

Volcano plot

まとめ、または図の補足説明

→ 3-1 (P.34~)

日本語マニュアルの該当ページ
×対応するスライド番号

←スライド番号

1



MASCOT SERVER PMF トレーニング資料

- MASCOT Serverとは
- 対応フォーマット、データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ
- [補足] MIS検索



■ MASCOT Serverとは

- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ

MASCOT Server とは

→ 1-1 (P.1~)

- ・質量分析装置のデータをもとに、**ペプチド配列あるいはタンパク質を同定**するソフトウェア(検索エンジン)
- ・入力と参照先、出力
 - **入力データ** : 質量分析装置から得られたスペクトル情報
 - **参照先** : 配列から計算された理論スペクトル (オプションで過去の実測スペクトルも可)
 - **出力データ** : 同定タンパク質の情報

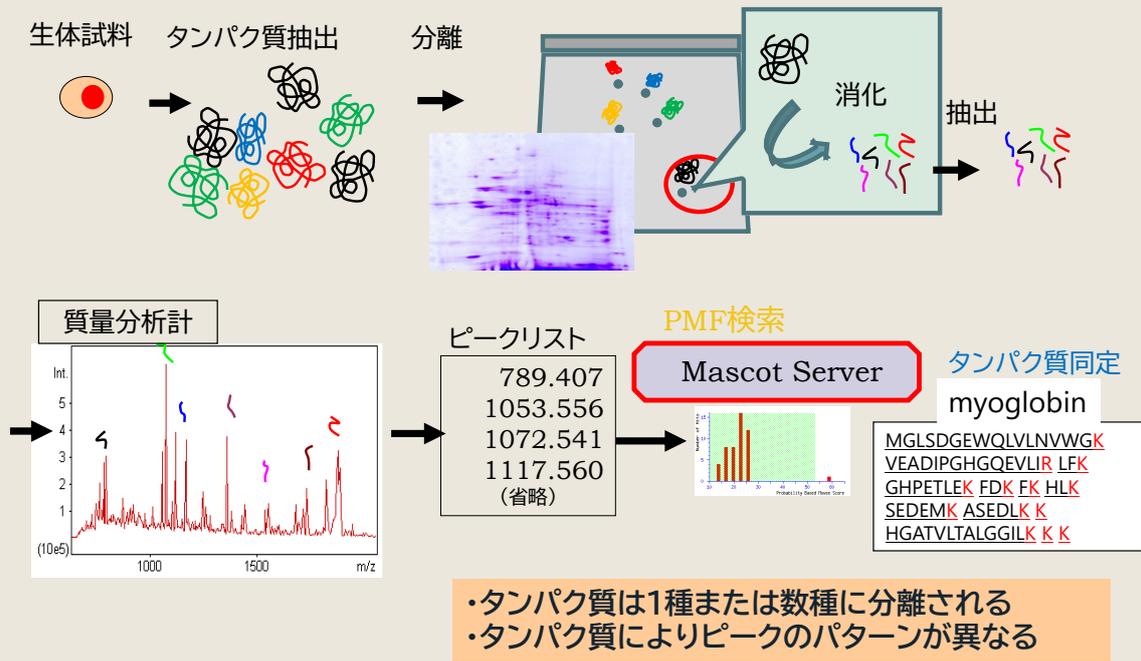
- MASCOT Serverでできる事・まとめ

- MASCOT Serverはウェブアプリケーション

MASCOT 3つの検索方法

2

補足説明：PMF解析の流れ



MASCOT Serverでできる事・まとめ①

→ 1-2 (P.1~)

定性 タンパク質の同定
 修飾の種類・付くペプチド

定量 PMFでは基本的に対応していない

以下の解析には **MIS**検索の実施が必要

- ペプチド内で修飾が付く位置
- ペプチド配列(確からしさの評価を含む)
- 定量 (ショットガンデータの解析)

MASCOT Serverとは

- MASCOT Serverでできる事・まとめ
- MASCOT Serverはウェブアプリケーション

■ 対応フォーマット・データ変換

■ ネットワーク・検索方法

■ 検索パラメーター

■ 結果画面

■ 同定タンパク質

■ 管理プログラム・カスタマイズ

補足説明 : MASCOT Serverでできる事・まとめ_図

→ 1-2 (P.2~)

1. [PML_HUMAN](#) Mass: 97489 Score: 185 Expect: 1.8e-013 Matches: 16
 Protein PML OS=Homo sapiens OX=9606 GN=PML PE=1 SV=3

[RECA_ROSCS](#) Mass: 37935 Score: 49 Expect: 6.7 Matches: 5
 Protein RecA OS=Roseiflexus castenholzii (strain DSM 13941 / HLO8) OX=383372 GN=recA PE=3 SV=1

[JF5A_PYRNV](#) Mass: 14588 Score: 47 Expect: 11 Matches: 4
 Translation initiation factor 5A OS=Pyrobaculum neutrophilum (strain DSM 2338 / JCM 9278 / V24Sta

[NADO_CHLL2](#) Mass: 22438 Score: 44 Expect: 20 Matches: 4
 Probable nicotinate-nucleotide adenyltransferase OS=Chlorobium limicola (strain DSM 245 / NBRC 10380

[RNS10_HORSE](#) Mass: 23926 Score: 42 Expect: 39 Matches: 4
 Inactive ribonuclease-like protein 10 OS=Equus caballus OX=9796 GN=RNASE10 PE=2 SV=2

[MURC_IDILO](#) Mass: 52994 Score: 42 Expect: 39 Matches: 5
 UDP-N-acetylmuramate--L-alanine liase OS=Idiomarina loihiensis (strain ATCC BAA-735 / DSM 15497 / L2-

[TRGV8_HUMAN](#) Mass: 13327 Score: 41 Expect: 11 Matches: 4
 Probable protein tyrosine phosphatase OS=Homo sapiens OX=9606 GN=PTPNS2 PE=1 SV=3

Unformatted sequence string: [882 residues](#) (for pasting into other applications).

Sort by residue number increasing mass decreasing mass
 Show matched peptides only predicted peptides also

Start - End	Observed	Mr (expt)	Mr (calc)	Delta M	Peptide
8 - 33	2882.5000	2881.4927	2881.3777	-0.1150	2 R.SPRQQDQPARPQEPMTMPPEPTEPSEGR.Q
34 - 44	1182.4400	1181.4327	1181.5677	-0.1349	0 R.QPSPSPSPPTER.A
45 - 56	1423.5200	1422.5127	1422.6779	-0.1652	0 R.APASSEEPQFLR.C
61 - 170	1191.5000	1190.4927	1190.6520	-0.1592	1 K.HEARPLAELR.N
08 - 315	1000.3300	999.3227	999.3967	-0.0740	0 R.DYEEMASR.L
319 - 325	814.4300	813.4227	813.4708	-0.0481	0 R.LDAVLQR.I
359 - 372	1624.7400	1623.7327	1623.8692	-0.1365	1 R.LRQEPQSLQAAVR.T
361 - 372	1355.5300	1354.5227	1354.6841	-0.1613	0 R.QEPQSLQAAVR.T
361 - 382	2550.3000	2549.2927	2549.2510	-0.0417	2 R.QEPQSLQAAVRTDGFDFEYR.L
373 - 380	958.3500	957.3427	957.4080	-0.0653	0 R.TDGFDFEK.V
491 - 500	1165.3900	1164.3827	1164.5081	-0.1253	1 K.MESEEGEAR.L
504 - 515	1300.4700	1299.4627	1299.6419	-0.1792	1 R.SSPQPRPSTSK.A
516 - 529	1426.5700	1425.5627	1425.7365	-0.1737	0 K.AVSPFHLDGPPSPR.S
530 - 555	2653.3900	2652.3827	2652.2780	-0.1048	0 R.SFVIGSEVLPNSNHVASGAGEAER.V
574 - 594	2265.1100	2264.1027	2264.0292	-0.0735	0 R.ELDDSSSSDQLLEGSTLR.V
595 - 616	2544.4100	2543.4027	2543.2908	-0.1120	1 R.VLDENLADPQAEPRPLVFFDLK.I

No match to: 1320.4000, 1348.4100

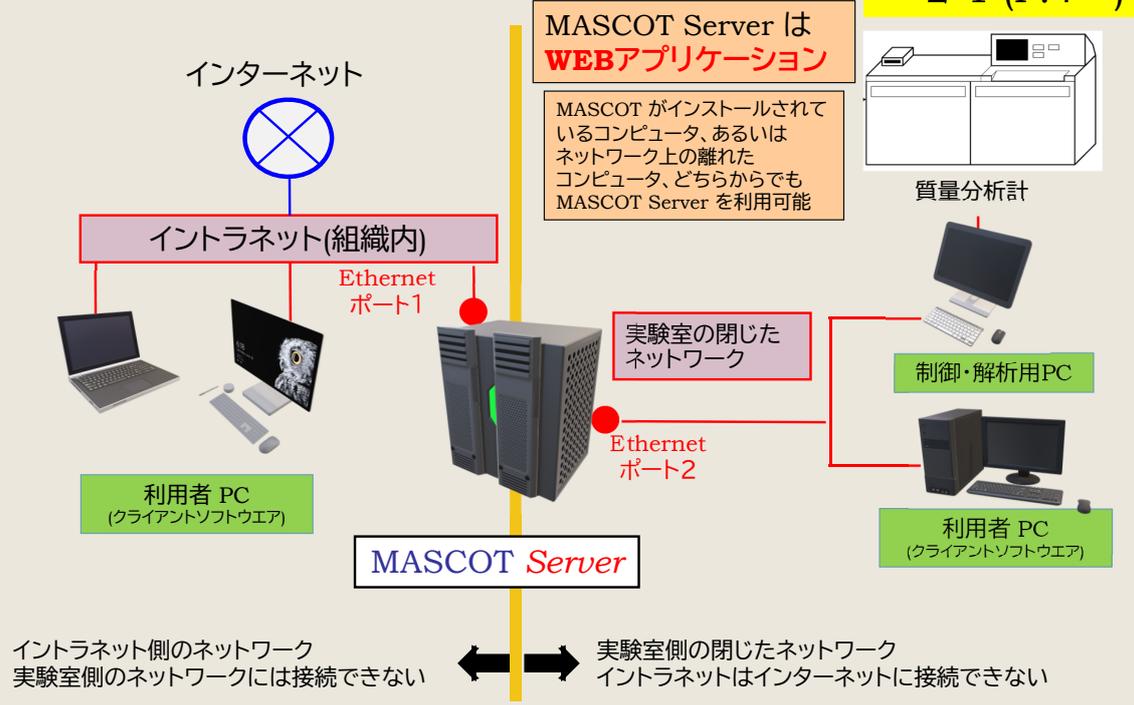
同定タンパク質のリスト

同定タンパク質にマッチしたピークの情報

-  Mascot Server
- MASCOT Serverとは
 - MASCOT Serverでできる事・まとめ
 - MASCOT Serverはウェブアプリケーション
 - MASCOT :3つの検索方法
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ

MASCOT はウェブアプリケーション

→ 2-1 (P.4~)



- MASCOT Serverとは
 - MASCOT Serverでできる事・まとめ
 - MASCOT Serverはウェブアプリケーション
 - MASCOT :3つの検索方法
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ

MASCOT : 3つの検索

→ 2-3 (P.6~)

比較的手軽に測定可能

■ PMF (Peptide Mass Fingerprinting)

MS1、ペプチドピークの組み合わせからタンパク質を同定

■ Sequence Query

MS1やMS2のピーク情報に各種絞り込み条件を追加して検索

例) 1234.2 seq(n-AC[DHK]) seq(c-HI)
1314.7 tag(513.3,T[I|L]SP,911.5)

■ MIS (Mascot Ions Search)

MS1情報とMS2情報を組み合わせてペプチドを同定。同定されたペプチドをもとにタンパク質を同定(推定)

・ペプチド単位での同定レベルを上げる
・修飾の位置を特定する

対応フォーマット・データ変換

→ 3 (P.7~)

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ

MASCOT 検索を行う前に、装置の**rawデータ**を判読可能で**必要な情報のみ抽出したデータに変換**する必要があります。MASCOT Server 自身ではrawデータを変換する事はできません。

対応するファイルフォーマット

Rawデータ変換プログラム

対応するフォーマット

→ 3-1-1 (P.7~)

■ MASCOT Serverとは

■ 対応フォーマット・データ変換
- 対応するファイルフォーマット
- rawデータ変換プログラム

■ ネットワーク・検索方法

■ 検索パラメーター

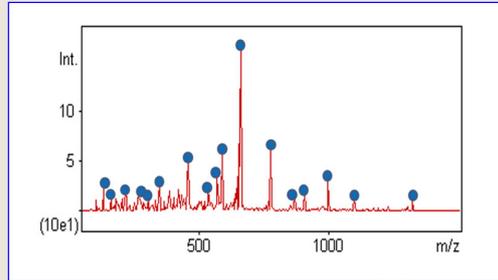
■ 結果画面

■ 同定タンパク質

■ 管理プログラム・カスタマイズ

9

ノイズをカット、ピークを抽出
(ペプチドのピークを反映していると考えられる箇所のみを抜き出して検索に使用)



テキストフォーマットであれば
ファイルの拡張子は問わない

```
832.662
903.342
1186.439
1373.681
1403.722
1515.444
1727.916
1743.951
1759.966
1788.721
1804.71
2174.812
2190.112
2256.871
2273.266
2288.489
```

```
361.21774 4838.6552
487.26656 5281.3009
494.28934 8868.6732
505.77755 16079.047
686.36334 22677.156
723.34797 36555.65
836.78635 6731.0498
955.4842 2890.8536
1002.4743 5553.321
1020.521 2661.7174
1263.679 1759.9609
1350.7063 3770.1817
1495.685 43476.619
1533.6332 3063.8439
1675.6148 3315.1174
```

1行につき1つの
m/zが先頭に
書かれていれば
読み込み可能

m/zが後ろにintensity
情報があれば入力データ
の再構成に利用
→intensityの有無で
結果が変わる事がある

rawデータ変換プログラム

→ 3-2 (P.11~)

■ MASCOT Serverとは

■ 対応フォーマット・データ変換
- 対応するファイルフォーマット
- rawデータ変換プログラム

■ ネットワーク・検索方法

■ 検索パラメーター

■ 結果画面

■ 同定タンパク質

■ 管理プログラム・カスタマイズ

10

ProteoWizard msconvert

- プロテオームデータ解析向け、オープンソースでクロスプラットフォームのツールまたはライブラリ
- msconvertGUI.exeというプログラムを起動してプログラム内でmgfを作成したり、MASCOT Daemonと組み合わせて利用する 事も可

MASCOT Distiller

- 弊社が販売しているソフトウェア。MASCOT Server/Daemonといった弊社取り扱いソフトウェアとの連携の良さや多価フラグメントピークデータも検索に組み込む事ができるのが特徴

質量分析装置付属のソフトウェア

- ほとんどの質量分析装置付属のソフトウェアではmgfファイル(ピークリストのテキストファイル)やmzMLフォーマットに出力が可能

ネットワーク・検索方法

→ 4 (P.12~)

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ

- ・MASCOT Serverをネットワーク上から指定し、検索を実行
- ・検索を実行する方法：以下3種類
 - webブラウザ経由
 - Daemon経由
 - 質量分析装置メーカーのソフトウェア経由

ネットワーク上でMASCOT Serverを指定

変換済みファイルをWEBブラウザで検索

Daemonを使ってrawデータを直接検索

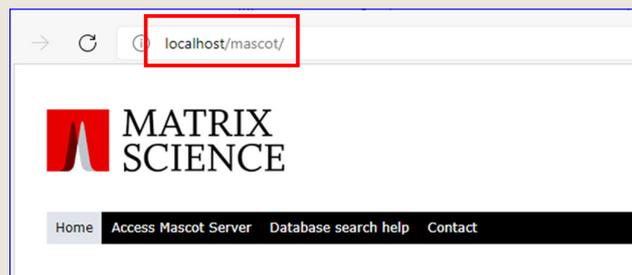
質量分析装置メーカーのソフトウェアから直接検索

ネットワーク上でMASCOT Serverを指定

→ 4-1 (P.12~)

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
 - ネットワーク上でMASCOT Serverを指定
 - 変換済みファイルをWEBブラウザで検索
 - Daemonでrawデータを直接検索
 - 質量分析装置メーカーのソフトウェアから直接検索
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ

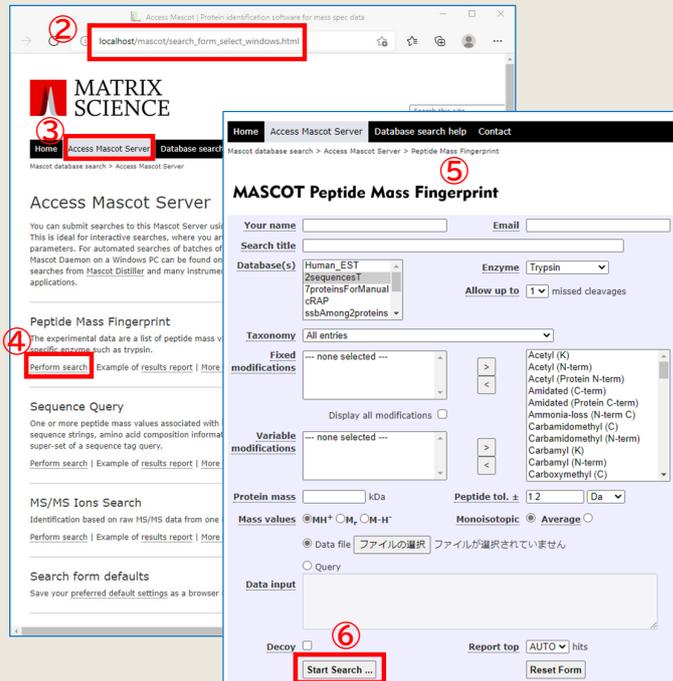
- ・自分自身あるいは別のコンピューターからMASCOT Server を**URL**で指定
- URLの記述：
- http://(computer 名)/mascot/ 例： <http://mascotserver/mascot/>
- http://(IP アドレス)/mascot/ 例： <http://192.168.100.222/mascot/>
- ・Serverがインストールされているコンピューターでは「**localhost**」を利用可能(下図)



- ネットワーク上でMASCOT Serverを指定
- 変換済みファイルをWEBブラウザで検索
- Daemonでrawデータを直接検索
- 質量分析装置メーカーのソフトウェアから直接検索

変換済みファイルをWEBブラウザで検索

→ 4-2-1 (P.13~)



- ① ピークリストファイルを準備
- ② ブラウザを開きMASCOT Serverへアクセス
- ③ Home -> Access MASCOT Server
- ④ 「Peptide Mass Fingerprint」の「Perform Search」
- ⑤ 検索パラメータ並びに入力データを指定
- ⑥ Start Search ボタンで検索開始

- ネットワーク上でMASCOT Serverを指定
- 変換済みファイルをWEBブラウザで検索
- Daemonでrawデータを直接検索
- 質量分析装置メーカーのソフトウェアから直接検索

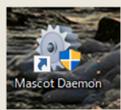
Daemonでrawデータを直接検索

→ 4-3 (P.15~)

