HVVF GK K.HVVF GK K	V	sp P45877	3
465	4484	2	1027.498	1027.497	1.17E-05	1.1676	0.4557	92.6	6	14	NSLF EYQK K.NSLF EYQ K	N	sp P02671	1
466	3518	2	748.4239	748.4232	1.32E-05	1.1051	0.0448	92.5	5	12	VAFSVAR R.VAFSVAIR	A	sp P39023	0
467	7581	2	1151.609	1151.609	2.78E-04	1.7439	0.5957	92.5	6	18	IANVFTNAR.IANVFTNR	Y	sp P05164	2
468	8489	2	989.5718	989.5698	4.56E-07	1.7155	0.4594	92.5	5	14	LGFQVWL R.LGFQVMR	N	sp Q01995	0
469	1843	2	784.4544	784.4443	1.02E-02	1.3165	0.2394	92	5	14	VLNPSGAIR.VLNPSG R	T	sp Q14315	1
470	11395	2	1320.671	1320.671	6.84E-05	1.6685	0.6403	92	7	26	STSGGTA K.STSGGT K	D	sp P01857	1
471	4086	2	1295.628	1295.626	2.50E-05	1.8041	0.6137	91.9	6	22	GSYVSIH K.GSYVSIH K	D	sp Q13838	2
472	7256	2	1418.743	1418.741	7.42E-08	1.6258	0.5878	91.8	10	22	QAQEYEA R.QAQEYER	V	sp P05783	3
473	2180	2	1030.567	1030.566	8.55E-06	1.4848	0.6746	91.7	5	18	LLEAAQSR.LLEAAAR	G	sp Q15149	8
474	6194	2	1091.565	1091.561	4.04E-07	1.6977	0.5773	91.5	5	16	AELNEFLTK.AELNEFL K	E	sp P23396	0
475	7813	2	1734.893	1734.879	1.47E-04	2.0746	0.6146	91.3	7	34	LGNTISSLF K.LGNTISS K	E	sp Q9Y4L1	0
476	6965	2	965.6288	965.6274	5.55E-06	1.5159	0.6849	91.1	6	16	LPLILVGNI R.LPLILVGIR	S	sp Q8IXI2	5
477	9717	2	1352.832	1352.839	8.60E-05	1.825	0.82	91.1	5	22	NLQNLLIL R.NLQNLLIR	A	sp Q00610	3
478	7836	2	1083.667	1083.665	1.25E-01	1.0124	0.1871	90.7	6	18	LVLPSLISS K.LVLPSLISS K	I	sp Q13228	1
479	9962	2	1360.692	1360.691	3.76E-07	1.7782	0.6347	90.7	6	20	SLFDYFLTK.R.SLFDYFL R	C	sp Q9NQ8	0
480	4987	2	1409.625	1409.625	2.17E-06	1.6197	0.589	90.6	6	18	YFYNQEEY R.YFYNQEIR	F	sp Q30134	1
481	3018	2	741.45	741.4497	3.08E-04	1.0032	0.1117	90.3	4	10	LNQLVR K.LNQLVR K	R	XXX_sp Q8	0
482	3485	2	837.5083	837.496	1.00E+00	0.9294	0.028	90.3	6	14	LPEGVVPK R.LPEGVVIR	S	sp Q9NVS	1
483	6356	2	1399.638	1399.641	1.33E-08	1.6915	0.2732	90.3	7	22	YGGDEIPF K.YGGDEIF K	V	sp P21333	1
484	8920	2	1518.828	1518.823	1.01E-02	1.5423	0.5992	90.3	6	28	FGANAILG K.FGANAIL K	A	sp P06733	5
485	9353	2	1502.789	1502.79	8.83E-09	2.221	0.6328	90.3	6	24	LVSDFMV K.LVSDFMK	N	sp P54819	5
486	6793	2	1176.638	1176.633	1.84E-02	1.0946	0.1355	90.2	5	18	VLYCAAQIR.VLYCAA R	A	XXX_sp Q8	1
487	8234	2	1239.684	1239.684	3.30E-02	1.2782	0.334	89.7	4	18	HRPYQVIT R.HRPYQVR	V	sp Q96MI9	2
488	3884	2	1549.825	1549.821	8.18E-06	2.1249	0.6669	89.6	8	30	LVGGPVA K.LVGGPV K	G	sp O15230	0
489	10963	2	1070.548	1070.547	3.73E-06	1.1392	0.4069	89.6	5	16	MFLSFPTT R.MFLSFPT R	T	sp P69905	0
490	8866	2	1557.806	1557.804	6.01E-07	2.0982	0.4725	89.4	7	26	FLSSSLYTA K.FLSSSLYTK	R	sp Q16706	0
491	2192	2	1235.507	1235.505	1.73E-07	1.8023	0.6682	89.3	5	22	DGEEAGA R.DGEEAGR	T	sp P30101	0
492	4105	2	972.5608	972.5604	4.75E-06	1.2805	0.4463	89.3	5	16	NIQVAITTK.NIQVAIT K	N	sp P12111	2

- TPP 실행
 - 바탕화면에 아이콘 실행



- TPP 실행

- 바탕화면에 아이콘 실행



- TPP 웹페이지에서 로그인

- User Name: guest
 - Password: guest

- Login 클릭

ISB/SPC Trans Proteomic Pipeline - login

User Name:

Password:

Login

INSTITUTE FOR
Systems Biology
Revolutionizing science. Enhancing life.

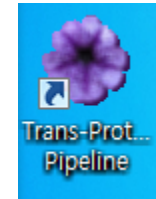
Seattle
Proteome
Center



TPP v4.8.0 PHILAE, Build 201411201551-6764 (mingw-i686)

- TPP 실행

- 바탕화면에 아이콘 실행



- TPP 웹페이지에서 로그인

- User Name: guest
 - Password: guest

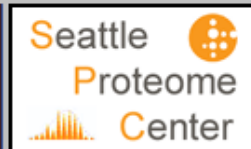
- Login 클릭

ISB/SPC Trans Proteomic Pipeline - login

User Name:

Password:

Login



TPP v4.8.0 PHILAE, Build 201411201551-6764 (mingw-i686)

- Analysis pipeline 확인
 - Comet으로 선택

The screenshot shows the ISB/SPC Trans Proteomic Pipeline (TPP) web interface. A red box highlights the dropdown menu for selecting the analysis pipeline, which currently shows 'Comet'. A red arrow points from this dropdown to a callout box. The callout box has a blue header 'Analysis Pipeline' and a list of options: 'Comet', 'Mascot', 'Sequest', 'SpectraST', and 'Tandem'. The 'Comet' option is highlighted in blue. The background interface includes a navigation bar with links like Home, Files, Account, Pre-Process, mzXML Utils, Analysis Pipeline (Comet), Decoy, Utilities, SpectraST Tools, and Jobs. A welcome message is displayed, and a list of steps for data analysis is provided: 1. RAW to mzML Conversion, 2. Peptide Database Search and Identification. The footer contains logos for the Institute for Systems Biology and the Seattle Proteome Center, along with version information: TPP v4.8.0 PHILAE, Build 201411201551-6764 (mingw-i686).

