

ペプチド/プロテインの同定精度向上を実現する成熟蛋白質構造と多様性を考慮したアミノ酸配列DB作成ソフトウェア

公共の蛋白質データベース (UNIPROT や NCBI Nr) のアミノ酸配列構造は、一般的に蛋白質の全長配列で構成されています。しかしながら、質量分析装置で取得されたスペクトラのもととなるペプチド/プロテインは、多様性や成熟蛋白質の構造をもっている場合があります。**Sequence Composer** は、1アミノ酸配列に付随するアノテーションを利用し、考える全てのアノテーションの組み合わせを反映させたアミノ酸配列を生成し DB 化します。本ソフトウェアで作成したデータベースを用いることで質量分析法による蛋白質同定精度の向上を実現します。

<特徴>

- ▶ 質量分析装置で取得したスペクトラとペプチド/プロテインを照合させるアルゴリズムに、本ソフトウェアで作成した配列データを用いと、低いスコアで同定されたペプチド/プロテインの同定スコアの向上を図ることができます。
- ▶ また、照合結果に含まれる擬陰性の疑いを見極めるアミノ酸配列構造の確認作業に費やす時間や労力の大幅な低減を図ることができます。
- ▶ 本ソフトウェアは、PC上で動作し操作性の良いGUIによりフレキシブルに利用目的の配列データを作成することができます。

<Sequence Composer>

● 多様性や成熟蛋白質の構造を考慮したアミノ酸配列を生成しデータベース化

1アミノ酸配列に付随する、遺伝的な多様性、転写産物の多様性、蛋白質の加工・修飾の全ての組み合わせを考慮したアミノ酸配列データを生成しデータベース化します。

● 加工対象のコンディションを選択。①

ダウンロードした UNIPROT のフラットファイルにアノテーションとして付随する情報をもとにした下記の加工条件を選択することができます。
 VAR_SEQ (Variants from the DNA sequence)
 VARIANT (Variants from the alternative splicing events)
 CHAIN, PEPTIDE, PROPEP = Matured Form (Main chain processing)

● 対象アミノ酸配列を Description や Accession 等で絞り込み可能。②

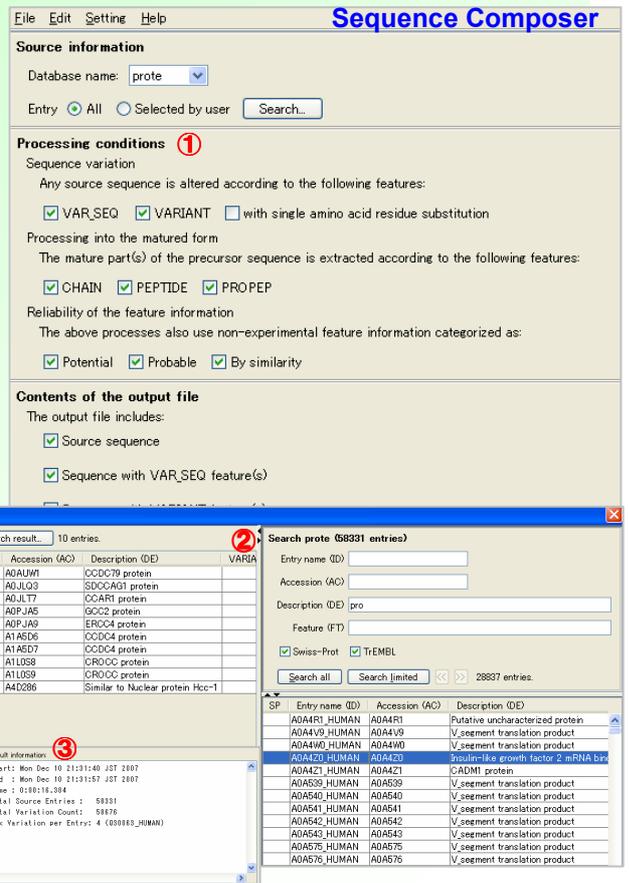
ユーザが、加工したい対象の蛋白質を Description や Accession で、検索し選択することができます。また、選択する際に指定した蛋白質の Feature 情報も参照することができます。

● 加工後の配列数を確認可能。③

加工し生成したアミノ酸配列数の総数を確認することができます (統計情報の機能拡張計画中)。

● 数十秒～数分でプロセッシング可能。

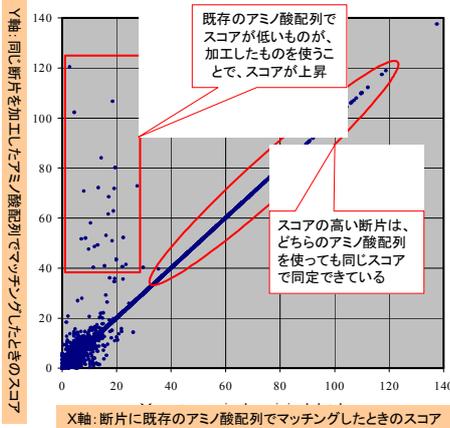
uniprot_sprot_human.dat をプロセッシングし、データベース化が完了するまでの所要時間は約 47 秒です。(メモリ 1GB の PC を利用)