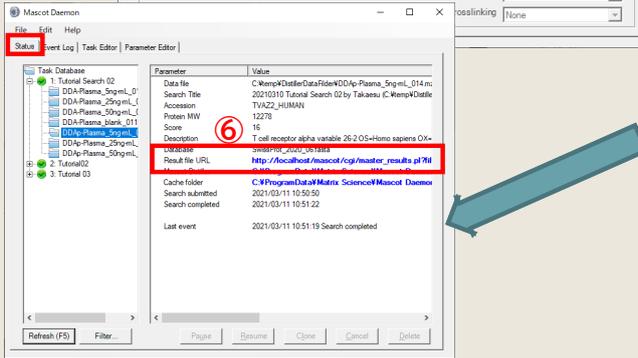
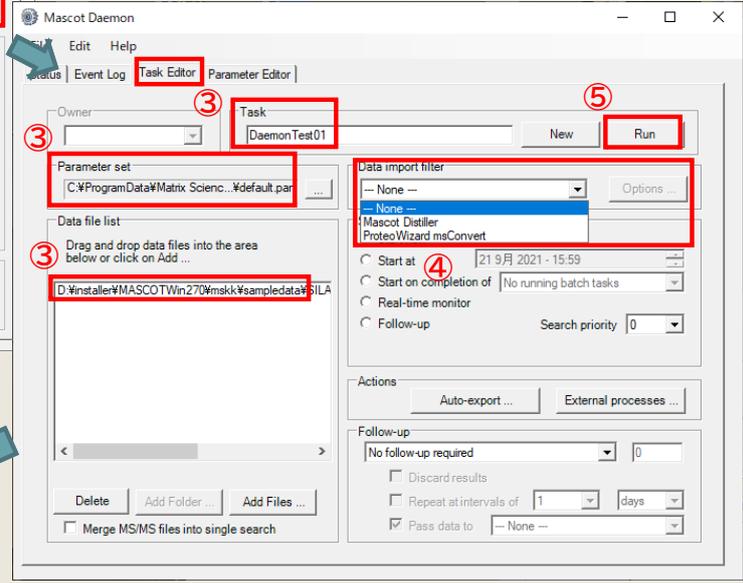
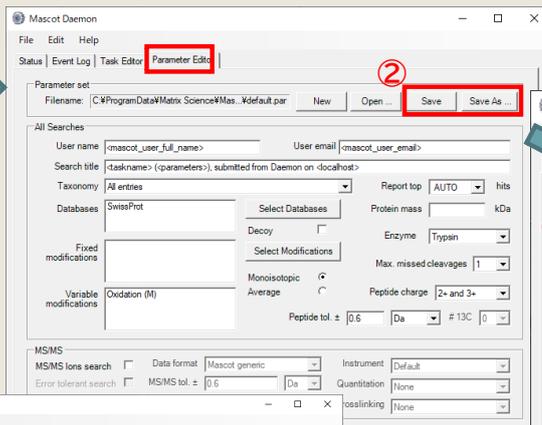
- ① Daemonを起動
- ② Parameter Editorタブで検索条件を指定し、条件をファイルで保存するため「Save」または「Save As」
- ③ Task Editor タブでタスク名、検索対象のファイル、パラメーターファイルを指定
- ④ 「Data import filter」で変換プログラム [MASCOT Distiller] または [ProteoWizard msConvert] を選択、隣の「options」ボタンでファイルフォーマットなどを指定
- ⑤ Task Editorで「Run」を押すと検索実行
- ⑥ Statusタブに進捗が表示。検索完了すると結果へのURLが表示されるのでクリック

補足説明：Daemonでrawデータを直接検索_図

→ 4-3 (P.16~)



①



Mascot Server

MASCOT Serverとは

- 対応フォーマット・データ変換
- ネットワーク・検索方法
 - ネットワーク上でMASCOT Serverを指定
 - 変換済みファイルをWEBブラウザで検索
 - Daemonでrawデータを直接検索
 - 質量分析装置メーカーのソフトウェアから直接検索
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ

メーカーのソフトウェアから直接検索

→ 4-4 (P.17~)

- ・ 質量分析装置には、概ね解析ソフトウェアがついています。
- ・ 装置メーカーソフトウェアとMASCOT Serverは基本的に別のコンピュータにインストールされています。
- ・ 装置メーカーソフトウェア上からMASCOT ServerをURL指定します。
- ・ 接続の前に、装置メーカーソフトウェアがインストールされているコンピュータからWEBブラウザでMASCOT Serverへアクセスできるかのテストを必ず行ってください。
- ・ 装置メーカーソフトウェアからMASCOT Serverに行った検索も必ずMASCOT Serverに検索ログが残ります。

ご利用方法の詳細は質量分析装置メーカーのサポートにお問い合わせください。

補足説明：具体的な検索方法

URL http :// /mascot/(index.html)

検索方法	データの種類	データ変換方法	その他
<ul style="list-style-type: none">□WEBブラウザ□Daemon□メーカーのソフトウェア ()	<ul style="list-style-type: none">□raw□mgf□その他 (mzMLなど)	<ul style="list-style-type: none">□最初からmgf等変換済み□ProteoWizard□Distiller□その他 ()	<ul style="list-style-type: none">• [注意事項など]

検索パラメーター

→ 5 (P.18~)

- ・ 参照先の配列データベースの各エントリーについて、理論値のピークリストを作成する際に検索パラメーターの情報を利用
- ・ MASCOT Server local版では各種カスタマイズが可能
- ・ 「データベースにないタンパク質・ペプチド」は見つからない
- ・ データベースの種類と特徴を把握して利用

検索パラメーター 一覧

カスタマイズ可能な項目

データベース

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ

■ MASCOT Serverとは

■ 対応フォーマット・データ変換

■ ネットワーク・検索方法

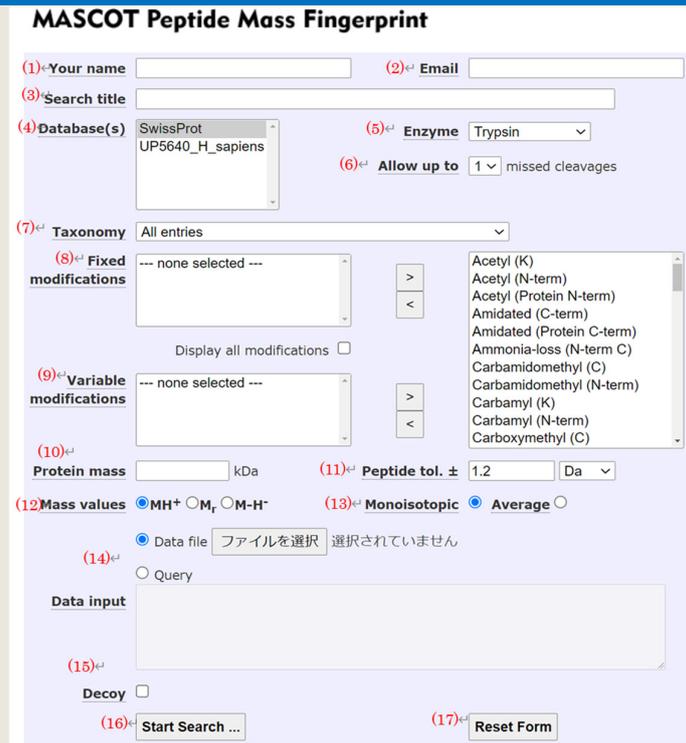
■ 検索パラメーター

- 検索パラメーター一覧
- カスタマイズ可能な項目
- データベース

■ 結果画面

■ 同定タンパク質

■ 管理プログラム・カスタマイズ



補足説明：検索パラメーター各項目について

No	項目名	説明	No	項目名	説明
1	Your name	ユーザー名	10	Protein mass	タンパク質質量の上限値
2	Email	電子メール。但しlocal版ではメモとして使用可能	11	Peptide tol. ±	実測値と理論値の誤差範囲
3	Search title	検索タイトル	12	Mass Vales	クエリーの各ピークが MH+かM-Hか、あるいはイオンが負荷していない質量に換算されたものなのか(Mr)を指定
4	Database	検索対象のデータベース	13	Monoisotopic / Average	アミノ酸の質量計算を Monoisotopicで行うか、Averageで行うか
5	Enzyme	タンパク質の切断パターン	14	Data file / Query Data input	検索queryとなるデータを指定
6	Allow up to	Enzyme設定について、切断箇所と認定された箇所を見逃し連結したペプチドを作成することができるが、何度まで見逃すことを許容するか	15	Decoy	[ver.3.0より廃止]
7	Taxonomy	生物種絞り込み	16	Start Search	検索開始
8	Fixed modifications	修飾、対象のすべてのアミノ酸について指定した内容に質量が変更	17	Reset Form	パラメーター設定をデフォルトに戻す
9	Variable modifications	修飾、修飾が つくパターンとつかないパターンの両方を考慮			

カスタマイズ可能な項目

→ 5-4 (P.33~)

■ MASCOT Serverとは

■ 対応フォーマット・
データ変換

■ ネットワーク・
検索方法

■ 検索パラメーター
- 検索パラメーター一覧
- カスタマイズ可能な項目
- データベース

■ 結果画面

■ 同定タンパク質

■ 管理プログラム・
カスタマイズ

Databases

Taxonomy

Enzyme

Quantitation

Crosslinking

Modification

Instrument

ほとんどの設定は**Configuration Editor**
(Home -> configuration Editor)で行います。各設定画面は
「**13. MASCOT Serverのカスタマイズ(P.112~)**」をご覧ください。

データベース

→ 5-5 (P.34~)

■ MASCOT Serverとは

■ 対応フォーマット・
データ変換

■ ネットワーク・
検索方法

■ 検索パラメーター
- 検索パラメーター一覧
- カスタマイズ可能な項目
- データベース

■ 結果画面

■ 同定タンパク質

■ 管理プログラム・
カスタマイズ

・ 大きく分けて2種類のデータベースを使用可能

- AA : Amino Acid : アミノ酸配列
- NA : Nucleic Acid : 核酸配列

・ 使用しているデータベースの更新や新規追加は「**Database Manager**」で

・ 「データベースにないタンパク質・ペプチド」は見つからない

・ ユーザーが適切なデータベースを選択しなければならない

・ Uniprotで生物種別に分けられたデータベース
Nは番号、Bは生物種名

・ 1st チョイスは「**SwissProt**」。→だめなら「UP_N_B」

・ MASCOTで準備しているPredefined データベース がある(次頁)

補足説明 : Predefined Database について

■ SwissProt : AA

<http://www.expasy.org/sprot/>

Uniprot データベースの中の1つ、タンパク質配列を集めたデータベース。各エントリーに対して、機能・ドメイン構造・修飾・バリエーション・論文情報・他データベースへのリンク、など、精査されたアノテーション情報が手動で付与されています。配列の冗長性はできるだけ無いように調整されていて、2025年2月版で **約57万** 件のデータが登録。 **最初に選ぶデータベースとして最適**です。

■ Uniprot : AA → MASCOTデータベースでは UP_N_B と表記 (Nは番号、Bは生物種)

