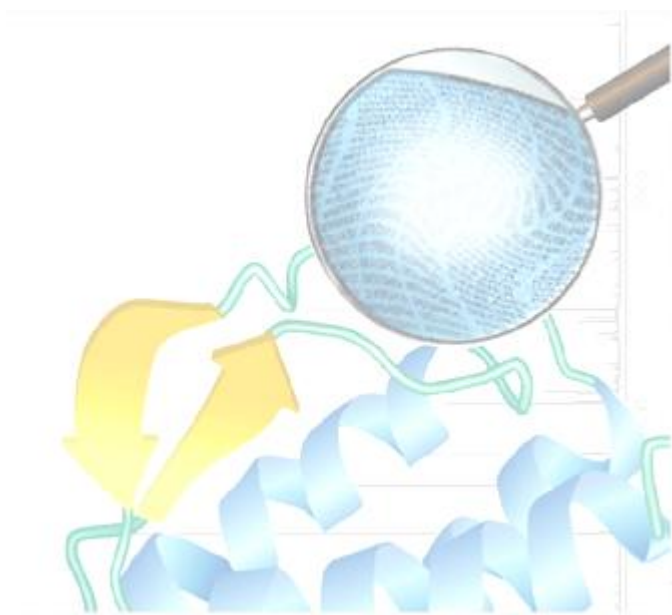

MASCOT *Distiller*

定量解析クイックスタート



1. 定量解析プロトコルの分類

質量データを利用した蛋白質の定量解析手法に関しては次のレビューに良くまとめられており、参考になります。

Ong, S. E. and Mann, M., Mass spectrometry-based proteomics turns quantitative, *Nature Chemical Biology* 1 252–262 (2005)

それぞれの手法で利用されている定量解析のメカニズムは様々ですが、これらは次に示す基本的な定量解析手順（プロトコル）に分類することができます。

(1) Reporter プロトコル

MS/MS スペクトルを構成する特定のイオンピークの強度比を利用して定量する方法です（代表例：iTRAQ、Tandem Mass Tag）。

(2) Precursor プロトコル

MS スペクトルを構成するプリカーサイオンの XIC の強度比を利用して定量する方法です。プリカーサイオンの質量変化を引き起こす化学的な処理を利用するため、多くの種類が存在し、最も広く使われている手法です（代表例： ^{18}O 、AQUA、ICAT、ICPL、Metabolic、SILAC）。

(3) Multiplex プロトコル

MS/MS スペクトルを構成するペプチド・フラグメントイオンの強度比を利用して定量する方法です。ペプチド末端に対する化学的処理による質量変化を利用します（代表例： ^{18}O 、SILAC）。

(4) Replicate プロトコル

複数の質量データ間におけるプリカーサイオンの XIC の強度比を利用する、ラベルフリーな定量方法です。

(5) emPAI プロトコル

Mascot 検索によって有意にヒットしたペプチドの数を利用する、ラベルフリーな定量方法です。

(6) Average プロトコル

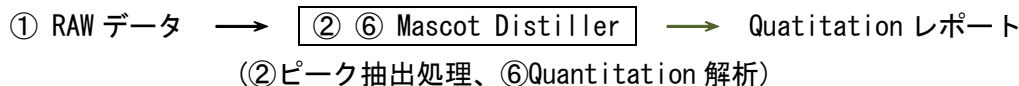
Mascot 検索によって有意にヒットしたペプチドに対応するプリカーサイオンの XIC の強度比を利用する、ラベルフリーな定量方法です。

これらのプロトコルのうち、(1) Reporter プロトコル、(3) Multiplex プロトコル、(5) emPAI プロトコルでは、定量解析に必要な情報がピークリストあるいは Mascot 検索結果に存在しますので、Mascot Server に組み込まれている定量解析プログラムを使って処理を行い、定量解析結果は検索結果ページに表示されます。また、これらのプロトコルでは、ノイズ成分を含まない、高品質なピークリストを必要としますので、RAW データから意味のあるピークを精度良く抽出することができる Mascot Distiller が必須となります。なお、emPAI プロトコル解析機能は常に ON の状態にあり、100 個以上の MS/MS スペクトルに対する Mascot 検索に対して自動的に emPAI の値を計算し、検索結果ページに表示します。

