

MASCOT Server Version 2.2

チュートリアル



マトリックスサイエンス株式会社 電子メール: support-jp@matrixscience.com 電 話:03-5807-7895 ファクシミリ:03-5807-7896 住 所:〒101-0021 東京都千代田区外神田 6-10-12 KNビル3 階

チュートリアル 目次

1.	Mascot Serverへのアクセス	
	1-1. Mascot Server トップページの表示	••••• 3
2.	PMF検索(Peptide Mass Fingerprint)	
	2-1. 検索条件設定画面へのアクセス	••••••
	2-2. 検索条件の種類	•••• 5
	2-3. 検索条件の設定例	••••• 6
	2-4. PMF 検索の実行	•••••
	2-5. 検索結果レポート	•••• 8
	2-6. 検索結果の出力	••••• 11
	2-7. 検索条件の変更と再検索	••••• 12
3.	MIS検索(MS/MS Ion Search)	
	3-1. 検索条件設定画面へのアクセス	••••• 13
	3-2. 検索条件の種類	••••• 13
	3-3. 検索条件の設定例	••••••• 15
	3-4. MIS 検索の実行	••••••• 15
	3-5. 検索結果レポート	••••••• 15
	3-6. 検索結果の出力	••••• 18
	3-7. 検索条件の変更と再検索	· · · · · · · · · 18
	3-8. Error Tolerant Search	••••• 19
4.	SQ検索(Sequence Query)	
	4-1. 検索条件設定画面へのアクセス	· · · · · · · · · · 21
	4-2. SQの構文と実行例	••••• 21
5.	Mascot Server の管理	
	5-1. 検索ログ	••••• 23
	5-2. 配列データベース管理画面	••••• 23
	5-3. 検索条件のデフォルト値の設定	•••••••• 24
	5-4. Mascot Server の設定値の変更	••••••• 25
	5-5. ヘルプへのアクセス	••••••• 26
	5-6. Mascot ファミリーに関する最新情報	••••• 28
6.	Mascot Daemon	
	6-1. インストール	· · · · · · · · · · 29
	6-2. チュートリアル	••••••• 29

7. 質量分析計との連携

	7-1.	Applied Biosystems	•	•	•	•	•	•	•	•	•	•	30
	7-2.	Waters	•	•	•	•	•	•	•	•	•	•	30
	7-3.	島津製作所	•	•	•	•	•	•	•	•	•	•	30
	7-4.	Thermo Fisher Scientific	•	•	•	•	•	•	•	•	•	•	30
	7-5.	日立ハイテクノロジーズ	•	•	•	•	•	•	•	•	•	•	31
	7-6.	Bruker Daltonics	•	•	•	•	•	•	•	•	•	•	31
8.	付録												
	8-1.	蛋白質同定のコツその1	•	•	•	•	•	•	•	•	•	•	32
	8-2.	蛋白質同定のコツその2	•	•	•	•	•	•	•	•	•	•	34

■ 1. Mascot Serverへのアクセス

Mascot Server はWebサーバ上に構築されたアプリケーションです。従いまして、Mascot Server とネットワーク接続されているコンピュータのWebブラウザからアクセスし、利用することができます。

1-1. Mascot Server トップページの表示

Webブラウザを起動し、アドレス入力欄にMascot Server 名(通常はコンピュータのホスト名)を含む次のURLを入力してください。

http://ホスト名/mascot/

Mascot Server に接続され、図1に示すトップページが表示されます。Mascot Server はひとつの独立したWebサイト になっており、トップページから Mascot Server が持つ様々なリソースにアクセスすることができます。

図1 M	ascot	Server	Ø	トッ	プペー	ジ
------	-------	--------	---	----	-----	---

Matrix Science - Home - Microsoft Internet Explorer	Mascot Server に関連したヘルプページにリンクして
	います。
	Mascot Server サイト内をキーワート検索することが
	(CEX9)
(MATRIX) (SCIENCE) HOME (MASCOT HELP) Search Go	
Home	次の Mascot 検索モード選択画面にリンクしています。
	1. Peptide Mass Fingerprint (PMF 検索)
Welcome	2. Sequence Query (SQ 検索)
This site featured Macont annwerful cearch engine that uses mass spectrometry data to identify proteins	3. MS/MS Ion Search (MIS 検索)
from primary sequence databases. To assist you, the <u>help text</u> for Mascot forms a substantial knowledge	
	Mascot Server を利用する上で必要なフラワサの機能
It this is your first visit, please check for proviser compatibility and read the small print. If you include results from Mascot in a publication, please cite either www.matrixscience.com or Electrophoresis, 20(18) 3551-67	(Java Script、Table、ファイルの読込、Gookle)を エーックナスニトがホキナナ
(1999) (abstract).	チェックすることができます。
Security	
Mascot incorporates a role based security model, which allows the system administrator to control which functionality is available to individual users. If security has been enabled, you may be prepared to be in if you	
attempt to access a page or perform a task that requires authorisation. If security is not enabled, which is the default the following links will simply confirm this	セキュリティ機能が有効な時、ログイン、ログアウト、
	パスワードの変更、登録内容の変更、現在のセッショ
Log in Log out	ン情報の確認、ユーザ権限の設定などを行う画面にリ
Change password Edit settings	ンクしています。
Current session information	
Whether or not Mascot security is enabled, in order to comply with the licensing conditions, access to this Mascot server must be restricted to authorised users. We strongly recommend that your server is protected by a secure firewall. For more information, start from the WWW Security FAQ.	
Mascot I Itilities	Mascot Server が利用する配列データベースの管理画
Muscot Guildes	面にリンクしています。
Your system administrator may have restricted access to these utilities	
Database Status View the status of all the sequence databases. Links to database statistics, search status, log files, etc.	Mascot 検索ログの画面にリングしています。
Search Log A tabular view of the search log. Can be filtered to find specific search results.	Massat Sarvar の記字店を答理するためのコーザイン
Configuration Editor Browser-based configuration editors, including security administration and sequence database	Mascoll Server の設定値を管理するためのユーサイン
management.	
Mascot Daemon	
Mascot Daemon is a client application which automates the submission of data files to a Mascot server. It runs	Mascot Daemon のインストールおよびアップグレード
Install Upgrade Troubleshoot	を行うためのページにリンクしています。
Mascot Documentation	
Help on using Mascot and understanding the results is contained in these HTML pages. Click on Help at the	Acrobat Reader がインストールされている場合は
mainly of interest to system administrators.	「Setup & Installation」マニュアルが表示されます。
Setup & Installation Manual	
Release Notes	Acrobat Reader がインストールされている場合は「リ
	リースノート」が表示されます。
Copyright of 2007 Wark's Science Los. All Right's Reserved. Last Oppared 03/27/2007 USIS2145	
3 1212491	

■ 2. PMF検索(Peptide Mass Fingerprint)

PMF 検索では、指定した配列データベースの中から、MS データ(蛋白質を酵素消化して得られたペプチドの 質量データのセット)にマッチする蛋白質を検索し、測定サンプルに含まれる蛋白質を同定します。

2-1 検索条件設定画面へのアクセス

Web ブラウザを起動し、アドレス入力欄に Mascot Server 名を含む次のURLを入力してください。

http://ホスト名/mascot/

Mascot Server に接続され、図1に示すトップページが表示されます。トップページ画面の Welcome の次 の行にある Mascot のリンク、またはページ最上段の HOME | MASCOT | HELP の中の MASCOT のリンクを クリックしてください。図2に示す Mascot Server の検索モード選択画面 (Mascot Search 画面) が表示され ます。

図2検索モード選択画面



Mascot Server は次の3つの検索モードをサポートしています。

- (1) **Peptide Mass Fingerprint**(PMF検索: MSデータを使って蛋白質を同定することができます)
- (2) Sequence Query (SQ検索: 配列タグ情報を利用して蛋白質を同定することができます)
- (3) MS/MS Ion Search (MIS検索: MS/MSデータを使って蛋白質を同定することができます)

図2に示す Mascot Search 画面の一番上の行にある Peptide Mass Fingerprint のリンクをクリックしてください。図3に示す PMF 検索検索条件設定画面が表示されます。

図3 PMF 検索の条件設定画面



2-2. 検索条件の種類

PMF 検索では、次の表に示す検索条件を入力または選択することができます。

質量データは Query ウインドウに数値を入力しますが、ファイルとして存在する場合は Data file の [参照] ボタンから直接読み込むこともできます。修飾は最大9個まで同時選択することができます。複数の修飾を選 択する場合および選択した修飾を解除する場合はキーボードの [Ctrl] キーを押しながらマウスでクリックして ください。

なお、各検索条件項目は Mascot Server のWebサイト内にあるヘルプページにリンクしており、その検索条件に関する意味や使い方をオンラインで参照することができます。

検索条件名	値と内容	設定例
Your name	名前(ユーザ名)を入力してください。	usami
Email	Emailアドレスを入力してください。	support-jp@matrixscience.com
	検索タイトルを入力してください。後日、	samplel, spot67, Trypsin,
Secret title	この検索結果にアクセスしやすくするた	50fm, 2457-9001-125-afd
Search title	めにできるだけ詳しい内容を入力するこ	
	とをお勧めします。	
	複数の配列データベースがセットアップ	SwissProt
Database	されている場合は、検索したい配列データ	
	ベースを選択します。	
Taxonomy	検索対象としたい生物種を選択します。	All entries
	蛋白質を消化するために使用した消化酵	Trypsin
Enzyme	素を選択します。必要であればユーザ定義	
	の消化酵素も定義することができます。	
Allowed on the [1 0] missed	ペプチドの消化状態を指定します。完全に	1
Allowed up to [1-9] missed	消化されていると考えられる場合は"0"	
cieavages	(ゼロ)を選択してください。	
Fixed modifications	修飾を選択します。指定したすべての(Fix	Carboxymethyl (C)