<https://www.uniprot.org/>

手動アノテーションされた上記 Uniprotに加え、自動かつ精査無し of データベース **TrEMBL** を併せたデータベース。TrEMBLの方がSwissProtに比べ圧倒的にエントリー数多く、2025年2月版で **253,206,171** 件のデータが登録。SwissProtのみに比べ配列のカバー範囲が広く、**SwissProtのみでマッチしなかった場合の次の選択肢に最適**です。ただしUniprotすべてのエントリーでは件数が多すぎるので、**生物種を限定したデータベースを準備してそれに対して検索をかける事を推奨**。

■ NCBIprot (旧名称 NCBInr) : AA

NCBI(National Center for Biotechnology Information)で公開されているタンパク質データベース「nr」。「nr」とは「non-redundant」の略ですが、実際にはほぼ同じ配列のデータが数多く登録されており、それが膨大なエントリー数の要因となっています。現在は使用をお勧めしていません。

結果画面

→ 6 (P.37~)

■ MASCOT Serverとは

■ 対応フォーマット・データ変換

■ ネットワーク・検索方法

■ 検索パラメーター

■ **結果画面**

■ 同定タンパク質

■ 管理プログラム・カスタマイズ

表示例の検索について

Summary画面

Protein View

結果のファイル出力

表示例の検索について

■ MASCOT Serverとは

■ 対応フォーマット・データ変換

■ ネットワーク・検索方法

■ 検索パラメーター

■ 結果画面
- 表示例の検索について
- Summary画面
- Protein View
- 結果のファイル出力

■ 同定タンパク質

■ 管理プログラム・カスタマイズ

[入力データ]

814.430
958.350
1000.330
1165.390
1182.440
1191.500
1300.470
1320.400
1348.410
1355.530
1423.520
1426.570
1624.740
2265.110
2544.410
2550.300
2653.390

→ 6-1 (P.37~)

設定項目	設定値
Database	SwissProt
Enzyme	Trypsin/P
Allow up to	2
Taxonomy	all
Peptide tol,±	0.2 Da
Mass Values	MH ⁺
Monoisotopic/Average	Monoisotopic

[公開サーバー]

http://www.matrixscience.com/cgi/master_results.pl?file=../data/F981122.dat

[ローカルサーバー]

http://localhost/mascot/cgi/master_results.pl?file=../data/F981122.dat

Summary画面

■ MASCOT Serverとは

■ 対応フォーマット・データ変換

■ ネットワーク・検索方法

■ 検索パラメーター

■ 結果画面
- 表示例の検索について
- Summary画面
- Protein View
- 結果のファイル出力

■ 同定タンパク質

■ 管理プログラム・カスタマイズ

→ 6-2 (P.37~)

検索直後に表示される結果画面

検索条件の確認や、タンパク質のスコア分布、並びにリストアップされたタンパク質に関する情報が表示されます

Mascot Search Results

User: ...
Email: ...
Search title: Peptide Mass Fingerprint Example
Database: SwissProt 2018_10 (561956 sequences; 201858228 residues)
Timestamp: 8 Jan 2020 at 11:25:28 GMT
Top Score: 185 For **PM_HUMAN**, Protein P1L:O=Homo sapiens (O19606 GN:P1L)

Score Histogram

Protein score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event. Protein scores greater than 70 are significant ($p < 0.05$).

Concise Protein Summary Report

Significance threshold $p < 0.05$ Max. number of hits AUTO
Preferred taxonomy All entries

タンパク質の情報

1. **PM_HUMAN** Mass: 97489 Score: 185 Expect: 1.3e-013 Matches: 18
Protein P1L:O=Homo sapiens (O19606 GN:P1L PE:1 SV:2)

==== 中略 ====

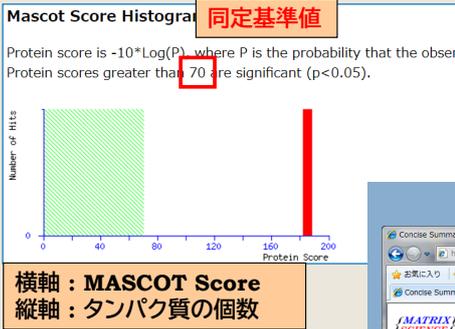
Search Parameters

Type of search : Peptide Mass Fingerprint
Enzyme : Trypsin/P
Mass values : Monoisotopic
Protein Mass : Unrestricted
Peptide Mass Tolerance : ± 0.2 Da
Peptide Charge State : 1+
Max Missed Cleavages : 2
Number of queries : 18

Mascot: <http://www.matrixscience.com/>

Summary画面 : Score Histogram

→ 6-2-3 (P.39~)

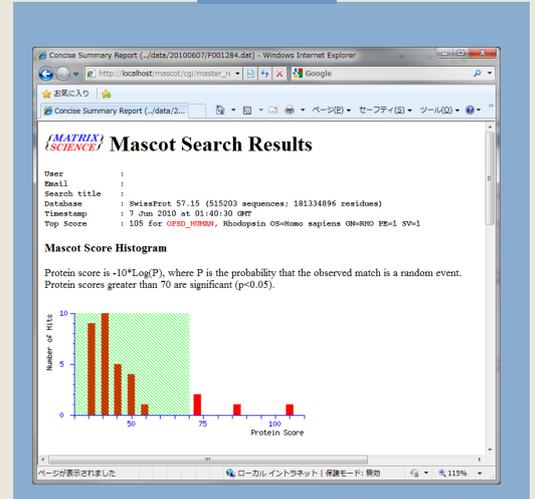
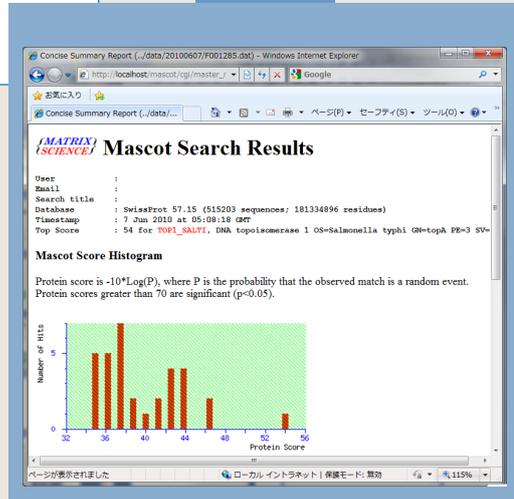


検索が成功/失敗したかの判断

緑の網掛けゾーンを超えたところにスコアを持つタンパク質があるかどうか

失敗

成功



Summary画面 : タンパク質の情報

→ 6-2-6 (P.41~)

- [PML_HUMAN](#) Mass: 97489 Score: **185** Expect: 1.8e-013 Matches: 16
Protein PML OS=Homo sapiens OX=9606 GN=PML PE=1 SV=3

[RECA_ROSCS](#) Mass: 37935 Score: 49 Expect: 6.7 Matches: 5
Protein RecA OS=Roseiflexus castenholzii (strain DSM 13941 / HLO8) OX=383372 GN=recA PE=3 SV=1

[IF5A_PYRNV](#) Mass: 14588 Score: 47 Expect: 11 Matches: 4
Translation initiation factor 5A OS=Pyrobaculum neutrophilum (strain DSM 2338 / JCM 9278 / V24Sta) OX=444157 GN=eIF5A PE=3 SV=

[NADD_CHLL2](#) Mass: 22438 Score: 44 Expect: 20 Matches: 4
Probable nicotinate-nucleotide adenyltransferase OS=Chlorobium limicola (strain DSM 245 / NBRC 103803 / 6330) OX=290315 GN=

[RNS10_HORSE](#) Mass: 23926 Score: 42 Expect: 39 Matches: 4
Inactive ribonuclease-like protein 10 OS=Equus caballus OX=9796 GN=RNASE10 PE=2 SV=2

Scoreが赤字で表示されているタンパク質が「同定タンパク質」

表示項目	説明
Accession	データベースのIDが表示。ハイパーリンクをクリック → 詳しい情報が記載されている「Protein View」
Mass	タンパク質の質量、データベースに登録されている配列情報から計算
Score	MASCOT Score。赤字の表示は同定基準を超えている事を表します。高いほど理論値と実測値がよりよくマッチしていることを示します。
Expect	Scoreと同定基準値をもとに算出された値。ランダムマッチだった場合、検索したデータベースからどれくらいの エントリーが見つかるかを表す「期待値」。同定基準を超えている時値が0.05(デフォルト設定の場合)より小さくなります。なお同定基準値は「Significance threshold $p < \dots$ 」の値と連動します。
Matches	マッチしたピーク数
Description	データベース各エントリーのヘッダー行に記載されている、タンパク質の機能に関する情報

- Mascot Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
 - 表示例の検索について
 - Summary画面
 - Protein View
 - 結果のファイル出力
- 同定タンパク質
- 管理プログラム・カスタマイズ

Protein View

→ 6-3 (P.42~)

ハイパーリンクになっている、タンパク質のAccession部分をクリックすると表示される画面

タンパク質について、マッチしたペプチドの情報などを確認できる

右図内の緑で囲われた部分について、以降でさらに詳細な説明があります

MASCOT Search Results
 Protein View: PML_HUMAN
 Protein PML OS=Homo sapiens OX=9606 GN=PML PE=1 SV=3
 Database: SwissProt
 Score: 185
 Expect: 1.8e-013
 Monoisotopic mass (M): 97480
 Calculated pI: 5.88
 Taxonomy: Homo sapiens
 Sequence similarity is available as an NCBI BLAST search of PML_HUMAN against .