(2) Precursor プロトコル、(4) Replicate プロトコル、(6) Average プロトコルでは、MS スキャン (Survey スキャン) のプリカーサイオン領域に対応する XIC 情報を利用しますので、質量データとしてピークリストではなく RAW データを使い、次のような処理プロセスで定量解析を行います。

④ Mascot Server

(③ ピークリストと検索条件) ↑ ↓ (⑤ Mascot 検索結果)



- ① 質量データとしてピークリストではなく RAW データファイルを準備します。
- ② Mascot Distiller で RAW データファイルを読み込み、ピークを抽出します。
- ③ Mascot Distiller から Mascot Server に検索投入します。
- ④ Mascot Server は Mascot 検索を実行します。
- ⑤ Mascot 検索終了後、Mascot Distiller は Mascot 検索結果を自動的に読み込みます。
- ⑥ Mascot Distiller で Quantitation 解析を実行します。

RAW データとしてプロファイルデータを使用してください。セントロイドデータの場合は同位体分布が正確に記録されていないことがあり、一般的に良い定量結果を得ることができません。なお、Orbitrap や FTMS などの解像度が高い質量分析計の場合はセントロイドデータであっても同位体分布が正確に記録されているため、問題ありません。

2. 定量解析前の準備

Precursor プロトコル、Replicate プロトコル、Average プロトコルでは RAW データと Mascot 検索結果を利用するため、Mascot Distiller (Core + Search Toolbox + Quantitation Toolbox) と Mascot Server 2.2.06 以降が必要です。また、Mascot Daemon を利用するとピーク抽出処理から Mascot 検索までのプロセスを自動化することができます (Daemon Toolbox が必要です)。

「Mascot Distiller Support」ページに Mascot Distiller のアップデートや Quantitation Method の更新ファイル等の情報を掲載していますので定期的にチェックしてください。

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

Mascot Distiller の使い方や解析結果レポートの見方を習得するための解析例を用意してあります。詳しくは、「3. Precursor プロトコル：SILAC の解析例」を参照してください。

定量解析プロセスは、ピークの抽出、定量プロトコルを加味した Mascot 検索、定量解析、解析結果レポート作成などの処理要素で構成されていますが、これらの処理要素は多数の処理パラメータで管理されています。処理パラメータは“Quantitation Method”としてまとめられており、“SILAC K+6 R+10 [MD]”のようなユニークな名前が付けられています。詳しくは Mascot Distiller のヘルプページ (メニューバー [Help] → [Mascot Distiller Help] → [Quantitation Toolbox] の [Method Configuration - Precursor] 等) を参照してください。

3. Precursor プロトコル : SILAC の解析例

3成分SILAC(Stable Isotopic Labeling using Amino Acids in Cell Culture)法による実験サンプルから得られた質量データ(Applied Biosystems社のQSTARで測定)の解析例を使って Mascot Distiller の使い方を説明します。次のファイルをダウンロードし、同じフォルダに解凍してください。

(1) 質量データ RAW ファイル(wiff 形式)

<http://www.pil.sdu.dk/1/MSQuant/QSTARSampleDataSet1,2003-09-15.zip>

(解凍すると「Triple Encoding SILAC sample data.wiff」ファイルになります)

(2) Mascot Distiller プロジェクトファイル(定量解析結果ファイル)

http://www.matrixscience.com/downloads/Triple_SILAC_Distiller_Project.zip

(解凍すると「Triple Encoding SILAC sample data.rov」ファイルになります)

Mascot Distiller を起動し、[File]→[Open Project]ダイログから Mascot Distiller のプロジェクトファイルである「Triple Encoding SILAC sample data.rov」ファイルを選択してください(「Triple Encoding SILAC sample data.rov」ファイルをダブルクリックしても開くことができます)。関連ファイルが読み込まれ、次のような画面構成になります。

The screenshot displays the Mascot Distiller software interface with three main windows highlighted by yellow boxes:

- Data Explorer:** Located on the left side, it shows a tree view of protein identifications and search results.
- Quantitation Table:** A central table displaying protein identification data. The table includes columns for Accession, Score, Mass, M/L, SD(geo), H/L, and Description. The first few rows are:

Accession	Score	Mass	M/L	SD(geo)	H/L	SD(geo)	Description
e 4607357	743	2248	1.0273	1.0874	7	0.8959	transin 2 [Homo sapiens]
e 119090	636	2278	0.9943	1.0744	8	0.8283	Peptide-prolyl 4c-methyl isomerase B precursor (PPIase)(Proteinase) (Cyclophilin B) ...
e 4606591	515	2014	0.9763	NN	3	0.8667	peroneus-1 [Homo sapiens]
e 254831	444	2385	0.9840	1.0295	3	0.9741	Chan A Crystall Structure Of Human Glutathione S-Transferase
e 294007	408	2395	0.9848	1.0295	3	0.9741	glutathione S-transferase [Homo sapiens]
e 3955007	337	2190	0.9730	1.1144	5	0.8699	Chan A, Thioester Peroxidase B From Red Blood Cells
e 4125506	280	2283	1.1077	1.0132	2	1.1077	transin [Homo sapiens]
e 1197408	263	2362	0.9570	1.1812	2	0.7660	lysophospholipase [Homo sapiens]
e 13159	232	2107	0.9553	1.0742	3	0.8384	neurotrophin-3 [human, brain, peptide, 186 aa]
e 3447913	229	2370	0.9160	1.0081	2	0.9503	RAB7, member RAB oncogene family [Homo sapiens]
e 301335	220	1919					collin 1 (non-muscle) [Homo sapiens]
e 478988	215	2281					RAB1, member RAB oncogene family [Homo sapiens]
- TIC Window:** A Total Ion Chromatogram (TIC) plot showing intensity versus time (minutes). The x-axis ranges from 0 to 600 minutes, and the y-axis represents intensity.
- Scan Window:** A mass spectrum plot showing relative intensity versus mass-to-charge ratio (m/z). The x-axis ranges from 450 to 1450 m/z, and the y-axis represents relative intensity from 0 to 70. A prominent peak is visible at approximately 550 m/z.

画面は4つのウィンドウで構成されています。左から時計回りに、[Data Explorer]、[Quantitation Table]、[TIC Window]、[Scan Window]です。

[Data Explorer] には複数のデータがまとめられており、下方にあるタブを選択することによりデータの表示を切り替えることができます。 [Proteins] タブには、Mascot 検索によってヒットしたタンパク質がプロテインスコア順に表示されています。また、先頭の (+) 印をクリックすると、そのタンパク質に帰属するペプチドリフトに展開され、Mascot 検索結果の Select Summary レポートと同じ内容が表示されます。タンパク質名を右クリックすることによりブラウザが起動し、Mascot 検索結果の Protein View ページを表示させることができます。同様に、タンパク質に帰属するペプチドを右クリックすることにより Peptide View ページにアクセスすることができます。これに対して、[Peak Lists] タブには、MS/MS スペクトルに対する検索結果(ペプチドのヒットリスト、イオンスコア順に最大 10 件)が表示されます。

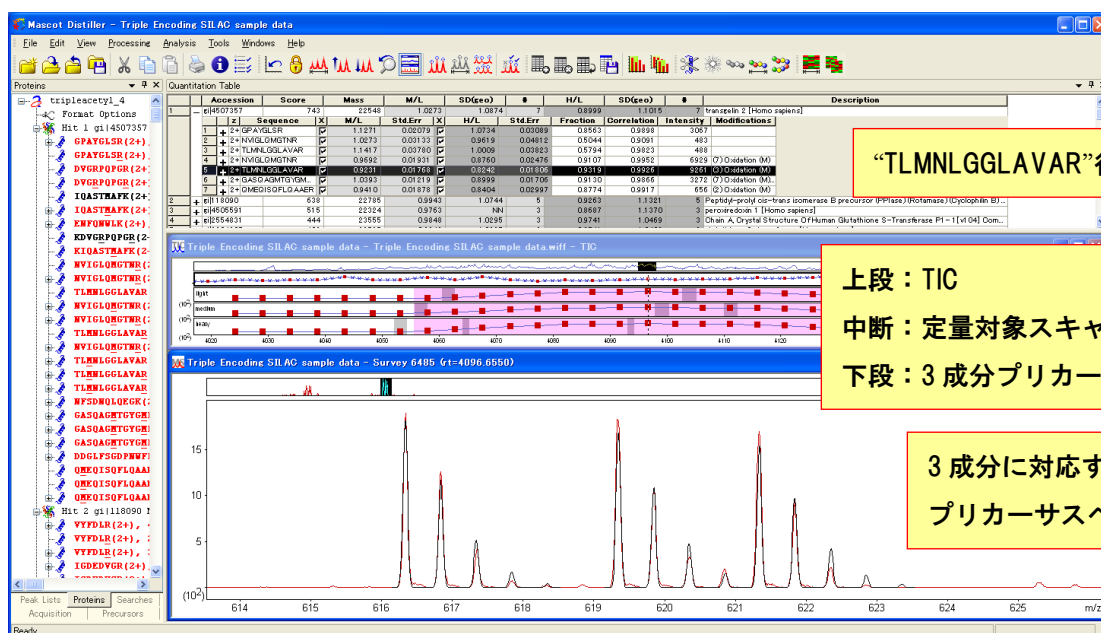
[Quantitation Table] には定量解析結果がスプレッドシートの形式で(ただし 2 次元ではなく多次元のデータを表示しています) ヒットタンパク質のプロテインスコア順に表示されます ([Data Explorer] の [Proteins] タブにおけるタンパク質と同じ順序です)。この例では SILAC 処理を行っていないサンプルと 2 種類の SILAC (R+6 と R+10) で処理したサンプルを混合した、3 成分サンプルを使用していますので、SILAC ラベルのないペプチドを基準にした 2 種類の成分比率、

