

タンパク質同定システム
MASCOT *Server*

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

タンパク質同定システム **MASCOT Server**

もくじ

1. 起動方法		
1-1. MASCOT Server へのアクセス	2
2. PMF (Peptide Mass Fingerprint) 検索		
2-1. 検索パラメータ設定画面へのアクセス	3
2-2. 検索パラメータの種類	4
2-3. 検索パラメータの設定	5
2-4. MASCOT検索の実行	5
2-5. 検索結果レポートの見方	7
2-6. 検索条件の変更と再検索	11
2-7. 検索ログ	11
2-8. データベース管理画面	12
2-9. 検索パラメータのデフォルト値の設定	13
2-10. ヘルプへのアクセス	14
2-11. MASCOT ファミリーに関する最新情報	15
3. MIS (MS/MS Ion Search) 検索		
3-1. 検索パラメータ設定画面へのアクセス	16
3-2. 検索パラメータの種類	17
3-3. 検索パラメータの設定	18
3-4. MASCOT 検索の実行	18
3-5. 検索結果レポートの見方	19
3-6. 検索条件の変更と再検索	21
3-7. Error Tolerant Search	21
(非特異的なペプチドの切断+1 塩基置換+全ての翻訳後修飾を考慮した再検索)		
4. LCQ_DTA Shell (Xcalibur RAWファイルに対する検索)		
4-1. 環境設定	24
4-2. DTA ファイルへの変換と MASCOT 検索の実行	24
5. 付録		
5-1. 蛋白質同定のコツその1	27
5-2. 蛋白質同定のコツその2	29

蛋白質同定システム MASCOT Server

1. 起動方法

MASCOT Server は Webサーバ上のアプリケーションです。従いまして、MASCOT Server がインストールされているコンピュータとネットワークで接続されているコンピュータから、Webブラウザを使ってアクセスすることができます。

1-1. MASCOT Server 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://MASCOT_Server名/mascot/

MASCOT Server に接続され、次の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。

The screenshot shows the MASCOT Server home page in a Microsoft Internet Explorer browser window. The page title is "Matrix Science - Home - Microsoft Internet Explorer". The address bar shows "http://iu/mascot/home.html". The page content includes a "Welcome" section, a "Security" section, "Mascot Utilities", "Mascot Daemon", and "Mascot Documentation".

Annotations (red boxes) point to the following elements:

- MASCOT Server に関連したヘルプ情報にリンクしています。
- 現在ログインしているユーザを示します。(セキュリティ機能が有効な時のみ)
- MASCOT Server サイト内をキーワード検索することができます。
- 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
 2. Sequence Query
 3. MS/MS Ion Search
- MASCOT Server を利用する上で必要なブラウザの機能 (Java Script、Table、ファイルの読込、Cookie) をチェックすることができます。
- 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、現在のセッション情報の確認、ユーザの権限の設定などを行う画面にリンクしています。
- MASCOT Server が利用する配列データベースの管理画面にリンクしています。
- MASCOT 検索のログ画面にリンクしています。
- MASCOT Server が利用する配列データベースを追加するための設定ユーザインターフェースにリンクしています。
- MASCOT Daemon のインストールおよびアップグレードを行うためページにリンクしています。
- Acrobat Reader がインストールされている場合は「Setup & Installation」マニュアルが表示されます。
- 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。

蛋白質同定システム **MASCOT Server**

2. PMF (Peptide Mass Fingerprint) 検索

MASCOT Server version 2.1のCDROMに含まれているサンプル質量データを使用して、MASCOT検索の方法およびMASCOT検索結果の見方について紹介します。

1-1. 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://MASCOT_Server名/mascot/`

MASCOT Server に接続され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。MASCOT Server の検索モード選択画面 (**Mascot Search** 画面)が表示されます。

The screenshot shows the MASCOT Server interface. The browser window title is "Matrix Science - Mascot - Microsoft Internet Explorer". The address bar shows "http://uu/mascot/search_form_select.html". The page content includes a search bar and a "Go" button. Below the search bar, there is a "Mascot Search" section with a list of search modes:

- Peptide Mass Fingerprint:** The experimental data are a list of peptide mass values from an enzymatic digest of a protein.
 - Example of results report
 - More information
- Sequence Query:** One or more peptide mass values associated with information such as partial or ambiguous sequence strings, amino acid composition information, MS/MS fragment ion masses, etc. A super-set of a sequence tag query.
 - Example of results report
 - More information
- MS/MS Ion Search:** Identification based on raw MS/MS data from one or more peptides.
 - Example of results report
 - More information
- LCQ DTA:** Convert Finnigan Xcalibur RAW file into peak lists using lq_dta.exe or extract_msn.exe and load into an MS/MS search form.

At the bottom of the list, there is a link for **Search Form Defaults:** Follow this link to save your preferred search form defaults as a browser cookie.

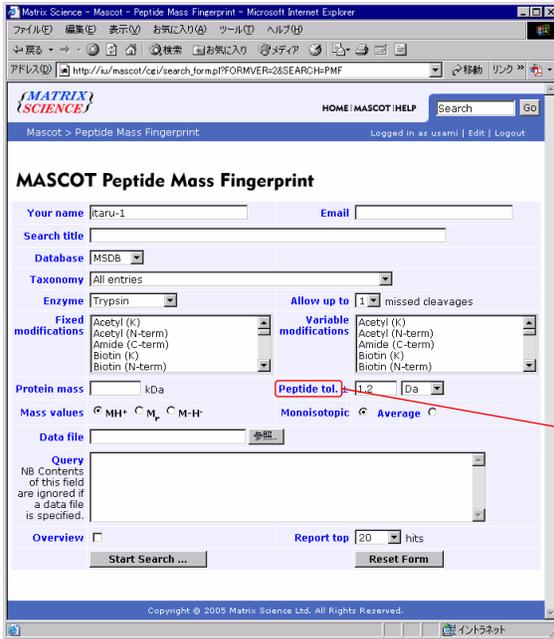
Red boxes and arrows highlight the search modes and their corresponding explanatory text:

- Red box: 各検索モード (Peptide Mass Fingerprint、Sequence Query、MS/MS Ion Search) にリンクしています。
- Red box: LCQ DTA プログラムのユーザインターフェースにリンクしています。
- Red box: 検索パラメータのデフォルト値を設定するためのユーザインターフェースにリンクしています。

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 (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 (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 (MS/MSデータを使って蛋白質を同定することができます)

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。Peptide Mass Fingerprint 検索のパラメータ設定画面が表示されます。



検索そのものに関するパラメータは、

1. Database (検索するデータベース)
2. Taxonomy (生物分類)
3. Enzyme (消化酵素)
4. Fixed / Variable modifications (翻訳後修飾)
5. Allow up to (missed cleavages)
6. Peptide tol +/- (ペプチドの質量誤差範囲)
7. Mass values (ペプチドの電荷数)
8. Monoisotopic/Average

などです。詳しくは次の検索パラメータの種類を参照してください。
 なお、各検索パラメータ項目はヘルプにリンクしており、その項目の意味や使い方を参照することができます。

検索パラメータの各項目はヘルプ書類にリンクされていますので、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。

2-2. 検索パラメータの種類

Peptide Mass Fingerprint 検索では、次の検索パラメータを入力または選択することができます。

質量データは [Query] ウィンドウに数値を入力しますが、質量データをファイルでお持ちの場合は [Data file] の [参照] から指定して読み込むことができます。翻訳後修飾パラメータは最大9個まで同時選択することができます。複数のパラメータを選択する場合および選択したパラメータを解除する場合はキーボードの [Ctrl] ボタンを押しながらマウスでクリックしてください。

なお、各検索パラメータ項目は MASCOT Server のWeb内にあるヘルプ書類にリンクしており、その項目に関する意味や使い方を参照することができます。

検索パラメータ	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
Search title	検索タイトルを入力してください。後日、この検索結果にアクセスしやすくするた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。	sample1, spot67, Trypsin, 50fm, 2457-9001-125-afd
Database	複数のデータベースがセットアップされている場合は、検索したいデータベースを選択します。	MSDB
Taxonomy	検索対象とする生物種を選択します。	All entries
Enzyme	蛋白質を消化するために使用した消化酵素を選択します。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。	Trypsin
Allowed up to [1-9] missed cleavages	未消化のペプチドの状態を指定します。完全に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。	2
Fixed modifications	翻訳後修飾を選択します。指定したすべての(Fixされた)アミノ酸に対して翻訳後修飾を考慮し検索します。	Carboxymethyl (C)
Variable modifications	翻訳後修飾を選択します。指定したアミノ酸に対して翻訳後修飾を考慮する場合としない場合の2通りの(Variableな)検索を行います。	Acetyl (K) Oxidation (M)
Protein mass	検索範囲ウィンドウを指定します。たとえば、"30"を指定した場合は30KDaの範囲に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。	30
Peptide tol. +/-	ペプチド質量の許容誤差をDa, mmu, %, ppmの単位で指定します。1.0 Daを指定した場合は(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。	1Da

