

「第八回 生物系研究者のための質量分析に関する教育セミナー」報告

日時：2010年8月26日（木）15:00～16:20

場所：生物科学研究棟 3階大セミナー室（S311）

講師：マトリックスサイエンス（株） 高江洲 宏智 氏

内容：MASCOTの便利な機能

- 1) 修飾・アミノ酸置換解析
- 2) 同定基準スコアの有効性検証
- 3) 定量解析（emPAI、Distillerを用いた各種定量手法）

セミナー参加者：26名

基幹研究所：4 研究室、3 チーム

脳科学総合研究センター：2チーム、2 ユニット

今回は基礎的だった第六回に比べて、一歩進んだ形で、やや専門的な内容となっていました。

内容はMASCOTの機能を生かしたいろいろな解析方法の紹介で、実際利用している研究者には有意義な内容であった。質問もより具体的な内容が多くて、非常に関心が多いセミナー内容となっていた。

以上

# ケミカルバイオロジー 研究基盤施設教育セミナー

## 第八回 生物系研究者のための 質量分析に関する教育セミナー

質量分析について、一歩進んだ教育セミナーを開催します。  
(やや専門的な内容になっています。)

日時 2010年8月26日（木）15:00～16:30

場所 生物科学研究棟 3階大セミナー室（S311）

講師 マトリックスサイエンス（株）

高江洲 宏智 氏

内容 MASCOTの便利な機能

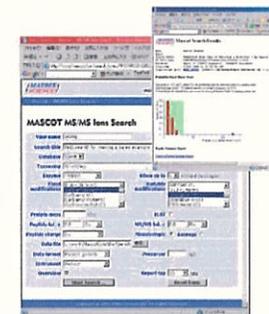
- 1) 修飾・アミノ酸置換解析
- 2) 同定基準スコアの有効性検証
- 3) 定量解析（emPAI、Distillerを用いた各種定量手法）

主催 ケミカルバイオロジー研究基盤施設  
バイオ解析チーム

連絡先 バイオ解析チーム

渡辺 ([kowashi@riken.jp](mailto:kowashi@riken.jp))

\* 参加に当たり事前予約が必要となりますので、  
連絡先（渡辺）までご連絡ください。



バイオ解析チームHPへの  
リンクはココをクリック

著作権はRIKENにあります。  
無断複製・転載を禁止します。

他の教育セミナーへの  
リンクはココをクリック

## 本日の内容

- 1) 修飾・アミノ酸置換解析 **Error Tolerant Search**
- 2) 同定基準スコアの有効性検証 **Decoy Search**
- 3) 定量解析 (**emPAI**、**Distiller**を用いた各種定量手法)
- 4) その他 技術トピックスとMASCOT(並びに関連製品)との繋がり

## 1) 修飾・アミノ酸置換解析 **Error Tolerant Search**

ペプチドとマッチしなかったスペクトルを詳しく解析する

# ペプチドとマッチしないスペクトル

Query	Observed	Mr(calc)	Mr(found)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
21	462.6807	923.3468	923.5116	-0.1649	0	33	1.1	1		R.PFYVADK.T
51	517.1760	1032.3375	1032.5404	-0.2029	0	70	0.0004	1		R.QNFIFGADK.R
65	564.6804	1127.3463	1127.5764	-0.2301	0	8	6.7e+02	1		R.GPFLPQDQK.I
65	567.6567	1133.2987	1133.5499	-0.2511	0	45	0.18	1		R.QNFVIRK.R.A + Oxidation (M)
88	614.2001	1224.3856	1224.6329	-0.2473	0	28	7.9	1	U	K.LGPRFLANR.P + Oxidation (M)
112	653.2161	1304.4037	1304.6437	-0.2380	0	87	1e+05	1		R.QNFVIRK.R.A.F
124	724.1806	1456.3465	1456.6477	-0.3011	0	69	0.00062	1		R.NMYRQNVVADK.Q
156	826.7283	1639.4420	1639.7763	-0.3343	0	104	1e+07	1		R.ALFYRQVVDK.R.A + Oxidation (M)
122	864.2888	1724.5429	1724.9294	-0.3864	0	43	0.07			R.QNFVIRK.R.A
228	954.2437	1916.4729	1916.8601	-0.3872	0	29	1			R.QNFVIRK.R.A
208	975.8100	1949.6055	1950.0245	-0.4130	0	86	1e+06			R.QNFVIRK.R.A
208	976.2348	1950.4534	1950.8555	-0.4021	1	27	7			R.QNFVIRK.R.A
212	1081.2027	2098.3908	2098.8058	-0.4150	0	163	0.001			R.QNFVIRK.R.A
211	667.8044	2008.3919	2008.8058	-0.4139	0	73	0.0002			R.QNFVIRK.R.A
222	681.8205	2042.4397	2042.8164	-0.3767	0	158	0.004			R.QNFVIRK.R.A
221	711.5744	2191.7013	2192.1340	-0.4327	1	10	3.5e+06			R.QNFVIRK.R.A
231	796.2187	2349.4341	2349.8295	-0.4054	0	93	1.9e+06			R.QNFVIRK.R.A
261	1078.6327	3232.8763	3233.5629	-0.6867	0	5	6.3e+02	1		R.QNFVIRK.R.A