M/L = Middle/Light = Label:13C (6) / ラベルなし

H/L = Heavy/Light = Label:13C (6) 15N (4) / ラベルなし

を表示しています。

タンパク質アクセッション番号左にある (+) 印をクリックすると、次の画面のようにそのタンパク質に帰属するペプチドに関する Quantitation 情報が表示されます。



“TLMNLGGLAVAR” 行を選択

上段 : TIC
 中断 : 定量対象スキャン領域
 下段 : 3 成分プリカーサの XIC

3 成分に対応する
 プリカーサスペクトル

ペプチド(アミノ酸配列)の右横にあるチェックボックス(□)にチェック(✓)が入っていない場合は、そのペプチドはタンパク質の比率計算には使われていないことを示しています。さらに、ペプチドの左にある(+)印をクリックすると、そのペプチドにマッチしたMS/MSデータ情報を確認することができます。

ペプチドの Quantitation 情報表示行(ペプチドのアミノ酸配列が表示されている行)をクリックすると[TIC Window]および[Scan Window]の表示は Quantitation モードとなり、次の画面のように選択したペプチドに対応するプリカーサ情報が表示されます。すなわち、定量対象となるスキャン領域、各スキャンにおいて定量対象となるプリカーサの XIC(Extracted Ion Chromatogram)のプロット、選択したスキャンにおいて定量対象となるプリカーサのスペクトルが表示されます。

Quantitation 解析結果は情報量が多く、表示されるウインドウの数が多いため、より大きなスクリーンを使用して、ウインドウのサイズを大きくしたり、ウインドウの配置を変更することにより、ストレスなく解析することができます。たとえば、[Windows]→[Auto Arrange TIC/Scan]を選択し、[TIC Window]と[Scan Window]の配置を自動的に調整する機能(デフォルトでは有効)を無効にした上で、[Windows]→[Tile Vertically]を選択してください。右図のように[TIC Window]と[Scan Window]が縦長に配置され、より見やすくなります。

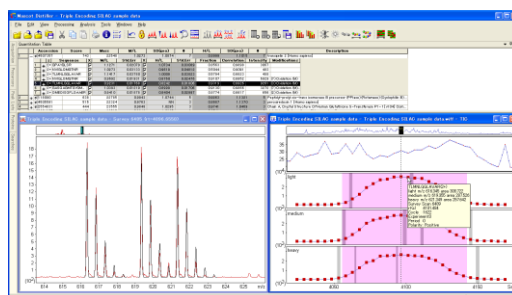
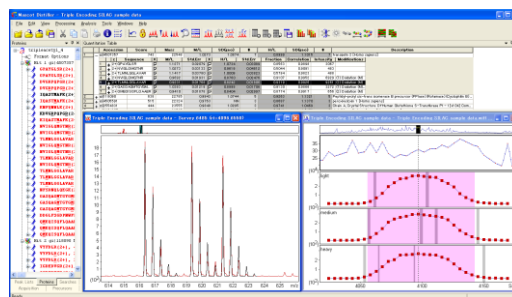
さらに、[Data Explorer]の右肩にあるピンのアイコンをクリックして画面の左端や左上のタブとして格納し、右図のようなウインドウ配置にすることもできます。