The screenshot shows the 'Sequence Composer' application window. It includes a menu bar (File, Edit, Settings, Help) and a title bar. The main interface is divided into several sections:

- Source information:** Database name: prote, Entry: All (Selected by user), Search...
- Processing conditions ①:**
 - Sequence variation: VAR_SEQ, VARIANT, with single amino acid residue substitution.
 - Processing into the matured form: CHAIN, PEPTIDE, PROPEP.
 - Reliability of the feature information: Potential, Probable, By similarity.
- Contents of the output file:** Source sequence, Sequence with VAR_SEQ feature(s).
- Entry selection ②:** A table with columns SP, Entry name (ID), Accession (AC), Description (DE), and VAR_SEQ. It lists various human proteins like ADALM1, ADJLQ3, ADJLT7, ADPJA5, ADPJA9, A1ASD6, A1ASD7, A1TOS8, A1TOS9, and A4C286.
- Search prote (68331 entries) ③:** Search fields for Entry name (ID), Accession (AC), and Description (DE). It shows search results for 'pro' with a table of protein entries including AD4A1R1, AD4A4V9, AD4A4W0, AD4A4Z1, AD4A539, AD4A540, AD4A541, AD4A542, AD4A543, AD4A575, and AD4A576.
- Result information ③:** Start: Mon Dec 18 21:31:40 JST 2007, End: Mon Dec 18 21:31:57 JST 2007, Time: 0:00:16.204, Total Source Entries: 58301, Total Variation Count: 58876, Max Variation per Entry: 4 (Q03083_HUMAN).

Improvement of peptide identification (1) Increases in Scores



Sample: Tryptic digest of human plasma proteins
Prior to digestion, standard human plasma solution (Sigma) was treated with an albumin/ immunoglobulin G depletion kit to reduce these abundant proteins.

LC-MS/MS conditions:
Reversed phase LC
Flow rate: Approx. 1μL/min
Mobile phases
A Formic acid/MeCN/Water = 0.1:2.98 (v/v/v)
B Formic acid/MeCN/Water = 0.1:90:10 (v/v/v)
Grad. 5-40%B in 70min, followed by 95%B in 5 min
ESI-Ion trap MS: Thermo Fisher Scientific LTQ

Number of peak lists (1): 19985
Database search parameters:
Mascof program, Version 2.1.04
Enzyme: Trypsin
Maximum Missed Cleavages: 1
Fixed modification: Carbamidomethyl (C) +57.0 Da
Peptide mass tolerance: ±2 Da
Fragment m/z tolerance: ±0.8
Mass values: Average
Instrument type: ESI-TRAP

(1) Bradshaw RA et al. Mol. Cell. Proteomics 5, 787, 2006.

ペプチドの同定精製度向上

同じ LC-MS/MS の結果に対して、2種類のアミノ酸配列 DB を使って、検索エンジンで検索した結果、公共のアミノ酸配列 DB (X 軸) では高いスコアをえられなかった25個のペプチドに対して、成熟構造と多様性を考慮したアミノ酸配列 DB (Y 軸) では、高いスコアを得ることができた。

Figure 8. Comparison of peptide ion scores between the original and altered databases.

プロセッシング後のアミノ酸配列構造

シグナルペプチドを削除したアミノ酸配列を使うと、N 末端のペプチドにヒットし、データベースのアノテーションを精査することなく α 鎖のみの配列にしたことで、スコアもアップした。

Identification of the terminal peptide of intact protein

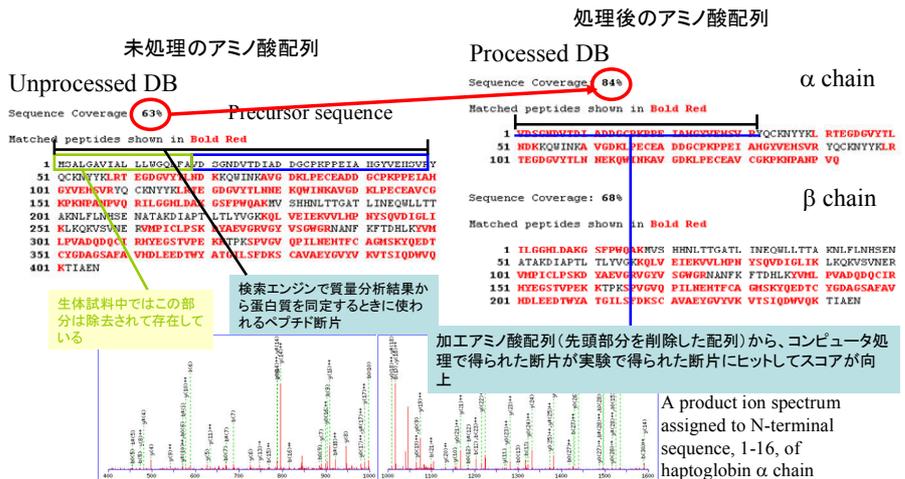


Figure 8. Comparison of identification results of haptoglobin between the original and altered databases.

開発元

株式会社メイズ

東京都渋谷区幡ヶ谷 3-20-2TSビル 101

Tel:03-3378-7430

e-mail: sales@maze.co.jp

URL: <http://www.maze.co.jp>