Mass values	ペプチドの電荷状態(MH+, Mr, M-H)を指定します。	MH+
Monoisotopic/Average	質量データの種類(MonoisotopicまたはAverage)を指定します。	Monoisotopic
Data file	読み込みたい質量データファイルを[ファイルの選択]ダイアログを通じて指定します。	データファイルを指定
Query	質量データを入力します。 PMF検索で使用する質量データは、ノイズがカットされたモノアイソトピック質量です。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ェアや、Mascot Distillerなどを利用してください。	1137.8 1209.4 1704.0 1862.0 1998.0 2037.0 2485.2 2923.6 3031.8
Overview	チェックボックスをチェックすることにより、検索結果ページに“Overview Table”を表示します。	ボックスをチェック
Report top	検索によりヒットした蛋白質の表示数を選択します。	20

2-3. 検索パラメータの設定

つぎの検索パラメータを設定してください。

検索パラメータ	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	
Database	MSDB	
Taxonomy	All entries	
Enzyme	Trypsin	
Allowed up to [1-9] missed cleavages	1	
Peptide tol. +/-	0.5 Da	
Mass values	MH+	
Monoisotopic/Average	Monoisotopic	
Query	1137.8 1209.4 1704.0 1862.0 1998.0 2037.0 2485.2 2923.6 3031.8	質量データを数値入力する代わりに、[Data file] の [参照] からMASCOT Server 2.1のCDROMの中のmskkフォルダにある“pmfSample.asc”ファイル(中身は、1137.8、12.19.4、1704.0、1862.1、1998.0、2.137.0、2485.2、2923.6、3031.8の質量データです)を読み込むことができます。
Overview	ボックスをチェック	
Report top	10	

2-4. MASCOT 検索の実行

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、検索が終了すると検索結果レポート画面が表示されます。

Matrix Science - Mascot - Peptide Mass Fingerprint - Microsoft Internet Explorer

HOME | MASCOT HELP Search Go

Mascot > Peptide Mass Fingerprint Logged in as usami | Edit | Logout

MASCOT Peptide Mass Fingerprint

Your name: [kuru-1] Email: [usami@matrixscience.com]

Search title: [PMF test Search #1]

Database: [MSDB]

Taxonomy: [All entries]

Enzyme: [Trypsin] Allow up to: [1] missed cleavages

Fixed modifications: [Acetyl (K), Acetyl (N-term), Amidate (C-term), Biotin (K), Biotin (N-term)]

Variable modifications: [Acetyl (K), Acetyl (N-term), Amidate (C-term), Biotin (K), Biotin (N-term)]

Protein mass: [] kDa Peptide tol. ±: [1.2] Da

Mass values: MH+ M_n M_r M_w M_v M_z Average

Monoisotopic Average

Data file: []

Query: [1137.800000
NB Contents of this field are ignored if a data file is specified.
1209.400000
1704.000000
1962.000000
1998.000000
2037.000000]

Overview Report top: [20] hits

Start Search ... Reset Form

Copyright © 2005 Matrix Science Ltd. All Rights Reserved.

検索パラメータの設定と質量データの入力または質量データを含むファイルの指定が完了したら、[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、検索が終了すると検索結果レポートが表示されます。

Mascot searching... - Microsoft Internet Explorer

アドレス: http://uu/mascot/cgi/nph-mascot.exe?1

MASCOT Mascot Search

Licensed to: Matrix Science in-house demonstration system (Tokyo), (4 processors).
Finished uploading search details...

Searching...

5% complete
...10% complete
...20% complete
...23% complete
...23% complete
...33% complete

ページ http://uu/mascot/cgi/nph-mascot/ インターネット

検索時間（検索に要する時間）は、検索に使用したデータベースのサイズ（データベースを構成するタンパク質エントリ数）、**missed cleavages** の設定値、**Fixed / 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、一般的には、検索時間は長くなります。

2-5. 検索結果レポートの見方

MASCOT検索が終了すると、次の Mascot Search Results 画面が表示されます。Mascot Search Results 画面は (1) Concise Protein Summary Report、(2) Protein Summary Report、(3) Export Search Results の3つの画面から構成されています。最初に表示されるのは (1) の Concise Protein Summary Report 画面です。

(1) Concise Protein Summary Report 画面

検索したデータベース (この場合は MSDB) のサイズ (蛋白質のエントリ数=2,011,572 件) です。

トップスコアを持つ蛋白質情報が表示されます。

統計的に有意な閾値スコア (有意水準を 5% として、この場合は 76) とスコアの計算式 ($P=1/2011572 \times 0.05$) を示しています。

Concise Protein Summary と Protein Summary 画面を切り替えることができます。

閾値を決める有意水準を変更することができます。

ヒットした蛋白質がスコア順に表示されます。なお、有意なスコアは太い赤字で表示されます。Accession 番号は詳細情報を表示する Protein View 画面にリンクしています。

(2) Protein Summary Report 画面

(1) の **Concise Protein Summary Report** 画面では、ヒットした蛋白質をスコア順に表示しますが、トップスコアを持つ蛋白質に帰属する質量データのうち、すべてまたは一部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、第1グループの有意なスコア **93** を持つ [Q6W897_HUMAN](#) は (1136.79、1208.39、1702.99、1860.99、2035.99、2484.19、2922.59) の質量データが帰属しますが、スコア 38 を持つ [Q72N89_LEPIC](#) は [Q6W897_HUMAN](#) に帰属する質量データの一部分に当たる (1136.79、1860.99、2484.19) の質量データが帰属しています。

1.	Q6W897_HUMAN	Mass: 57382	Score: 93	Expect: 0.001	Queries matched: 7
	HSD-34.- Homo sapiens (Human).				
	Q86WT6_HUMAN	Mass: 57411	Score: 93	Expect: 0.001	Queries matched: 7
	RNF36 protein.- Homo sapiens (Human).				
	.				
	.				
	(省略)				
	.				
	.				
	Q83B90_COXBU	Mass: 23825	Score: 38	Expect: 3e+002	Queries matched: 3
	IcmJ protein.- Coxiella burnetii.				
	Q72N89_LEPIC	Mass: 24272	Score: 38	Expect: 3e+002	Queries matched: 3
	Flagellin protein.- Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar Copenhageni).				

2.	AAL82101	Mass: 34383	Score: 51	Expect: 17	Queries matched: 4
	AE010291 NID: - Pyrococcus furiosus DSM 3638				
	AAL821011	Mass: 34583	Score: 51	Expect: 18	Queries matched: 4
	AE009950 NID: - Pyrococcus furiosus DSM 3638				
	Q74A00_GEOSL	Mass: 29253	Score: 38	Expect: 3.3e+002	Queries matched: 3
	ISGsu5, transposase.- Geobacter sulfurreducens.				

3.	Q6RUD2_SIVCZ	Mass: 16695	Score: 48	Expect: 29	Queries matched: 3
	Envelope glycoprotein (Fragment).- Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).				
	.				
	.				
	(省略)				
	.				
	.				

16.	Q6DBN8_ARATH	Mass: 40326	Score: 38	Expect: 3.3e+002	Queries matched: 3
	At5g23610.- Arabidopsis thaliana (Mouse-ear cress).				

ページの最後にはこの検索で使った検索パラメータが表示されます。

Search Parameters

Type of search : Peptide Mass Fingerprint
 Enzyme : Trypsin
 Mass values : Monoisotopic
 Protein Mass : Unrestricted
 Peptide Mass Tolerance : ± 0.5 Da
 Peptide Charge State : 1+
 Max Missed Cleavages : 1

Number of queries : 9

(2) の **Protein Summary Report** 画面では、ヒットした蛋白質を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)、未消化状態数 (Miss) などの情報を表示します。Concise Prptein Summary Report の時のように、グループにまとめません。

	Accession	Mass	Score	Description
1.	Q6W897_HUMAN	57382	93	HSD-34.- Homo sapiens (Human).
2.	Q86WT6_HUMAN	57411	93	RNF36 protein.- Homo sapiens (Human).
3.	Q8WY16_HUMAN	32047	83	TRIMLESS beta.- Homo sapiens (Human).
4.	Q8WY17_HUMAN	34011	83	TRIMLESS alpha.- Homo sapiens (Human).
5.	Q81YY3_HUMAN	39229	79	Ring finger protein 36, isoform b.- Homo sapiens (Human).
6.	Q8QHA5_XENLA	119656	54	Helicase.- Xenopus laevis (African clawed frog).
7.	AAL82101	34383	51	AE010291 NID: - Pyrococcus furiosus DSM 3638
8.	AAL821011	34583	51	AE009950 NID: - Pyrococcus furiosus DSM 3638
9.	Q6RUD2_SIVGZ	16695	48	Envelope glycoprotein (Fragment).- Chimpanzee immunodeficiency virus (SIV)
10.	Q8HRQ8_9ROSI	38256	46	NADH dehydrogenase subunit F (Fragment).- Morisonia americana.

Results List

1.	Q6W897_HUMAN	Mass: 57382	Score: 93	Expect: 0.001	Queries matched: 7		
HSD-34.- Homo sapiens (Human).							
Observed	Mr (expt)	Mr (calc)	Delta	Start	End	Miss	Peptide
1137.80	1136.79	1136.58	0.21	361	- 371	0	R. FDSSVAVLGSR. G
1209.40	1208.39	1208.59	-0.19	378	- 386	0	K. WYWEVEVAK. K
1704.00	1702.99	1702.85	0.14	403	- 417	0	K. GSCPLTPEQGFWLLR. L
1862.00	1860.99	1860.98	0.01	426	- 442	0	K. ALDLPSFSLTLTNNLDK. V
2037.00	2035.99	2035.99	0.00	443	- 460	0	K. VGIYLDYEGGQLSFYNAK. T
2485.20	2484.19	2484.05	0.14	50	- 70	0	R. DPLMLSCGHNFCEACIQDFWR. L
2923.60	2922.59	2922.44	0.15	261	- 285	1	K. TEQQNSFDLKDITLLHSLEGGMK. V
No match to: 1998.00, 3031.80							

タンパク質のアクセッション番号 (そのタンパク質のユニークな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、そのタンパク質の詳細情報 (全体アミノ酸配列と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: [Q6W897_HUMAN](#) Score: 93 Expect: 0.001

HSD-34.- Homo sapiens (Human).