	された) アミノ酸に対して修飾を考慮して	
	検索を行います。システインの誘導体化の	
	ような意図された化学修飾に対応します。	
	修飾を選択します。指定したアミノ酸に対	Acetyl (K)
	して修飾を考慮する場合としない場合の	Oxidation (M)
Variable modifications	2通りの(Variableな)検索を行います。	
	サンプルの前処理の段階で発生する偶発	
	的な修飾や翻訳後修飾に対応します。	
	検索範囲ウインドウを指定します。たとえ	30
	ば、″30″を指定した場合は30KDaの範囲に	
Protein mass	ある連続したアミノ酸配列を対象として	
	検索を行います。	
	ペプチド質量の許容誤差をDa、mmu、%、ppm	1Da
	の単位で指定します。1.0 Daを指定した場	
Peptide tol. +/-	合は (ペプチド質量-1.0) Daから (ペプチ	
	ド質量+1.0) Daの範囲でヒットするペプチ	
	ドを検索します。	
	ペプチドの電荷状態 (MH+, Mr, M-H ⁻)を	MH+
Mass values	指定します。	
Monoisotopic/Average	質量データの種類 (Monoisotopicまたは	Monoisotopic
	Average)を指定します。	-
	読み込みたい質量データファイルを[ファ	データファイルを指定
Data file	イルの選択]ダイアログを通じて指定しま	
	+	
	y .	
	9 o	594.5
	У о	594. 5 649. 5
	У о.	594. 5 649. 5 693. 3
		594. 5 649. 5 693. 3 756. 7
	質量データを入力します。	594. 5 649. 5 693. 3 756. 7 998. 8
	す。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使	594. 5 649. 5 693. 3 756. 7 998. 8 1060. 8
Query	す。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽	594. 5 649. 5 693. 3 756. 7 998. 8 1060. 8 1079. 8
Query	す。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア	594. 5 649. 5 693. 3 756. 7 998. 8 1060. 8 1079. 8 1116. 9
Query	 「質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア や、Mascot Distillerなどを利用してくだ 	594.5 649.5 693.3 756.7 998.8 1060.8 1079.8 1116.9 1365.9
Query	す。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア や、Mascot Distillerなどを利用してくだ さい。	594. 5 649. 5 693. 3 756. 7 998. 8 1060. 8 1079. 8 1116. 9 1365. 9 1726. 1
Query	す。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア や、Mascot Distillerなどを利用してくだ さい。	594. 5 649. 5 693. 3 756. 7 998. 8 1060. 8 1079. 8 1116. 9 1365. 9 1726. 1 1748. 1
Query	す。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア や、Mascot Distillerなどを利用してくだ さい。	594. 5 649. 5 693. 3 756. 7 998. 8 1060. 8 1079. 8 1116. 9 1365. 9 1726. 1 1748. 1 1851. 2
Query	質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア や、Mascot Distillerなどを利用してくだ さい。	594. 5 649. 5 693. 3 756. 7 998. 8 1060. 8 1079. 8 1116. 9 1365. 9 1726. 1 1748. 1 1851. 2 2561. 8
Query	す。	594.5 649.5 693.3 756.7 998.8 1060.8 1079.8 1116.9 1365.9 1726.1 1748.1 1851.2 2561.8 ボックスをチェック
Query	 す。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア や、Mascot Distillerなどを利用してくだ さい。 チェックボックスをチェックすることに 上り Decovデータベースの同時検索を実 	594.5 649.5 693.3 756.7 998.8 1060.8 1079.8 1116.9 1365.9 1726.1 1748.1 1851.2 2561.8 $\vec{x} \cdot y \cdot j \cdot x \cdot \hat{x} \cdot y \cdot j$
Query Decoy	 「す。 「質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア や、Mascot Distillerなどを利用してくだ さい。 チェックボックスをチェックすることに より、Decoyデータベースの同時検索を実 行します。 	594.5 649.5 693.3 756.7 998.8 1060.8 1079.8 1116.9 1365.9 1726.1 1748.1 1851.2 2561.8 ボックスをチェック
Query Decoy	 す。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使用します。質量データからこれらの値を抽出するには、各装置の解析ソフトウエアや、Mascot Distillerなどを利用してください。 チェックボックスをチェックすることにより、Decoyデータベースの同時検索を実行します。 検索結果レポートに表示する蛋白質の数 	594.5 649.5 693.3 756.7 998.8 1060.8 1079.8 1116.9 1365.9 1726.1 1748.1 1851.2 2561.8 ボックスをチェック
	 質量データを入力します。	594. 5 649. 5 693. 3 756. 7 998. 8
Query	す。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア や、Mascot Distillerなどを利用してくだ さい。	594. 5 649. 5 693. 3 756. 7 998. 8 1060. 8 1079. 8 1116. 9 1365. 9 1726. 1 1748. 1 1851. 2 2561. 8
Query	 「す。 「雪量データを入力します。 PMF検索ではモノアイソトピック質量を使用します。 質量データからこれらの値を抽出するには、各装置の解析ソフトウエアや、Mascot Distillerなどを利用してください。 チェックボックスをチェックすることに 	594.5 649.5 693.3 756.7 998.8 1060.8 1079.8 1116.9 1365.9 1726.1 1748.1 1851.2 2561.8 ボックスをチェック
Query Decoy	 「す。 「質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア や、Mascot Distillerなどを利用してくだ さい。 チェックボックスをチェックすることに より、Decoyデータベースの同時検索を実 行します。 	594.5 649.5 693.3 756.7 998.8 1060.8 1079.8 1116.9 1365.9 1726.1 1748.1 1851.2 2561.8 ボックスをチェック
Query Decoy	 ダ。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使 用します。質量データからこれらの値を抽 出するには、各装置の解析ソフトウエア や、Mascot Distillerなどを利用してくだ さい。 チェックボックスをチェックすることに より、Decoyデータベースの同時検索を実 行します。 体索結果レポートに表示する蛋白質の教 	594.5 649.5 693.3 756.7 998.8 1060.8 1079.8 1116.9 1365.9 1726.1 1748.1 1851.2 2561.8 ボックスをチェック
Query Decoy Report top	 す。 質量データを入力します。 PMF検索ではモノアイソトピック質量を使用します。質量データからこれらの値を抽出するには、各装置の解析ソフトウエアや、Mascot Distillerなどを利用してください。 チェックボックスをチェックすることにより、Decoyデータベースの同時検索を実行します。 検索結果レポートに表示する蛋白質の数を選択します。 	594.5 649.5 693.3 756.7 998.8 1060.8 1079.8 1116.9 1365.9 1726.1 1748.1 1851.2 2561.8 ボックスをチェック

2-3. 検索条件の設定例

つぎの検索条件を設定してください。

検索条件	設定値	摘要
Your name	任意	ユーザ名をアルファベットで入力
Search title	PMF test search # 1	Eメールアドレスをアルファベットで入力
Database	SwissProt	
Taxonomy	All entries	
Enzyme	Trypsin	
Allowed up to [1-9] missed	1	
cleavages	1	

MASCOT Server : f_{1} - h_{1} - h_{2} - h_{2}

Peptide tol. +/-	1.0 Da	
Mass values	MH+	
Monoisotopic/Average	Monoisotopic	
	594. 5	数値を入力する代わりに、Data file の [参照]
	649.5	ボタンから直接"mixtureSample.asc"データファ
	693. 3	イル (Mascot のインストール CD の中の
	756.7	"mskk¥sampledata"フォルダまたは
	998.8	"C:¥inetpub¥mascot¥mskk¥sampledata"フォル
	1060.8	ダにあります)を読み込むことができます。
Query	1079.8	"mixtureSample.asc"データファイルには、
	1116.9	594. 5, 649. 5, 693. 3, 756. 7, 998. 8, 1060. 8,
	1365.9	1079.8、1116.9、1365.9、1726.1、1748.1、
	1726.1	1851.2、2561.8の質量値が含まれています。
	1748.1	
	1851.2	
	2561.8	
Decoy	チェック無し	
Report top	AUTO	

2-4. PMF 検索の実行

検索条件設定画面左下にある [Start Search ...] ボタンを押してください。Mascot 検索が始まり、検索の進捗 状況が%表示されます。

図4 PMF 検索の実行と進捗状況表示



検索時間(検索に要する時間)は、検索に使用した配列データベースのサイズ(配列データベースを構成す る蛋白質エントリの総数)、missed cleavagesの設定値、Variable modificationsの設定値、Peptide tol +/-の設 定値などにより異なります。これらの設定値が大きくなるほど検索空間が広がるために、検索時間は長くなり ます。

2-5. 検索結果レポート

PMF検索が終了すると、検索結果を示す Mascot Search Results 画面が表示されます。Mascot Search Results 画面には、図5の Concise Protein Summary Report 画面、図6の Protein Summary Report 画面の2つの表示モ ードがあり、[Fomat As] ボタンの右にあるプルダウンメニューから選択することができます。最初に表示され るのは図5の Concise Protein Summary Report 画面です。

図5 Concise Protein Summary Report 画面

4555 150 5000 740

alf 2015 Backweit (4 A. geboon (2003) gebon Understern, allegened into Kastatik Dase-Fail, 01 Mil 157, majore werden 1 Understern, majore werden 1 Understerne gebonen,

SARASSAR'S



図 5 の Concise Protein Summary Report 画面では、ヒットした蛋白質をプロテインスコア順に表示します が、各ヒットランクの代表蛋白質に帰属する質量データのうち、すべてまたは一部分の質量データを持つ蛋白 質をひとつのグループにまとめて表示します。たとえば、ヒットランク第1位の、有意なプロテインスコア 201 を 持つ Mixture 1 には(594.5、649.5、693.3、756.7、998.8、1060.8、1079.8、1116.9、1365.9、1726.1、1748.1、1851.2、 2561.8)の質量データが帰属しますが、プロテインスコア 80 を持つ RS14Z_MYCPA は Mixture 1 に帰属する質量デ ータの一部分に当たる (594.5、1060.8、1079.8、1365.9、1726.1、1851.2)の質量データが帰属しています。

- Mixture 1 Total score: 201 Expect: 2.1e-015 Queries matched: 13 Components (only one family member shown for each component): LAP_BOVIN Mass: 7037 Score: 99 Expect: 3.3e-005 Queries matched: 7 Lingual antimicrobial peptide precursor - Bos taurus (Bovine) RS14Z_MYCPA Mass: 6889 Score: 80 Expect: 0.003 Queries matched: 6 30S ribosomal protein S14 type Z - Mycobacterium paratuberculosis
- LAP_BOVIN 2 Mass: 7037 Score: 99 Expect: 3.3e-005 Queries matched: 7 Lingual antimicrobial peptide precursor - Bos taurus (Bovine) LAP BUBBU Mass: 7037 Score: 99 Expect: 3.3e-005 Queries matched: 7 Lingual antimicrobial peptide precursor - Bubalus bubalis (Domestic water buffalo) KV112 HUMAN Mass: 12119 Score: 42 Expect: 17 Queries matched: 4 lg kappa chain V-I region Kue - Homo sapiens (Human) (省略)
- 3 RS14Z_MYCLE Mass: 6892 Score: 95 Expect: 8.8e-005 Queries matched: 7 30S ribosomal protein S14 type Z - Mycobacterium leprae Score: 80 RS14Z_MYCPA Mass: 6889 Expect: 0.003 Queries matched: 6 30S ribosomal protein S14 type Z - Mycobacterium paratuberculosis RS14Z_MYCSS Mass: 6984 Score: 65 Expect: 0.09 Queries matched: 5 30S ribosomal protein S14 type Z - Mycobacterium sp. (strain MCS) (省略)
- RK14 COFAR Expect: 15 Queries matched: 4 4 Mass: 13568 Score: 42 Chloroplast 50S ribosomal protein L14 - Coffea arabica (Coffee) RK14_NICSY Mass: 13569 Score: 42 Expect: 15 Queries matched: 4 Chloroplast 50S ribosomal protein L14 - Nicotiana sylvestris (Wood tobacco) RK14_SOLBU Mass: 13569 Score: 42 Expect: 15 Queries matched: 4 Chloroplast 50S ribosomal protein L14 - Solanum bulbocastanum (Wild potato) RK14_SOLLC Mass: 13569 Score: 42 Expect: 15 Queries matched: 4 Chloroplast 50S ribosomal protein L14 - Solanum lycopersicum (Tomato) (Lycopersicon esc (省略)

ページの最後にはこの検索で使用した検索条件(Search Parameters)が表示されます。

Search Parameters		
Type of search	:	Peptide Mass Fingerprint
Enzyme	:	Trypsin
Mass values	:	Monoisotopic
Protein Mass	:	Unrestricted
Peptide Mass Tolerance	:	± 1 Da
Peptide Charge State	:	1+
Max Missed Cleavages	:	1
Number of queries	:	13

図6の Protein Summary Report 画面では、ヒットした蛋白質をプロテインスコア順に表示します。また、 ヒットした各々の蛋白質に関して、帰属している質量データ(Observed/Mr(expt))とその理論値(Mr(calc)) および差分(Delta)、質量データに対応するペプチドのアミノ酸配列(Peptide)と蛋白質内の位置(Start/ End)、未切断サイト数(Miss)などの情報を表示します。Concise Prptein Summary Report 画面のように、 蛋白質グループにはまとめません。