[TIC Window]には上から順に、「TIC スペクトル」、「定量対象なるスキャン領域スペクトル」、3つの成分に対応するプリカーサイオンの「XICプロット」が表示されています。各 XICプロットは左肩にある「light」、「medium」、「heavy」のラベルで区別されています。スキャンは■印で表示しています(それに対してMS/MS スキャンは×印で表示しています)。

[定量対象なるスキャン領域スペクトル]の■をクリックすると、そのスキャンに対応する Survey scan スペクトルが[Scan Window]に表示されます。また、■にマウスマウスカーソルを合わせるとそのスキャンの情報を示すウインドウ(薄黄色地の Tooltip)が表示されます。

同様に、[XICプロット]の■をクリックすると、そのスキャンにおける3成分のプリカーサイオンのスペクトルが[Scan Window]に表示されます。また、■にマウスマウスカーソルを合わせるとその3成分のプリカーサイオンの XIC 情報を示すウインドウ(薄黄色地の Tooltip)が表示されます。

XICプロットのピンク領域は XIC のピークを示しており、Quantitation 解析に使われたスキャ



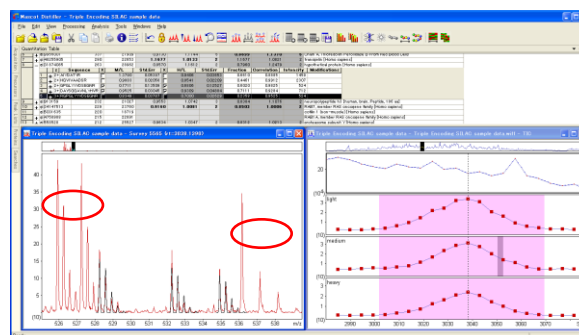
ン領域を示しています。また、グレーの帯(ピンク領域の中では紫色に見えます)は Mascot 検索でヒットした MS/MS スキャンを示しています。定量解析の対象領域に属する全ての MS/MS スペクトルに対して Mascot 検索でヒットする必要はありません。Mascot 検索でヒットしたプリカーサ成分がひとつでも存在すればそれと組をなす他のプリカーサ成分を自動的に探し出します。

[Scan Window]の赤線はRAW データを示しています。それに対して、黒線はレポートされているペプチドの比率に基づく予測ピークを示しています。なお、この例では、「medium」と「heavy」に対応するプリカーサの同位体分布が互いに重なっていますが、このような場合はそれらの重なりを考慮して各プリカーサの占有領域を計算しています。

実験から得られたピークと理論的に計算したピークが完全に同じになることはなく、通常は余計なピークが混じった状態でスペクトルが構成されています。質の悪いスペクトルから計算されたペプチドの比率はタンパク質の比率計算の際に無視することが必要ですが、Mascot Distiller では次の3つの閾値を使って無視するかどうかを判定しています。

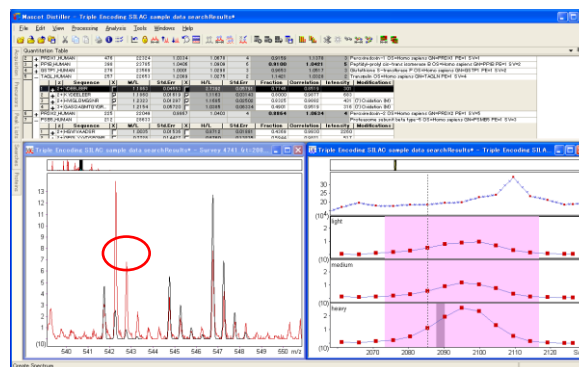
(1) ピーク領域の面積比(Fraction of peak area)

[Quantitation Table]では「Fraction」項として表示されます。[Scan Window]のピーク領域全体面積に対するプリカーサピーク部分の面積比です。右図の例では、定量対象となっているプリカーサピークの前後に強度の強い余計なピーク(右図の赤丸で示すピーク領域)が存在するため、ピーク領域全体面積に占める定量対象プリカーサピーク部分の面積の比率は小さく、XIC プロット全体で積算しても 25%程度にしかありませんので、「Fraction」閾値を 0.5(50%)に設定した場合、このペプチドの比率はタンパク質の比率計算から除外されます。すなわち、ペプチドの右横にあるチェックボックス(□)にはチェック(✓)が入りません。



(2) 相関係数(Correlation coefficient)