Search parameters
 Enzyme: Trypsin/P: cuts C-term side of KR.
 Mass values searched: 18
 Mass values matched: 16

Protein sequence coverage: 23%

Matched peptides shown in **bold red**.

```

1 MEAPAPASPR PQQDPARPQE PTMPPPETPS EGRQPSPPS PTERAPASEE
51 EFQFLRCQQC QAEAKCPKLL PCLHTLCSCG LEASGMQCPI CQAPWPLGAD
101 TPALDNVFFE SLQRRLSVYR QIVDAQAVCT RCKESADFWC FECEQLLCAK
151 CFEAHQWFLK HEARPLAELR NQSVREFLDG TRKTNINFCS NPNHRTPLT
201 SIYCRGCSKP LCCSCALLDS SHSELKCDIS AEIQQRQEEL DAMTQALQEQ
251 DSFAFQAVHAQ MHAAVGQLGR ARAETEELIR ERVRQVVAHV RAQERELLEA
301 VDARYQRDYE EMASRLGRLD AVLQRIITGS ALVQRMKCYA SDQEVLDMHG
351 FLRQALCRLR QEEPQSLQAA VRTDGFDEFK VRLQDLSSCI TQKDAAVSK
401 KASPEAASP RDPIDVDLPE EAERVKAQVQ ALGLAEAQPM AVQVSPVGAH
451 FVPVYAFSIK GPSYGEDVSN ITTAQKRKCS QTQCPKRVIK MESEEGKEAR
501 LARSSPEQPR PSTSKAVSPP HLDGPPSPRS PVIGSEVFLP NSNHVASGAG
551 EAEERVVVIS SSEDSDAENS SSELDDSSS ESSDLQLEGP STLRVLDENL
601 ADPQAEADRL VFFDLKIDNE TQKISQLAAV NRESKFRVVI QPEAFFSIYS
651 KAVSLEVGLQ HFSLFLSSMR RPILACYKLW GPGLPNFFRA LEDINRLWEF
701 QEAISGFLAA LPLIRERVPG ASSFKLKNLA QTYLARNMSE RSAMAAVLAM
751 RDLCLRLEVS PGPQLAQHVY PFSSLQCFAS LQPLVQAAVL PRAEARLLAL
801 HNVSMFELLS AHRRDRQGGI KKYSTRYLSLQ TITLPPAQP A FNLAALGTYF
851 EGLLEGPALA RAEGVSTPLA GRGLAERASQ QS
  
```

Start - End	Observed	Mr(expt)	Mr(calcd)	Delta M	Peptide
8 - 33	2882.4000	2881.4927	2881.3777	-0.1150	R.SFRQDQPARPQETMPPPETSEGR.Q
34 - 44	1192.4400	1181.4327	1181.2677	-0.1549	R.QPSFSPPTER.LA
45 - 56	1463.5000	1462.5127	1462.4779	-0.1658	R.AHARSEEPER.LC
161 - 170	1191.3000	1190.4927	1190.6520	-0.1592	R.HEARPLAELR.N
359 - 315	1000.3300	999.3227	999.3967	-0.0740	R.DIEMAR.L
319 - 325	814.4100	813.4627	813.4708	-0.0481	R.LAVQR.L
359 - 372	1624.7400	1623.7327	1623.8692	-0.1365	R.LRQEPQIQAAV.T
361 - 372	1330.3300	1324.3227	1324.6841	-0.1613	R.QEPQIQAAV.T
361 - 389	2250.3000	2249.3927	2249.2510	0.0417	R.QEPQIQAAVTDPEERV.L
373 - 380	958.3500	957.3427	957.4080	-0.0653	R.TDGFDEFK.V
491 - 500	1165.3900	1164.3927	1164.5081	-0.1153	R.MESERGER.L
504 - 512	1300.4700	1299.4627	1299.5419	-0.1791	R.SVDFPPTER.LA
516 - 529	1426.5700	1425.5627	1425.7365	-0.1737	R.KAVSFLDQPSPR.S
530 - 535	2633.3900	2632.3927	2632.2780	0.1048	R.SVDFPPTER.LA
574 - 594	2265.1100	2264.1027	2264.0292	0.0735	R.HDSKESDQLQEPFL.V
595 - 616	2544.4100	2543.4027	2543.2908	0.1120	R.VLDENLADQPSDFLVPDLK.I

No Match: 60: 1240-1300, 1340-1400

Error (Da) vs Mass (Da) plots for matched and predicted peptides.

Protein View: Sequence coverage

Protein sequence coverage: **23%**

Matched peptides shown in **bold red**.

```

1 MEAPAPASPR PQQDPARPQE PTMPPPETPS EGRQPSPPS PTERAPASEE
51 EFQFLRCQQC QAEAKCPKLL PCLHTLCSCG LEASGMQCPI CQAPWPLGAD
101 TPALDNVFFE SLQRRLSVYR QIVDAQAVCT RCKESADFWC FECEQLLCAK
151 CFEAHQWFLK HEARPLAELR NQSVREFLDG TRKTNINFCS NPNHRTPLT
201 SIYCRGCSKP LCCSCALLDS SHSELKCDIS AEIQQRQEEL DAMTQALQEQ
251 DSFAFQAVHAQ MHAAVGQLGR ARAETEELIR ERVRQVVAHV RAQERELLEA
301 VDARYQRDYE EMASRLGRLD AVLQRIITGS ALVQRMKCYA SDQEVLDMHG
351 FLRQALCRLR QEEPQSLQAA VRTDGFDEFK VRLQDLSSCI TQKDAAVSK
401 KASPEAASP RDPIDVDLPE EAERVKAQVQ ALGLAEAQPM AVQVSPVGAH
451 FVPVYAFSIK GPSYGEDVSN ITTAQKRKCS QTQCPKRVIK MESEEGKEAR
501 LARSSPEQPR PSTSKAVSPP HLDGPPSPRS PVIGSEVFLP NSNHVASGAG
551 EAEERVVVIS SSEDSDAENS SSELDDSSS ESSDLQLEGP STLRVLDENL
601 ADPQAEADRL VFFDLKIDNE TQKISQLAAV NRESKFRVVI QPEAFFSIYS
651 KAVSLEVGLQ HFSLFLSSMR RPILACYKLW GPGLPNFFRA LEDINRLWEF
701 QEAISGFLAA LPLIRERVPG ASSFKLKNLA QTYLARNMSE RSAMAAVLAM
751 RDLCLRLEVS PGPQLAQHVY PFSSLQCFAS LQPLVQAAVL PRAEARLLAL
801 HNVSMFELLS AHRRDRQGGI KKYSTRYLSLQ TITLPPAQP A FNLAALGTYF
851 EGLLEGPALA RAEGVSTPLA GRGLAERASQ QS
  
```

タンパク質全長に対してマッチしたペプチドがどの部位にあたるのか、並びにその割合についての情報を表示

Protein sequence coverage :
 全長に対するマッチペプチド残基数の割合

Protein View: matched peptides

Unformatted sequence string: [882 residues](#) (for pasting into other applications).

Sort by residue number increasing mass decreasing mass
 Show matched peptides only predicted peptides also

Start - End	Observed	Mr (expt)	Mr (calc)	Delta M	Peptide
8 - 33	2882.5000	2881.4927	2881.3777	0.1150 2	R.SFRPQQDPAAPQEPMTMPPPTPSEGR.Q
34 - 44	1182.4400	1181.4327	1181.5677	-0.1349 0	R.QPSFSPSPPTER.A
45 - 56	1423.5200	1422.5127	1422.6779	-0.1652 0	R.APASEEEFQFLR.C
161 - 170	1191.5000	1190.4927	1190.6520	-0.1592 1	K.HEARPLAELR.N
308 - 315	1000.3300	999.3227	999.3967	-0.0740 0	R.DYEEMASR.L
319 - 325	814.4300	813.4227	813.4708	-0.0481 0	R.LDAVLQR.I
359 - 372	1624.7400	1623.7327	1623.8692	-0.1365 1	R.LRQEEPOSQAQAVR.T
361 - 372	1355.5300	1354.5227	1354.6841	-0.1613 0	R.QEEPOSQAQAVR.T
361 - 382	2550.3000	2549.2927	2549.2510	0.0417 2	R.QEEPOSQAQAVRTDGFDFKVR.L
373 - 380	958.3500	957.3427	957.4080	-0.0653 0	R.TDGFDFK.V
491 - 500	1165.3900	1164.3827	1164.5081	-0.1253 1	K.MESEEGKEAR.L
504 - 515	1300.4700	1299.4627	1299.6419	-0.1792 1	R.SSPQPRPSTSK.A
516 - 529	1426.5700	1425.5627	1425.7365	-0.1737 0	K.AVSPPHLDGPPSPR.S
530 - 555	2653.3900	2652.3827	2652.2780	0.1048 0	R.SPVIGSEVFLPNSNHVASGAGEAER.V
574 - 594	2265.1100	2264.1027	2264.0292	0.0735 0	R.ELDDSSSESSLQLEGPSTLR.V
595 - 616	2544.4100	2543.4027	2543.2908	0.1120 1	R.VLDENLADPQAEADRLVFFDLK.I

No match to: 1320.4000, 1348.4100

マッチしたペプチドについて、
アミノ酸残基順 (デフォルト設定)
に並べたリストを表示

表示項目	説明
Unformatted sequence string	ハイパーリンクをクリックすると、配列をコピーしやすくなるページが開きます。他プログラムで配列を使用したい場合に便利
Sort by	リストの並び順を指定。残基番号、質量の昇順/降順 が選択可
Show	理論値と実測値がマッチしたペプチドのみをリストに表示させるか、マッチしなかった理論ピークも表示させるかを選択
Start-End	タンパク質全長におけるアミノ酸残基番号
Observed	ピークリストファイルの m/z
Mr(expt)	ピークリストの値から計算されたペプチドの質量
Mr(calc)	配列から計算されたペプチドの質量
Delta	Mr(expt) - Mr(calc)
M	Missed cleavage
Peptide	ペプチド配列。修飾も含まれる場合は併せて表示

結果のファイル出力

→ 6-4 (P.47~)

Concise Protein Summary Report

Format As **Export Search Results** [Help](#)

Significance threshold p < Max. number of hits

Preferred taxonomy

「Format As」の選択肢で
「Export Search Results」
を選択したうえで
「Format As」ボタンを押します

出力のファイルフォーマットや条件、
出力項目を選択し、画面下の「Export
Search Results」ボタンを押すと
ファイル出力が実行されます

MATRIX SCIENCE

Home Access Mascot Server Database search help Contact

Mascot database search > Access Mascot Server > Export search results

Export search results

Export format

Significance threshold p <

Max. number of hits

Min. number of sig. unique sequences

Include same-set protein hits (additional proteins that span the same set of peptides)

Include sub-set protein hits (additional proteins that span a sub-set of peptides)

Preferred Taxonomy*

* Occasionally requires information to be retrieved from external utilities, which