ISB/SPC Trans Proteomic Pipeline - home

Home | Files | Account | Pre-Process | mzXML Utils | Analysis Pipeline (Comet) | Decoy | Utilities | SpectraST Tools | Jobs

You are logged in as guest Log Out

Home FILES ACCOUNT PRE-PROCESS mzXML UTILS ANALYSIS PIPELINE DECOY UTILITIES SPECTRAST TOOLS JOBS

Messages [Show / Hide]

Welcome, guest.

Welcome

Welcome to the Trans-Proteomic Pipeline (TPP) web interface. These tools and interfaces were developed and are being maintained at the [Institute for Systems Biology](#) (ISB) under a grant from [NIGMS](#). Please visit [www.isb-spi.org](#) and [www.proteomecenter.org](#) for more information.

Please select analysis pipeline you want to use: Comet

Analysis Pipeline

Follow these steps to convert, search, and analyze your data:

1. RAW to mzML Conversion
Convert original .RAW files to the standard mzML input format used by the tools
2. Peptide Database Search and Identification

tions using PeptideProphet and/or use ASAPRatio or XPRESS to calculate the relative abundances

perienced users should modify.

uss mailing list. Click [here](#) to find out how to join this list.

Please select analysis pipeline you want to use: Comet

Analysis Pipeline

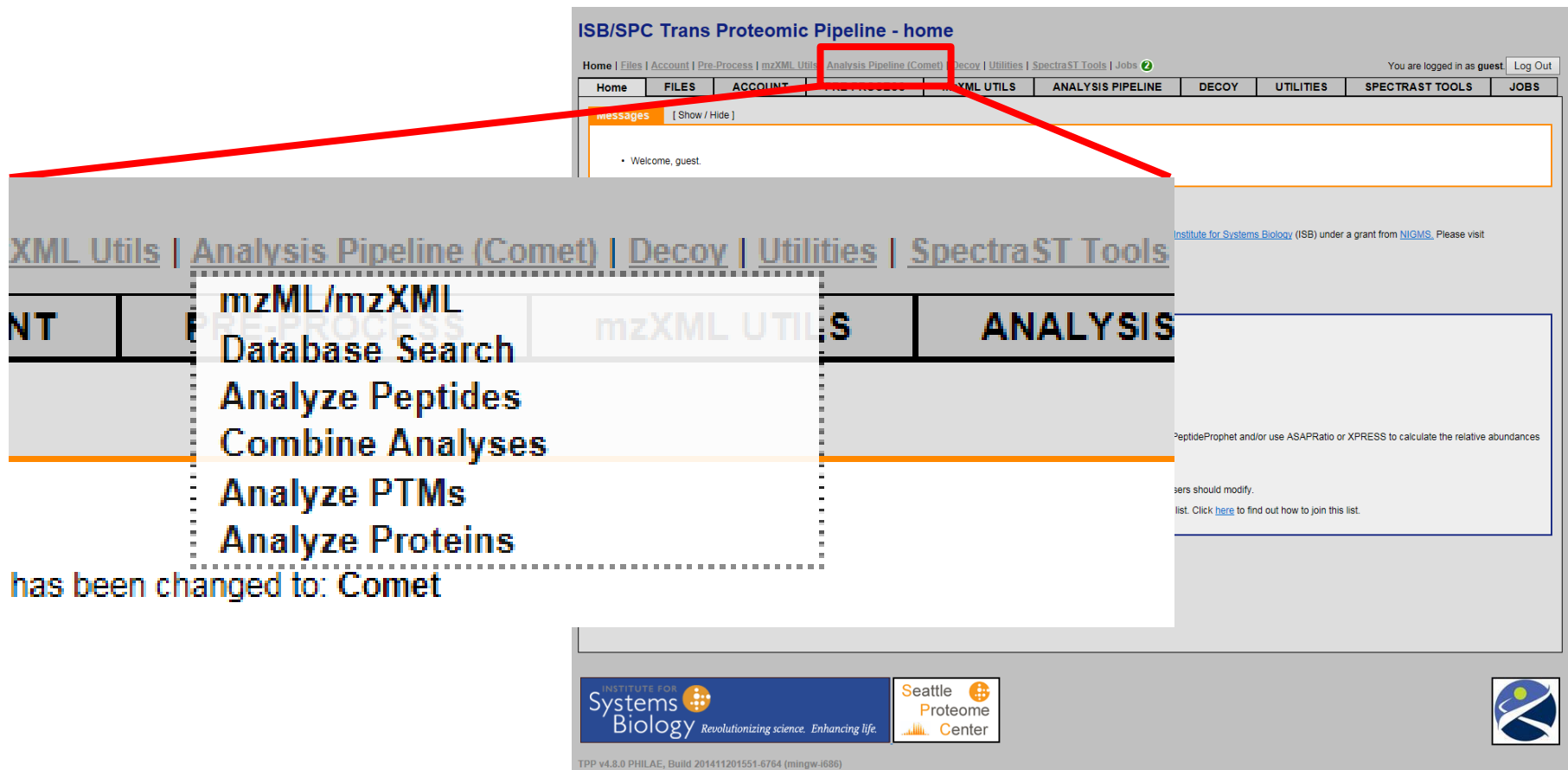
Comet
Mascot
Sequest
SpectraST
Tandem

INSTITUTE FOR Systems Biology Revolutionizing science. Enhancing life.

Seattle Proteome Center

TPP v4.8.0 PHILAE, Build 201411201551-6764 (mingw-i686)

- Analysis pipeline 확인
 - Analysis Pipeline(Comet)



ISB/SPC Trans Proteomic Pipeline - home

Home | Files | Account | Pre-Process | mzXML Utils | **Analysis Pipeline (Comet)** | Decoy | Utilities | SpectraST Tools | Jobs

You are logged in as guest. Log Out

Home FILES ACCOUNT PRE-PROCESS **ANALYSIS PIPELINE** DECOY UTILITIES SPECTRAST TOOLS JOBS

messages [Show / Hide]

Welcome, guest.