Total peptides 283  
 Peptides above homology threshold 71  
 peptides unmatched 212

なぜマッチしないの？

## ペプチド配列とマッチしないスペクトル：理由

[装置の精度]

- ペプチドの質量とデータベースのペプチド理論質量が誤差範囲内でマッチしない
- フラグメントイオンの質量と理論質量が誤差範囲内でマッチしない

[スペクトルの質 (精度以外)]

- プロダクトイオンスペクトルがブアで十分なマッチングが見られない

- データベースとサンプルの配列が一致しない。SNPs, データベース側のエラーなど
  - 予期せぬ修飾
  - 予期せぬ (非特異的な) ペプチド切断
- } Error Tolerant 検索対応可能

## Error Tolerant Search 手順

- 2回の検索を行う
  - 1回目は通常を検索
  - 1回目の検索で、a) 気になったタンパク質を自身で選択  
b) 同定基準を超えるタンパク質を自動的に選択 のいずれかの方法でタンパク質を選択
  - 2回目の検索で、選ばれたタンパク質について、アミノ酸配列の置換 / 非特異的な切断 / 翻訳後修飾を網羅的 \* に考慮した再検索を実施

➤ Creasy, D. M. and Cottrell, J. S., Error tolerant searching of uninterpreted tandem mass spectrometry data, *Proteomics* 2 1426-1434 (2002)

## Error Tolerant Search 実施前

Peptide Summary Report (Before Automatic error tolerant search example) - Windows Internet Explorer

1. **PP21\_HUMAN** Mass: 58259 Score: 806 Matches: 18 (9) Sequences: 16 (7) emPAI: 0.92  
alkaline phosphatase, placental type OS=Homo sapiens GN=ALPP PF=1 SV=2

Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calcd)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/> 27	462.6807	923.3468	923.5116	-0.1649	0	33	1.1	1		R.FPTVALSK.F
<input checked="" type="checkbox"/> 41	517.1760	1032.3375	1032.5604	-0.2229	0	70	0.0004	1		R.GSRIFGLAAGK.A
<input checked="" type="checkbox"/> 62	564.6804	1127.3463	1127.5764	-0.2301	0	8	6.7e+02	1		R.GFFLPVGGK.I
<input checked="" type="checkbox"/> 65	567.6567	1133.2987	1133.5499	-0.2511	0	45	0.18	1		R.GNEVIVMNR.A + Oxidation (M)
<input checked="" type="checkbox"/> 86	614.2001	1226.3856	1226.6329	-0.2473	0	28	7.9	1	U	K.LGPEIPLAMDR.F + Oxidation (M)
<input checked="" type="checkbox"/> 100	653.2101	1304.4057	1304.6837	-0.2780	0	87	1e-05	1		K.GNPFQIGLSAAR.F
<input checked="" type="checkbox"/> 126	726.1806	1450.3465	1450.6477	-0.3011	0	69	0.00062	1		R.NWYSADAVPAAR.Q
<input checked="" type="checkbox"/> 156	820.7283	1639.4420	1639.7763	-0.3343	0	106	1e-07	1		R.ALVTETIFDDAIER.A + Oxidation (M)
<input checked="" type="checkbox"/> 170	864.2888	1726.5629	1726.9294	-0.3664	0	43	0.21	1		K.AYTVLLVGGVGVK.D
<input checked="" type="checkbox"/> 204	956.2437	1910.4729	1910.8601	-0.3872	0	29	5.2	1	U	R.DSTLDPSIMRTEAALR.L + 2 Oxidation (M)
<input checked="" type="checkbox"/> 208	975.8100	1949.6055	1950.0245	-0.4190	0	86	1e-05	1		K.NLIIFLDGEGVSTVAAR.I + Oxidation (M)
<input checked="" type="checkbox"/> 209	976.2340	1950.4534	1950.8555	-0.4021	1	27	7.3	1		K.DGARPDVTESESGSPETR.Q
<input checked="" type="checkbox"/> 216	1001.2027	2000.3908	2000.8058	-0.4150	0	(65)	0.0012	1	U	R.MCTFPDPEYDDYSQOQTR.L + Oxidation (M)
<input checked="" type="checkbox"/> 217	667.8046	2000.3915	2000.8058	-0.4139	0	73	0.00022	1	U	R.MCTFPDPEYDDYSQOQTR.L + Oxidation (M)
<input checked="" type="checkbox"/> 222	681.8205	2042.4397	2042.8164	-0.3767	0	(58)	0.0063	1	U	R.MCTFPDPEYDDYSQOQTR.L + Acetyl (N-term); Oxidation (M)
<input checked="" type="checkbox"/> 227	711.5744	2131.7013	2132.1340	-0.4327	1	19	3.5e+02	2	U	K.LGPEIPLAMDRFPVALSK.F + Oxidation (M)
<input checked="" type="checkbox"/> 252	790.2187	2367.6341	2368.1295	-0.4954	0	93	1.9e-06	1	U	R.QQSNVLDSETHAGEVYAFAR.G
<input checked="" type="checkbox"/> 281	1078.6327	3232.8763	3233.5629	-0.6867	0	5	6.3e+02	1		R.AQQLTSEDFVLSVTDHSHVFSFGQYPLR.G