Search Information

Header

Modification deltas

Search parameters

Format parameters

Residue masses

Protein Hit Information

Score

Significance threshold

Expectation value

MASCOT Serverとは

対応フォーマット・データ変換

ネットワーク・検索方法

検索パラメーター

結果画面

- 表示例の検索について
- Summary画面
- Protein View
- 結果のファイル出力

同定タンパク質

管理プログラム・カスタマイズ

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ

同定タンパク質

・MASCOT 検索結果で、基本的には1種類、→ 8 (P.71~)
同定基準値を超えるスコアを持つタンパク質が同定タンパク質

・類似配列のタンパク質の扱いに注意

PMFタンパク質同定のまとめ

入力データの調整

配列から行う理論値計算

マッチングとスコア、同定基準値、期待値

ユニーク/シェア ペプチド、タンパク質のグループ化

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
 - タンパク質同定のまとめ
 - 入力データの調整
 - 配列から計算される理論ピーク
 - マッチングとスコア、同定基準値、期待値
 - ユニーク/シェアペプチド、タンパク質のグループ化
- 管理プログラム・カスタマイズ

PMFタンパク質同定のまとめ

→ 8-1 (P.71~)

- PMFでは1度の検索で**基本的に1種類のタンパク質が同定**される
- 入力データはピーク作成時とマッチング時の**2段階で選別・組み換え**される
- **スコア**とは理論値と実測値のマッチング度合いを評価した数値で、**高いほどよくマッチ**している
- 検索規模をもとに同定基準値が計算される。**スコアが同定基準値を超えた時タンパク質を同定とみなす(信頼度は95%)**
- タンパク質間でシェアされているペプチド情報に十分注意して結果を解釈しなければならない
- 「検索対象のデータベースの選択」も結果に大きく影響する。
 - 「配列のないタンパク質はマッチしない」
 - 「データベースが大きすぎると同定しにくい」

入力データの調整

→ 8-2 (P.71~)

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
 - タンパク質同定のまとめ
 - 入力データの調整
 - 配列から計算される理論ピーク
 - マッチングとスコア、同定基準値、期待値
 - ユニーク/シェアペプチド、タンパク質のグループ化
- 管理プログラム・カスタマイズ

入力データは2段階で調整

1段階目：

検索にかける前段階でペプチドのピークを抽出しノイズをカットしたものを入力データとして受け付ける

2段階目：

intensity情報がピークリストに含まれている時、MASCOT Serverプログラムにおいて、入力データから強度情報に基づいて**10種類のサブセット入力データ**を作成し、それぞれのサブセットに対して理論ピークとのマッチングを行って最もスコアが高くなったサブセットを採用する

- MASCOT Serverプログラムは2段階目の処理を行う事ができますが、あくまでも**1段階目の前処理としてピーク抽出が行われていることを前提**です

配列から行う理論値ピーク

→ 8-3 (P.72~)

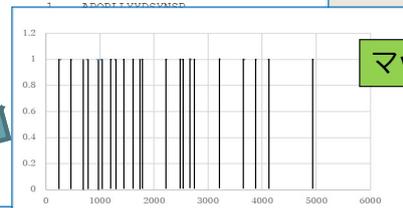
- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
 - タンパク質同定のまとめ
 - 入力データの調整
 - 配列から計算される理論値ピーク
 - マッチングとスコア、同定基準値、期待値
 - ユニーク/シェアペプチド、タンパク質のグループ化
- 管理プログラム・カスタマイズ

検索時に指定したデータベースに登録されているタンパク質一つ一つに対して、パラメーターをもとに理論ペプチドを作成しその質量を計算

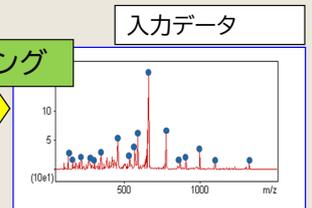
```
>sp|A0A0C4DH27|TRGV8_HUMAN T cell receptor gamma variable 8
MLLALALLLAFLLPPASQKSSNLEGRTKSVTRPTGSSAVITCDLPVENAVYTHWYHLHQEGKAPQRL
LYYDSYNSRVVLESGISREKYHTYASTGKSLKFLLENLIERDSGVYYCATWDR
```

2 - 18	1778.1070	0	LLALALLLAFLLPPASQK
2 - 25	2521.4632	1	LLALALLLAFLLPPASQKSSNLEGR
2 - 27	2750.6058	2	LLALALLLAFLLPPASQKSSNLEGRK
19 - 25	761.3668	0	SSNLEGR
19 - 27	990.5094	1	SSNLEGRK
19 - 31	1433.7587	2	SSNLEGRKSVTR
26 - 27	247.1532	0	SVTR
26 - 31	690.4024	1	SVTRK
26 - 60	3886.9312	2	SVTRPTGSSAVITCDLPVENAVYTHWYHLHQEGK
28 - 31	461.2598	0	SVTR
28 - 60	3657.7886	1	SVTRPTGSSAVITCDLPVENAVYTHWYHLHQEGK
28 - 64	4110.0382	2	SVTRPTGSSAVITCDLPVENAVYTHWYHLHQEGK
32 - 60	3214.5394	0	PTGSSAVITCDLPVENAVYTHWYHLHQEGK
32 - 64	3666.7889	1	PTGSSAVITCDLPVENAVYTHWYHLHQEGKAPQF
32 - 74	4941.3821	2	PTGSSAVITCDLPVENAVYTHWYHLHQEGKAPQF
61 - 64	470.2601	0	APQF
61 - 74	1744.8533	1	APQFPTGSSAVITCDLPVENAVYTHWYHLHQEGK
61 - 83	2685.3875	0	SVTR
65 - 74	1292.6037	0	SVTR
65 - 83	2233.1379	0	SVTR
65 - 85	2490.2755	0	SVTR
75 - 83	958.5447	0	SVTR
75 - 85	1215.6877	0	SVTR
75 - 94	2224.1488	0	SVTR
84 - 85	275.1481	0	SVTR
84 - 94	1283.6146	0	SVTR
84 - 97	1611.8257	0	SVTR
86 - 94	1026.4771	0	SVTR

タンパク質の、「理論値ペプチドピーク」セットを作成



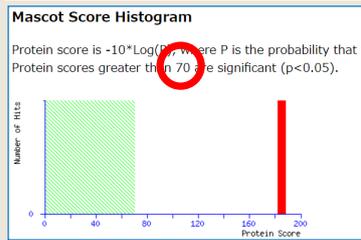
マッチング



マッチングとスコア、同定基準値、期待値

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
 - タンパク質同定のまとめ
 - 入力データの調整
 - 配列から計算される理論ピーク
 - マッチングとスコア、同定基準値、期待値
 - ユニーク/シェアペプチド、タンパク質のグループ化
- 管理プログラム・カスタマイズ

→ 8-4 (P.73~)



PML_HUMAN	Mass: 97489	Score: 185	Expect: 1.8e-013	Matches: 16
Protein PML OS=Homo sapiens OX=9606 GN=PML PE=1 SV=1				
RECA_ROSCS	Mass: 37935	Score: 49	Expect: 6.7	Matches: 5
Protein RecA OS=Roseiflexus castenholzii (strain DSM 13941 / HL08) OX=383372				

スコア (S)

実測スペクトルと理論スペクトルとのマッチング度合いを表し、**スコアが高いほど両スペクトルが良くマッチ**。実測スペクトルと理論スペクトルのマッチングがランダムな事象である確率をProとした場合、**スコアは $-10 \log_{10}(\text{Pro})$** と表すことができます。

同定基準値 (St)

検索毎に計算される同定基準。同定基準は信頼度**95%**を満たす点が提示され、主に検索対象となったデータベースの**エントリー数を基に算出**。スコア1位の結果でも同定基準値を超えない限りそのタンパク質が同定されたとはみなしません。

期待値 (Expect)

検索対象のデータベース中に、同様のランダムマッチをする事が期待されるタンパク質の数。**期待値が 0.05より小さい時同定**とします。

スコアと同定基準値、期待値の関係

$$\text{Expect} = 0.05 \cdot 10^{-(S-St)/10}$$

ユニーク/シェアペプチド、タンパク質のグループ化

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
 - タンパク質同定のまとめ
 - 入力データの調整
 - 配列から計算される理論ピーク
 - マッチングとスコア、同定基準値、期待値
 - ユニーク/シェアペプチド、タンパク質のグループ化
- 管理プログラム・カスタマイズ

→ 8-5 (P.74~)

- ・ 類似配列のタンパク質に注意する
- ・ 最終的に類似タンパク質の中のどれが答えかは、判断しにくい/できない 事も多い

・MASCOTの中では類似タンパク質をどのように処理しているか、次スライド以降で説明します

補足説明：4タンパク質のペプチドシェア状況と結果画面 1

ピーク	OPSD_HUMAN ●	OPSD_PHOVI ○	OPSD_MACFA ●	OPSD_CRIGR ●
832.662	■	■	■	■
903.342	■	■	■	■
1186.439	■		■	■
1403.722	■		■	
1617.857	■		■	■
1727.916	■	■		
1743.951			■	
1759.966				■
1788.721				
1818.963	■	■		■
2159.143	■	■		
2174.812			■	
2256.871	■		■	■