[Institute for Systems Biology \(ISB\)](#) under a grant from [NIGMS](#). Please visit

[XML Utils](#) | [Analysis Pipeline \(Comet\)](#) | [Decoy](#) | [Utilities](#) | [SpectraST Tools](#)

ANALYSIS PIPELINE

- mzML/mzXML
- Database Search
- Analyze Peptides
- Combine Analyses
- Analyze PTMs
- Analyze Proteins

has been changed to: Comet

PeptideProphet and/or use ASAPRatio or XPRESS to calculate the relative abundances

users should modify.

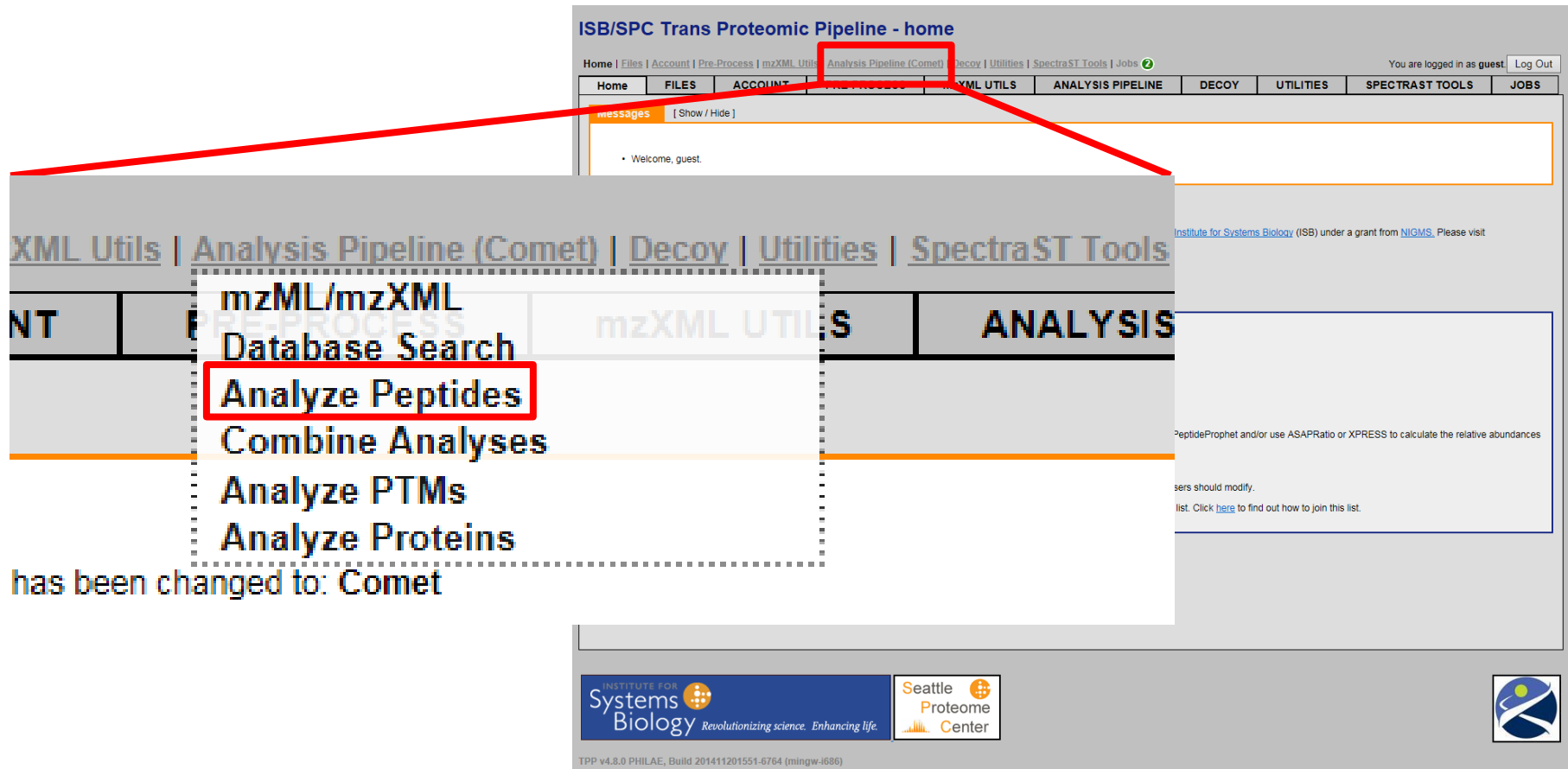
list. Click [here](#) to find out how to join this list.

Institute for Systems Biology Revolutionizing science. Enhancing life.

Seattle Proteome Center

TPP v4.8.0 PHILAE, Build 201411201551-6764 (mingw-i686)

- Analysis pipeline 확인
 - Analysis Pipeline(Comet)
 - Analyze Peptides 클릭



ISB/SPC Trans Proteomic Pipeline - home

Home | Files | Account | Pre-Process | mzXML Utils | **Analysis Pipeline (Comet)** | Decoy | Utilities | SpectraST Tools | Jobs

You are logged in as guest. Log Out

Home FILES ACCOUNT ANALYSIS PIPELINE DECOY UTILITIES SPECTRAST TOOLS JOBS

messages [Show / Hide]

Welcome, guest.

XML Utils | Analysis Pipeline (Comet) | Decoy | Utilities | SpectraST Tools

NT PRE-PROCESS mzXML UTILS ANALYSIS

mzML/mzXML
Database Search
Analyze Peptides
Combine Analyses
Analyze PTMs
Analyze Proteins

has been changed to: Comet

Institute for Systems Biology (ISB) under a grant from NIGMS. Please visit

PeptideProphet and/or use ASAPRatio or XPRESS to calculate the relative abundances

users should modify.
list. Click [here](#) to find out how to join this list.

INSTITUTE FOR
Systems Biology Revolutionizing science. Enhancing life.

Seattle
Proteome
Center

TPP v4.8.0 PHILAE, Build 201411201551-6764 (mingw-i686)

- Peptide identification 결과 선택
 - Add files 클릭

Select File(s) to Analyze [Show / Hide]

No files selected yet.