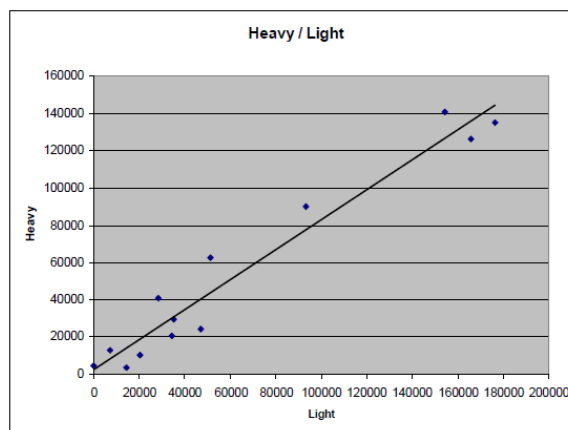
[Quantitation Table]では「Correlation」項として表示されます。右図の例のように、プリカーサピークの同位体分布に重なるようにして余計なピークが存在する(右図の赤丸で示すピーク領域)場合は(1)の「Fraction」条件では判定できませんので、ピークの形(実験と理論値の同位体分布)に対する相関係数を「Correlation」閾値として採用しています。



この例では相関係数は0.85になりますが、「Correlation」 閾値を0.9に設定した場合、このペプチドの比率はタンパク質の比率計算から除外されます。

(3) 標準誤差 (Standard error)

[Quantitation Table]では「StdErr」 項として表示されます。計算されたペプチドの比率に対する標準誤差です。ペプチドの比率はXICピークにおける各プリカーサイオンの強度に対して最小自乗法フィッティングにより求めています。右の例では、各ポイントはXICピークを構成するスキャンのLight及びHeavyに対応するプリカーサイオン強度をプロットしています。最小二乗法によりフィッティングして得られた直線の傾きがペプチドの比率の最良推



定値になりますので、「StdErr」 閾値を0.1に設定した場合、このフィッティングに対する標準誤差が0.1よりも大きいペプチドの比率はタンパク質の比率計算から除外されます。

メニューバー[Analysis]→[Quantitation Reports]のサブメニューから定量解析結果レポートを項目別に表示させたり、XMLファイルとして出力することができます。また、メニューバー[Edit]→[Copy Quantitation Table]を選択すると、[Quantitation Table]のデータがクリップボードにコピーされますので、表計算ソフトウェアのスプレッドシートにペーストすることができます。これらのデータの取扱等に関しては、Mascot Distillerのヘルプページ(メニューバー[Help]→[Mascot Distiller Help]→[Quantitation Toolbox])の[Reporting and Exporting]を参照してください。

4. 定量解析の流れ

Mascot Distiller と Mascot Server を利用した定量解析プロセスは単純ですが、次のことに注意する必要があります。

1. 質量 RAW データの品質
2. ピーク抽出条件
3. 定量解析条件
4. Mascot 検索条件

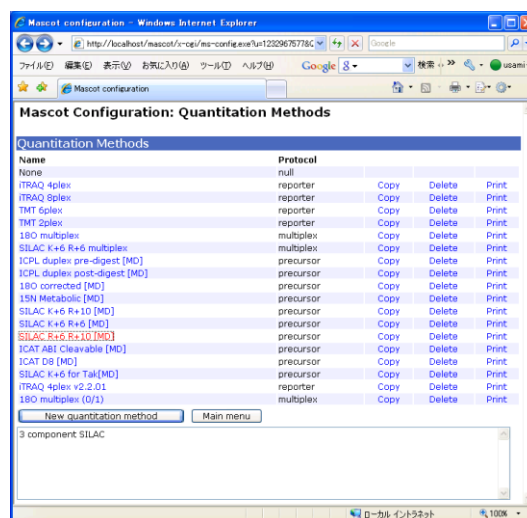
定量解析プロセスの中で一つでも不適切な条件が含まれると、信頼できる定量解析結果を得ることはできません。

RAW データから定量解析結果レポート得るまでのプロセスは概ね次のようになります。

(1) Quantitation Method(定量解析条件)の設定

定量解析処理パラメータは「Quantitation Method」としてまとめられており、“SILAC K+6 R+10 [MD]”のような名前が付けられています。定量解析を加味した Mascot 検索を行う際、Mascot 検索条件設定ページの [Quantitation] 選択リストから適当な「Quantitation Method」を選択してください。