「OPSD_HUMAN」が答えだが、類似配列のタンパク質が MASCOTの中でどのようにレポートされるか？→次スライド

39

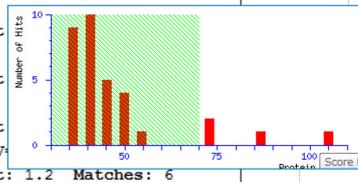
補足説明：4タンパク質のペプチドシェア状況と結果画面 2

1.	OPSD_HUMAN Mass: 38866 Score: 105 Expect: 1.6e-05 Matches: 9 Rhodopsin OS=Homo sapiens GN=RHO PE=1 SV=1
	OPSD_PHOVI Mass: 38947 Score: 45 Expect: 16 Matches: 5 Rhodopsin OS=Phoca vitulina GN=RHO PE=1 SV=1
	OPSD_MOUSE Mass: 39002 Score: 43 Expect: 28 Matches: 5 Rhodopsin OS=Mus musculus GN=Rho PE=1 SV=1
	OPSD_SHEEP Mass: 38866 Score: 43 Expect: 28 Matches: 5 Rhodopsin OS=Ovis aries GN=RHO PE=1 SV=2
2.	OPSD_MACFA Mass: 39036 Score: 88 Expect: 0.00075 Matches: 8 Rhodopsin OS=Macaca fascicularis GN=RHO PE=2 SV=1
	LUXS2_LACDB Mass: 17500 Score: 45 Expect: 17 Matches: 4 S-ribosylhomocysteine lyase 2 OS=Lactobacillus delbrueckii subsp. bulga
3.	OPSD_CRIGR Mass: 39071 Score: 73 Expect: 0.027 Matches: 7 Rhodopsin OS=Cricetulus griseus GN=RHO PE=1 SV=1
	OPSD_CANFA Mass: 38936 Score: 56 Expect: 1.2 Matches: 6 Rhodopsin OS=Canis familiaris GN=RHO PE=1 SV=1
	OPSD_FELCA Mass: 39023 Score: 56 Expect: 1.2 Matches: 6 Rhodopsin OS=Canis familiaris GN=RHO PE=1 SV=1
	OPSD_RABIT Mass: 38968 Score: 56 Expect: 1.2 Matches: 6 Rhodopsin OS=Lepus sylvaticus GN=RHO PE=1 SV=1
	OPSD_SMICR Mass: 39037 Score: 56 Expect: 1.2 Matches: 6 Rhodopsin OS=Oryctolagus cuniculus GN=RHO PE=1 SV=1

○であるOPSD_HUMANと○のOPSD_PHOVIは1位のグループとしてまとめられる (OPSD_PHOVIはOPSD_HUMANのsub-set)

●のOPSD_MACFAとOPSD_CRIGRは2位、3位として別に報告。しかし実際に**OPSD_HUMANの他にOPSD_MACFAやOPSD_CRIGRが** 本当に別に含まれていたかについてはケースバイケースで、確率はあまり高くない

* 「■」のピークの部分がランダムマッチである可能性は否定できない



区別するためには？

MALDIの測定でポイントとなるペプチドを取り出してMS2データを測定してMISを行ったり、最初からショットガンなどMIS検索を実施する必要がある

40

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ

- ・MASCOTで使用しているデータベース、検索結果などを管理する事ができる画面があります
- ・MASCOTで使用している各種項目をカスタマイズする事ができます

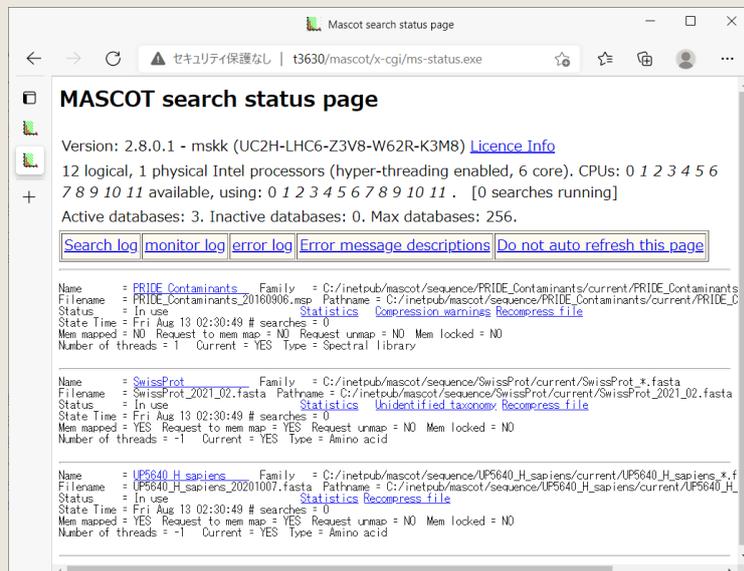
Database Status

Search log

Configuration Editor

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ
 - Database Status
 - Search log
 - Configuration Editor

- ・Home → Database Status で開く画面
- ・各種動作のログに関する情報へのリンクや、MASCOT Server上で現在使用可能なデータベースに関する情報について表示されます



補足説明：各データベースの表示内容

```
Name = SwissProt Family = C:/inetpub/mascot/sequence/SwissProt/current/SwissProt_*.fasta
Filename = SwissProt_2021_02.fasta Pathname = C:/inetpub/mascot/sequence/SwissProt/current/SwissProt_2021_02.fasta
Status = In use Statistics Unidentified taxonomy Recompress file
State Time = Fri Aug 13 02:30:49 # searches = 0
Mem mapped = YES Request to mem map = YES Request unmap = NO Mem locked = NO
Number of threads = -1 Current = YES Type = Amino acid
```

項目名	説明
Name	データベースの名称
Family	MASCOT Server側で管理している、ファイルのpath並びにファイルの名称ルールの情報
Filename	現在認識されているfastaファイルの名称。 データベースのバージョン もわかります。
Pathname	現在認識されているfastaとファイルが置かれているpath
Status	データベースの現在の状況
State Time	現在示されている Status を認識した日時
Mem mapped	データベースがメモリ上にマップされた状態であるか

項目名	説明
Request to mem map	MASCOT Serverの設定でデータベースをメモリにマッピングする事を試みる設定であるか
Request unmap	MASCOTのプログラムがメモリ上へのマッピングを解除する命令を下した状態であるか
Mem locked	データベースのメモリ上へのマッピングを固定(lock)する状態になっているかどうか
Number of thresholds	検索に使用可能なコア数の設定。通常は、最適設定を自動適用する設定であることを表す「-1」と表示
Current	現在データベース関連のファイルが正しく認識されているか
Type	配列の種類が Amino AcidかNucleic Acidか

43



Search log

→ 12-2 (P.111~)

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ
 - Database Status
 - Search log
 - Configuration Editor

- ・Home → Search log で開く画面
- ・MASCOT Serverで行ったすべての検索のログを開き、結果画面を開く事ができます
- ・項目別に並び替えやフィルターリング表示させることも可能

MASCOT search log

Version: 2.8.0 - mskk (YRNB-5YZ8-GFBC-T9W9-CYNQ)

Log File: logs\searches.log Start at: (-1=end, 1=start) how many: 50 273 in log, 273 after filters. Data dir: GETS?:

Job#	PID	dbase	User Name	Email	Ti	In	start time	Durati	Status	Prio	Type	Enzyme	IP	User ID	Peak list data file
1511	5436	UP2195_D	Monitor Test DB 0		MS	..	Fri Aug 6 11:57:49 2021	1	No email setu	0	MIS	Yes		0	test_search.mgf
1510	12840	UP5640_H			be	..	Wed Aug 4 00:00:25 2021	410	User read res	0	MIS	Yes		0	C:\temp\mascotsearchtest202
1509	14116	UP5640_H			be	..	Tue Aug 3 23:49:12 2021	415	User read res	0	MIS	Yes		0	C:\temp\mascotsearchtest202
1508	2452	UP5640_H			be	..	Tue Aug 3 23:38:02 2021	414	User read res	0	MIS	Yes		0	C:\temp\mascotsearchtest202
1507	6932	SwissPro	Monitor Test DB 0		MS	..	Tue Aug 3 17:05:40 2021	3	No email setu	0	MIS	Yes		0	test_search.mgf
1506	9944	SwissPro	takaesu		Ly	..	Wed Jul 14 10:35:59 2021	23	User read res	0	MIS	Yes	19	0	Lysozyme_2p2mJ_1minute.ten
1505	716	SwissPro			IT	..	Thu Jul 8 02:01:25 2021	17	User read res	0	MIS	Yes		0	C:\temp\ITRAQ8plex6data\wfr
1504	18348	SwissPro			IT	..	Thu Jul 8 01:57:29 2021	17	User read res	0	MIS	Yes		0	C:\temp\ITRAQ8plex6data\wfr

44

- MASCOT Serverとは
- 対応フォーマット・データ変換
- ネットワーク・検索方法
- 検索パラメーター
- 結果画面
- 同定タンパク質
- 管理プログラム・カスタマイズ
 - Database Status
 - Search log
 - Configuration Editor

- ・Home → Configuration Editor で開く画面
- ・MASCOT Serverで使用する各種設定変更・カスタマイズ

Mascot Configuration

Amino Acids	Amino Acid Data
Modifications	Modification definitions
Symbols	Symbols used in chemical formulae
Linkers	Linker definitions
Enzymes	Enzyme definitions
Instruments	Fragmentation Rules
Quantitation	Quantitation Methods
Crosslinking	Crosslinking Methods
Configuration Options	Global Options in mascot.dat
Database Manager	Sequence databases, Parse Rules and automated downloads

補足説明 : Configuration Editorの各項目とつかさどる内容

項目名	説明
Amino Acids	アミノ酸の質量。U,J,O に オリジナル設定も可
Modifications	修飾、アミノ酸の質量の変化
Symbols	元素や分子の質量に割り当てた記号(Symbol)とその質量を確認できる画面
Linkers	クロスリンクの設定に関連するLinker
Enzymes	タンパク質からペプチドに切断するパターン
Instruments	理論的なMS2フラグメントピーク作成の際、考慮するイオンシリーズ
Quantitation	タンパク質の定量解析
Crosslinking	クロスリンクペプチド検索
Configuration Options	MASCOT Server のオプション
Database Manager	MASCOTで使用しているデータベース

[補足] MIS検索

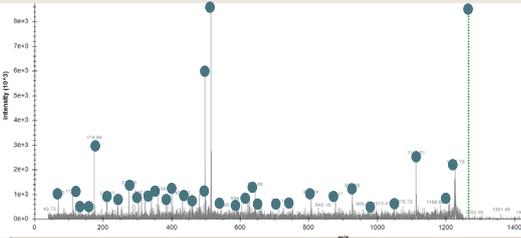
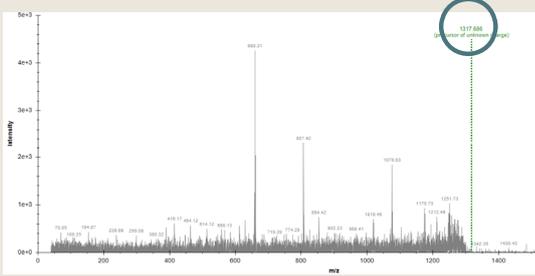
47

MISを実施して、「ペプチドを同定」したタンパク質を特定する

ピーク	OPSD_HUMAN	OPSD_PHOVI	OPSD_MACFA	OPSD_CRIGR
832.662	■	■	■	■
903.342	■	■	■	■
1186.439	■		■	■
1403.722	■		■	
1617.857	■		■	■
1727.916	■	■		
1743.951				
1759.966				■
1788.721				
1818.963	■	■		■
2159.143	■	■		
2174.812			■	
2256.871	■		■	■

1403.722と2159.143、2つのピークが確かに存在すると確認できれば、「OPSD_HUMAN」の存在も特定できる

MS1 : ペプチドの質量(m/zとcharge)



MS2 : ノイズをカットし、フラグメント由来のピークを選択

BEGIN IONS

TITLE=1: Scan 10 (rt=182.28)

PEPMASS=1403.722 123456

CHARGE=1+

566.70161 445127.71

734.56885 253205.82

1092.2019 445679.33

.....