	Accession	n Ma	ass Score	Descrip	otion					
1.	Mixture	1	201	LAP_BO\	/IN +	RS	14Z_	MYCPA		
2.	LAP_BOVII	N 70	037 99	Lingual	ant	imi	crob	oial pep [.]	tide pre	ecursor - Bos taurus (Bovine)
3.	LAP_BUBB	<mark>U</mark> 7(037 99	Lingual	ant	imi	crob	oial pep [.]	tide pre	ecursor - Bubalus bubalis (Domestic water
4.	RS14Z_MY	CLE 68	392 95	30S rib	osoma	al	prot	ein S14	type Z	- Mycobacterium leprae
5.	RS14Z_MY	CPA 68	389 80	30Sribo	osoma	lр	rote	in S14	type Z -	- Mycobacterium paratuberculosis
6.	RS14Z_MY	CSS 69	984 65	30Sribo	osoma	lр	rote	ein S14	type Z -	– Mycobacterium sp. (strain MCS)
Re	sults Lis [.]	t				~ ~				
1.	Mixture 1	lota	score: 201	Expect:	2.1e-	-01	5 (lueries i	matched:	13
	Component	s∶1. LAH	P_BOVIN Lingua	il antimicr	obia	l p	epti	de preci	ursor –	Bos taurus (Bovine)
		2. RS1	14Z_MYCPA 30S	ribosomal	prote	ein	\$14	type Z	- Mycob	pacterium paratuberculosis
	Observed	Mr(exp	t) Mr(calc)	Delta	Star	t	End	Miss	Comp	Peptide
	594. 5000	593.492	27 594. 2948	-0. 8021	37	-	41	0	2	K. FGLCR. I
	649. 5000	648.492	649. 2676	-0.7749	43	-	48	0	1	R. CPGSMR. Q
	693. 3000	692.292	693. 2864	-0. 9937	27	-	32	0	1	R. NSQSCR. R
	756. 7000	755. 692	27 756. 4316	-0. 7389	36	-	42	0	1	K. GICVPIR. C
	998. 8000	997.792	998. 5695	-0. 7768	34	-	42	1	1	R. NKGI CVP I R. C
1	060. 8000	1059.792	27 1060. 5712	-0. 7785	27	-	35	1	2	K. CGRPRAVFR. K
1	079.8000	1078.792	27 1079. 5732	-0. 7805	37	-	45	1	2	K. FGLCRICLR. E
1	116.9000	1115.892	27 1116. 5961	-0. 7034	49	-	59	0	1	R. QIGTCLGAQVK. C
1	365. 9000	1364.892	27 1365. 6711	-0. 7783	46	-	58	0	2	R. EMAHAGELPGVQK. S
1	726. 1000	1725.092	27 1725. 8144	-0. 7217	46	-	61	1	2	R. EMAHAGELPGVQKSSW
1	748. 1000	1747.092	27 1747.8532	-0. 7604	43	_	59	1	1	R. CPGSMRQIGTCLGAQVK. C
1	851.2000	1850. 192	1850. 9495	-0. 7568	42	-	58	1	2	R. ICLREMAHAGELPGVQK. S
2	561.8000	2560. 792	27 2561. 4846	-0. 6919	3	-	26	0	1	R. LHHLLLALLFLVLSAGSGFTQGVR. N
0		M	- 7007 - 6		F			2- 00F	0	
Ζ.	LAP_DOVIN Lingual a	mass ntimicrol	ial peptide r	pre: 99 precursor -	- Bos	ta	. ວ. urus	Se-005 (Bovin	oueries e)	
	Observed	Mr (exp	t) Mir(calc)	Delta	Star	t	Fnd	Miss	Peptide	
	649 5000	648 492	27 649 2676	-0 7749	43	_	48	0	R CPGSN	/ IR Q
	693 3000	692 292	27 693 2864	-0 9937	27	_	32	0	R NSQSC	ж. R
	756 7000	755 692	27 756 4316	-0 7389	36	_	42	0	K GICVE	
	998 8000	997 792	008 5695	-0 7768	34	_	42	1		
1	116 9000	1115 804	27 1116 5061	-0 7034	<u>4</u> 0	_	50	0		CI GAOVK C
'				V / / / / / H	- TV		00	~		
1	748 1000	1747 003	7 1747 8532	-0 7604	43	_	50	1		

No match to: 594.5000, 1060.8000, 1079.8000, 1365.9000, 1726.1000, 1851.2000 No match to: 1998.00, 3031.80

蛋白質のアクセッション番号(その蛋白質のユニークなID)のリンクをクリックすると Protein View 画面が 表示されます。Protein View 画面では、その蛋白質の詳細情報(全体アミノ酸配列とヒットした質量データが カバーしているアミノ酸配列部分、等電点理論値、アノテーション情報、NCBIサイトへのリンクなど)を表示 します。

Protein View

Match to: LAP_BOVINE Score: 99 Expect: 3.3e-005

Lingual antimicrobial peptide precursor - Bos Taurus (Bovine)

Nominal mass (Mr): 7037; Calculated pl value: 11.25 NCBI BLAST search of LAP_BOVIN against nr Unformatted sequence string for pasting into other applications

Taxonomy: Bos taurus

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P Number of mass values searched: 13 Number of mass values matched: 7 Sequence Coverage: 87%

Matched peptides shown in Bold Red

```
1 MRLHHLLLAL LFLVLSAGSG FTQGVRNSQS CRRNKGICVP IRCPGSMRQI
51 GTCLGAQVKC CRRK
```

Start – End	Observed Mr	(expt) Mr (ca	alc) Delta	Miss Sequ	ence	
3 - 26	2561.8000	2560. 7927	2561. 4846	-0. 6919	0	R. LHHLLLALLFLVLSAGSGFTQGVR. N
27 - 32	693, 3000	692. 2927	693. 2864	-0. 9937	0	R. NSQSCR. R
34 - 42	998. 8000	997. 7927	998. 5695	-0. 7768	1	R. NKGICVPIR. C
36 - 42	756. 7000	755. 6927	756. 4316	-0. 7389	0	K. GICVPIR. C
43 - 48	649. 5000	648. 4927	649. 2676	-0. 7749	0	R. CPGSMR. Q
43 - 59	1748. 1000	1747. 0927	1747. 8532	-0. 7604	1	R. CPGSMRQIGTCLGAQVK. C
49 - 59	1116. 9000	1115. 8927	1116. 5961	-0. 7034	0	R. QIGTCLGAQVK. C

No match to: 594.5000, 1060.8000, 1079.8000, 1365.9000, 1726.1000, 1851.2000

また、質量データ(実験値)とその理論値との質量差をプロットした質量誤差グラフは、質量分析計固有の キャリブレーションカーブ(質量校正曲線)に対応します。質量誤差グラフと実際のキャリブレーションカー ブが異なる場合は、この蛋白質は擬陽性(False Positive)の可能性がありますので、質量データや検索条件 を吟味する必要があります。



2-6. 検索結果の出力

検索結果の内容を XML または CSV 形式のファイルに出力することができます。図8に示すように、[Format As] ボタンの右にあるプルダウンメニューから [Export Search Results] を選択し、[Format As] ボタンを押し てください。



Protein Summary Report



図9に示すような出力設定画面が表示されます。[Export format] のプルダウンメニューから出力したいファイル形式 (XMLまたはCSV) を選択した後、出力する蛋白質を絞るための数値 ([Signifucance threshold p ζ , [Max. number of hits]など)を入力し、画面最下方にある [Export search results] ボタンを押してください。出力ファイルは C:¥temp フォルダに検索IDを持つファイル名 (F001234. csv や F001234. xml) で保存されます。

Matrix Science - Mascot - Export s	earch results - Microsoft Inter	net Explorer	
ファイル(E) 編集(E) 表示(V) お気に入り(A) ツール(工) ヘルプ(田)		
🔾 🖏 🔹 🕑 🖌 🗾 🕄 🎸	検索 🥎 お気に入り 🚱 🔗・	🎍 🛛 • 📒 🖗	
: アドレス(1)) 🍓 http://saki/mascot/cgi/export_c	at_2.pl?file=%2Fdata%2F20070601%2FF00	2509.dat&REPTYPE=export	&_sigt 🔽 🔁 移動 🕴 リンク 🎽
Google G - ✔ 校索 ♦	🧔 👺 🗸 🟠 ブックマークマ 🛃 ポッ	プアップを許可 🏾 🍄 チェック	▼ 》 🔘 設定▼ 🔹 🏂 ▼
{MATRIX } {SCIENCE		HOME MASCOT HELP	Search Go
Mascot > Export search results			
Export search results	Help		
Export format	XML 🚩		
Significance threshold p<	0.05		
Max. number of hits	AUTO		
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)	0		
Search Information	V		
Header	✓		
Variable mod. info.	✓		
Search parameters	✓		
Format parameters			~
🕘 ページが表示されました			😌 イントラネット 🛛 📑

図9 検索結果出力設定画面

2-7. 検索条件の変更と再検索

図5の Concise Protein Summary Report 画面あるいは図6の Protein Summary Report 画面の [Re-Search All] または [Search Unmatched] ボタンを押すことにより、PMF検索条件設定画面に戻りますので、検索条件を 変更して再検索することができます。

■■■ 3. MIS検索 (MS/MS Ion Search)

MIS 検索では、指定した配列データベースの中から、MS/MS データ(蛋白質を酵素消化して得られたペプチドに由来するフラグメントイオン質量データのセット)にマッチするペプチドを検索し、マッチしたペプチドが帰属する蛋白質を整理することにより、サンプルに含まれる蛋白質を同定します。

3-1. 検索条件設定画面へのアクセス

Webブラウザを起動し、アドレス入力欄に Mascot Server 名を含む次のURLを入力してください。

http://ホスト名/mascot/

Mascot Server に接続され、図1に示すトップページ が表示されます。トップページ画面の Welcome の次の 行にある Mascot のリンク、またはページ最上段の HOME | MASCOT | HELP の中の MASCOT のリンクをク リックしてください。図2に示す Mascot Server の検 索モード選択画面 (Mascot Search 画面) が表示されま す。

図2の Mascot Search 画面の3番目にある MS/MS Ion Search のリンクをクリックしてください。図10に 示す MIS 検索のパラメータ設定画面が表示されます。 PMF 検索パラメータ設定画面と同様に、各パラメータ項

図 10 MIS 検索の条件設定画面

🗿 Natrix Scienc	s - Mascol - MS/MS Tons Se	arch – Nicrosoft Interr	net Explorer 📃 🗖 🔀
ファイル(日) 編集(D 表示(V) お気に入り(A) ツール	(D) ~117(H)	12
G #8 · 🖸) - 🖹 🗟 🏠 🔎 🗮 🧯	📩 🐼 (V.S.Wes	• 🎍 🖬 • 🛄 🍋
PEDROD AL HOLD	//saki/mascot/cgi/search_formpl?FC	RMVER=28SEARCH=MIS	🛩 🛃 154b 955 🕬
Google C+	💌 検索 🕫 💕	 	57アップを許可 🎽 🔘 決定・ 抱・
(MATRIX)		HOW	
tornaters	Are ton count		
Mascot > Ms	/MS Ions Search		
MASCOT	MS/MS lons Sear	ch	
Your name	Itaru Usami 👻	Email	usami@matrixscience.com 👱
Search title			
Database	Sprot 💌		
Taxonomy	All entries		*
Enzyme	Trypsin 💌	Allow up to	1 missed cleavages
Fixed modifications	Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term) Amidated (Protein C-term)	Variable modifications	Acetyl (K) Acetyl (N-berm) Acetyl (Protein N-term) Amdated (C-term) Amdated (Protein C-term)
Quantitation	None	~	
Peptide tol. ±	2 Da 💌 # 13C 0	MS/MS tol. ±	0.8 Da 💌
Peptide charge	1+ 💌	Monoisotopic	⊙ Average ○
Data file		李照	
Data format	Mascot generic	Precursor	m/z
Instrument	ESI-TRAP	Error tolerant	
Decoy	0	Report top	AUTO Mits
	Start Search		Reset Form
	Copyright @ 2006 M	atrix Science Ltd. All Rights	Received.
6)			くいたうネット
			-