Found in search of C:\inetpub\wwwroot\mascot\masks\pmfSample.asc

Nominal mass (M_r): 57382: Calculated pI value: 6.04

NCBI BLAST search of [Q6W897_HUMAN](#) against nr

Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Homo sapiens](#)

Links to retrieve other entries containing this sequence from NCBI Entrez:

[AAQ75551](#) from [Homo sapiens](#)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

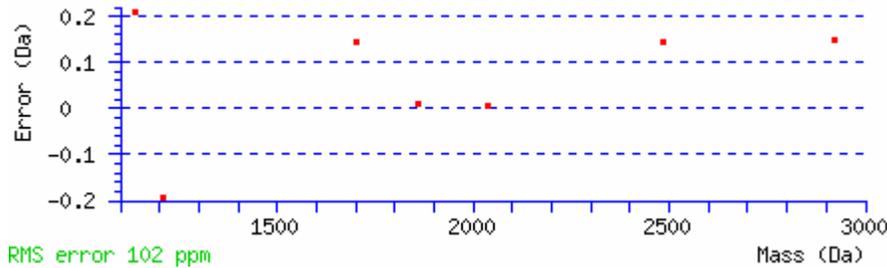
Number of mass values searched: 9
 Number of mass values matched: 7
 Sequence Coverage: 23%

Matched peptides shown in **Bold Red**

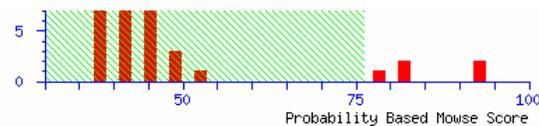
1 MEVSTNPSSN IDPGNYVEMN DSITHLPSKV VIQDITMELH CPLCNDWFRD
 51 **PLMLSCGHNF CEACIQDFWR** LQAKETFCPE CKMLCQYNNC TFNPVLDKLV
 101 EKIKKLPLLK GHPQCPEHGE NLKLFSPDG KLICFQCKDA RLSVGQSKEF
 151 LQISDAVHFF MEELAIQGGQ LETTLKELQT LRNMQKEAIA AHKENKLHLQ
 200 QHVSMEFLKL HQFLHSKEKD ILTELREEGK ALNEEMELNL SQLQEQLLA
 251 KMLVSIQAK **TEQQNSFDL KDITLLHSL EQGMK**VLATR ELISRKLNLG
 301 QYKGP IQYMV WREMQDTLCP GLSPLTLDPK TAHPNLVLSK SQTSVWHGDI
 351 KKIMPDDPER **FDSSVAVLGS** RGFSTSGK**WYW** **EVEVAK**TKW TVGVVRESII
 401 RK**GSCPLTPE** **QGFWLLR**LRN QTDLK**ALDLP** **SFSLTLTNNL** **DKVGIYLDYE**
 451 **GGQLSFYNAK** TMTHIYTFSN TFMEKLYPYF CPCLNDGGEN KEPLHILHPQ
 501

Start - End	Observed	Mr (expt)	Mr (calc)	Delta	Miss	Sequence
50 - 70	2485.20	2484.19	2484.05	0.14	0	R. DPLMLSCGHNFCEACIQDFWR. L
261 - 285	2923.60	2922.59	2922.44	0.15	1	K. TEQQNSFDLKDITLLHSL EQGMK. V
361 - 371	1137.80	1136.79	1136.58	0.21	0	R. FDSSVAVLGS. G
378 - 386	1209.40	1208.39	1208.59	-0.19	0	K. WYWEVAK. K
403 - 417	1704.00	1702.99	1702.85	0.14	0	K. GSCPLTPEQGFWLLR. L
426 - 442	1862.00	1860.99	1860.98	0.01	0	K. ALDLP SFSLTLTNNL DK. V
443 - 460	2037.00	2035.99	2035.99	0.00	0	K. VGIYLDYEGGQLSFYNAK. T

また、質量データとその理論との質量誤差グラフは、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）に対応しますので、この質量誤差グラフと実際のキャリブレーションカーブが大きく異なる場合は、質量データや検索条件を吟味する必要があります。



(3) の **Export Search Results** を選択すると検索結果をXMLまたはCSV形式のファイルに出力するための設定画面が表示されますので、出力項目を指定して [Export search results] ボタンを押してください。



Protein Summary Report

Format As [Help](#)
 Protein Summary
 Concise Protein Summary
 Export Search Results
 Max. number of hits

2-6. 検索条件の変更と再検索

(1) の **Concise Protein Summary Report** 画面の [Re-Search All] または [Search Unmatched] ボタン、あるいは、(2) の **Protein Summary Report** 画面の [Search Selected] ボタンを押すことにより、PMF検索パラメータ設定画面に戻りますので、パラメータ設定を変更して再検索することができます。

2-7. 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://MASCOT_Server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 MASCOT検索毎に付与する固有IDで、**Mascot Search Results** の **Concise Protein Summary Report** 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、[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ィ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MASCOT検索結果がリスト表示されます。

Job#	PID	dbase	Us	Em	Title	In	start time	Dur	Status	Pr	Typ	Enzyme	IP	User ID
1590	1676	MSDB	it	us	PMF test Search #1		Tue Jul 12 15:57:37 2005	104	User read res		PMF	Yes	19	0
1589	1512	MSDB	it	us	PMF test Search #1		Tue Jul 12 15:04:41 2005	122	User read res		PMF	Yes	19	0
1588	444	MSDB	it	us	PMF test Search #1		Tue Jul 12 15:01:45 2005	103	User read res		PMF	Yes	19	0
1587	756	MSDB	it	us	PMF test Search #1		Tue Jul 12 14:59:04 2005	103	User read res		PMF	Yes	19	0
1586	120	MSDB	it	us	PMF test search #1		Tue Jul 12 14:41:28 2005	110	User read res		PMF	Yes	19	0
1585	848	MSDB	it	us	PMF test search #1		Tue Jul 12 14:35:13 2005	110	User read res		PMF	Yes	19	0
1584	848	MSDB	it	us	PMF test search #1		Tue Jul 12 14:27:53 2005	109	User read res		PMF	Yes	19	0
1583	1016	Sprot	Ma		Submitted from msms by Mascot		Sun Jul 03 19:58:58 2005	4	User read res		MIS	Yes	19	0
1582	544	Sprot	Ma		Submitted from msms by Mascot		Sun Jul 03 19:58:26 2005	3	User read res		MIS	Yes	19	0
1581	796	Sprot	Ma		Submitted from msms by Mascot		Sun Jul 03 19:57:51 2005	3	User read res		MIS	Yes	19	0
1580	1028	Sprot	Ma		Submitted from msms by Mascot		Sun Jul 03 19:57:19 2005	4	User read res		MIS	Yes	19	0
1579	1656	Sprot	Ma		Submitted from msms by Mascot		Sun Jul 03 19:56:46 2005	4	User read res		MIS	Yes	19	0
1578	1512	Sprot	Ma		Submitted from msms by Mascot		Sun Jul 03 19:56:14 2005	4	User read res		MIS	Yes	19	0
1577	1692	Sprot	us		Submitted from axima test 3 by		Mon Jun 27 11:09:32 2005	11	User read res		PMF	Yes	19	0