「Quantitation Method」は Quantitation Methods エディタ (Mascot Server のトップページにある [Configuration Editor] リンク→[Quantitation] リンク) を使って定義・編集することができます。代表的な定量解析手法に対応する「Quantitation Method」はすでに登録されていますが、新規に登録する場合は内容が似ている「Quantitation Method」のコピーを作成し、それを編集すると簡便です。



(2) ピーク抽出

Mascot Distiller の「Processing Option」→[Processing Options] でピーク条件を設定するか、[Load ...] ボタンから適当なピーク抽出条件ファイルを読み込み、ピーク抽出処理を実行してください。ピーク抽出条件の設定に関しては、Mascot Distiller のヘルプページ(メニューバー [Help] → [Mascot Distiller Help] → [Preference] → [Options Dialogs] → [Processing Options])を参照してください。

(3) Mascot 検索の実行

ピーク抽出処理が終了したら、メニューバー[Analysis]→[Mascot Search]→[All Peaklists]から Mascot 検索条件設定ページを開き、[Quantitation]選択リストから適当な「Quantitation Method」を選択した上で Mascot 検索を実行してください。

(4) 定量解析

Mascot 検索が終了すると、Mascot 検索結果は自動的に Mascot Distiller に読み込まれ、[Data Explorer]に格納されます。メニューバー[Analysis]→[Quantitate]から定量解析したいタンパク質([All hits]を選択または[Range]を選択して「3, 4, 6-10, 13」のように指定します)に対して定量解析処理を実行してください。

(5) 結果レポート作成

メニューバー[Analysis]→[Quantitation Reports]のサブメニューから定量解析結果レポートの項目別表示、XML ファイル出力など実行することができます。

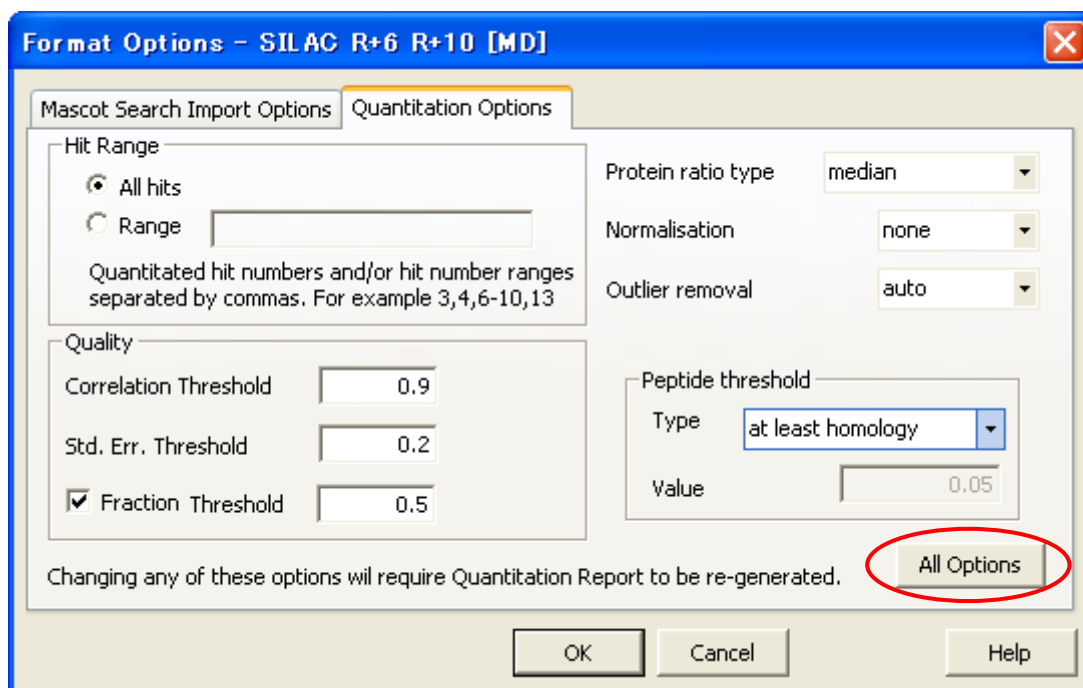
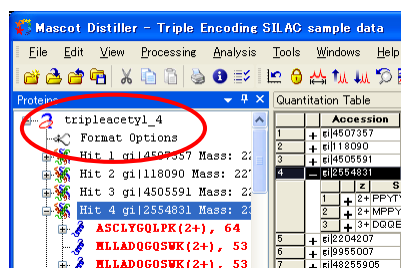
5. 定量解析処理パラメータの設定確認・変更