目はヘルプ書類にリンクしていますので、その使い方や意味をオンラインで参照することができます。

3-2. 検索条件の種類

MIS検索では、次の表に示す検索パラメータを入力または選択することができます。

質量データファイルは Data file の [参照] ボタンから指定して読み込みます。修飾は最大9個まで同時選択するこ とができます。複数の修飾を選択する場合や選択した修飾を解除する場合はキーボードの [Ctrl] キーを押しながらマ ウスでクリックしてください。

検索条件	値と内容	設定例
Your name	名前(ユーザ名)を入力してください。	usami
Email	Emailアドレスを入力してください。	support-jp@matrixscience.com
	検索タイトルを入力してください。後日、	sample1, spot67, Trypsin,
Soorah titla	この検索結果にアクセスしやすくするた	50fm, 2457-9001-125-afd
Search title	めにできるだけ詳しい内容を入力するこ	
	とをお勧めします。	
	複数の配列データベースがセットアップ	SwissProt
Database	されている場合は、検索したい配列データ	
	ベースを選択します。	
Taxonomy	検索対象としたい生物種を選択します。	All entries
	蛋白質を消化するために使用した消化酵	Trypsin
Enzyme	素を選択します。必要であれば消化酵素を	
	定義することができます。	
Allowed up to [1-9] missed	ペプチドの消化状態を指定します。完全に	1

cleavages	消化されていると考えられる場合は"0"	
	(ゼロ)を選択してください。	
	修飾を選択します。指定したすべての(Fix	Carboxymethyl (C)
Dine d me diffication a	された) アミノ酸に対して修飾を考慮して	
Fixed modifications	検索を行います。システインの誘導体化の	
	ような意図された化学修飾に対応します。	
	修飾を選択します。指定したアミノ酸に対	Acetyl (K)
	して修飾を考慮する場合としない場合の	Oxidation (M)
Variable modifications	2通りの(Variableな)検索を行います。	
	サンプルの前処理の段階で発生する偶発	
	的な修飾や翻訳後修飾に対応します。	
	サンプルに含まれる蛋白質の量を解析す	iTRAQ 4plex
	るために使用した試薬あるいは手法名を	
Quantitation	選択します。最後に[MD]がついているエン	
	トリは Mascot Distiller と連動して動作	
	するようになっています。	
	¹³ C同位体ピークを検出した可能性がある	1
11100	場合に選択します。[Peptide tol. +/-]	
#13C	で指定した範囲に加え、プラス側に1また	
	は2Daの範囲を拡張して検索します。	
	プリカーサイオン質量の許容誤差をDa、	1Da
	mmu、%、ppmの単位で指定します。1.0 Da	
Peptide tol. +/-	を指定した場合は (ペプチド質量-1.0) Da	
1	から (ペプチド質量+1.0) Daの範囲でヒッ	
	トするペプチドを検索します。	
	フラッグメントイオン質量の許容誤差を	0.5Da
MS/MS tol. +/-	Da、mmuの単位で指定します。	
	プリカーサイオンの電荷の値を指定しま	2+
Peptide charge	す。データファイルに電荷の値が記載され	
	ている場合は無視されます。	
Monoisotopic/Average	質量データの種類(Monoisotopicまたは	Monoisotopic
	Average) を指定します。	-
	読み込みたい質量データファイルを「ファ	データファイルを指定
Data file	イルの選択]ダイアログを通じて指定しま	
Data format	質量データのフォーマットを選択します。	Mascot generic
	プリカーサイオンの質量値を指定します。	ブランク
	データファイルにプリカーサイオンの質	
Precursor	量の値が記載されている場合は無視され	
	ます。	
	プリカーサイオンの開裂様式(発生するフ	EST-QUAD-TOF
	ラグメントイオンの種類)は装置によって	
Instrument	異なるため、測定に用いた質量分析計の種	
	類を選択します。	
	チェックボックスをチェックすることに	チェック
Decov	より、Decovデータベースの同時検索を実	
	行します。	
	チェックボックスをチェックすることに	チェック
Error tokerant	より、自動的にError tolerant検索を実行	
	します。	
	検索によりヒットした蛋白質の表示数を	20
Report top	選択します。	

3-3. 検索条件の設定例

次の検索パラメータを設定してください。

検索条件	設定値	摘要
Your name	任意	ユーザ名をアルファベットで入力してくださ い。
Search title	任意	後々判別がつくような名称、たとえば、MIS search #1 を入力してください。
Database	SwissProt	またはSprot
Taxonomy	All entries	デフォルト値
Enzyme	Lys-C/P	
Allowed up to [1-9] missed cleavages	1	デフォルト値
Fixed modifications	なし	
Variable modifications	なし	
Peptide tol. +/-	0.3 Da	
#13C	0	
MS/MS tol. +/-	0.3	
Monoisotopic/Average	Monoisotopic	デフォルト値
Data file	errorTolSample.mgf	[参照]ボタンを押して、 "C:¥inetpub¥mascot¥mskk¥sampledata"フォル ダまたはVersion 2.2のインストールCDの "mskk¥sampledata"フォルダにある "errorTolSample.mgf"データファイルを指定 してください。このデータファイルには3個の MS/MSデータが含まれています。
Data format	Mascot generic	デフォルト値
Instrument	ESI-QUAD-TOF	
Report top	AUTO	

3-4. MIS 検索の実行

検索条件設定画面(図10)左下にある[Start Search ...]ボタンを押してください。MIS検索が始まり、検索の進捗状況 が%表示されます。MIS検索では2段階の検索が行われます。最初にプリカーサイオン質量に対して Peptide tol +/-および #13C で指定された質量誤差範囲に含まれるペプチドを検索します。この段階で、検索条件にマッチしたペプ チドのアミノ酸配列がわかりますので、次の段階でフラグメントイオンの質量データがアミノ酸配列に一致するかどうかを 検索します。

検索時間(検索に要する時間)は、検索に使用した配列データベースのサイズ(配列データベースを構成す る蛋白質エントリの総数)、missed cleavagesの設定値、Variable modificationsの設定値、Peptide tol +/-の設 定値などにより異なります。これらの設定値が大きくなるほど検索空間が広がるために、検索時間は長くなり ます。

PMF検索ではペプチドの質量スペクトルに一致する蛋白質を検索・同定するのに対し、MIS検索ではフラグメントイオンの質量スペクトルに一致するペプチドを検索・同定します。

3-5. 検索結果レポート

MIS検索が終了すると、Mascot Search Results 画面に切り替わり、ヒットした蛋白質がプロテインスコア順に表示され ます(図11)。Mascot Search Results 画面には、Peptide Summary Report 画面と Select Summary Report 画面の 2つの表示モードがあり、データファイルに含まれるMS/MSデータの数 (クエリー数) によって自動的に切り替 わります。クエリー数が1000未満の場合は Peptide Summary Report 画面が、1000以上の場合は表示内容がよ りコンパクトまとめられた Select Summary Report 画面が最初に表示されます。このMIS検索例では3個の MS/MSデータを含むデータファイルを使用していますので、最初に Peptide Summary Report 画面が表示されま す。なお、[Fomat As] ボタンの右にあるプルダウンメニューからこれら2つの表示モードを互いに切り替える ことができます。





Mascot Search Results 画面では、ヒットした蛋白質はプロテインスコア順にリスト表示され、図12に示す 内容を含んでいます。

図 12 ヒットした蛋白質の情報

1. ELM1_YEAST Mass: 72105 Score: 96 Queries matched: 1
 Serine/threonine-protein kinase ELM1 - Saccharomyces cerevisiae (Baker's yeast)
 Check to include this hit in error tolerant search or archive report

 Query
 Observed
 Mr(expt)
 Mr(calc)
 ppm
 Miss
 Score
 Expect
 Rank
 Peptide

 2
 1084.56
 1083.55
 1083.53
 16.4
 0
 96
 2e-008
 1
 K.RDFVTDGFK.L

図12の1行目の表示内容は、1. (リスト番号)、ELM1_YEAST (アクセッション番号)、Mass: 72105 (分子 量)、Score: 96 (プロテインスコア、イオンスコアの積算値)、Queries matched: 1 (蛋白質に帰属するク エリー数)、2行目は蛋白質情報 (Serine/threonine-protein kinase ELM1 - Saccharomyces cerevisiae (Baker's yeast))、4行目以降は、Query (クエリー番号、2)、Observed (プリカーサイオン質量の実験値、 1084.56)、Mr(exppt)(プリカーサのMr実験値、1083.55)、Mr(calc)(プリカーサのMr理論値、1083.53)、 ppm (プリカーサMrの実験値と理論値の差(ppm表示)、16.4)、Miss (Missed cleavage数、0)、Score (イオン スコア、96)、Expect (期待値、2e-008)、Rank (イオンスコア順位、1)、Peptide (アミノ酸配列、K. RDFVTDGFK.L) です。

1行目のアクセッション番号(ELM1_YEAST)にはProtein View画面にリンクが張られており、蛋白質に関す るより詳しい情報を見ることができます。また、Query(クエリー番号、2)はPeptide View画面にリンクが張 られており、フラグメントイオン質量ピークに対応するイオンシリーズの一致状況などをMS/MSスペクトルグラ フやフラグメント表を見ながら確認することができます。



図 13 Protein View(左)と Peptide View(右)

クエリー番号(2) にマウスカーソルを合わせると図14に示すようなポップアップ画面が表示され、そのクエ リーがヒットしたペプチドのうち、イオンスコア上位10個のペプチドに関する情報(MS/MSデータに付随するコ メント、イオンスコア閾値、イオンスコアと期待値、ペプチドが帰属する蛋白質のリスト番号やアクセッショ ン番号、アミノ酸配列など)を確認することができます。

図 14 クエリーのポップアップ画面

1.	ELM1_YEAST	Mass: 72105	Score: 96	Queries matched: 1	
	Serine/threo	nine-protein k	inase ELM1 - S	accharomyces cerevisia	e (Baker's yeast
	Check to inc	lude this hit :	in error toler	ant search or archive	report

	Query	0bserved	Mr(ex	pt) Mr(calc)	ppm M	liss	Score	Expect	Rank	Peptide	
[2	1084.56	1083	.55 10	83.53	16.4	0	96	2e-008	1	K.RDFVTDGFK	. L
												-
		Top scori	ng pept	ide matc	hes t	o query 2						H
Pept:	ide mat	CID Spect	rum of a	a Proton	ated	Peptide (MH+=	748.43) which	n is a	Candidate	
		of Phosph	opeptid	e by Dr.	Ken-	ichi Yosh	ino					
	Query	Score gre	ater th	an 31 in	dicat	es identi	ty					
Г	v 1											
ſ		Score	Expect	թթա	Hit	Protein		Peptio	le			2
L	<u> </u>	95.8	2e-008	16.4	1	ELM1_YE	AST	K.RDFV	/TDGFK.I			r.
		22.1	0.48	16.5				K.RVYI	JEEFK.L			Г
	_	11.4	5.6	16.5				K.AEAJ	AEFRYK. M	r.		
Sear	ch Par	9.9	8	-40.32				K.ALSH	ALTYAK.	K		
		9.9	8	16.4				K.DRF:	SFDLGK.C	;		
Туре	of sea	8.8	10	39.7				K.RDDI	JDIHK.T			
Enzyr	ne	8.8	10	-36.62				K.RDLV	/PEDLK.P	ζ		
Mass	values	8.8	10	-23.77				K.RDS:	SKLNHK.B	5		
Prote	ein Mas	8.7	10	19.4				K.RSCI	ONVCFK.C	;		
Pept:	ide Mas	7.4	14	22.5				K.ASLI	INHMFK.7	Г		
Frag	nent Ma	1										
Max 1	Missed	CLEAVAGES	: 1									