# Error Tolerant Search - 非特異的ペプチド切断

2. **FPRL\_HUMAN** Mass: 58259 Score: 450 Matches: 31(8) Sequences: 19(7) eMFAI: 0.92  
 Alkaline phosphatase, placental type precursor (EC 3.1.3.1) (PLAP-1) (Regan isozyme) - Homo sapiens  
 Check to include this hit in error tolerant search

<input checked="" type="checkbox"/>	<b>176</b>	<b>879.2425</b>	<b>1756.4705</b>	<b>1756.8420</b>	<b>-0.3715</b>	<b>0</b>		
		<b>83</b>		<b>1</b>			<b>G IIPVEEENPDFWNR.E</b>	
トリプシン断片ではない								
<input type="checkbox"/>	103	499.1349	1494				L.DFSLRHTFAALR.L + 2 Oxidation (M)	
<input type="checkbox"/>	145	526.1538				(68)	R.ALTEYIMDDAER.A + [-48.0000 at P8]	
<input type="checkbox"/>	156	820.7281				106 5.2e-08	R.ALTEYIMDDAER.A + Oxidation (M)	
<input type="checkbox"/>	165	841.0000				(75)	R.ALTEYIMDDAER.A + Oxidation (M); [+41.0266 at N-term A]	
<input type="checkbox"/>	170	867.0000	1726.9294	-0.3644	0	43 0.11	K.ATYVLLVNGPQTVLR.D	
<input type="checkbox"/>	173	867.0000	1764.4635	1756.8420	-0.3786	0	G.IIPVEEENPDFWNR.E	
<input type="checkbox"/>	175	879.2425	1756.4705	1756.8420	-0.3715	0	G.IIPVEEENPDFWNR.E	
<input type="checkbox"/>	204	956.2437	1910.4729	1910.8601	-0.3872	0	U R.DSTLQSIAMHTFAALR.L + 2 Oxidation (M)	
<input type="checkbox"/>	208	975.8100	1949.6055	1950.0245	-0.4190	0	K.NLIFLGGQEVSTFAAR.I + Oxidation (M)	
<input type="checkbox"/>	209	976.2340	1950.4534	1950.8555	-0.4021	1	(27)	3.7 1 K.DGARSVTEREAGSFEYR.Q
<input type="checkbox"/>	211	656.1752	1965.5039	1964.8712	0.6327	1	(68)	1 K.DGARSVTEREAGSFEYR.Q + [+14.0157 at K11]
<input type="checkbox"/>	213	664.5518	1990.6336	1991.0510	-0.4174	0	(56)	1 K.NLIFLGGQEVSTFAAR.I + Oxidation (M); [+41.0266 at N-term N]
<input type="checkbox"/>	216	1001.2027	2000.3908	2000.8058	-0.4150	0	(65)	0.00061 1 U R.HQTFPEYFDYSGQGTFR.L + Oxidation (M)
<input type="checkbox"/>	217	667.8046	2000.3919	2000.8058	-0.4139	0	73 0.00011 1 U R.HQTFPEYFDYSGQGTFR.L + Oxidation (M)	
<input type="checkbox"/>	218	670.1541	2007.4464	2007.8770	-0.4304	1	75 1 K.DGARSVTEREAGSFEYR.Q + [+57.0215 at N-term D]	
<input type="checkbox"/>	222	681.8205	2042.4397	2042.8164	-0.3767	0	(62)	1 U R.HQTFPEYFDYSGQGTFR.L + Acetyl (N-term); [+15.9949 at P6]
<input type="checkbox"/>	227	711.5744	2131.7013	2132.1340	-0.4327	1	10 1.7e+02 2 U K.LGPEFLAMR.F + Oxidation (M)	
<input type="checkbox"/>	252	784.5440	2350.6103	2351.1030	-0.4927	0	(68)	1 U R.QQAVFLDETHAGEDVAVFAR.G + [-17.0265 at N-term Q]
<input type="checkbox"/>	253	790.2187	2367.6341	2368.1295	-0.4954	0	93 5.8e-07 1 U R.QQAVFLDETHAGEDVAVFAR.G	
<input type="checkbox"/>	260	809.2208	2424.6406	2425.1510	-0.5104	0	(66)	1 U R.QQAVFLDETHAGEDVAVFAR.G + [+57.0215 at N-term Q]
<input type="checkbox"/>	275	920.5878	2758.7415	2759.3582	-0.6167	0	90 1 R.QEQDQIATQLISNDIVILGGGR.K + Acetyl (N-term); Oxidation (M); [-0.9476 at E2]	
<input type="checkbox"/>	281	1078.6327	3232.8763	3233.5629	-0.6867	0	5 3.2e+02 1 R.AQLTSEETLALVADSHRHFVFGTFLR.G	