Add Files

ISB/SPC Trans Proteomic Pipeline - xinteract

[Home](#) | [Files](#) | [Account](#) | [Pre-Process](#) | [mXML/mXML](#) | [mXML/SPC](#) | [Analysis Pipeline \(Current\)](#) | [Display](#) | [Utilities](#) | [Support](#) | [Tools](#) | [Help](#) | [Log Out](#)

[Home](#) | [mXML/mXML](#) | [Database Search](#) | [Analyze Peptides](#) | [Combine Analyses](#) | [Analyze PTMs](#) | [Analyze Proteins](#)

Select File(s) to Analyze [Show / Hide]

No files selected yet.

Add Files

Output File and Filter Options

Please select input files first.

PeptideProphet Options [Show / Hide]

☒ RUN PeptideProphet
☐ Use accurate mass binning, using Da
☐ Use pI information
☐ Use Phospho information
☐ Use hydrophobicity / RT information
☐ Use isotopic information
☐ Do not use isotopic information
☐ Do not use the NIST model
☐ Do not use the NIST model
☐ Use only use Expect Score as the dominant - helpful for data with homologous top hits, e.g. phospho or glyco (Tandem and Comet only)
☐ Use Gamma distribution test for the negative (Tandem only)
☐ Force the fitting of the mixture model (overrides automatic mixture model checks)
☐ Use decoy hits to pin down the negative distribution (decoy Protein names begin with (antipeptide not allowed))
☐ Use Non-parameterized (can only be used with decoy option)
☐ Report decoy hits with computed probability (based on the model learned)
☐ Ignore charge states ☐ +1 ☐ +2 ☐ +3 ☐ +4 ☐ +5
☐ Run ProteinProphet afterwards
 Enter additional options to pass directly to the command-line (expert use only)

InterProphet Options [Show / Hide]

☐ RUN InterProphet
☐ Activate InterProphet on these results
☐ do NOT use number of replicate spectra (NIST) model
☐ do NOT use number of sibling ions (NIST) model
☐ do NOT use number of sibling modifications (NIST) model
☐ do NOT use number of sibling searches (NIST) model
☐ do NOT use number of sibling experiments (NIST) model
☐ do NOT use number of sibling peptides (NIST) model

PTMProphet Options [Show / Hide]

☐ RUN PTMProphet
 Specify modifications:
 Residue(s): Mass shift:
 Residue(s): Mass shift:
 Residue(s): Mass shift:
 Residue(s): Mass shift:
 Residue(s): Mass shift:
 min tolerance: daltons
☐ Do not update modification_info tags in pepXML.

APRIS Options [Show / Hide]

☐ RUN APRIS
 Mass tolerance: Da (Default)
☐ For ratio, set to light to 1, very heavy
☐ For ratio, set to heavy to 1, very light
☒ Default (set less abundant to zero to 1)
☐ Heavy labeled peptide elutes before light labeled partner
 Residue 1 and mass difference:
 Residue 2 and mass difference:
 Residue 3 and mass difference:
☐ Fix elution peak area as scans from peak apex
 Minimum number of chromatogram points needed for quantitation:
 Number of isotopic peaks to sum, use narrow tolerance:
☐ Heaviest Labeling [4:1:1:1:1] - ignore all other parameters, and
☐ assume CIs are normal and quantify corresponding 15N or 13C heavy pair
☐ assume CIs are 15N or 13C heavy and quantify corresponding 14N or 12C light pair

ASAPRatio Options [Show / Hide]

☐ RUN ASAPRatio
☐ Static modification quantification (i.e. each run is either all light or all heavy)
 Labeled residues:
☐ Heavy labeled peptide elutes before light labeled partner
☐ Use fixed scan range for light and heavy
☐ Quantitate only the charge state where the CID was made
 Set area/flag to: (ratio display option)
☐ Zero out all background
☐ Quantitate despite high background
 min range to include in summation of peak:
 Residue 1 mass:
 Residue 2 mass:
 Residue 3 mass:
 Residue 4 mass:
 Residue 5 mass:
☐ specify monoisotopic masses

Laba Quantification Options [Show / Hide]

☐ RUN Laba
 Condition File: [Click here to generate a condition file](#)

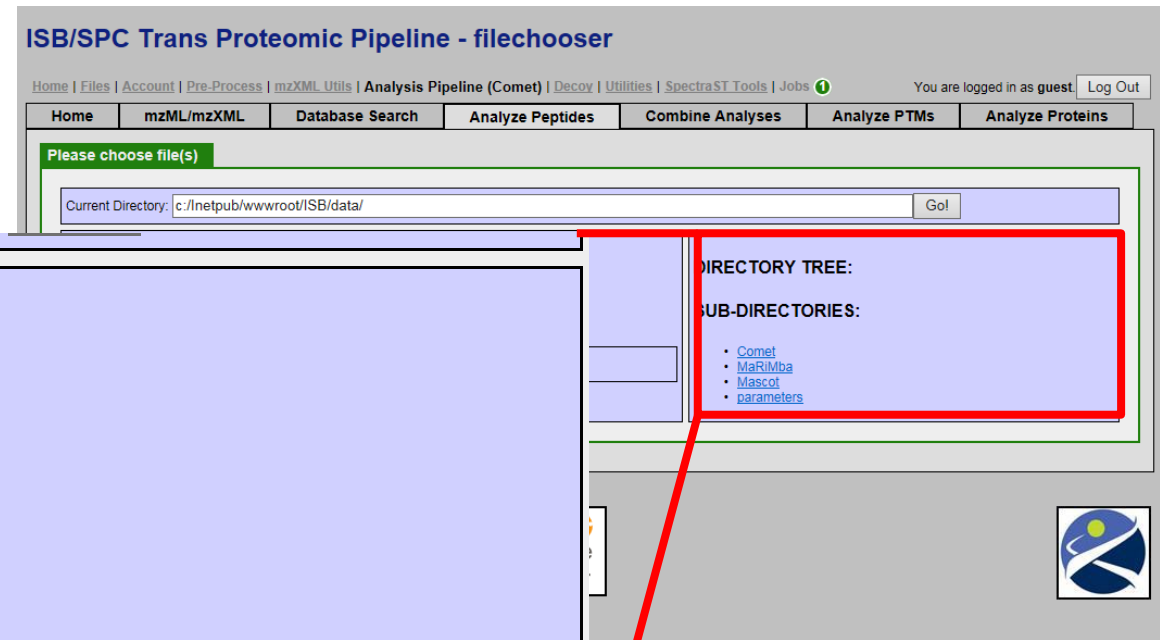
Run Analysis!