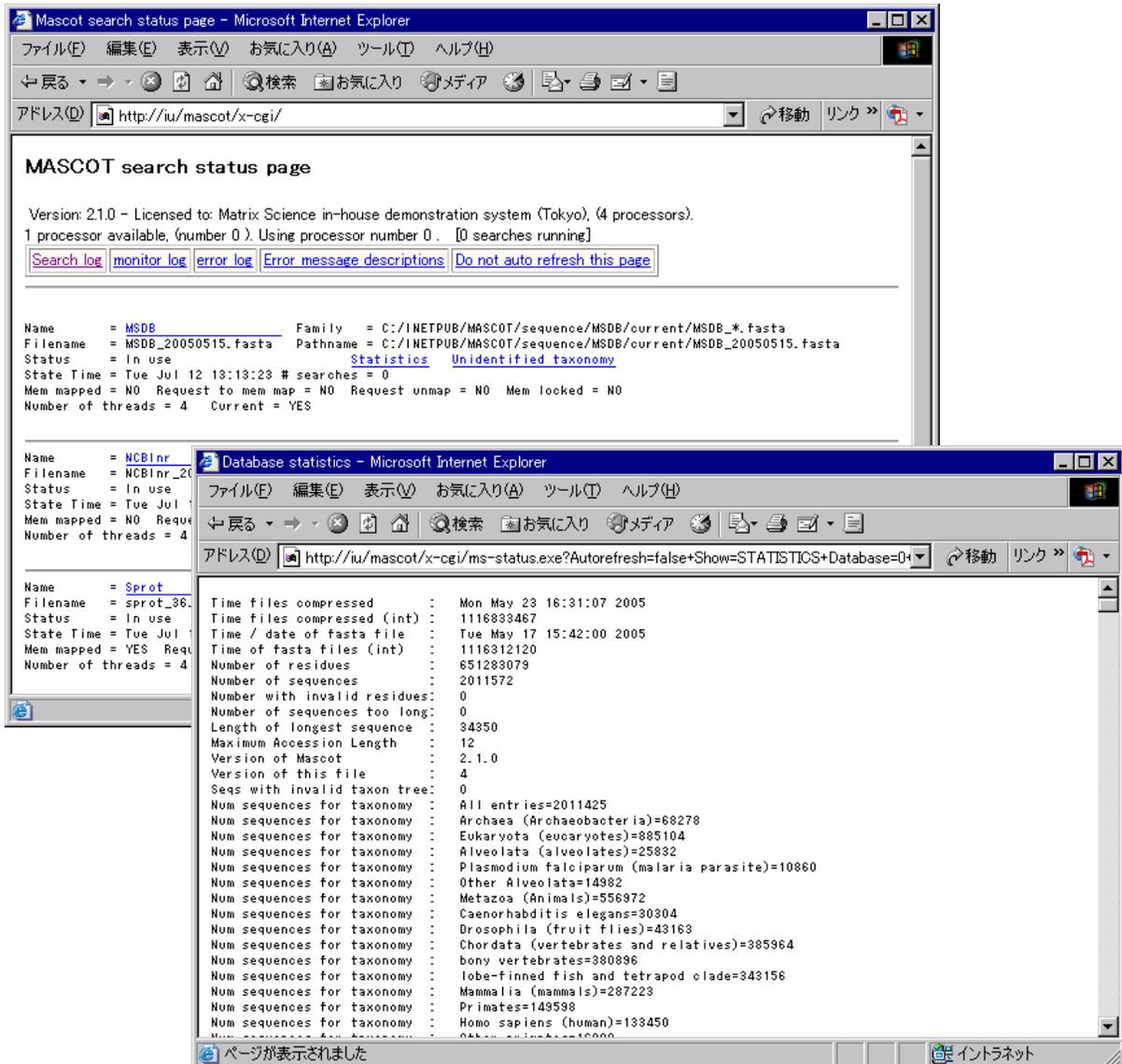
8. データベース管理画面

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://MASCOT_Server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 に登録したデータベースの状態を表示する他、検索ログへのリンク、MASCOT 検索ジョブ等が実行された際に記録される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、現在進行中の検索数、エラーログへのリンクを含め、様々な情報も確認することができます。



2-9. 検索パラメータのデフォルト値の設定

Mascot Search 画面の [Search Form Defaults] リンクをクリックするか、URLとして、

http://MASCOT_Server名/mascot/search_form_select.html

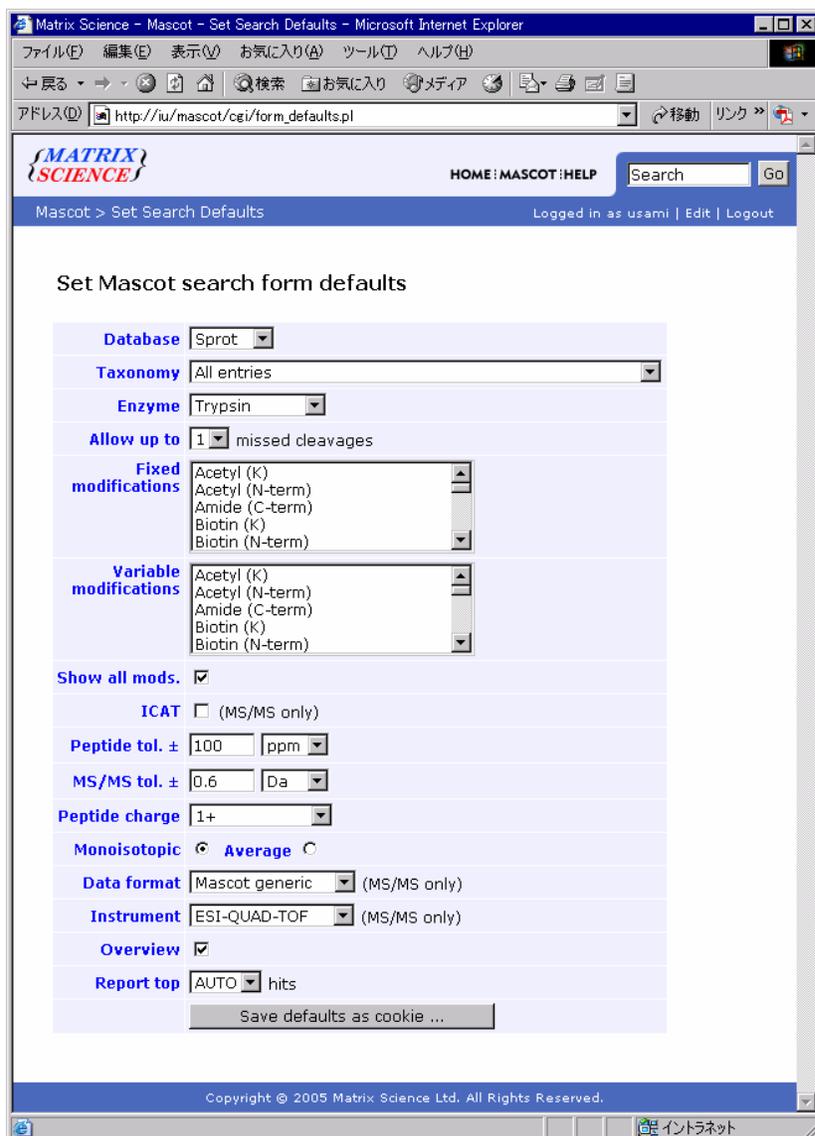
を指定すると **Set Mascot search form defaults** 画面が表示されます。

この画面では、検索パラメータの初期値（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、検索パラメータ設定画面中でパラメータ初期値として表示されます。

初期値として設定したい検索パラメータを指定し、[**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、パラメータ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PC上で起動したWebブラウザでは反映されません）。



2-10. 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://MASCOT_Server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ので、関係するリンクをクリックしてください。

MASCOT Server – 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
- ASMS 2001 User Meeting presentations
- ASMS 2002 User Meeting presentations
- ASMS 2003 User Meeting presentations
- ASMS 2004 User Meeting presentations
- Mascot FAQ's

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

- Search parameter reference
- Error tolerant search
- Data file format
- Scoring algorithm
- Results format
- Result Interpretation
- Sharing result reports
- Exporting result reports

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
- 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
- Bibliography

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
- Bibliography

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
Bibliography

Miscellaneous : その他の関連事項についてまとめてあります。

Mascot brochure (PDF 1.8 Mb)
Privacy FAQ
Browser Check
Useful links

2-11. 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。

蛋白質同定システム **MASCOT Server**

3. MIS (MS/MS Ion Search) 検索

MASCOT Server version 2.1のCDROMに含まれているサンプル質量データを使用して、MS/MS質量データに対するMASCOT検索の方法およびMASCOT検索結果の見方について紹介します。

3-1. 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://MASCOT_Server名/mascot/

MASCOT Serverのトップページに接続されます。表示された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。MASCOT Serverの検索モード選択画面 (**Mascot Search** 画面) が表示されます。

The screenshot shows the MASCOT Server Mascot Search page. The page title is "Mascot Search" and it is logged in as "usami". The main content area lists four search modes with links to "Example of results report" and "More information":

- Peptide Mass Fingerprint:** The experimental data are a list of peptide mass values from an enzymatic digest of a protein.
- Sequence Query:** One or more peptide mass values associated with information such as partial or ambiguous sequence strings, amino acid composition information, MS/MS fragment ion masses, etc. A super-set of a sequence tag query.
- MS/MS Ion Search:** Identification based on raw MS/MS data from one or more peptides.
- LCQ DTA:** Convert Finnigan Xcalibur RAW file into peak lists using lq_dta.exe or extract_msn.exe and load into an MS/MS search form.

Red callout boxes provide additional information:

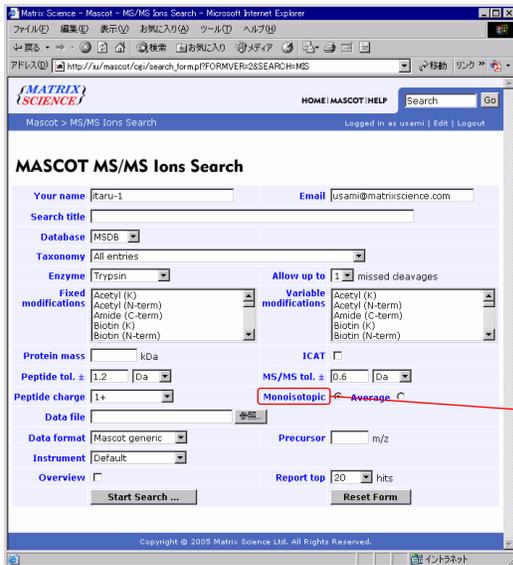
- 各検索モード (Peptide Mass Fingerprint、Sequence Query、MS/MS Ion Search) にリンクしています。
- LCQ_DTA プログラムのユーザーインターフェースにリンクしています。
- 検索パラメータのデフォルト値を設定するためのユーザーインターフェースにリンクしています。

At the bottom of the page, there is a link for "Search Form Defaults" with the text: "Follow this link to save your preferred search form defaults as a browser cookie."