# Error Tolerant Search - アミノ酸置換、修飾

2. **FPRL\_HUMAN** Mass: 58259 Score: 450 Matches: 31(8) Sequences: 19(7) eMFAI: 0.92  
 Alkaline phosphatase, placental type precursor (EC 3.1.3.1) (PLAP-1) (Regan isozyme) - Homo sapiens  
 Check to include this hit in error tolerant search

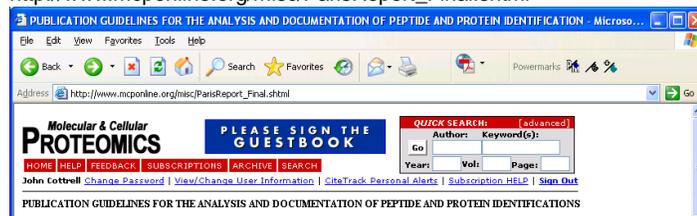
Query	Observed	Mr (exp1)	Mr (calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/>	<b>222</b>	<b>681.8205</b>	<b>2042.4397</b>	<b>2042.8164</b>	<b>-0.3767</b>	<b>0</b>	<b>(62)</b>	<b>1</b>	<b>1</b>	<b>U R.MGTFDEYPPDDYSQGQTR.L + Acetyl (N-term) [+15.9949 at P6]</b>
質量のズレから、修飾・アミノ酸の置換の両方を検証 Possible Assignments: Oxidation (P) [+15.9949] Pro->Leu (P) [+16.0313]										
<input type="checkbox"/>	208	975.8100	1949.6055	1950.0245	-0.4190	0	(68)	1	U R.NLIFLGGQEVSTFAAR.I + Oxidation (M)	
<input type="checkbox"/>	209	976.2340	1950.4534	1950.8555	-0.4021	1	(27)	3.7 1	K.DGARSVTEREAGSFEYR.Q	
<input type="checkbox"/>	211	656.1752	1965.5039	1964.8712	0.6327	1	(68)	1	K.DGARSVTEREAGSFEYR.Q + [+14.0157 at K11]	
<input type="checkbox"/>	213	664.5518	1990.6336	1991.0510	-0.4174	0	(56)	1	K.NLIFLGGQEVSTFAAR.I + Oxidation (M); [+41.0266 at N-term N]	
<input type="checkbox"/>	216	1001.2027	2000.3908	2000.8058	-0.4150	0	(65)	0.00061 1	U R.HQTFPEYFDYSGQGTFR.L + Oxidation (M)	
<input type="checkbox"/>	217	667.8046	2000.3919	2000.8058	-0.4139	0	73 0.00011 1	U R.HQTFPEYFDYSGQGTFR.L + Oxidation (M)		
<input type="checkbox"/>	218	670.1541	2007.4464	2007.8770	-0.4304	1	75 1	K.DGARSVTEREAGSFEYR.Q + [+57.0215 at N-term D]		
<input type="checkbox"/>	222	681.8205	2042.4397	2042.8164	-0.3767	0	(62)	1	U R.HQTFPEYFDYSGQGTFR.L + Acetyl (N-term); [+15.9949 at P6]	
<input type="checkbox"/>	227	711.5744	2131.7013	2132.1340	-0.4327	1	10 1.7e+02 2	U K.LGPEFLAMR.F + Oxidation (M)		
<input type="checkbox"/>	252	784.5440	2350.6103	2351.1030	-0.4927	0	(68)	1	U R.QQAVFLDETHAGEDVAVFAR.G + [-17.0265 at N-term Q]	
<input type="checkbox"/>	253	790.2187	2367.6341	2368.1295	-0.4954	0	93 5.8e-07 1	U R.QQAVFLDETHAGEDVAVFAR.G		
<input type="checkbox"/>	260	809.2208	2424.6406	2425.1510	-0.5104	0	(66)	1	U R.QQAVFLDETHAGEDVAVFAR.G + [+57.0215 at N-term Q]	
<input type="checkbox"/>	275	920.5878	2758.7415	2759.3582	-0.6167	0	90 1	R.QEQDQIATQLISNDIVILGGGR.K + Acetyl (N-term); Oxidation (M); [-0.9476 at E2]		
<input type="checkbox"/>	281	1078.6327	3232.8763	3233.5629	-0.6867	0	5 3.2e+02 1	R.AQLTSEETLALVADSHRHFVFGTFLR.G		