3-6. 検索結果の出力

検索結果の内容をXMLまたはCSV形式のファイルに出力することができます。操作方法には「2-6. 検索結果の出力」をご覧ください。

3-7. 検索条件の変更と再検索

図15に示す Peptide Summary Report 画面の[Search Selected]ボタンを押すことによりMS/MS Ion Searchの検索条件設定画面に戻りますので、検索条件を変更して再検索することができます。

図 15 [Search Selected]ボタンが表示された Peptide Summary Report 画面

Peptide Summary Report

Format As	Peptide Summary		Help				
	Significance threshold p< 0.05	Max. number of hits AUTO					
	Standard scoring 💿 MudPIT scoring 🔘	Ions score or expect cut-off 0	Show sub-sets 0				
	Show pop-ups 💿 Suppress pop-ups 🔘	Sort unassigned Decreasing Score 🛛 👻	Require bold red 📃				
Import results in	to MI						
Select All	Select None Search Selected	Error tolerant Archive Report					
1. ELM1_YEAST Mass: 72105 Score: 96 Queries matched: 1 Serine/threonine-protein kinase ELM1 - Saccharomyces cerevisiae (Baker's yeast) Check to include this hit in error tolerant search or archive report							
Query 2	Observed Mr(expt) Mr(calc) <u>r</u> 1084.56 1083.55 1083.53 1	opm Miss Score Expect Rank Pept .6.4 0 96 2e-008 1 K.RI	tide)FVTDGFK.L				

なお、Protein Summary Report 画面には[Search Selected]ボタンがありませんので、Peptide Summary Report 画面 に切り替えてください。

3-8. Error Tolerant Search

図11に示すように、Mascot Search Results 画面に表示されたヒット蛋白質に帰属しないMS/MSデータは "Peptide matches not assigned to protein hits" として、ページの下方にリスト表示されます。 MS/MSデータが表示されているヒット蛋白質に帰属しない理由としては、次のような項目が考えられます。

- (1) 質量ピークの数あるいは強度が不十分のためスコアが低い(十分なマッチングが得られない)
- (2) 測定における質量誤差が大きい
- (3) プリカーサイオンの電荷が正しくない
- (4) 蛋白質処理に用いた消化酵素以外の非特異的な切断が起きている
- (5) 検索条件として指定した以外の化学修飾あるいは翻訳後修飾が存在する
- (6) 求める配列が配列データベース中に存在しない

(1)の場合は論外ですが、(2)の測定の質量誤差が大きい場合は、Peptide View および Protein View 画面 に表示される質量誤差グラフ(ペプチドあるいはフラッグメントイオン質量の測定値と理論値の差をプロット したグラフ)で誤差の大きさを確認し、検索パラメータとして再設定・再検索を実行することにより、対処で きることがあります。

通常、プリカーサイオンの電荷は、プリカーサイオンのピークを検出する際に決定されますが、(3)のよう に信頼性が低いと思われる電荷が得られた場合は、考えられる全ての電荷各々に対してピークリストを作成し、 Mascot 検索を実行することで有意なヒットが得られる場合があります。

残る(4)、(5)、(6)の項目に対しては、Error Tolerant Search を使って対応できる場合があります。Error Tolerant Search では、指定した蛋白質に対して、"全ての酵素切断特性"、"全ての修飾"、"1アミノ酸置換" を考慮して検索を実行します。

【Error Tolerant Search の実行方法】

Mascot Search Results 画面の表示モードを **Peptide Summary Report** 表示にしてください。Error tolerant のチェックボックスをチェックすると Error Tolerant Search が有効になりますので、続いて、Error Tolerant Search を実行したい蛋白質を選択するために、蛋白質情報の下にある、Check to include this hit in error tolerant search or archive report のチェックボックスをチェックしてください。

図 16 Error tolerant Se	earch を実行するためのチ:	ェックボックス
------------------------	------------------	---------



[Search Selected] ボタンを押してください。検索パラメータ設定画面と同様な画面が表示されますが、Error Tolerant Search のモードになっていますので、Enzyme (None) や Database 、Fixed / modifications など の検索条件は変更せずに、そのまま[Start Search…]ボタンを押してください。Error Tolerant Search が開始され、検索が終了すると検索結果が表示されます(図17)。

図 17 Error Tolerant Search の検索結果								
Select All Select None Search Selected Internet Archive Report								
 ELMI_YEAST Mass: 72105 Score: 232 Queries matched: 3 Serine/threonine-protein kinase ELMI - Saccharomyces cerevisiae (Baker's yeast) 								
\overline{arphi} Check to include this hit in error tolerant search or archive report								
Query Observed Mr(expt) Mr(calc) ppm Miss Score Expect Rank Peptide								
▼ 1 748.43 747.42 747.39 39.5 0 53 1 K.IPTPIK.A +	[+79.97 at T3]							
	+ [-0.04 at C-term K]							
▼ 3 1367.63 1366.62 1366.63 -6.49 0 118 1 K.PSSPLMDRTVG	K.R + [<u>+79.97</u> at S3]							
	Possible Assignments:							
Search Parameters	Sulfo (S) [+79.96] Phospho (ST) [+79.97]							
Type of search : MS/MS Ion Search								
Error tolerant search : "ELM1_YEAST"								
Enzyme : None Mass values : Monoisotopic								
nass : nunusucupit Protein Mass : Unrestricted								
Peptide Mass Tolerance : ± 50 ppm								
Fragment Mass Tolerance: \pm 0.3 Da								

Error Tolerant Search の結果、修飾基の導入や配列の変更を施されたペプチドについては、検索前後の質 量差とその発生場所を表示します。質量差を示す青数値にマウス・カーソルを乗せると、より詳しい情報がポ ップアップ表示されます。図17の検索結果例の Query1 および Query 3 では、それぞれスレオニン(T)とセ リン(S)が質量差 +79.97 Da の増加をもたらすスルホン酸(Sulfo)またはリン酸(Phospho)によって修飾され た場合により高いスコアが得られることを示唆しています。Query1 および Query2 のスペクトルを見ると、プ リカーサイオン由来のピークよりも 98Da 小さいところに強度の強いピークが見られ、98Da のニュートラルロ スが起こっていると考えた場合のスコアが最も高くなるため、これらのペプチドはリン酸化されている可能性 が高いと判断することができます。

■■■■ 4. SQ検索 (Sequence Query)

プリカーサイオン質量とアミノ酸配列に関する情報(アミノ酸配列、アミノ酸構成、フラグメントイオン質 量など)が組み合わされたデータに一致するペプチドを検索することによりペプチドあるいはタンパク質を同 定する方法を SQ 検索と呼んでいます。アミノ酸配列に関する情報は MS/MS スペクトルを解析して取得します (完全に正しいアミノ酸配列を決定するのは非常に難しく、通常は3から4残基程度の配列にとどまります)。

SQ 検索手法は EMBL で研究活動をしていた Mann の研究グループによって一般化されました。彼らはペプチ ドの質量と数残基のアミノ酸配列に関する情報の組み合わせ、すなわち、MS/MS スペクトルを解析して得られ た次の4つの情報の組、

- (1) プリカーサイオン質量
- (2) 解明されたアミノ酸配列に対応するピーク群の中の最初のピーク質量
- (3) 解明されたアミノ酸配列情報
- (4) 解明されたアミノ酸配列に対応するピーク群の中の最後のピーク質量

を「配列タグ (sequence tag)」と定義し、これらの情報に一致するペプチドあるいは蛋白質の検索方法と して標準配列タグ (standard sequence tag)検索および Error Tolerant 配列タグ (error tolerant sequence tag)検索のアルゴリズムを開発しました。

Mascot は標準配列タグ (standard sequence tag) 検索および Error Tolerant 配列タグ検索 (error tolerant sequence tag) 検索の両方をサポートしています。さらにアミノ酸配列に関する情報として、アミノ酸配列だ けではなく、アミノ酸構成、フラグメントイオン質量値を任意に組み合わせて入力データとして利用できるように機能を拡張しています。

4-1. 検索パラメータ設定画面へのアクセス

Webブラウザを起動し、アドレス入力欄に Mascot Server 名を含む次のURLを入力してください。

http://ホスト名/mascot/

Mascot Server に接続され、図1に示すトップページ が表示されます。トップページ画面の Welcome の次の 行にある Mascot のリンク、またはページ最上段の HOME | MASCOT | HELP の中の MASCOT のリンクをク リックしてください。図2に示す Mascot Server の検索 モード選択画面 (Mascot Search 画面) が表示されます。

図2の Mascot Search 画面の2番目にある Sequence Query のリンクをクリックしてください。図18 に示す SQ検索のパラメータ設定画面が表示されます。各パラメ ータ項目はヘルプ書類にリンクしていますので、その使 い方や意味をオンラインで参照することができます。



4-2. SQの構文と実行例

検索パラメータ設定画面のクエリー入力窓へ入力するクエリーは次のように1行で表現され、プリカーサイ

オンの質量とアミノ酸配列に関する情報から構成されます。プリカーサイオンの質量は必須入力データ、それ に続くアミノ酸配列に関する情報はオプション入力データです。クエリー入力窓へは複数のクエリーを入力す ることができます。

 $M \operatorname{seq}(\dots) \operatorname{comp}(\dots) \operatorname{ions}(\dots) \operatorname{tag}(\dots) \operatorname{etag}(\dots)$

ここで、M は実験から得られたプリカーサイオンの質量、seq(…) はアミノ酸配列情報、comp(…) はアミノ 酸構成情報、ions(…) は フラグメントイオン情報 (MS/MS フラグメントイオンの質量とオプションとしてそ の信号強度)、tag(…) は標準配列タグ (sequence tag)、etag(…) は Error tolerant 配列タグです。

オプション入力データであるアミノ酸配列に関する情報は0個以上いくつでもかまいません。もし複数の標 準配列タグ "tag(…)" と一つ以上の Error torelant 配列タグ "etag(…)" が混在する場合は全て Error tolerant 配列タグ "etag(…)" として処理されます。

オプション入力データ ions(…)、tag(…)、etag(…)に対しては確率的なスコアリングアルゴリズムが適応 されます。すなわち、一致するオプション入力データの数が多いペプチドほどスコアは高くなります。また、 そのペプチドに一致しないオプション入力データは無視されます。一方、seq(…) と comp(…) はフィルタと して機能します。すなわち、これらのオプション入力データに一致するペプチドが存在しない場合、このクエ リーは破棄されます。従いまして、seq(…) と comp(…)を含むクエリーを実行する際は注意が必要です。な お、seq(…)は指定したアミノ酸配列に一致するペプチドを検索しますが、Blast 検索とは異なります。

SQの構文などに関する詳しい内容はヘルプページ (http://**ホスト名**/mascot/help/sq_help.htmlまたは http://www.matrixscience.com/help/sq_help.html) を参照してください。

検索条件設定画面のクエリー入力窓に次の内容を入力し、SQ検索を実行してください。

TAXONOMY=.....lobe-finned fish and tetrapod clade REPTYPE=Peptide TOL=0.03 TOLU=% ITOL=0.5 ITOLU=Da CHARGE=2+ INSTRUMENT=ESI-TRAP 877.4 tag(376.2, [IL][QK][IL], 730.2) 687.3 etag(782.3, NG[IL], 1066.1)

ここで使用した2つの配列タグから Lysozyme がヒットします。

 <u>1A2YC</u> Mass: 14260 Score: 76 Peptides matched: 2 lysozyme (EC 3.2.1.17) mutant (D18A), chain C - chicken Check to include this hit in error tolerant search or archive report