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 (MS質量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 (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 (MS/MS質量データを使って蛋白質を同定することができます)

MS/MS Ion Search: のリンクをクリックしてください。MS/MS Ion Search検索のパラメータ設定画面が表示されます。



検索そのものに関するパラメータは、

1. Database (検索するデータベース)
2. Taxonomy (生物分類)
3. Enzyme (消化酵素)
4. Fixed / Variable modifications (翻訳後修飾)
5. Allow up to (missed cleavages)
6. Peptide tol +/- (ペプチドの質量誤差範囲)
7. Mass values (ペプチドの電荷数)
8. Monoisotopic/Average

などです。詳しくは次の検索パラメータの種類を参照してください。
 なお、各検索パラメータ項目はヘルプにリンクしており、その項目の意味や使い方を参照することができます。

検索パラメータの各項目はヘルプ書類にリンクされていますので、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. 検索パラメータの種類

MS/MS Ion Search検索では、次の検索パラメータを入力または選択することができます。

質量データは **Data file** 項の [参照] ボタンから指定して読み込みます。翻訳後修飾パラメータは最大9個まで同時選択することができます。複数のパラメータを選択する場合および選択したパラメータを解除する場合はキーボードの [Ctrl] ボタンを押しながらマウスでクリックしてください。

なお、各検索パラメータ項目はMASCOT Server のWeb内にあるヘルプ書類にリンクしており、その項目に関する意味や使い方を参照することができます。

検索パラメータ	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
Search title	検索タイトルを入力してください。後日、この検索結果にアクセスしやすくするた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。	sample1, spot67, Trypsin, 50fm, 2457-9001-125-afd
Database	複数のデータベースがセットアップされている場合は、検索したいデータベースを選択します。	MSDB
Taxonomy	検索対象とする生物種を選択します。	All entries
Enzyme	蛋白質を消化するために使用した消化酵素を選択します。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。	Trypsin
Allowed up to [1-9] missed cleavages	未消化のペプチドの状態を指定します。完全に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。	2
Fixed modifications	翻訳後修飾を選択します。指定したすべての(Fixされた)アミノ酸に対して翻訳後修飾を考慮して検索を行います。	Carboxymethyl (C)
Variable modifications	翻訳後修飾を選択します。指定したアミノ酸に対して翻訳後修飾を考慮する場合としない場合の2通りの(Variableな)検索を行います。	Acetyl (K) Oxidation (M)
Protein mass	検索範囲ウインドウを指定します。例えば"30"を指定した場合は30KDaの範囲に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。	ブランク
ICAT	ICATによる修飾が存在する場合にチェックします。	チェック無し
Peptide tol. +/-	ペプチド質量の許容誤差をDa, mmu, %, ppmの単位で指定します。1.0 Daを指定した場合は(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。	1Da

MS/MS tol. +/-	フラグメントの許容誤差をDa、mmuの単位で指定します。	0.5Da
Peptide charge	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。	2+
Monoisotopic/Average	質量データの種類 (MonoisotopicまたはAverage) を指定します。	Monoisotopic
Data file	読み込みたい質量データファイルを[ファイルの選択]ダイアログを通じて指定します。	データファイルを指定
Data format	質量データのフォーマットを選択します。	Mascot generic
Precursor	データファイルにプリカーサ質量の値が記載されている場合は無視されます。	ブランク
Instrument	フラグメント開裂様式は装置によって異なるため、測定に用いた装置の種類を選択します。	ESI-QUAD-TOF
Overview	チェックボックスをチェックすることにより、検索結果ページに“Overview Table”を表示します。	チェック無し
Report top	検索によりヒットした蛋白質の表示数を選択します。	20

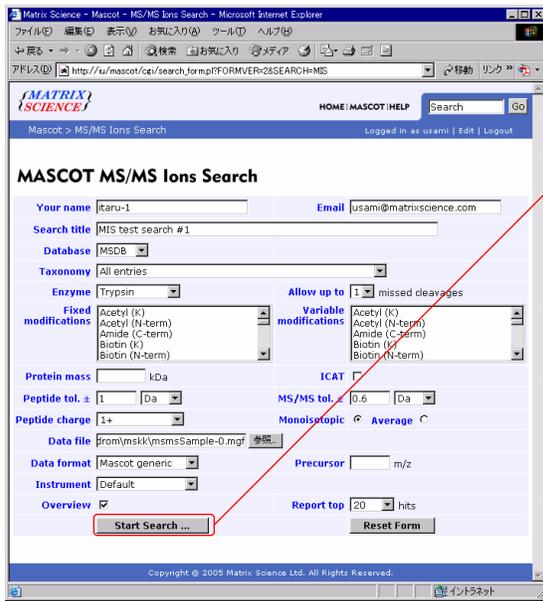
3-3. 検索パラメータの設定

次の検索パラメータを設定してください。

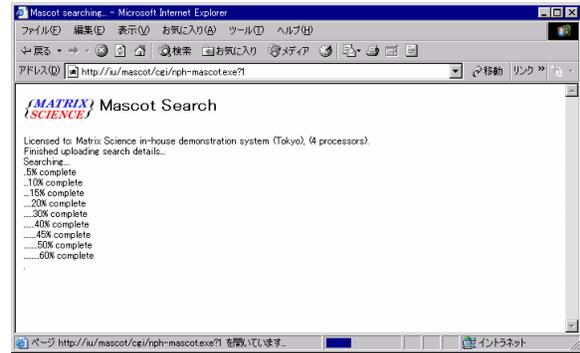
検索パラメータ	設定値	摘要
Your name	任意	ユーザ名をアルファベットで入力してください。
Search title	MIS test search # 1	
Database	MSDB	
Taxonomy	All entries	デフォルト値
Enzyme	Trypsin	
Allowed up to [1-9] missed cleavages	1	デフォルト値
Peptide tol. +/-	1.0 Da	
MS/MS tol. +/-	0.6	
Peptide charge	1+	デフォルト値
Monoisotopic/Average	Monoisotopic	デフォルト値
Data file	[参照]ボタンを押して msmsSample-0.mgfを選択	Version 2.1のCDROMのmskkフォルダにあります。5つのMS/MSデータから構成されています。
Data format	Mascot generic	デフォルト値
Instrument	Default	デフォルト値
Overview	ボックスをチェック	
Report top	20	

3-4. MASCOT 検索の実行

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、検索が終了すると検索結果レポート画面が表示されます。MS/MS Ion Searchでは2段階の検索が行われます。最初に「Peptide tol +/-」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、ヒットしたペプチドのアミノ酸配列がわかりますので、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。



検索パラメータの設定と質量データファイルの指定が完了したら、[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、検索が終了すると検索結果レポートが表示されます。



3-5. 検索結果レポートの見方

MASCOT検索が終了すると、Mascot Search Results 画面が表示されます。[Query] 番号は Peptide View 画面に、ヒットした蛋白質のアクセッション番号は Protein View 画面リンクしています。

Peptide Summary 画面にはヒットした蛋白質のスコアの状況を示すグラフに続いて、ヒットしたペプチドとそれが帰属する蛋白質がスコア順に表示されます。

蛋白質のスコアはペプチドのスコアの積算値として表示されます。スコアは統計的な意味を持ち、ヒットしたペプチド(すなわちそれが帰属する蛋白質)が有意かどうかを判定するための数値として、デフォルトでは5%の出現率をもって有意性を判断するための閾値としています。閾値以上のスコア(5%以下の出現率)を持つペプチドあるいは蛋白質が統計的に有意であり、また逆に、5%以上の出現率を持つペプチドあるいは蛋白質は統計的に有意ではない(ランダムヒットである)、という読み方をします。蛋白質のスコア状況グラフを参照することにより、有意にヒットしている蛋白質の存在を直感的に把握することができます。

Peptide View 画面を参照することにより、ヒットした各々のペプチドに関して、そのフラグメントとアミノ酸配列の一致状況を確認することができます。

< (1) Peptide Summary Report 画面 >

Mascot Search Results

User: Test
 Email: support-3p@matrixscience.com
 Search title: MS test search #1
 MS data file: C:\wmp1\data\msms\sample-0.msf
 Database: MSDB 20050227 (1942918 sequences: 629040812 residues)
 Time stamp: 30-Jun-2005 at 14:38:05 GMT

Significant hits:
 RNBPK1 DNA-directed RNA polymerase (EC 2.7.7.6) - phage K11
 QSRCJ3_PONPY Hypothetical protein DKFP4690910 - Pongo pygmaeus (Orangutan).
 LCEZA bacteriophage T7 RNA polymerase (EC 2.7.7.6), chain A - phage T7 (fragments)

Probability Based Mouse Score

Ions score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event. Individual ions scores > 48 indicate identity or extensive homology (p<0.05). Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.

閾値 48 のライン

有意なヒットが 3 件存在することを示しています。

Peptide Summary Report

Format As: Peptide Summary
 Significance threshold p<: 0.05
 Max. number of hits: 20
 Standard scoring: MudPIT scoring
 Ions score cut-off: 0
 Show sub-sets:
 Show pop-ups: Suppress pop-ups: Sort unassigned: Decreasing Score
 Require bold red:

Overview Table

Click on column header to jump to entry in results list.
 Move mouse over any indicator to highlight identical peptides.
 Click on an indicator to see details of individual match.
 Use check boxes to select sub-set of queries for new search.