## 2) 同定基準スコアの有効性 検証 Decoy Search

MASCOTの同定基準を保証する

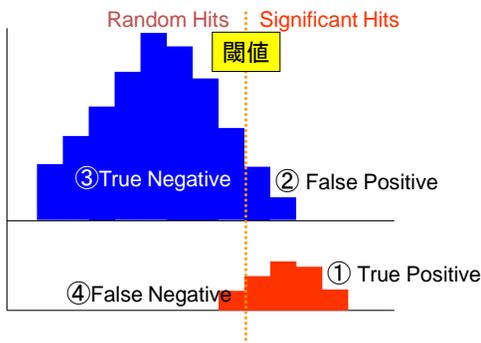
## Molecular & Cellular Proteomics

[http://www.mcponline.org/misc/ParisReport\\_Final.shtml](http://www.mcponline.org/misc/ParisReport_Final.shtml)



For large scale experiments, provide the results of any additional statistical analyses that indicate or establish a measure of identification certainty, or allow a determination of false-positive rate, e.g., the results of randomized database searches or other computational approaches.

# 用語の確認 (Terminology)



The MS/MS spectrum comes from a peptide sequence in the DB

False Discovery Rate  
=  $FP / (FP + TP)$

True Positive Rate  
=  $TP / (TP + FN)$

False Positive Rate  
=  $FP / (FP + TN)$

Search reports a match to the correct sequence

	True	False
True	True positive	False positive
False	False negative	True negative

## FDR 計算方法

	Sprot	Decoy	False discovery rate
Peptide matches above identity threshold	3290	8	0.24 %
Peptide matches above homology or identity threshold	6037	224	3.71 %

- 同定基準を超えるスペクトル (FP + TP) 3290
- 不正解なのに、基準を超えているスペクトル (本来のFP) わからない
- でたらめな配列のデータベースに対して検索を行い、同定基準を超えたスペクトル 8  
→ この数字を FP として計算に利用

**FDR : 8 / 3290 = 0.24 %**

[Elias, J. E., et al., Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations, Nature Methods 2 667-675 \(2005\).](#)

## FDR トピックス

- 検証時の目安は  $FDR < 5\%$  (or  $1\%$ ) 以下
- FDR 計算方法は千差万別
  - FP, TP の数え方に多様性。標準手法なし
  - MASCOT の方法は比較的シンプル
  - 論文などでは都合良い方法を各自勝手に選択しているのが現状
- FDR を同定基準として利用する方法も。FDR を基準とすることで複数 DB 検索結果の結合も容易に

## 3) 定量解析 emPAI, Distiller を用いた 各種定量手法

変動解析

# Spectral Counting

- 量が多いタンパク質は、検索後アサインされるペプチド数も多い、という経験が基となっている
- 各種補正 (タンパク質の大きさ、実験差 etc)
- 適用範囲、用法、誤差など要検討

## emPAI (Exponentially Modified Protein Abundance Index)

$$emPAI = 10^{\frac{N_{observed}}{N_{observable}} - 1}$$

- $N_{observed}$  アサインされたペプチドの数  
scoreがhomology threshold 以上でユニークなペプチドであること。  
電荷が別なら配列同じでも OK [論文とMASCOTで同じルール]
- $N_{observable}$  理論的に生成されたペプチドの数
  - 理論ペプチド生成時消化酵素設定 [MASCOT: 特異性無視]
  - mass range 制限 [MASCOT: Serverでの制限のみ]
  - retention time 制限 [MASCOT: 使用しない]
  - length 制限 [MASCOT: Serverでの制限のみ]

Ishihama, Y., et al., Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein, Molecular & Cellular Proteomics 4 1265-1272 (2005)

# emPAI 表示例

RAFT-8 (Mascot Search Results) - Windows Internet Explorer

http://localhost/mascot.cgi/master\_results\_2.pl?File=20100820%2F501332.dat#0?show=quantitation

Filter: Significance threshold p < 0.05 Max. number of families AUTO [help]  
Ions score or expect cut-off 0

Proteins (17) Quantitation (21) Unassigned (195) [link to]