 Query
 Observed
 Mr (expt)
 Mr (calc)
 Delta
 Miss
 Score
 Expect
 Rank
 Peptide

 1
 877.40
 1752.79
 1752.83
 -0.04
 0
 35
 0.0021
 1
 NTDGSTDYGILQINSR

 2
 687.30
 1372.58
 1267.60
 104.98
 0
 42
 0.31
 1
 GYSLGNWVCAAK

この検索では、質量が 687.3 のペプチドが 104.98 Da の質量差でヒットするため、システインがピリジルエ チル化されていることが考えられます。

■■■■■ 5. Mascot Serverの管理

Mascot Serverは検索結果や配列データベースなどを管理するためのユーティリティを装備しています。

5-1. 検索ログ

Mascot Serverのトップページの [Search Log] リンクをクリックするか、URLとして

http://ホスト名/mascot/x-cgi/ms-review.exe

を指定すると、MASCOT検索ログのリスト画面が表示されます。

画面中の [Job#] はMascot Server が検索毎に付与する固有IDで、それに対応する検索結果画面にリンクしています。

ラジオボタン(○)は、昇順に並べ替える際に選択します。選択後に[Sort / filter]ボタンを押してください。チェックボックス(□)は、チェックされたタイトルの内容を表示させる際に選択してください。

チェックボックスの下にある枠は文字列検索の入力ウインドウになっています。検索したい文字列を入力し、 [Sort / filter]ボタンを押してください。入力文字列にヒットした検索結果がリスト表示されます。

🗿 Ma	scot Searc	:hLog - I	Microsoft Inter	net Explorer														\mathbf{X}
77	(ル(<u>F</u>) 編集	(E) 表示	(V) お気に入り(A)) ツール(エ) ヘル	プ(日)												1	
: 0	6					0				600								
G	戻る 🔻 🥃) - 💌	🖻 🚺 🏸	/ 検索 🏑 お気し	EAU	G	🎽 🚳	. 🕅 .		19								
: アドレ	ス(D) 🍯 http	⊳://saki/ma	ascot/x-cgi/ms-re	view.exe												🗸 🔁 移動	カ 目りング	5 »
: Co			🗸 検索 💧	📾 🎼 🚽 🕎	ブッカマ		□ ポップア・	ップを許可	ABC T	- тур 🚽	- 🔼 次に送信	- A				◎設定▼	: 🛖	Ŧ
:			in period				E-1077		• •	100		·				- BAAL	: •	1000
MAS	COT sear	ch log																^
		011 100																
Vers	ion: 2.2.1 ·	Licensed	l to: Matrix Scie	nce K. K., KN-Bu	ild. 6	8-10-	12 Sotokand	a, Chiyoda	-ku,	Tokyo,	(1 processor).	Logged in	as ho	st name	is smile -	logout	
Sc	rt / filter	Log Fil	e:/logs/searc	heslog Start at:	(-1=	end,	1=start) -	1 how m	any:	50 1	182 in log,	1182 aft	er filter	s. Data	dir:	GET	s?: 🔲	
Job	PID	dbase	User Name	Email	Ti	In	start tim	e	C)ur	Status	Pr Ty	p Enzy	ie IP	User I	D		
۲	\circ	\circ	\circ	\circ	\circ	\bigcirc	\circ			\circ	\circ	\circ	\circ	\circ	\circ			
	~	~	V	V		~				~	✓		· ·		V			=
																1		
2512	988	NCBIor	Monitor Test D	3			Mon Jun 04	13:52:23	2007 2	58	No email set		S Yes		3]		
2511	5416	Sprot	dog	-		Ξ.	Fri Jun 01	17:54:25	2007 1	0	User read re	s PN	F Yes	19	1028			
2509	2080	Sprot	torobee	takaesu®matrixs	,		Fri Jun 01	16:56:21	2007 1	1	User read re	s PM	F Yes	19	1020			
2508	5180	Sprot	W3110_R		,		Fri Jun 01	16:55:54	2007 1	1	User read re	s PM	F Yes	19	1020			
2507	5052	Sprot	W3110_R		,		Fri Jun 01	16:50:14	2007 1	1	User read re	s PN	F Yes	19	1020			
2506	5920	Sprot	dog		,		Fri Jun 01	16:48:46	2007 1	1	User read re	s PM	F Yes	19	1020			_
2505	4300	Sprot	saki		,		Fri Jun 01	16:47:18	2007 1	0	User read re	s PN	F Yes	19	1020			
2504	5492	Sprot	tanaka	tanaka@matrixsc	,		Fri Jun 01	16:45:42	2007 1	7	User read re	s PM	F Yes	19	1020			
2503	544	NCBInr	Mascot Wizard		,		Fri Jun 01	16:36:57	2007 2	48	User read re	s PN/	F Yes	19	1			
2502	5048	Sprot	usami	usami@matrixsci	,	•• -	Fri Jun 01	16:21:41	2007 1	8	User read re	s PN/	F Yes	19	1020			
2501	6020	Sprot	W3110_R		,	· · ·	Fri Jun 01	16:17:09	2007 3	0	User read re	s PN≀	F Yes	19	0			
2500	5820	Sprot	usami	usami@matrixsci		••	Thu May 31	18:14:49	2007 1	1	User read re	s PN	F Yes	19	0			
2499	4352	NCBInr	usami	usami®matrixsci	PM	••	Thu May 31	18:04:16	2007 1	71	User read re	s PN	F Yes	19	0			
2498	4532	Sprot	usami	usami®matrixsci	PM	÷ .	Thu May 31	18:03:53	2007 1	2	User read re	s P∦r	F Yes	19	0			
2497	5636	NUBInr	usami			÷ .	Thu May 31	17:33:01	2007.2	20	User read re	s PN	F Yes	19	0			
2496	4/96	Sprot	usami			÷ .	Thu May 31	17:21:48	2007 6		User read re	s PN	⊢ Yes	19	U			
2495	0848	Sprot	usamı			••	Thu May 31	17:19:45	2007 6		User read re	s PN - DN	⊢ ĭes	19	U			
2484	4440	Sprot	usami				Thu May 31	17:18:10	2007 b 2007 5		user read re	s PN - DN	r ïes F V	10	0			
2483	0004 5856	Sprot	usanii				Thu May 31	17-12-59	2007 F		User read re	S PN e Pk	r ies F Yee	13	0			
2491	5524	Sprot	usami			÷.,	Thu May 31	17.10.53	2007 6		liser read re	s PN	F Yes	19	ñ			~
<u>ک</u> ا	・ジが表示され:	ました					a may or		2001 0			- ""				🧐 イントラネット	r	.:

5-2. データベース管理画面

Mascot Server のトップページの [Database Status] リンクをクリックするか、URLとして、

http://ホスト名/mascot/x-cgi/ms-status.exe

を指定すると、MASCOT search status page が表示されます。

この画面では、Mascot Server に登録した配列データベースの状態を表示する他、検索ログへのリンク、検 索にともなって発生したエラーログへのリンクがあります。また、データベースの内容(残基数、エントリ数、 最も大きなエントリの残基数、残基の出現数、各生物種に含まれるエントリ数など)や、(入力間違いによる スペルミスや生物種辞書に登録されていないことなどが原因の)不明な生物種、Mascot Serverのバージョン、 ライセンス情報、現在進行中の検索ジョブ数などの情報も確認することができます。

🗿 Mascot search status page – Microsoft Internet Explorer			X
: ファイル(E) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)		A	1
😋 戻る 🔹 💿 - 💌 😰 🏠 🔎 検索 🌟 お気に入り 🍕	3 🔗 🌺 🛚 · 🗾 🕅	3	
アドレス① 🕘 http://saki/mascot/x-cgi/ms-status.exe		💌 🄁 移動 🗍 リンク	»
🛛 Google 🕞 🗸 🥣 🕐 検索 🖓 🗸 😭 ブックマー	り 🚽 ポップアップを許可 🏾 🏷 チェック	🔻 🚔 次に送信 🗸 🖉 🔵 設定 🔫 🦓 🔪	·
MASCOT search status page			
Version: 2.2.1 - Licensed to: Matrix Science K. K., KN-Build. 8-1 processor). 2 logical, 1 physical Intel processors (hyper-threading disabled Search log monitor log error log Error message descriptions Do	0-12 Sotokanda, Chiyoda-ku, Tokyo in bios, dual core). CPUs: 0 / av not auto refresh this page	p, (1 Logged in as host name is smile - logout ailable, using: 0 /. [0 searches running]	
Name = NCBInr Family = C:/INETPUB/MASCOT/s Filename = NCBInr_20070506.fasta Pathname = C:/INETPUB/MASCOT/ Status = Not in use StateTime Mon Jun 04.13:56:47 # searches = 0	equence/NCBInr/current/NCBInr_* f sequence/NCBInr/current/NCBInr_20	asta 070506.fasta	
Mem mapped = NO Request to mem map = YES Request unmap = NO Me Number of threads = 2 Current = NO	🕙 Database statistics - Micr	osoft Internet Explorer	
Name = NCBInr Family = C:/INETPUB/MASCOT/s Filename = NCBInr_20070603.fasta Pathname = C:/INETPUB/MASCOT/ Status = In use Statistics Compressio	ファイル(E) 編集(E) 表示(V)	お気に入り(A) ツール(D) ヘルプ(H)	🦧 🛛 - 🗾 🎽
State Time = Mon Jun 04 13:55:49 # searches = 0 Mem mapped = YES Request to mem map = YES Request unmap = NO M Number of themester = 2 Suprement = VES	: : アドレス(D) 🕋 http://saki/mascot/	x-cgi/ms-status.exe?Autorefresh=falsetShow=STAT	
Nomber of Lifeaus - 2 Corrent - 165	Google G-	▼ 検索 🐗 💕 ▼ 🟠 ブックマーク▼ 🌺	
Name = Sprot Family = C://NETPUB/MNSCOT/s Filename = Sprot_52.4.fasta Pathname = C://NETPUB/MNSCOT/s Status = In use In use Sproke Status = Sproke	Time files compressed : Time files compressed (int) : Time / date of fasta file : Time of fasta files (int) : Number of residues : Number of sequences :	Mon May 14 17:30:38 2007 1173131438 Tue May 01 13:36:00 2007 1177934100 97521944 265950	
Name = IPI_human Family = C:/INETPUB/MASCOT/s Filename = IPI_human_20070403.fasta Pathname = C:/INETPUB/MASC Status = In use Statistics Status = In use Statistics Status = No Jun 04.03:52:43 # searches = 0 Mem.paped = YES Request to mem map = YES Request unmap = N0 M Number of threads = 2 Current = YES	Number with invalid residues: Number of sequences too long: Length of longest sequence Maximum Accession Length Version of Mascot Version of this file Seqs with invalid taxon tree Num sequences for taxonomy	0 34350 11 2.2.1 4 0 All entries=265002	
🧟 ページが表示されました	Num sequences for taxonomy : Num sequences for taxonomy :	Archaes (Archaeobacteria)=10941 Eukaryota (eucaryotes)=15386 Alueolata (alueolates)=685 Plasmodium faloiparum (malaria parasite)=198 Other Alueolata=486 Metazoa (Animsia)=74735 Caenorhabditis elegans=3021 Drasonbila (fruit filios)=3148	
	Num sequences for taxonomy Num sequences for taxonomy	Chord at Curtebrates and relatives)=63257 bony vertebrates=62720 loberfinned fish and tetrapod clade=59621 Mammalia (mammals)=52413 Primates=21605 Homo sapiens (human)=16356	~
	🕘 ページが表示されました		イントラネット

5-3. 検索条件のデフォルト値の設定

図2の検索モード設定画面の [Search Form Defaults] リンクをクリックするか、URLとして、

http://ホスト名/mascot/search_form_select.html

を指定すると Set Mascot search form defaults 画面が表示されます。

この画面では、検索条件の初期値(デフォルト値)を指定することができます。ここで指定された値は、検 索条件設定画面中反映されます。

初期値として設定したい検索条件を指定し、[Save defaults as cookie …]ボタンを押してください。「Your search form defaults have been saved as cookie」のメッセージが表示され、検索条件の初期値が設定されます。