Mouse over: -Query-
 -Accession- -Sequence-

Hit:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
414.56 (3+)																				
484.54 (3+)																				
490.24 (3+)																				
547.68 (3+)																				
586.99 (3+)																				

Select All Select None Search Selected Error tolerant Archive Report

1. **RNBPK1** Mass: 101064 Score: 300 Queries matched: 3
 DNA-directed RNA polymerase (EC 2.7.7.6) - phage K11
 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	414.56	1240.66	1240.66	0.00	0	44	0.19	1	F.ARNLAGHTFWK.V
2	490.24	1467.70	1467.71	-0.01	0	82	2.4e-005	1	K.APKQVVEADMISK.G
3	547.68	1639.96	1639.86	0.10	1	175	1e-014	1	R.SVNTLAYSKISLVR.Q

2. **QSRCJ3_PONPY** Mass: 194302 Score: 145 Queries matched: 1
 Hypothetical protein DKFP4690910 - Pongo pygmaeus (Orangutan).
 Check to include this hit in error tolerant search or archive report

Hitした蛋白質がスコア順に表示されます。Accession 番号は蛋白質の詳細情報を表示する Protein View 画面にリンクしています。また、[Query]番号の上のマウスカーソルを合わせると右のようなポップアップ画面が表示され、このQuery にヒットしている蛋白質一覧を見ることができます。また、[Query]番号は Peptide View 画面にリンクしています。

検索したデータベース (この場合は MSDB) のサイズ (蛋白質のエントリ数=1,942,918 件) です。

有意なスコアを持つ蛋白質情報が表示されます。

統計的に有意な閾値スコア (有意水準を 5% として、この場合は 48) を示しています。

有意水準やペプチドのスコアを変更して再表示することができます。

< (2) Peptide View 画面 >

フラグメントイオンのアサイン状況と質量誤差 (Error graph) を確認することができます。質量誤差グラフは質量分析計の質量誤差特性を表していると考えられますので、高いスコアでヒットした蛋白質でも、質量分析計の質量特性から外れている場合は注意が必要です。

Click mouse within plot area to zoom in by factor of two about that point
 Or, Plot from [100] to [1400] Da

Monoisotopic mass of neutral peptide M(calc): 1467.71
 Zine Score: 92 Expect: 1.4e-005
 Matches (Bold Red): 11/102 fragment ions using 12 most intense peaks

#	a	a++	a+	a+++	b	b++	b+	b+++	Seq.	y	y++	y+	y+++	#
1	44.05	22.53			72.04	36.53			A					13
2	191.12	96.06			219.11	110.06			F	1397.68	699.34	1380.65	690.83	12
3	322.16	161.58			330.15	175.58			M	1250.61	625.81	1233.59	617.30	11
4	450.22	225.61	433.19	217.10	478.21	239.61	461.19	231.10	Q	1119.57	560.29	1102.54	551.78	10
5	49.25	275.15	532.26	266.63	577.28	289.14	560.25	280.63	V	991.51	496.26	974.49	487.75	9
6	48.35	324.68	631.33	316.17	676.33	338.68	659.32	330.16	V	892.44	446.73	875.42	438.21	8
7	77.40	389.20	760.37	380.69	805.39	403.20	788.36	394.69	E	793.38	397.19	776.35	388.68	7
8	48.43	424.72	831.41	416.21	876.43	438.72	859.40	430.20	A	664.33	332.67	647.31	324.16	6
9	63.46	482.23	946.43	473.72	991.46	496.23	974.43	487.72	D	593.30	297.15	576.27	288.64	5
10	94.50	547.75	1077.47	539.24	1122.50	561.75	1105.47	553.24	M	478.27	239.64	461.24	231.13	4
11	07.58	604.30	1190.56	593.78	1235.58	618.29	1218.55	609.78	I	347.23	174.12	330.20	165.60	3
12	94.62	647.81	1277.59	639.30	1322.61	661.81	1305.59	653.30	S	234.14	117.58	217.12	109.06	2
13									K	147.11	74.06	130.09	65.55	1

蛋白質のアクセッション番号は Protein View 画面にリンクしています。Protein View 画面では、蛋白質の詳細情報 (全体アミノ酸配列、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、質量誤差グラフ、アノテーション情報、NCBIサイトのBLAST検索やTaxonomy Browserへのリンクなど) を表示する他、Peptide View 画面へのリンクも張られています。

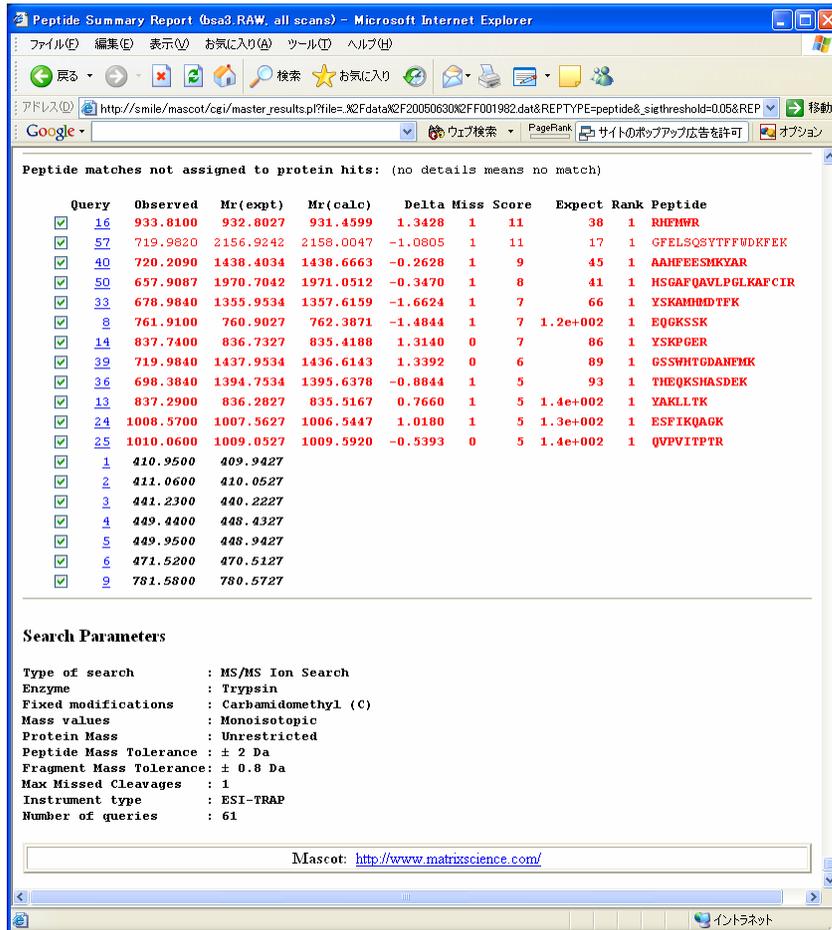
3-6. 検索条件の変更と再検索

Peptide Summary Report 画面の [Search Selected] ボタンを押すことにより、MS/MS Ion Searchの検索パラメータ設定画面に戻りますので、パラメータ設定を変更して再検索することができます。

3-7. Error Tolerant Search

(非特異的なペプチドの切断+1塩基置換+全ての翻訳後修飾を考慮した再検索)

複数のMS/MSデータで構成された質量データに対してMASCOT検索を実行した場合、ヒットした蛋白質に帰属しないMS/MSデータは“Peptide matches not assigned to protein hits”としてリスト表示されます。



リスト表示されたMS/MSデータが、MASCOT検索によってヒットした蛋白質に帰属しない理由としては、

- (1) ペプチド処理に用いた消化酵素以外の非特異的な切断が起きた
- (2) サンプル蛋白質にSNPs など1塩基置換・欠失・挿入が起きていて、DBに登録されている配列と異なる
- (3) 検索時に指定しなかった翻訳後修飾を持っている
- (4) スペクトルが不十分で、検索に十分なマッチングが得られない
- (5) サンプルがまったく新規のタンパク質由来のペプチドで、マッチングが得られない