定量解析処理パラメータである「Quantitation Method」は Mascot Server がインストールされている PC 上の次の XML ファイルにまとめられており、Mascot Server および Mascot Distiller が参照しています。

C:\inetpub\mascot\config\quantitation.xml

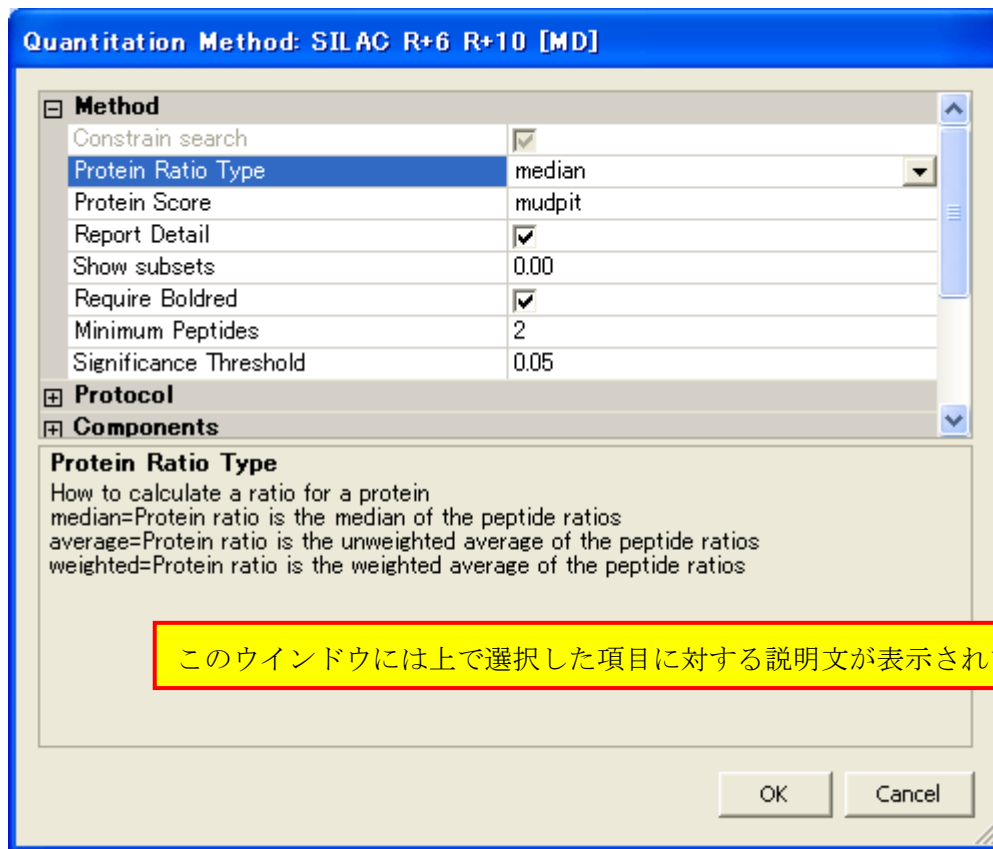
quantitation.xml ファイルに設定されている内容はブラウザベースの Configuration エディタ (Mascot Server トップページでの [Configuration Editor] リンク → [Quantitation] リンク) を使って新規登録、削除、コピーおよび編集することができます。

Mascot Distiller からは [Data Explorer] の [Proteins] タブに表示されているサンプル名の下での [Format Options] をクリックするか (右図の例ではサンプル名は「tripleacetyl_4」であり、その下に [Format Options] があります)、メニューバー [Tools] → [Format Options] を選択すると次の図に示す Format Options ダイアログが表示されますので、[Quantitation Options] タブから現在の設定値を変更することができます。



ダイアログの右下にある [All Options] ボタンを押すと、次の図に示すように、すべての定量解析処理パラメータにアクセスすることができる [Quantitation Method] ダイアログが表示され

ます。



6. お問い合わせ

何かお困りのことがありましたら弊社技術サポートにご連絡ください。



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

Mascot Distiller Quantitation Toolbox Quick Start

2009年10月 第1版

Copyright 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。