なお、この設定はCookieを利用しているため、Webブラウザ毎の設定値になります(異なるログオンユーザや 異なるコンピュータ上で起動したWebブラウザでは反映されません)。

🕘 Matrix Science - I	Aascot – Set Search Defaults – Microsoft Internet Exp	olorer 📃 🗖 🔀
- ファイル(E) 編集(E) ま	表示(V) お気に入り(A) ツール(T) ヘルプ(H)	A.
🔾 戻る 🔹 🕥 🐇	🛐 🛃 🏠 🔎 検索 📌 お気に入り 🚱 🔗 🍓	🛛 • 🔜 📬
アドレス(D) 🙋 http://sak	i/mascot/cgi/form_defaults.pl	💙 🄁 移動 🕴 リンク 🎽
Google G-	💙 検索 🗄 🌮 🚼 👻 🏠 ブックマークマ 🌺	🔘 設定 🗸 🍡 🔹
{MATRIX}		
(SCIENCE)	HOME: MASCOT HELP	Search 60
Mascot > Set Sea	rch Defaults	
Set Mascot s	search form defaults	
Database	NCBInr	
Taxonomy	All entries	*
Enzyme	Trypsin V	
Allow up to	1 v missed cleavages	
Fixed modifications	Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term) Amidated (Protein C-term)	
Variable modifications	Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term) Amidated (Protein C-term)	
Show all mods.		
Quantitation	None	
Peptide tol. ±	1.2 Da 💌 # 13C 0 💌	
MS/MS tol. ±	0.6 Da 💌	
Peptide charge	1+	
Monoisotopic	⊙ Average ○	
Data format	Mascot generic 🛛 🖌 (MS/MS only)	
Instrument	Default 📉 (MS/MS only)	
Decoy		
Error tolerant		
Report top	AUTO 💙 hits	
	Save defaults as cookie	
	Copyright © 2006 Matrix Science Ltd. All Rights Reserved.	✓
é		🧐 イントラネット 🔡

5-4 Mascot Server の設定値の変更

Mascot Serverは、原子量、修飾、消化酵素、イオンシリーズ、配列データベース、定量解析手法、セキュリティ機能な どに関係する様々な設定値を持っています。これらの設定値は、C:¥inetpub¥mascot¥config フォルダに存在する設定フ ァイルの中で定義されおり、Mascot Configuration ユーティリティを利用して閲覧・編集することができます。Mascot Server のトップページにある Configuration Editor リンクをクリックするか、URLとして、

http://ホスト名/mascot/x-cgi/ms-config.exe

を指定すると図22に示す Mascot Configuration ページが表示されます。

図 22 Mascot Configuration ページ Mascot configuration - Microsoft Internet Explorer ファイル(E) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H) 🔇 戻る - 🕥 - 💌 😰 🏠 🔎 検索 🌟 お気に入り 🚱 😒 - 🌄 🏰 アドレス(D) 🗃 http://saki/mascot/x-cgi/ms-config.exe 🗸 ラ 移動 **Mascot Configuration** Administrator Logout Elements Element masses Amino Acids Amino Acid Data Modification definitions Modifications Symbols used in chemical formulae Symbols Enzymes Enzyme definitions Instruments Fragmentation Rules Quantitation Quantitation Methods Database Maintenance Sequence databases, Parse Rules and Options Security Security Administration Pages

たとえば、新たな修飾を追加したい場合は Modifications リンクをクリックしてください。Mascot Configuration: Modification ページが表示されますので、[Add new modification] ボタンを押し、修飾情報を定義してください。 Mascot Configuration ユーティリティを使って次の設定値を閲覧・編集することができます。

(1) Elements

原子の名称とモノアイソトピックおよび平均原子量を編集するこができます。

(2) Amino Acids

アミノ酸情報を閲覧することができます。

(3) Modifications

修飾情報の閲覧・編集の他、追加・削除することができます。

(4) Enzymes

消化酵素情報の閲覧・編集の他、追加・削除することができます。

(5) Instruments

MIS検索で使用するイオンシリーズの組の閲覧・編集の他、追加・削除することができます。

(6) Quantitation

定量解析手法の閲覧・編集の他、追加・削除することができます。

(7) Database Maintenance

配列データベースの設定内容の閲覧・編集の他、追加・削除することができます。また、mascot.datのOptionsセ クションの設定値を閲覧・編集することができます。

(8) Security

ロールに基づくユーザ情報の閲覧・編集の他、追加・削除することができます。

5-5. ヘルプへのアクセス

各画面の上方にある HOME | MASCOT | HELP の中の HELP リンクをクリックするか、 URLとして、

http://ホスト名/mascot/help_index.html

を指定すると つぎのような Help Topic Index 画面が表示されます。各々のタイトルはリンクされています。

MASCOT - General: ASMSなどで発表したプレゼンテーション資料などを参照することができます。

Mascot search overview

•A History of Mascot and Mowse

 $\bullet \text{ASMS}$ 2001 User Meeting presentations

•ASMS 2002 User Meeting presentations

•ASMS 2003 User Meeting presentations

 $\bullet \mathrm{ASMS}$ 2004 User Meeting presentations

 $\bullet \mathrm{ASMS}$ 2005 User Meeting presentations

•ASMS 2006 User Meeting presentations

•ASMS 2007 User Meeting presentations

•Mascot FAQ's

Using Mascot : MASCOTの使い方をまとめてあります。

•Search parameter reference

- •Error tolerant search
- •Data file format
- Scoring algorithm
- Results format
- •Result Interpretation
- •Sharing result reports
- •Exporting result reports
- •Decoy Database
- Quantitation

Mascot System Administration : DBのセットアップ方法をまとめてあります。

•Sequence database setup

Instrument Specific Tips : 質量分析計のデータ処理システムとMASCOTの連携方法をまとめてあります。

- •Applied Biosystems | MDS Sciex Analyst
- •Applied Biosystems Data Explorer
- •Applied Biosystems 4000 Series (TOF/TOF)
- Micromass Masslynx
- •Thermo Finnigan Xcalibur

Protein Identification : 蛋白質同定の原理などをまとめてあります。

- Peptide Mass Fingerprint
- •Sequence Query
- •MS/MS Ion Search

Protein chemistry: 蛋白質同定に関連する蛋白質科学に関してまとめてあります。

- •Post translational modifications
- •Enzymes
- Autolysis
- •Sequence databases
- Nucleic acid translation
- •Amino acid reference
- Contaminants

•BLAST & FastA

Mass spectrometry : MASCOTの使い方に関してまとめてあります。

•MS/MS fragmentation

Accuracy & resolution

•Quadrupole Ion Trap MS

Miscellaneous: その他の関連事項についてまとめてあります。

•Mascot brochure (PDF 2.7 Mb)

•PC Hardware for Mascot

Privacy FAQ

•Browser Check

•Useful links

5-6. Mascot ファミリーに関する最新情報

弊社のWebサイトにアクセスしてください。Mascot Serverのトップページ左上にある弊社のロゴをクリックするか、アドレス欄に次のURLを入力してください。

http://www.matrixscience.com/

WHAT'S NEW

バージョンアップ情報やプレスリリース等に関する最新情報を掲載しています。

MASCOT

弊社が試用公開しているMASCOT Serverです。時差の関係で日本では昼間の時間が比較的空いているようです。

PRODUCTS

弊社の製品情報を入手することができます。

SUPPORT

製品別の最新技術情報を入手することができます。また、製品試用版をダウンロードすることができます。

■■■■■■ 6. Mascot Daemon

Mascot ServerはWebサーバ上に構築されたアプリケーションですので、Webブラウザを利用して検索条件の設 定や検索の実行はWebブラウザを利用してインタラクティブに操作することができます。一方で、100件のデ ータファイルを同じ検索条件で検索したいとき、インタラクティブなWebブラウザを利用する場合では、PC画面 に付きっきりで、100回の同じような操作を行う必要があります。

Mascot Daemonは検索作業を自動化するためのクライアントソフトウエアです。100件のMSデータをバッチ 処理したり、指定した時間に検索を開始することができます。Mascot Daemon は Mascot Server (サーバソフ トウエア) に対するクライアントソフトウエアとして動作しますので、通常は、Mascot Server がインストー ルしてあるPCとネットワークで接続されたPCにインストールします。Mascot Daemon をインストールできるOS はWindows 2000、Windows XP Professional、Windows 2003です(最新のサービスパックが適用されている必要 があります)。

6-1. インストール

Mascot Daemon のインストール関連ファイルは Mascot Server 内にあります。Mascot Server のトップページ (http://ホスト名/mascot/) を開いてください。トップページの中程にある、Mascot Daemon のインストール に関する install リンクをクリックしてください。

Mascot Daemon

Mascot Daemon is a client application which automates the submission of data files to a Mascot server. It runs under Microsoft Windows NT 4 (SP6), 2000, XP, or 2003. Install | Upgrade | Troubleshoot

Installing Mascot Daemon のページが開きますので、インストール手順を読み進み、「5. Install Mascot Daemon by clicking on this link」の次にある Daemon.msi リンクをクリックしてください。[ファイルのダ ウンロード]のダイアログが表示されますので、[実行]ボタンを押し、新たに表示されるダイアログの内容を読みながらインストール進めてください。

インストールが終了したら、[スタート]→[すべてのプログラム]→[Mascot]→[Mascot Daemon]をクリックし て Mascot Daemon を起動してください。初めて Mascot Daemon を起動すると、Mascot Server の URL を指定す るための [Preferences] ダイアログが表示されますので、例にならって Mascot Server の URL (http://ホス ト名/mascot/cgi/) を入力し、[Save]ボタンを押してください。Mascot Server と通信ができることが確認で きた時点で Mascot Daemon が起動します。

6-2. チュートリアル

日本語チュートリアルは次の資料を参照してください。

http://www.matrixscience.jp/pdf/jap_2.2_daemon_manual.pdf

なお、英文ヘルプページは、Mascot Daemon を起動し、[Help]メニューから[Mascot Daemon Help …]を選択 することにより参照することができます。

■■■■■■■ 7. 質量分析計との連携

質量分析計各社のデータ処理ソフトウエアと Mascot Server を連携させることができます。詳しくは、Mascot Server のヘルプページにある Instrument Specific Tips の項を参照してください。

7-1. Applied Biosystems

(1) Analyst

Analyst の[Script]メニューに登録されている [mascot.dl1] を使用して、ピークリストを作成し、検索 を実行することができます。なお、mascot.dl1の最新版は、

http://www.matrixscience.com/help/instruments_analyst.html

からダウンロードすることができます。

(2) Data Explorer

Data Explorer のサポートサイトからダウロードできる Visual Basic マクロを使って Mascot Server と接 続することができます。詳しくは、ヘルプページを参照してください。

http://ホスト名/mascot/help/instruments_data_explorer.html http://www.matrixscience.com/help/instruments_data_explorer.html

(3) GPS Explorer

TS2Mascot ユーティリティを使って、GPS Explorer のデータベースへアクセスし、ピークリストを作成した後、検索を実行することができます。詳しくは、ヘルプページを参照してください。

http://ホスト名/mascot/help/instruments_4000.html http://www.matrixscience.com/help/instruments_4000.html

7-2. Waters

MassLynx の Sample List 機能を利用して Mascot Server に接続することができます。詳しくはヘルプページを参照してください。

http://ホスト名/mascot/help/instruments_masslynx.html http://www.matrixscience.com/help/instruments_masslynx.html