### Quantitation overview (21 proteins)

	Protein	Score	Mass	Matches	Sequences	emPAI
1.1	PPB1_HUMAN	430	58259	17 (11)	16 (10)	0.92 Alkaline phosphatase, placental type OS=Homo sapiens GN=ALPL...
1.2	PPBN_HUMAN	298	57626	12 (9)	12 (9)	0.72 Alkaline phosphatase, placental-like OS=Homo sapiens GN=ALPL...
1.3	PPB1_HUMAN	68	57119	2 (2)	2 (2)	0.13 Intestinal-type alkaline phosphatase OS=Homo sapiens GN=ALPL...
2	TRY1_BOVIN	355	26453	8 (7)	5 (5)	1.17 Cationic trypsin OS=Bos taurus PE=1 SV=3
3.1	K2C1_HUMAN	228	66170	11 (9)	10 (8)	0.60 Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1...
3.2	K2Z2_HUMAN	104	65678	4 (4)	4 (4)	0.23 Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=K...
4.1	LYN_HUMAN	224	61276	14 (9)	13 (8)	0.66 Proto-oncogene tyrosine-protein kinase Yes OS=Homo sapiens G...
4.2	LYN_HUMAN	66	58993	4 (2)	4 (2)	0.12 Tyrosine-protein kinase Lyn OS=Homo sapiens GN=LYN PE=1 SV=3
5	BAIP2_HUMAN	197	61115	10 (8)	10 (8)	0.57 Brain-specific angiogenesis inhibitor 1-associated protein 2 OS...
6	RPN2_HUMAN	96	69355	2 (2)	2 (2)	0.11 Dolichyl-diphosphooligosaccharide--protein glycosyltransferase s...
7	K1C9_HUMAN	87	62255	5 (4)	5 (4)	0.25 Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1...
8	KPYM_HUMAN	85	58470	3 (3)	3 (3)	0.19 Pyruvate kinase isozymes M1/M2 OS=Homo sapiens GN=PKM2 PE...
9	BAS1_BOVIN	83	22333	3 (3)	3 (3)	0.58 Basigin (Fragment) OS=Bos taurus GN=BSG PE=2 SV=1
10	CBPM_HUMAN	74	50938	3 (2)	3 (2)	0.15 Carboxypeptidase M OS=Homo sapiens GN=CPM PE=1 SV=2
11	GTR1_BOVIN	69	54439	1 (1)	1 (1)	0.07 Solute carrier family 2, facilitated glucose transporter member 1...
12	B1Z1_HUMAN	69	57189	3 (2)	3 (2)	0.13 Brain-specific angiogenesis inhibitor 1-associated protein 2-like...
13	K2C5_HUMAN	57	62568	1 (1)	1 (1)	0.06 Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1...

©2010 Matrix Science K.K. 17

## SILAC解析例

Mascot Distiller - Triple Encoding SILAC sample data

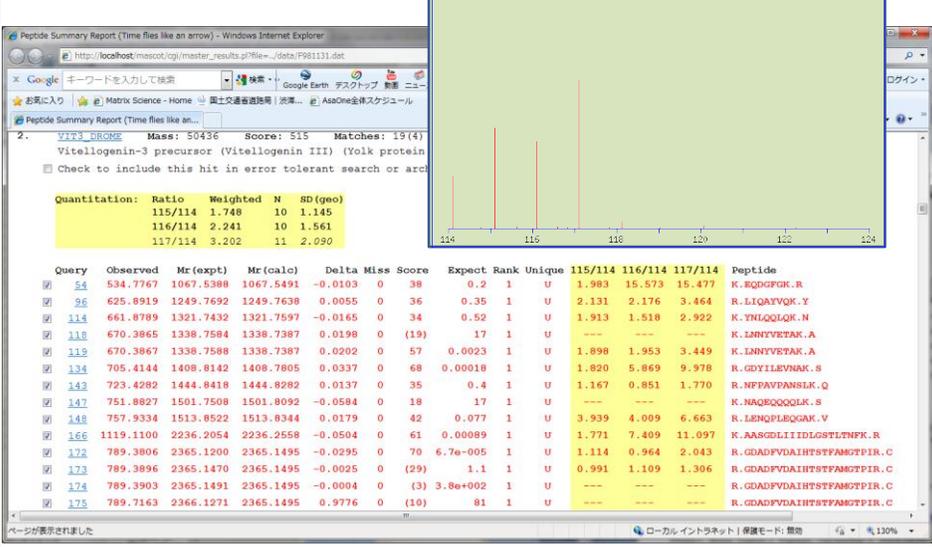
Proteins Quantitation Table

Accession	Score	Mass	M/L	SD(geo)	H/L	SD(geo)	Description
1	42	17149	0.011	0.014	0.009	0.008	1-oxo-5-oxo-L-proline (Homo sapiens)
2	17201	22098	1.2697	0.0088	0.014	0.011	48S ribosomal protein L16 (Homo sapiens)
3	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
4	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
5	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
6	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
7	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
8	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
9	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
10	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
11	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
12	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
13	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
14	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
15	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
16	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
17	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
18	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
19	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
20	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)
21	10333	10333	0.2113	0.0113	0.004	0.004	48S ribosomal protein L16 (Homo sapiens)