1226.779 1569294.9

1347.8916 3574730.8

END IONS

パラメーターについては次スライド

① ピークリストファイルを準備
 ② ブラウザを開きMASCOT Serverへアクセス
 ③ Home -> Access MASCOT Server
 ④ 「MS/MS Ions Search」の「Perform Search」
 ⑤ 検索パラメータ並びに入力データを指定
 ⑥ Start Search ボタンで検索開始

MASCOT MS/MS Ions Search

(1) Your name: Taklaesu (2) Email: [input]
 (3) Search title: search test 01 20250301
 (4) Database(s): SwissProt (AA)
 (5) Taxonomy: All entries
 (6) Enzyme: Trypsin/P (7) Allow up to: 1 missed cleavages
 (8) Quantitation: None
 (9) Crosslinking: None
 (10) Fixed modifications: --- none selected ---
 (11) Variable modifications: --- none selected ---
 (12) Error tolerant: Automatic second pass search of selected modification classes
 (13) Peptide tol. ±: 1.2 Da #¹³C: 0 MS/MS tol. ±: 0.6 Da
 (14) (15)
 (16) Data file: ファイルを選択 (選択されていません)
 (17) Data format: Mascot generic
 (18) Instrument: Default
 (19) Target FDR: 1%
 (20) Machine learning: Refine results with machine learning (Percolator)
 (21) Use features calculated by Mascot
 (22) DeepLC model for retention times: full_hc_unmod_fixed_mods
 (23) MS2PIP model for spectral similarity: CID
 (24) Start Search ... Reset Form

- MASCOT Serverとは
- 検索方法
- 対応フォーマット・データ変換
- 検索パラメーター
 - 検索パラメーター一覧
 - カスタマイズ可能な項目
 - データベース
- 結果画面
- ペプチド同定
- タンパク質同定
- MIS 検索のオプション
- 管理プログラム・カスタマイズ

補足説明：検索パラメーター各項目について

No	項目名	説明	No	項目名	説明	No	項目名	説明
1	Your name	ユーザー名	10	Fixed modifications	修飾、対象のすべてのアミノ酸について指定した内容に質量が変更	19	Target FDR	同定基準となるFDR値 (q値)。通常はDefault値である1%
2	Email	電子メール。但しlocal版ではメモとして使用する事も可能	11	Variable modifications	修飾、修飾がつくパターンとつかないパターンの両方を考慮	20	Machine learning	機械学習によるrefinementを実施するかどうか (Percolator並びにMS2Rescoreの実施)
3	Search title	検索タイトル	12	Error tolerant	拡張2段階検索の実施	21	DeepLC model	保持時間予測の使用、鍛えたデータセット群(モデル)選択
4	Database	検索対象のデータベース	13	Peptide tol. ±	ペプチドの実測値と理論値の誤差範囲	22	MS2PIP model	MS2ピーク予測の使用、鍛えたデータセット群(モデル)の選択
5	Taxonomy	生物種絞り込み	14	# 13C	13Cを含むペプチドの補正	23	Start Search	検索の開始
6	Enzyme	タンパク質の切断パターン	15	MS/MS tol. ±	フラグメントの実測値と理論値の誤差範囲	24	Reset Form	パラメータ設定を初期値に戻す
7	Allow up to	Enzyme設定について、切断箇所と認定された箇所を何度まで見逃す事を許容するか	15	Peptide charge	[通常は使用されない]ピークリストに電荷情報がない場合に割り当てる電荷の情報			
8	Quantitation	タンパク質の定量計算	16	Data file	検索queryとなる入力データ			
9	Crosslinking	リンカー又は共有結合で結合したペプチドの検索	17	Data format	入力データのフォーマット			
			18	Instrument	理論フラグメント作成を行うイオンシリーズの定義セット			

好ましいデフォルト設定
→ 口頭で説明

Proteins (1) [Report Builder](#)

Protein family 1 (out of 1)

10 per page 1 [Expand all](#) [Collapse all](#)

Accession contains Find Clear

▼1 **OPSD_HUMAN** 295 Rhodopsin OS=Homo sapiens OX=9606 GN=RHO PE=1 SV=1

	Score	Mass	Matches	Sequences
1.1 OPSD_HUMAN	295	38866	2 (2)	2 (2) Rhodopsin OS=Homo sapiens OX=9606 GN=RHO PE=1 SV=1

▼2 peptide matches (2 non-[Protein view](#))

Auto-[Peptide view](#)

Query Dupes	Observed	Mr (expt)	Mr (calc)	Delta M	Score	Expect	Rank	U	Peptide
#1	1403.2000	1402.1927	1402.6576	-0.4649	0 146	1.9e-12	▶1	U	K.NPLGDDEASATVSK.T
#2	2159.2000	2158.1927	2158.1067	0.0860	1 189	5.1e-17	▶1	U	K.SAAIYNPVIYIMMNKQFR.N

・タンパク質にアサインされたペプチドそれぞれに、スコアや期待値が表示される
 ・**Expect(期待値)が0.05**以下のペプチドはMASCOTで**同定**できたものとみなしてよい

MASCOT Search Results

Protein View: OPSD_HUMAN

Rhodopsin OS=Homo sapiens OX=9606 GN=RHO PE=1 SV=1

Database: SwissProt
 Score: 295
 Monoisotopic mass (M₀): 38866
 Calculated pI: 6.21
 Taxonomy: Homo sapiens

Sequence similarity is available as an [NCBI BLAST search](#)

Search parameters

MS data file: opsdhuman2seq.txt
 Enzyme: Trypsin: cuts C-term side of KR unless ne

Protein sequence coverage: 9%

Matched peptides shown in **bold red**.

1 MDTGEGNFY VEPSSATGV RSPFEVFOY LAEPHQFSM
 51 GFTNIFLTV VVQHGKLR PLNYLNLIA VADLRYLVLR
 101 GYVFFGFTG NLEGFATLG GEIALMSLVV LAIERYVWC KFNSEFRGE
 151 NHAIVSWAF WQGLACAP FLAHSRYIP EGIQCSGID YTLKFEVNN
 201 ERFVYHFWV HTIPIHIF FCYGLVTV KEAAQQQS ATTQKAEKV
 251 IRWYHIVIA FLIWFVFA VAFIFFDG SNFFPFIPI FAFKASAAI
 301 IYVYIIMAN KQFRNKLIT ICGRNPFD DEASATVSKI ETSQVAP

Unformatted sequence string: **348 residues** (for pasting into other applications).

Sort by residue number increasing mass decreasing mass
 Show matched peptides only predicted peptides also

Query Start - End	Observed	Mr (expt)	Mr (calc)	Delta M	Score	Expect	Rank	U	Peptide
#2 297 - 314	2159.2000	2158.1927	2158.1067	0.0860	1 189	5.1e-17	1	U	K.SAAIYNPVIYIMMNKQFR.N
#1 326 - 339	1403.2000	1402.1927	1402.6576	-0.4649	0 146	1.9e-12	1	U	K.NPLGDDEASATVSK.T

MS error 236 ppm

ID OPSD_HUMAN Reviewed: 348 AA.
 AC P08100; Q16414; Q2M249;
 DT 01-AUG-1998, integrated into UniProtKB/Swiss-Prot.
 DI 01-AUG-1988, sequence version 1.
 DR 02-JUN-2021, entry version 227.
 DE Function: Full=Rhodopsin;
 DE AltName: Full=Opsin-2;
 GN Name=RHO; Synonyms=OPN2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
 OC Primate; Euarchonta; Primates; Haplorhina; Catarrhini; Hominoidea;

Protein view :
 Protein のAccessionを
 クリックすると開く。
 PMFと同じく同定タンパク質に
 関するより詳しい情報

Mascot Search Results

Peptide View

MS/MS Fragmentation of **NPLGDDEASATVSK**
 Found in **OPSD_HUMAN** in SwissProt, Rhodopsin OS=Homo sapiens OX=9606 GN=RHO PE=1 SV=1
 Match to Query 1: 1402.192724 from(1403.20000,14) index(0)
 Data file opsdhuman2seq.txt

Peptide view :
 Query番号をクリックすると開く。
 ピークリストと理論値のマッチングを
 確認可能

Label all possible matches Label matches used for scoring

#	a	a*	b	b*	Seq.	y	y*	#
1	87.0553	70.0287	115.0502	98.0237	N			14
2	184.1081	167.0615	212.1030	195.0764	P			13
3	297.1921	280.1658	325.1870	308.1605	L			12
4	354.2136	337.1870	382.2085	365.1819	G			11
5	469.2405	452.2140	497.2354	480.2089	D			10
6	584.2675	567.2409	612.2624	595.2358	D			9
7	713.3101	696.2835	741.3050	724.2784	E			8
8	784.3472	767.3206	812.3421	795.3155	A			7
9	871.3792	854.3527	899.3741	882.3476	S			6
10	942.4163	925.3898	970.4112	953.2847	A			5
11	1043.4640	1026.4374	1071.4589	1054.4324	T			4
12	1142.5324	1125.5059	1170.5273	1153.5008	V			3
13	1229.5644	1212.5379	1257.5594	1240.5328	S			2
14					K			1

MS error 0 ppm