No files selected yet.

Systems Biology Revolutionizing science. Enhancing life.

Seattle Proteome Center

- Peptide identification 결과 선택
 - Comet 클릭

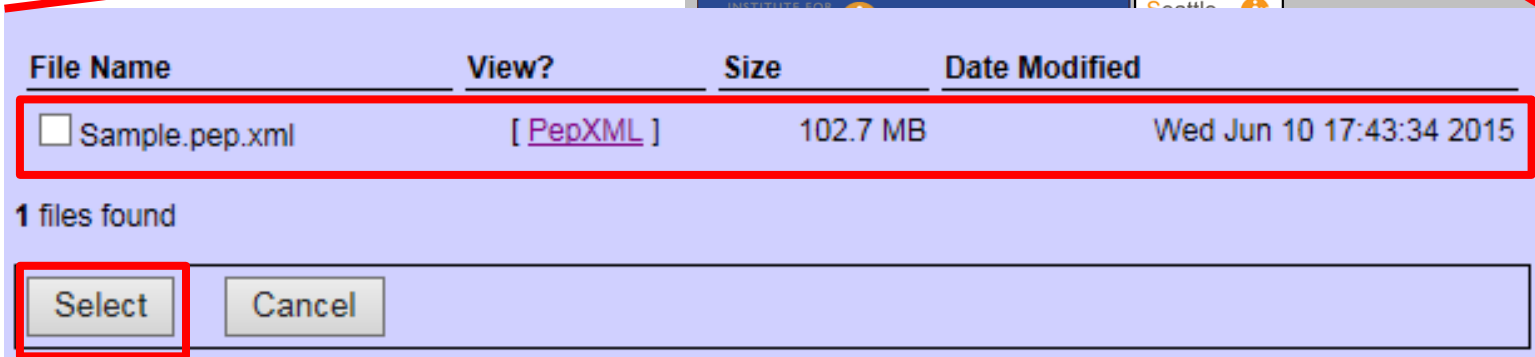
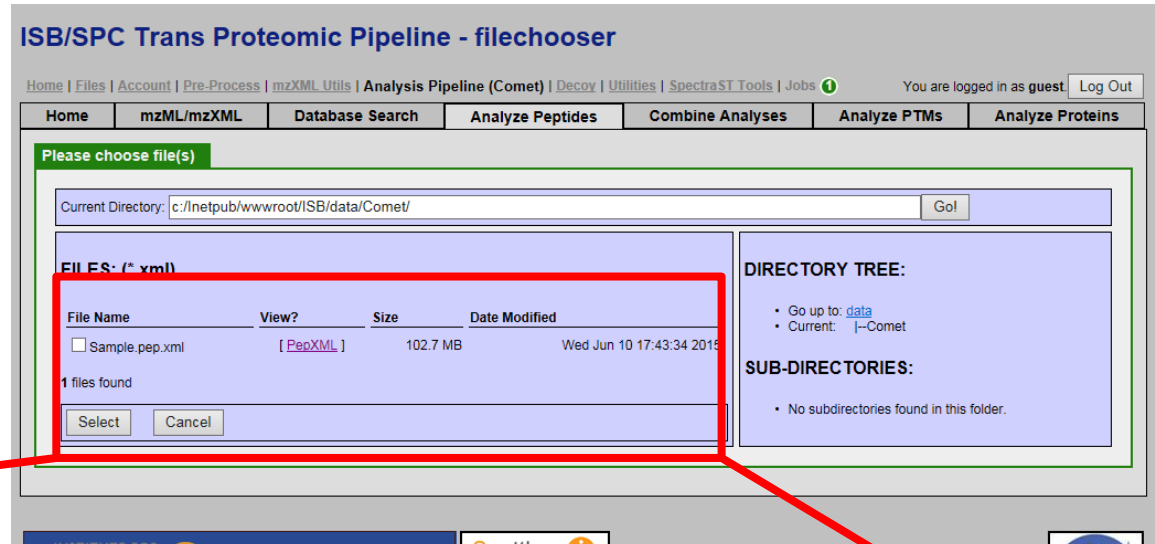


DIRECTORY TREE:

SUB-DIRECTORIES:

- [Comet](#)
- [MaRimba](#)
- [Mascot](#)
- [parameters](#)

- Peptide identification 결과 선택
 - Sample.pep.xml 선택
 - Select 클릭



• Peptide identification 결과 확인

Select File(s) to Analyze [Show / Hide]

`c:/inetpub/wwwroot/ISB/data/Comet/Sample.pep.xml`

Add Files Or choose from list: `c:/inetpub/wwwroot/ISB/data/Comet/` [xml] ▼

ISB/SPC Trans Proteomic Pipeline - xinteract

Home | Files | Accounts | Data Processes | Results | Analysis Pipeline (Current) | Query | Settings | Search | Tools | Help | You are logged in as guest | Log Out

Messages [Show / Hide]

- Your files have been added.
- Sample.pep.xml

Select File(s) to Analyze [Show / Hide]

`c:/inetpub/wwwroot/ISB/data/Comet/Sample.pep.xml`

Add Files Or choose from list: `c:/inetpub/wwwroot/ISB/data/Comet/` [xml] ▼

PeptideProphet Options [Show / Hide]

☒ Run PeptideProphet

☐ Use accurate mass timing, using [0.5] Da

☐ Use pI information

☐ Use Phospho information

☐ Use tryptic motif information

☐ Use hydrophobicity / RT information

☐ Use salt information

☐ Do not use salt information

☐ Do not use the N-T1 model

☐ Do not use the NAC model

☐ MALDI data

☐ Only use Exact mass for the discriminant: helpful for data with homologous top hits, e.g. phospho or glyco (Tandem and Comet only)

☐ Use Gamma distribution to model the negatives (Tandem only)

☐ Force the fitting of the mixture model (bypass automatic mixture model checks)

☐ Use decoy hits to show the negative distribution (Comet: Protein names begin with [] whitespace not allowed)

☐ Use Non-parametric model (can only be used with decoy option)

☐ Report decoy hits with a computed probability (based on the model learned).