7-3. 島津製作所

Kompact から直接Mascot serverを利用することができます。

7-4. Thermo Fisher Scienctific

Mascot Serverはextract_msn. exe(またはlcq_dta. exe) プログラムを利用するためのユーザインターフェース (トップページの Mascot リンク → 検索モード選択画面 → LCQ_DTA リンク) を持っています。設定方法 および使用方法に関してはヘルプページを参照してください。

http://ホスト名/mascot/help/instruments_xcalibur.html http://www.matrixscience.com/help/instruments_xcalibur.html 日本語マニュアル:http://www.matrixscience.jp/pdf/ jap_MascotAndDaemonSettingsForXcaliburBioworksInstalligOrUpgrading.pdf

7-5. 日立ハイテクノロジーズ

データ処理ソフトウエアによってMascot Serverの標準的な入力ファイル形式である mgf ファイルが作成されますので、そのファイルを使って検索を実行してください。

7-6. Bruker Daltonics

BioToolsのメニューから直接 Masoct Server を利用することができます。

■■■■■■■■ 8. 付録

8-1. 蛋白質同定のコツ その1

8-1-1. はじめに

質量分析計を利用して、試料に含まれるペプチドに由来するイオンの質量を測定し、NCBInr等の配列データ ベースに対して測定データに一致する配列を検索することにより、試料に含まれる蛋白質やペプチドの有力候 補を特定することができる。

通常、検索作業は「検索エンジン」と呼ばれるソフトウエアを利用する。「消化酵素」、「生物種」、「質 量誤差範囲」などの検索条件の下で、アミノ酸配列や塩基配列の情報を含む配列データベースの中から測定デ ータである質量値に一致する配列、すなわちその配列を含む蛋白質群を探し出す。配列に一致する質量値の数 が多くなるほど、それらを同時に満足する候補蛋白質は絞られ、統計的な明確さをもって蛋白質を同定するこ とができる。

検索エンジンは機能が単純であるため、検索結果は質量データや検索条件に強く影響される。ここでは弊社 が提供している蛋白質同定ソフトウエアツールを例にして、蛋白質同定をより確実に行うためのコツを紹介す る。

なお、蛋白質同定精度をより高めるためには、質量分析計の生データからのモノアイソトピックなピークの 同定、配列データベース内に存在する質量データに一致する蛋白質の同定、プロテオーム実験プロセスから発 生する様々なデータやアノテーション情報を加味したデータマイニングによる同定を組み合わせることがより 重要である。

8-1-2. 翻訳後修飾は控えめに

MASCOT検索では同時に最大9個の翻訳後修飾を指定することができるが、翻訳後修飾設定数の増加は検索空間 の増大を意味し、ひとつの質量データに一致する配列数は増加するため、相対的にスコアは低下し、検索時間 は増加する。「0xdation(M)」の条件では10秒程度で有意なヒットが得られる質量データセットでも、翻訳後 修飾設定数を9個にして検索すると、有意なヒットは得られず、検索時間も100秒以上を要する。したがって、 翻訳後修飾は必要最低限の設定数でMASCOT検索を実行することが望ましい。

8-1-2. 質量誤差範囲グラフの有効活用

検索結果ページに表示される質量誤差グラフ(Error graph)は、配列に一致した質量データに関して実験値 と理論値の誤差を表示する。質量誤差グラフは質量分析計の特性を表しているのとも考えられるので、質量誤 差グラフが質量分析計の質量誤差特性から外れている場合は、有意なヒットと思われる結果についても検証す る必要がある。また、「Peptide tol. +/-」や「MS/MS tol. +/-」の値を質量分析計の質量誤差特性に応じた 値に変更し、再検索すると有意なヒットが得られることがある。

8-1-4. Noneを使うのも一案

「enzymes」ファイルで定義した消化酵素の切断特性が反映されていないような疑惑がある場合は、「Enzyme」 の中から「None」を選択し、再検索することにより有意なヒットが得られることがある。

8-1-5. 正確なピーク検出

正確なモノアイソトピック質量データを使って検索を行うことが結果の信頼度を上げると同時に検索時間の 節約にもなる。また、LC-MS/MSデータの場合は、同じ質量値を持つプリカーサのMS/MSスペクトルデータを互い に積算した上でピーク検出処理を実行することにより、重複配列のヒットを減らすことができる。

MASCOT Distillerは最新のピーク検出技術を利用したモノアイソトピック同定ソフトウエアツールであり、 蛋白質同定の信頼度、検索時間を改善する。試用版は日本語のインストレーションマニュアルと共に、次のサ イトから入手できる。

http://www.matrixscience.com/distiller_support.html

8-1-6. Taxonomyフィルタ

実験対象とした生物種に含まれているはずの蛋白質がデータベースの中に存在せず、有意なヒットが得られ なかった場合は、Taxonomyのエントリ数を確認し、その生物種の近傍を含めて検索してみる。2.104/6/3版の NCBInrデータベースにおけるPrimatesのエントリ数は、

```
Primates = 126,605
Homo sapiens (human) = 116,140
```

Other primates = 10,680

である。たとえば、チンパンジーに対する検索は「Other primates」を選択するのではなく、「Primates」 を選択して検索すると目的の蛋白質がヒットする可能性が高くなる。

8-1-7. 翻訳後修飾 — Unimodの利用 (http://unimod.org/)

「mod_file」に定義されていない翻訳後修飾を使いたい場合は「Unimod」のサイトを利用すると良い (http://www.unimod.org/)。Unimodは翻訳後修飾に関するデータベースを提供しており、新規な翻訳後修飾 を登録したり、他のMASCOTユーザが登録した翻訳後修飾を共有することができる。また、弊社が提供している MASCOT検索Webサービス(http://www.matrixscience.com/search_form_select.html)ではUnimodに登録されて いる翻訳後修飾を毎週更新しており、最新の翻訳後修飾を利用することができる。

8-1-8. データマイニング

蛋白質の同定精度をより高めるためには、質量分析計が出力する生データに含まれるモノアイソトピックピークの同定情報、ピークリストに対応する蛋白質の同定情報と共に、プロテオーム実験に関わる様々な情報やアノテーション情報を蓄積し、データマイニングすることが重要となる。弊社ではこのような環境を実現するためのデータ管理ソフトウエア「MASCOT Integra」をこの秋にリリースする予定である。

〈 参考文献 〉

(1) Matrix Science Ltd. 資料, Top 10 Tips for Successful Searching, 2003

8-2. 蛋白質同定のコツ その2

8-2-1. はじめに

昨年に引き続き、弊社が提供している蛋白質同定ソフトウエアツール MASCOT ファミリー を例にして、蛋白 質同定をより確実に行うためのコツを紹介する。

8-2-2. エラー・トレラント検索: Error Tolerant Search

Error Tolerant Search は、MASCOT 検索を実行して得られたヒット蛋白質群の中で、指定した蛋白質に対し て全ての酵素切断特性 (Enzyme = none)、全ての翻訳後修飾および全てのポイントミューテーションを考慮し た再検索機能である。

最初の MASCOT 検索で、(a) 1084.56 のプリカーサ値を持つ MS/MS データが有意なスコアで ELM1_YEAST にヒットし、(b) 748.43 および(c) 1367.63 のプロカーサ値を持つ MS/MS データは有意なスコアでヒットしなかったとする。

Select All Select None Search Selected Error tolerant Archive Report Mass: 72105 ELM1 YEAST Queries matched: 1 Score: 96 1. (P32801) PROTEIN KINASE ELM1 (EC 2.7.1.-) Check to include this hit in error tolerant search or archive report Query Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Peptide <u>2</u> 1084.56 1083.55 1083.53 96 5.6e-009 1 K.RDFVTDGFK.L 0.02 0 Peptide matches not assigned to protein hits: (no details means no match) Ouery Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Peptide × 748.43 747.42 747.42 0.01 0 11 2.7 1 DLYPIK 1 1367.63 1366.62 **~** 3

(b)、(c)の MS/MS データがヒットしない理由として、1. これらのデータに対応するアミノ酸配列が検索対象 としたデータベースに存在しない、2. 特定のアミノ酸が翻訳後修飾されている、3. ミューテーションが存在 するなどの理由が考えられるが、ELM1_YEAST に対して Error Tolerant Search を実行するとセリンとトレオニ ンに対するリン酸化を検出し、次の結果が得られる。

1. ELM1 YEAST Mass: 72105 Score: 234 Oueries matched: 3 (P32801) PROTEIN KINASE ELM1 (EC 2.7.1.-) Check to include this hit in error tolerant search or archive report Query Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Peptide \checkmark 1 748.43 747.42 747.39 0.03 0 53 8.7e-005 1 K.IPTPIK.A + [+79.97 at T3] 1084.56 1083.55 1083.53 0.02 0 96 9.3e-009 1 K.RDFVTDGFK.L \checkmark 2 1367.63 1366.62 1366.63 -0.01 0 118 7.8e-011 1 K.PSSPLMDRTVGK.R + [+79.97 at S3] ¥ 3 Possible Assignments: Sulfation (S) [+79.96] Phospho-NL (S) [+79.97] Phospho (STY) [+79.97] Phospho (ST) [+79.97]

MASCOT 検索では同時に最大9個の翻訳後修飾を指定することができるが、翻訳後修飾設定数の増加は検索空間の増大を意味し、プリカーサ質量値に一致するペプチド数は増加するため、相対的にスコアは低下し、検索時間も長くなる。Error Tolerant Search は MS/MS データセットから特定の蛋白質を切り分けた後、その蛋白質の存在可能性をより確実なものにしたり、その蛋白質の構造を検討するための有力な検索方法である。

8-2-3. エラー・トレラント・タグ検索: Error Tolerant Tag Search

MIS (MS/MS Ion Search)検索では、プリカーサ質量値に一致するペプチド群の中から、フラグメントイオン 質量値に一致するペプチドを検索する。すなわち、プリカーサ質量値を構成するアミノ酸の質量値と翻訳後修 飾などの質量値に対する検索方法である。

一方、SQ (Sequence Query) 検索の Error Tolerant Tag Search モードでは、プリカーサ質量値の情報を使 わずに、指定した酵素でデータベースを「消化」して得られたペプチド群の中から、ペプチドを構成するアミ ノ酸の種類とフラグメントイオン質量値にマッチするペプチドを検索することができる。したがって、翻訳後 修飾やミューテーションに対応する質量値が誤差(実験値と理論値の差:Delta)として表示される。

次の条件(tag および etag キーワード)を使って SQ 検索を実行すると、プリカーサ値 687.30 のペプチドは 104.98 の質量誤差でヒットするため、システインがピリジルエチル化されている可能が示唆される。

TAXONOMY=....lobe-finned fish and tetrapod clade REPTYPE=Peptide TOL=0.03 TOLU=% ITOL=0.5 ITOLU=Da CHARGE=2+ INSTRUMENT=ESI-TRAP 877.4 tag(376.2, [IL][QK][IL], 730.2) 687.3 etag(782.3, NG[IL], 1066.1)

1. LYC_CHICK Mass: 16228 Score: 76 Queries matched: 2 (P00698) LYSOZYME C PRECURSOR (EC 3.2.1.17) (1,4-Check to include this hit in error tolerant search or archive report

 Query
 Observed
 Mr(expt)
 Mr(calc)
 Delta Miss
 Score
 Expect
 Rank
 Peptide

 Image: Score and Scor

MASCOT Distiller の Tag 抽出および De novo の機能を組み合わせることにより、より現実的な検索手段となりうる。

〈 参考文献 〉

- (1) Matrix Science Ltd. 資料, Mascot Search Results FAQ, 2005
- (2) Matrix Science Ltd. 資料, http://www.matrixscience.com/help/sq_help.html
- (3) 私信, 吉野健一

2009年3月 第6版 Copyright 2004 - 2009 Matrix Science Ltd., マトリックスサイエンス株式会社 本書の一部あるいは全部について、マトリックスサイエンス株式会社から文書による許可を得ずに、いかなる 方法においても無断で複写、複製することを禁じます。