Triple Encoding SILAC sample data - Survey 6005 (t=3526:1440)

Scan 5981 - 3487.870

## iTRAQ 解析例



Peptide Summary Report (Time files like an arrow) - Windows Internet Explorer

http://localhost/mascot/cgi/mascot\_results.pl?file=.../data/P091131.dat

2. **VIT3\_PRCURSE** Mass: 50436 Score: 515 Matches: 19(4)  
 Vitellogenin-3 precursor (Vitellogenin III) (Yolk protein)

Quantitation:

Ratio	Weighted	N	SD (geo)
115/114	1.748	10	1.145
116/114	2.241	10	1.561
117/114	3.202	11	2.090

Query	Observed	Mr (expt)	Mr (calc)	Delta	Miss	Score	Expect	Rank	Unique	115/114	116/114	117/114	Peptide
54	534.7767	1067.5388	1067.5491	-0.0103	0	38	0.2	1	U	1.983	15.573	15.477	K.EQDGFQK.R
96	625.8919	1249.7692	1249.7638	0.0055	0	36	0.35	1	U	2.131	2.176	3.464	R.LIQAYVQK.Y
114	661.8789	1321.7432	1321.7597	-0.0165	0	34	0.52	1	U	1.913	1.518	2.922	K.YNIQQIQK.N
118	670.3865	1338.7584	1338.7387	0.0198	0	(19)	17	1	U	---	---	---	K.LNNYVETAK.A
119	670.3867	1338.7588	1338.7387	0.0202	0	57	0.0023	1	U	1.898	1.953	3.449	R.LNNYVETAK.A
134	705.4144	1408.8142	1408.7805	0.0337	0	68	0.00018	1	U	1.820	5.869	9.978	R.GDYILEVNAK.S
143	723.4282	1444.8418	1444.8282	-0.0137	0	35	0.4	1	U	1.167	0.851	1.770	R.NFPAVFNLSK.Q
147	751.8827	1501.7508	1501.8092	-0.0584	0	18	17	1	U	---	---	---	K.NAGEQQQLK.S
148	757.9334	1513.8522	1513.8344	0.0179	0	42	0.077	1	U	3.939	4.009	6.663	R.LENQPLQGGAK.V
166	1119.1100	2236.2054	2236.2558	-0.0504	0	61	0.00089	1	U	1.771	7.409	11.097	K.AASGDLIIIDLGSLTINPK.R
172	789.3806	2365.1200	2365.1495	-0.0295	0	70	6.7e-005	1	U	1.114	0.964	2.043	R.GDADFVDIAHTSTFAMGTPPIR.C
173	789.3896	2365.1470	2365.1495	-0.0025	0	(29)	1.1	1	U	0.991	1.109	1.306	R.GDADFVDIAHTSTFAMGTPPIR.C
174	789.3903	2365.1491	2365.1495	-0.0004	0	(3)	3.8e+002	1	U	---	---	---	R.GDADFVDIAHTSTFAMGTPPIR.C
175	789.7163	2366.1271	2365.1495	0.9776	0	(10)	81	1	U	---	---	---	R.GDADFVDIAHTSTFAMGTPPIR.C

19

## 4) その他 技術トピックと MASCOT (並びに関連製品) との繋がり

## de novo Sequencing

### ■ 特徴とメリット

- MS/MS のProduct Ion マスペクトルを利用
- データベース検索を経ずペプチドのアミノ酸配列を決定
- MS, MS/MS において、データベース検索できないデータに対する唯一の配列決定法
- 部分配列のみの決定でも使い道
  - \*\* Sequence Tag検索
  - \*\* 配列相同性検索
- 提示された結果の検証法などが不十分。ユーザー目視の補助的なツールという位置づけが妥当

### ■ 必要な関連製品

- Mascot Distiller [Search Toolbox]

## de novo Sequencing

### ■ Sequence Tag

- MS/MS のProduct Ion スペクトルを利用しているが、MIS とは異なるアプローチ
  - > ペプチドの質量 (換算)、部分配列、部分配列を特定したスペクトル上の位置 (m/z 左端、右端) の3つの情報を利用

例)

1890.2 tag(1004.1, LSADTG, 1548.5)

- 「最初に配列を特定する」 (上記例なら「LSADTG」が確かに存在すると決めつける) 部分に、リスクがある半面、絞込みには強力な情報になる

# de novo Sequencing

## ■ 部分配列の相同性検索

- Sequence Tagと比べると、特定された配列部分にミス（入れ替え、末端の誤った決定）などがあっても対処可能  
例) LSADTGと LASDTG,  
LSADTG と VSADTG など。
- Sequence Tagと比べ、配列以外の情報（ペプチドの質量（換算）、部分配列を特定したスペクトル上の位置（m/z 左端、右端））が使えない分、絞込みが弱い。