Ignore charge states: ☐ +1 ☐ +2 ☐ +3 ☐ +4 ☐ +5

☐ Run PeptideProphet afterwards

Enter additional options to pass directly to the command-line (expert use only):

IdentProphet Options [Show / Hide]

☐ Run IdentProphet

☐ May run ProteinProphet on these results

☐ do NOT use number of replicate spectra (NRS) model

☐ do NOT use number of sibling ions (NSI) model

☐ do NOT use number of sibling modifications (NSM) model

☐ do NOT use number of sibling searches (NSS) model

☐ do NOT use number of sibling experiments (NSE) model

☐ do NOT use number of sibling peptides (NSP) model

PTMProphet Options [Show / Hide]

☐ Run PTMProphet

Specify modifications:

Residue(s)	CTY	Mass shift
Residue(s)		Mass shift
Residue(s)		Mass shift
Residue(s)		Mass shift
Residue(s)		Mass shift
Residue(s)		Mass shift

MS tolerance: [0.1] Da

☐ Do not update modification_info tags in peptide.

XPRESS Options [Show / Hide]

☐ Run XPRESS

Mass tolerance: [0.5] Da

☐ For ratio, weight light to 1, very heavy

☐ For ratio, weight heavy to 1, very light

☒ Default (set less abundant in ratio to 1)

☐ Heavily labeled peptide elutes before light labeled partner

Residue 1 and mass difference: [0.5] Da

Residue 2 and mass difference: [0.5] Da

Residue 3 and mass difference: [0.5] Da

☐ Fix elution peak area as +/- [5] scans from peak apex

Minimum number of chromatogram points needed for quantification: [5]

Number of isotopic peaks to sum, use narrow tolerance: [7]

☐ Metabolic Labeling [14C15N] -- ignore all other parameters, and

☐ assume ICRs are normal and quantify with corresponding 12N or 13C heavy pair

☐ assume ICRs are 12N or 13C heavy and quantify with corresponding 14N or 15C light pair

ASAPRatio Options [Show / Hide]

☐ Run ASAPRatio

☐ Blank modification quantification (i.e. each hit is either at light or at heavy)

Labelled residues: [C] [L] [V] [W] [Y] []

☐ Heavily labeled peptide elutes before light labeled partner

☐ Use fixed scan range for light and heavy

☐ Quantitate only the charge state where the CID was made

Set analysis to [] (info display option)

☐ Zero out all background

☐ Quantitate despite high background

Set range to include in normalization of peak: [0.5]

Residue 1 mass: []

Residue 2 mass: []

Residue 3 mass: []

Residue 4 mass: []

Residue 5 mass: []

☐ specify monoisotopic masses

Label Quantification Options [Show / Hide]

☐ Run Label

Condition file: [] [Click here to generate a condition file](#)

Run Analysis!

Systems Biology **Seattle Proteome Center**

PeptideProphet v4.2.0 (PILAT, Build 201410101015 0164 000000-0000)

- PeptideProphet option

PeptideProphet Options [Show / Hide]

☒ RUN PeptideProphet

☐ Use accurate mass binning, using: PPM ▼

☐ Use pI information

☐ Use Phospho information

☐ Use N-glyc motif information

☐ Use Hydrophobicity / RT information

☐ Use icat information

☐ Do not use icat information

☐ Do not use the NTT model

☐ Do not use the NMC model

☐ MALDI data

☐ Only use Expect Score as the discriminant - helpful for data with homologous top hits, e.g. phospho or glyco (Tandem and Comet only)

☐ Use Gamma distribution to model the negatives (Tandem only)

☐ Force the fitting of the mixture model (bypass automatic mixture model checks)

☐ Use decoy hits to pin down the negative distribution. Decoy Protein names begin with: (whitespace not allowed)

☐ Use Non-parametric model (can only be used with decoy option)

☐ Report decoy hits with a computed probability (based on the model learned).

Ignore charge states: ☐ +1 ☐ +2 ☐ +3 ☐ +4 ☐ +5

☐ Run ProteinProphet afterwards

Enter additional options to pass directly to the command-line (expert use only!)

ISB/SPC Trans Proteomic Pipeline - xinteract

Home | Data Management | Data Process & QC Tools | **Analyze Pipeline (Current)** | Search | Settings | Downloads | Links | You are logged in as guest. Log Out

Home	Data Mgmt	Tools	Database Search	Analyze Peptides	Combine Analyses	Analyze PTMs	Analyze Proteins
------	-----------	-------	-----------------	------------------	------------------	--------------	------------------

Messages [Show / Hide]

- Your file has been added:
[Sample.msml](#)

Select Files to Analyze [Show / Hide]

c:\peptide\newworld\SS\data\sample.pep.xml

[Add Files] or choose from set c:\peptide\newworld\SS\data\Comet [xml ▼] [Remove]

Output File and Filter Options

File path (Folder) c:\peptide\newworld\SS\data\Comet
Write output to file interact.pep.xml

Filter out results below this PeptideFragProphet probability 0.05
Minimum peptide length considered in the analysis 7

PeptideFragProphet Options [Show / Hide]

☒ RUN PeptideFragProphet

☐ Use accurate mass binning, using CPM

☐ Use all information

☐ Use Prospector information

☐ Use N-glycan modification information

☐ Use hydrophobicity or RT information

☐ Use cal information

☐ Do not use all information

☐ Do not use the NTT model

☐ Do not use the NAC model

NAC/NLQ score

☐ Only use Expect Score as the discriminant - helpful for data with homologue top hits, e.g. prosequin or glyco (Tandem and Comet only)

☐ Use Gamma distribution to model the negatives (Tandem only)

☐ Force the fitting of the modified model (Scout automatically modifies model checks)

☐ Use decoy hits to pin down the negative distribution (Deconv Protein names begin with _____ (interspace not allowed))

☐ Use Non-parametric model (can only be used with decoy options)

☐ Report decoy hits with a computed probability (based on the model learned)

Ignore charge states: ☐ +1 ☐ +2 ☐ +3 ☐ +4 ☐ +5

☐ Run PeptideFragProphet afterwards

Enter additional options to pass directly to the command-line (expect user input):

Run Parameters [Show / Hide]

☐ RUN PeptideFragProphet

☐ Also run ProstagProphet on these results

☐ do NOT use number of replicate spectra (NRS) model

☐ do NOT use number of sibling ions (NSI) model

☐ do NOT use number of sibling modifications (NSM) model

☐ do NOT use number of sibling sequences (NSS) model

☐ do NOT use number of sibling experiments (NSE) model

☐ do NOT use number of sibling peptides (NSP) model

PTM/Prophet Options [Show / Hide]

☐ RUN PTM/Prophet

Spectro modification:

Residue(s):	STY	Mass shift:	73.066
Residue(s):		Mass shift:	
Residue(s):		Mass shift:	
Residue(s):		Mass shift:	
Residue(s):		Mass shift:	
Residue(s):		Mass shift:	

no tolerance 0.1 ignore

☐ Do not update modification_info tags in pepXML

XPRESS Options [Show / Hide]

☐ RUN XPRESS

Mass tolerance 0.5 Custom

☐ For ratio, settings light to 1, very heavy

☐ For ratio, settings heavy to 1, very light

* Default (set less abundant to rate to 1)

☐ Heavy labeled peptide elutes before light labeled partner

Reaction 1 and mass difference:	-10.0
Reaction 2 and mass difference:	+10.0
Reaction 3 and mass difference:	-10.0

☐ Fix elution peak area at +/- % scans from peak apex

Minimum number of chromatogram points needed for quantitation: 5

Number of isotopic peaks to sum, use narrow tolerance: 1

ASAP/Ratio Options [Show / Hide]

☐ RUN ASAP/Ratio

☐ Static modification quantification (ie. each run is either all or heavy)

Labeled molecules C V V V V V V V V V

☐ Heavy labeled peptide elutes before light labeled partner

☐ Use fixed scan range for Light and heavy

☐ Quantitate only the charge state where the CID was made

Set averaging to: (static charge option)

☐ Zero out all background

☐ Quantitate despite high background

net range to include in summation of peak: 0.5

Reaction 1 mass:	V	
Reaction 2 mass:	V	
Reaction 3 mass:	V	* specify monoisotopic masses
Reaction 4 mass:	V	
Reaction 5 mass:	V	

Litara Quantification Options [Show / Hide]

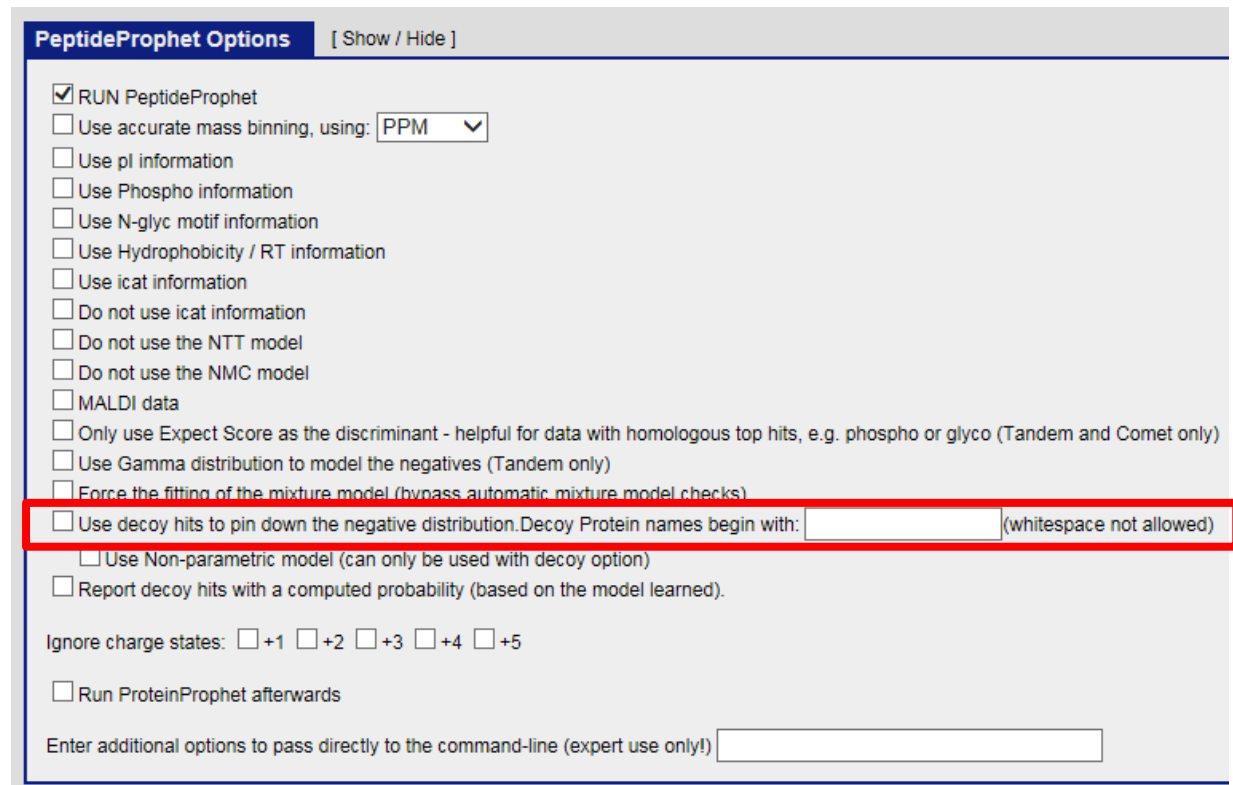
☐ RUN Litara

Condition File Condition.xml [Click here to generate a condition file](#)

Run Analysis!

[Run Xinteract]

- PeptideProphet option
 - Use decoy hits to pin down the negative distribution.
 - Decoy Protein names begin with: **XXX_**



PeptideProphet Options [Show / Hide]

- ☒ RUN PeptideProphet
- ☐ Use accurate mass binning, using: PPM
- ☐ Use pI information
- ☐ Use Phospho information
- ☐ Use N-glyc motif information
- ☐ Use Hydrophobicity / RT information
- ☐ Use icat information
- ☐ Do not use icat information
- ☐ Do not use the NTT model
- ☐ Do not use the NMC model
- ☐ MALDI data
- ☐ Only use Expect Score as the discriminant - helpful for data with homologous top hits, e.g. phospho or glyco (Tandem and Comet only)
- ☐ Use Gamma distribution to model the negatives (Tandem only)
- ☐ Force the fitting of the mixture model (bypass automatic mixture model checks)
- ☐ Use decoy hits to pin down the negative distribution. Decoy Protein names begin with: (whitespace not allowed)
- ☐ Use Non-parametric model (can only be used with decoy option)
- ☐ Report decoy hits with a computed probability (based on the model learned).

Ignore charge states: ☐ +1 ☐ +2 ☐ +3 ☐ +4 ☐ +5

☐ Run ProteinProphet afterwards

Enter additional options to pass directly to the command-line (expert use only!)