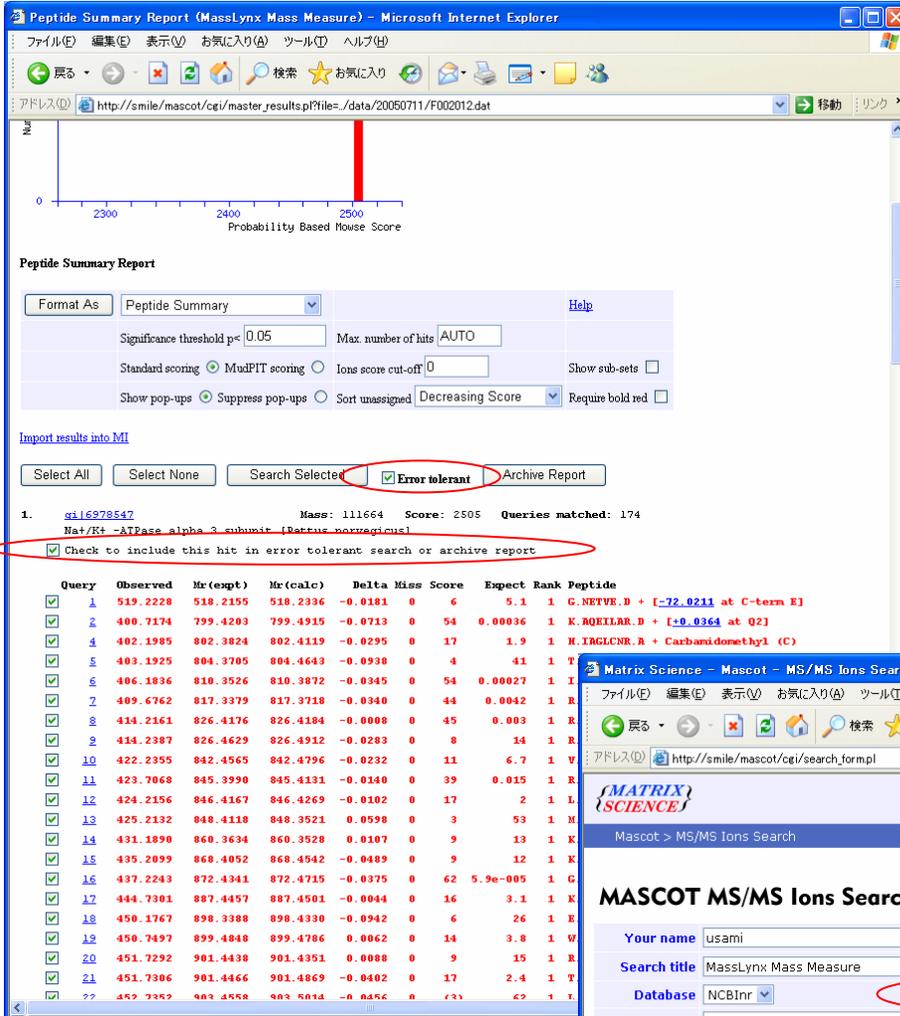
が上げられますが、MASCOT Serverでは、選択したヒット蛋白質に対して、(1)、(2)、(3) を考慮した再検索、 Error Tolerant Search を実行することができます。

【Error Tolerant Search の実行】

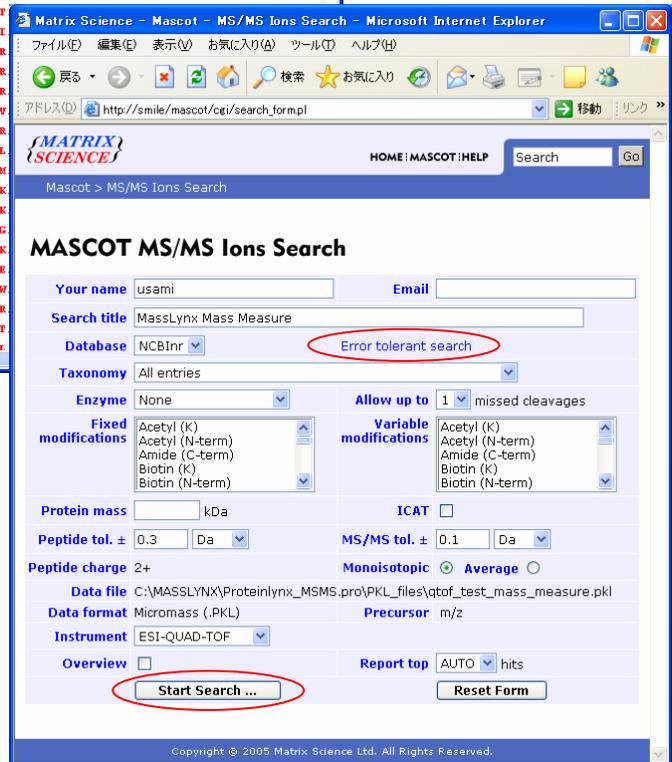
Mascot Search Results 画面を Peptide Summary 表示にした後、Error tolerant のチェックボックスをチェックしてください。Error Tolerant Search が有効になります。ヒット蛋白質名の下にある、

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、Error Tolerant Search を実行したい蛋白質を選択してください。複数の蛋白質を選択することができます。[Search Selected] ボタンを押してください。Error Tolerant Search 画面に切り替わります。



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、Databaseやmodificationsなどの検索条件は変更せずに、[Start Search...]ボタンを押してください。Error Tolerant Search が開始されます。



1塩基置換によるアミノ酸の変化、MASCOTの設定ファイルに含まれる翻訳後修飾、非特異的な切断を考慮した Error Tolerant Search の結果が表示されます。各MS/MSデータに対するヒット情報を示す行の後半に示されている青色の数値はデータベース中のアミノ酸配列に対する質量差を表しています。また、青色数値の上にマウスカーソルを合わせると、詳細情報がポップアップ表示されます。

Peptide Summary Report

Format As: Peptide Summary | Help

Significance threshold p < 0.05 | Max. number of hits 20

Standard scoring | MudPIT scoring | Ions score cut-off 0 | Show sub-sets

Show pop-ups Suppress pop-ups Sort unassigned: Decreasing Score | Require bold red

[Import results into MI](#)

Select All | Select None | Search Selected | Error tolerant | Archive Report

1. [DPOD_HUMAN](#) | Mass: 123557 | Score: 884 | Queries matched: 1126
 (P28340) DNA POLYMERASE DELTA CATALYTIC CHAIN (E)

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 1	354.7800	707.5454	707.3020	0.2434	0	11	17	1	C.QRCQGS.L + [+30.0106 at G5]
<input checked="" type="checkbox"/> 2	354.7991	707.5836	709.3707	-1.7870	0	(15)	7.3	1	R.KGVARYM.K + [-29.0027 at C-term M]
<input checked="" type="checkbox"/> 3	354.8300	707.6454	709.3871	-1.7416	0	11	17	1	R.QVPVGGP.P + [+59.0483 at P2]
<input checked="" type="checkbox"/> 4	354.8510	707.6873	706.3608	1.3266	0	15	7.2	1	Q.TGRRDT.K + [+2.0042 at N-term T]
<input checked="" type="checkbox"/> 5	354.9399	707.8653	707.3466	0.5187	0	9	32	1	R.ELTGPAV.L + [+21.9819 at C-term V]
<input checked="" type="checkbox"/> 6	354.9600	707.9054	707.3272	0.5783	0	15	8.1	1	E.LCSRES.M + [+14.0156 at N-term L]
<input checked="" type="checkbox"/> 7	354.9882	707.9619	709.2999	-1.3380	0	(10)	24	1	R.KRRNCC.I + [-69.0691 at P2]
<input checked="" type="checkbox"/> 8	355.0394	708.0643	709.2999	-1.2356	0	14	9.9	1	R.KRRNCC.I + [-69.0691 at P2]
<input checked="" type="checkbox"/> 9	355.0496	708.0847	709.4213	-1.3366	0	(4)	1e+002	1	I.LGAKVQS.Y + [+6.0156 at P2]
<input checked="" type="checkbox"/> 10	355.0530	708.0915	707.1968	0.8947	0	11	17	1	F.YTPRP.P + [+79.0156 at P2]
<input checked="" type="checkbox"/> 11	355.2352	708.4558	708.4507	0.0052	0	4	90	1	G.KRRPGP.G + [-0.0156 at P2]
<input checked="" type="checkbox"/> 12	356.7731	711.5316	712.3867	-0.8552	0	9	33	1	L.KEKATQ.C + [+9.0156 at P2]
<input checked="" type="checkbox"/> 13	356.8100	711.6054	712.3431	-0.7377	0	12	16	1	G.DEFVK.T + [+76.0313 at C-term K]
<input checked="" type="checkbox"/> 14	356.8400	711.6654	711.3187	0.3467	0	10	26	1	P.EGAVRHA.Q + [+58.0055 at R5]
<input checked="" type="checkbox"/> 15	356.8701	711.7257	710.2040	1.5217	0	10	27	1	G.PPGPE.R + [+214.9711 at N-term P]
<input checked="" type="checkbox"/> 16	356.8714	711.7282	713.3820	-1.6538	0	11	21	1	S.APSLGR.V + [-0.9840 at D6]
<input checked="" type="checkbox"/> 17	356.8962	711.7779	712.3292	-0.5514	0	15	7.6	1	T.VLSHQG.R + [+72.9952 at L2]
<input checked="" type="checkbox"/> 18	356.9307	711.8469	710.3347	1.5122	0	10	24	1	G.DPPSRGS.V + [+13.9793 at P2]

Possible Assignments:

- Methyl (N-term) [+14.0156]
- N-Methyl (Protein) [+14.0156]
- L->Q [+14.9745]

蛋白質同定システム **MASCOT Server**

4. LCQ_DTA Shell (Xcalibur RAW ファイルに対する検索)

サーモエレクトロン社は、Xcaliburが出力する質量データファイルである*.RAW (バイナリ形式) を*.DTA ファイル(テキスト形式)に変換するためのプログラム “extract_msn.exe (LCQ_DTA.EXE) ” を提供しています。*.DTAファイルを使ってMASCOT検索を実行することもできますが、MASCOT Serverはextract_msn.exe (LCQ_DTA.EXE) プログラムのインターフェースである “LCQ_DTA Shell” を内蔵しており、次のような流れで*.RAWファイルから直接MASCOT検索を実行することができます。