# 比較Proteomics

## ■ 特徴

- サンプル間の発現タンパク質の違い、量の変動を比較

## ■ 必要な製品

- Scaffold (Proteome Software 社製品、日本では弊社が代理店)  
-> MASCOT では、「単一サンプルに由来するタンパク質のリスト」を一画面で表示するのが上限

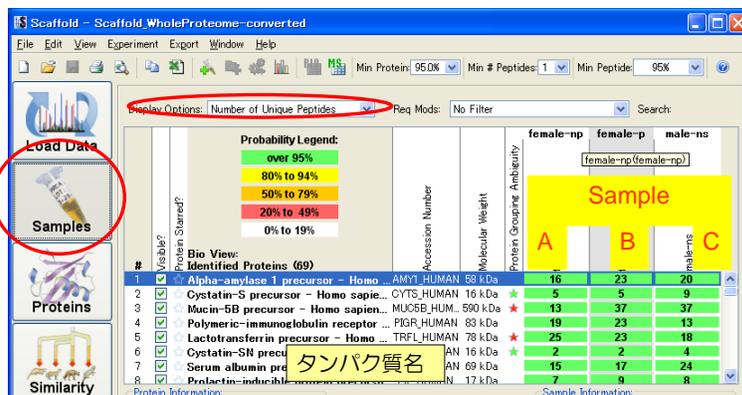
# Scaffold

- 結果まとめを行うソフトウェア
- 複数サンプル間のタンパク質リストを比較
- 複数検索エンジンの結果取り込み
- アサインされたペプチドやMS/MSのスペクトル確認も可能
- 同定タンパク質比較のベン図、GO情報円グラフ

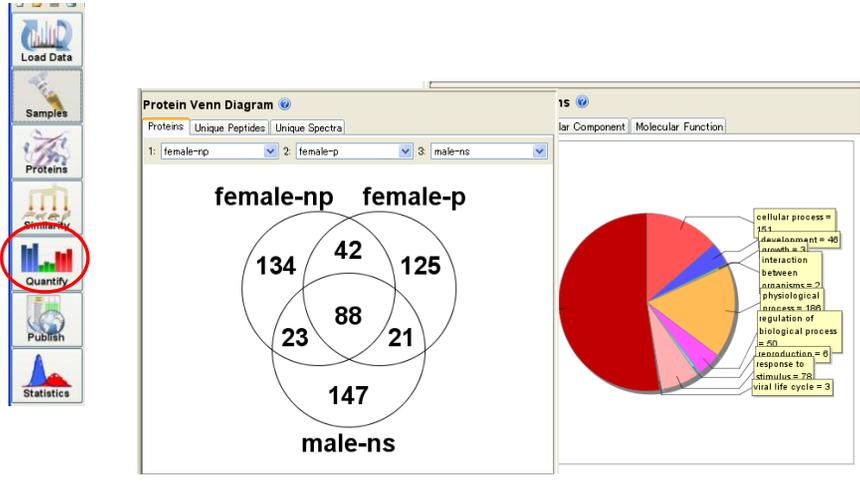
Mascot は Matrix Science 社  
Scaffold は Proteome Software 社の製品です

## サンプル間の「タンパク質同定リスト」 を並べられます

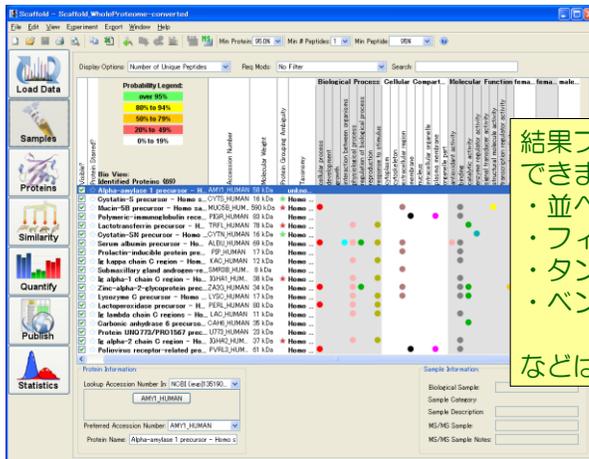
各サンプルにおけるタンパク質の同定確率、タンパク質にアサインされたペプチド / スペクトル の数、などを並べて表示します



# ベン図、Gene Ontology の円グラフ



# 取り込ませた結果を共同研究先に配り、フリーのViewer で閲覧することも可能



結果ファイル の取り込みは  
できませんが、  
・並べ替え  
・フィルターリング  
・タンパク質リストの整理  
・ベン図、GO表示  
などはすべてできます。

[http://www.proteomesoftware.com/Proteome\\_software\\_prod\\_Scaffold\\_download-main.html](http://www.proteomesoftware.com/Proteome_software_prod_Scaffold_download-main.html)