- PeptideProphet 실행
– Run XInteract 클릭

The image shows the PeptideProphet XInteract web interface. On the left, a large blue box contains the text "Run Analysis!" and a button labeled "Run XInteract". A red box highlights this button. A red arrow points from this button to the "Run XInteract" button in the "Run Analysis!" section at the bottom of the page. The right side of the image shows the main configuration area of the web interface, which includes sections for "Select Files to Analyze", "Output File and Filter Options", "PeptideProphet Options", "InterProphet Options", "PTMProphet Options", and "ASAPRatio Options". The "Run XInteract" button is located at the bottom of the "Run Analysis!" section.

- PeptideProphet 실행
 - running → finished
 - Refresh 클릭

ISB/SPC Trans Proteomic Pipeline - jobs

Home | Files | Account | Pre-Process | mzXML Utils | Analysis Pipeline (Comet) | Decoy | Utilities | SpectraST Tools | Jobs

You are logged in as guest Log Out

Home Jobs

Commands Status [Show / Hide]

Status as of: Thu Jun 11 16:14:07 2015

All Jobs [Show / Hide]


Session ID	Job	Location	Start date / time	Actions	Status	Output
FLT7ILNQR	xinteract	localhost	2015-06-11 16:14:06	Kill job	running	Refresh
QPZW0ZA9A	runcomet	localhost	2015-06-10 17:40:23	Delete Log	viewed	View

Output for job id FLT7ILNQR 20150611-161406

Session ID	Job	Location	Start date / time	Actions	Status	Output
FLT7ILNQR	xinteract	localhost	2015-06-11 16:14:06	Kill job	* finished	Refresh
QPZW0ZA9A	runcomet	localhost	2015-06-10 17:40:23	Delete Log	viewed	View

Output for job id FLT7ILNQR 20150611-161406

pep.xml



- PeptideProphet 실행 결과 확인
 - PepXML 클릭

ISB/SPC Trans Proteomic Pipeline - jobs

Home | Files | Account | Pre-Process | mzXML Utils | Analysis Pipeline (Comet) | Decoy | Utilities | SpectraST Tools | Jobs

You are logged in as guest. Log Out

Commands Status [Show / Hide]

Status as of: Thu Jun 11 16:20:55 2015
Output from all commands has been viewed; auto-refresh is now OFF.

All Jobs [Show / Hide]

Session ID	Job	Location	Start date / time	Actions	Status	Output
FLT7ILNQR	xinteract	localhost	2015-06-11 16:14:06	Delete Log	viewed	Refresh
QPZWQZAGA	runcomet	localhost	2015-06-10 17:40:23	Delete Log	viewed	View

Output for job id FLT7ILNQR_20150611-161406

Output Files

- c:/inetpub/wwwroot/ISB/data/Comet/interact.pep.xml [[PepXML](#)]

EXECUTING: cd c:/inetpub/wwwroot/ISB/data/Comet& c:/inetpub/pp/bin/xinteract_Ninteract.pep.xml -p0.05 -f7 -PPM -O -dXXX_Sample.pep.xml

- Linking duplicate entries... - Printing results...

05 PPM DECOY=XXX_

Analysed: UNDECOYED, Detector: UNDECOYED

ler@ISB

All commands finished at Thu Jun 11 16:14:36 2015

Output Files

- c:/inetpub/wwwroot/ISB/data/Comet/interact.pep.xml [[PepXML](#)]

INSTITUTE FOR
Systems
Biology Revolutionizing science. Enhancing life.

Seattle
Proteome
Center

TPP v4.8.0 PHILAE, Build 201411201551-6764 (mingw-686)

• PeptideProphet 실행 결과 확인 – PROB 열

Summary

Display Options

Pick Columns

Filtering Options

Other Actions

[Hide options]

Sorting: spectrum desc asc

PeptideProphet: min max

Update Page

pepXML file: c:/inetpub/wwwroot/ISB/data/Comet/interact.pep.xml

trypsin digest, COMET search engine, quantitation: [none]

displaying 499 of 499 total spectra, page 1 of 10

444 unique peptides, 444 unique stripped peptides, 432 unique proteins, 389 single hits

PepXML Viewer, 2006 SPC188

Page 1 of 10

1 FIRST 1 2 3 4 5 6 NEXT LAST

PROB	SPECTRUM	SSCAN	EXPECT	IONS2	PEPTIDE	PROTEIN	CALO_MASS
0.7738	Sample.00032.00032.4 ^{SR}	32	1.00E+000	4/84	K.KPVLEEKPAVPYPER.A ^{BA}	spIQ8WZ42ITITIN_HUMAN +6	1686.966856
0.8699	Sample.00111.00111.4 ^{SR}	111	4.17E+001	5/84	R.LPLLRSANHTVTIR.V ^{BA}	spIQ86W56-2IPARG_HUMAN +2	1686.989323
0.3381	Sample.00279.00279.2 ^{SR}	279	4.22E-008	6/24	R.GGNIGDGGGAADR.V ^{BA}	spIP55072ITERA_HUMAN	1115.495563
0.9212	Sample.00363.00363.2 ^{SR}	363	3.62E-016	11/24	R.GGNIGDGGGAADR.V ^{BA}	spIP55072ITERA_HUMAN	1115.495563
0.5868	Sample.00447.00447.2 ^{SR}	447	3.76E-011	6/24	R.GGNIGDGGGAADR.V ^{BA}	spIP55072ITERA_HUMAN	1115.495563
0.6452	Sample.02251.02251.2 ^{SR}	2251	3.46E-005	6/16	K.NDSWONNMR.S ^{BA}	spIQ9UPQ9ITNR6B_HUMAN +1	1133.448613
0.4036	Sample.02492.02492.4 ^{SR}	2492	7.74E+000	7/90	R.RPPLAEL AALNLSGSR.L ^{BA}	spIQ13641ITPBG_HUMAN	1663.936953
0.8445	Sample.02865.02865.2 ^{SR}	2865	9.51E-007	11/16	R.TDTGPMGR.G ^{BA}	spIP07900IHS90A_HUMAN +1	962.412745
0.8348	Sample.02953.02953.2 ^{SR}	2953	6.18E-011	11/16	R.TDTGPMGR.G ^{BA}	spIP07900IHS90A_HUMAN +1	962.412745
0.9421	Sample.03039.03039.2 ^{SR}	3039	6.02E-015	12/16	R.TDTGPMGR.G ^{BA}	spIP07900IHS90A_HUMAN +1	962.412745
0.9146	Sample.03122.03122.2 ^{SR}	3122	5.29E-012	12/16	R.TDTGPMGR.G ^{BA}	spIP07900IHS90A_HUMAN +1	962.412745