- (1) LCQ_DTA Shellユーザインターフェースの表示
- (2) *.RAWファイルの指定
- (3) LCQ_DTA.EXEプログラムの条件設定
- (4) RAW → DTAへの変換実行
- (5) *.DTAファイル生成 (中間ファイルとして存在)
- (6) MASCOT ServerのMS/MS Ion Search検索設定画面の[Data file]欄への自動入力
- (7) MASCOT検索実行
- (8) MASCOT検索結果画面の表示

4-1. 環境設定

MASCOT ServerをインストールしたPCに次のフォルダを作成してください。

```
C:¥LCQ¥system¥programs
C:¥temp
```

BioWorksのCDROMまたはBioworksがインストールされているPCから次のファイルを検索し、MASCOT ServerをインストールしたPCの “C:¥LCQ¥system¥programs” フォルダにコピーしてください。

```
extract_msn.exe (*注)
Registry.dll
Fileio.dll
```

(*注) 存在しない場合は 同じフォルダ内のLCQ_DTA.EXEを使用してください。

LCQ_DTA.EXEプログラムにはいくつかのバージョンが存在し、上記の2つのファイルだけでは動作しない場合があります。この場合は、“Fglobal.dll”、“Fcontrol2.dll”、“MFC42U.dll” を検索し、MASCOT ServerをインストールしたPCの “C:¥LCQ¥system¥programs” フォルダにコピーしてください。

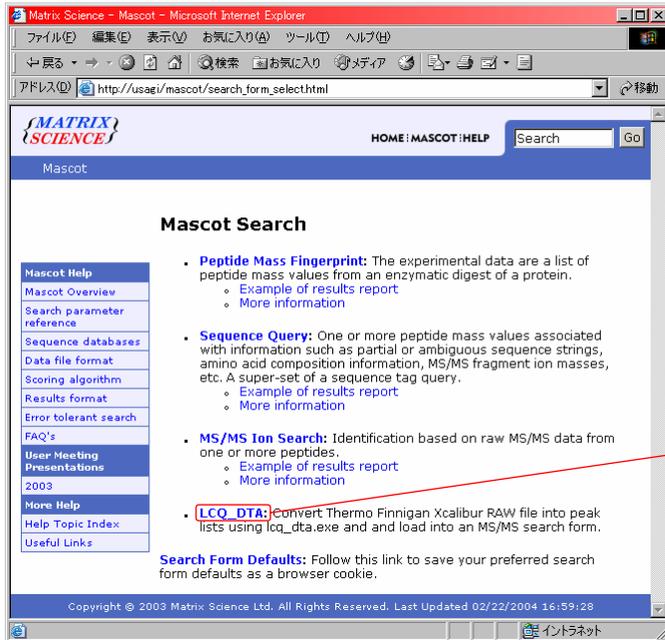
4-2. DTA ファイルへの変換と MASCOT 検索の実行

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://MASCOT_Server名/mascot/
```

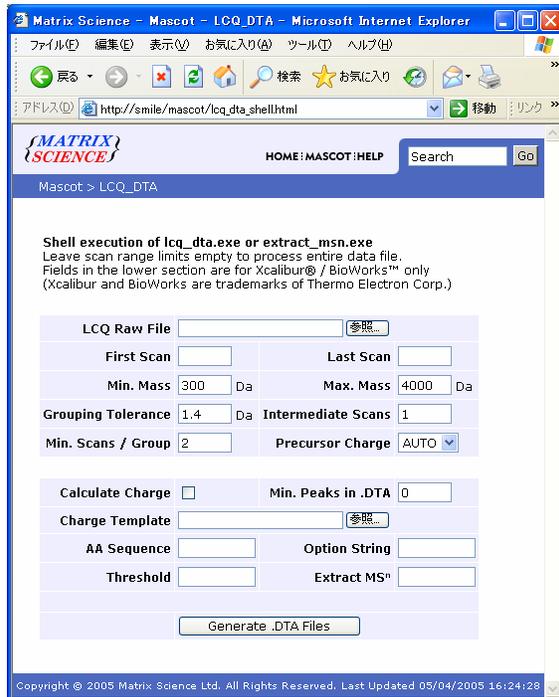
MASCOT Serverのトップページに接続されます。表示されたトップページ画面の **Welcome** の次の行にある **Mascot** のリンク、またはページ最上段の HOME | MASCOT | HELP の中の **MASCOT** のリンクをクリックしてくだ

さい。MASCOT Serverの検索モード選択画面 (Mascot Search画面) が表示されます。



LCQ_DTA プログラムのユーザインターフェース画面にリンクしています。

LCQ_DTAのリンクをクリックしてください。LCQ_DTA.EXEプログラムを動作させるために必要な変換パラメータ値を入力できるユーザインターフェース (LCQ_DTA Shell画面) が表示されます。



LCQ Raw File項の[参照]ボタンを押し、RAWファイルデータを選択してください。次に、LCQあるいはLTQの測定条件に応じて、RAWファイルからDTAファイルへの最適な変換パラメータ値を設定してください。「Min. Scan / Group」のデフォルト値は2ですが、測定条件によっては1を設定しないとLCQ_DTA.EXEプログラムがエラーを返すことがあります。

LCQ_DTA.EXEプログラムの変換パラメータ設定に関しては、

http:// MASCOT_Server名/mascot/help/instruments_xcalibur.html

に若干の説明がありますが、詳しい内容に関しましては、LCQ_DTA.EXEプログラムの供給元であるサーモエレクトロン株式会社にお問い合わせください。

変換パラメータ値の設定が終わりましたら、画面の最下段にある [Generate .DTA Files] ボタンを押してください。RAWファイルがDTAファイルに変換され、MS/MS Ion Search画面に切り替わりますので、検索パラメータを設定し、[Start Search ...] ボタンを押してMASCOT検索を実行してください。

The screenshot shows the MASCOT MS/MS Ions Search web interface. The search parameters are as follows:

- Your name: test
- Email: test@matrixscience.com
- Search title: CytC_test.RAW, all scans
- Database: MSDB
- Taxonomy: All entries
- Enzyme: Trypsin
- Allow up to: 1 missed cleavages
- Fixed modifications: Biotin (K), Biotin (N-term), Carbamidomethyl (C), Carbamyl (K), Carbamyl (N-term)
- Variable modifications: NIPCAM (C), O18 (C-term), Oxidation (M), Oxidation (HW), PEO-Biotin (C)
- Protein mass: [] kDa
- ICAT:
- Peptide tol. ±: 1.5 Da
- MS/MS tol. ±: 0.8 Da
- Peptide charge: 2+
- Monoisotopic: Average:
- Data file: C:\LCQ\system\programs\CytC_test.RAW
- Data format: Mascot generic
- Precursor: m/z
- Instrument: ESI-TRAP
- Overview:
- Report top: AUTO hits

Two red callout boxes provide additional information:

- One box points to the "Search title" field, stating: "RAW ファイル名と変換された Scan 範囲に関する情報が自動的に挿入されます。"
- Another box points to the "Data file" field, stating: "LCQ_DTA.EXE プログラムで変換された DTA ファイルが自動的に読み込まれます。"

なお、LCQ_DTA.EXEプログラムはモノアイソトピックなピークを検出するためのソフトウェアではありません。RAWファイルのプロファイルがそのままDTAファイルとして作成されることにご注意ください。

蛋白質同定システム MASCOT Server

5. 付録

5-1. 蛋白質同定のコツその1

5-1-1. 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ェアを利用する。「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ェアツールを例にして、蛋白質同定をより確実にを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。

5-1-2. 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。「Oxidation (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。

5-1-2. 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。

5-1-4. 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。

5-1-5. 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ェアツールであり、

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ラク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

5-1-6. Taxonomyフィルタ

実験対象とした生物種に含まれていない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NCBIInr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。

5-1-7. 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。

5-1-8. 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ェア「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

5-2. 蛋白質同定のコツその 2

5-2-1. 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ェアツール MASCOT ファミリー を例にして、蛋白質同定をより確実にを行うためのコツを紹介する。

5-2-2. 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。

Error tolerant

1. [ELM1_YEAST](#) **Mass:** 72105 **Score:** 96 **Queries matched:** 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
<input checked="" type="checkbox"/> 2	1084.56	1083.55	1083.53	0.02	0	96	5.6e-009	1	K.RDFVTDGFK.L

Peptide matches not assigned to protein hits: (no details means no match)

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 1	748.43	747.42	747.42	0.01	0	11	2.7	1	DLYPIK
<input checked="" type="checkbox"/> 3	1367.63	1366.62							

(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 **Queries 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
<input checked="" type="checkbox"/> 1	748.43	747.42	747.39	0.03	0	53	8.7e-005	1	K.IPTPIK.A + [+79.97 at T3]
<input checked="" type="checkbox"/> 2	1084.56	1083.55	1083.53	0.02	0	96	9.3e-009	1	K.RDFVTDGFK.L
<input checked="" type="checkbox"/> 3	1367.63	1366.62	1366.63	-0.01	0	118	7.8e-011	1	K.PSSPLMDRTVGK.R + [+79.97 at S3]

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。

5-2-3. 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
<input checked="" type="checkbox"/> 2	687.30	1372.59	1267.60	104.98	0	42	0.018	1	R.GYSLGNWVCAAK.F
<input checked="" type="checkbox"/> 1	877.40	1752.79	1752.83	-0.04	0	35	0.00019	1	R.NTDGSTDYGILQINSR.W

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

2005年7月 第3版

Copyright 2004, 2005 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。