

# Mass++ : 差異解析のための統計・多変量解析機能

## Mass++ : Statistical and multivariate analysis functions for differential analysis

○藤田雄一郎<sup>1</sup>、船越なつ美<sup>1</sup>、田中聡<sup>1</sup>、山田賢志<sup>1</sup>、岩本慎一<sup>1</sup>、宇都宮真一<sup>1</sup>、梶原茂樹<sup>1</sup>、青島 健<sup>2</sup>、小田吉哉<sup>2</sup>、田中耕一<sup>1</sup> 島津製作所 田中最先端研究所<sup>1</sup> エーザイ株式会社<sup>2</sup>  
○Y.Fujita<sup>1</sup>, N.Funakoshi<sup>1</sup>, S.Tanaka<sup>1</sup>, Y.Yamada<sup>1</sup>, S.Iwamoto<sup>1</sup>, S.Utsunomiya<sup>1</sup>, S.Kajihara<sup>1</sup>, K.Aoshima<sup>2</sup>, Y.Oda<sup>2</sup>, K.Tanaka<sup>1</sup> Koichi Tanaka Laboratory of Advanced Science and Technology, Shimadzu Corp.<sup>1</sup>, Eisai Co., Ltd.<sup>2</sup>

### 1. Introduction

質量分析を利用したラベルフリー差異解析は、多数のサンプルのMSスペクトルから1つのピークマトリクスを作成し、サンプル間のピーク強度(ないし面積)。本ポスターでは両方を合わせて単に「ピーク値」と記す)を比較することにより行われる。ここで、ピークマトリクスとはMS<sup>1</sup>のピーク位置(Retention Time(RT), m/z)と、そのピーク位置におけるピーク値の行列である(LC-MALDIデータの場合)。ここで、例えば健康者のサンプルグループには「共通して存在しない」が、ガン患者のサンプルグループには「共通して存在する」ピークがあれば、そのピークが表す物質が健康者-ガン患者間の差異を表わす物質(バイオマーカー)の候補である。

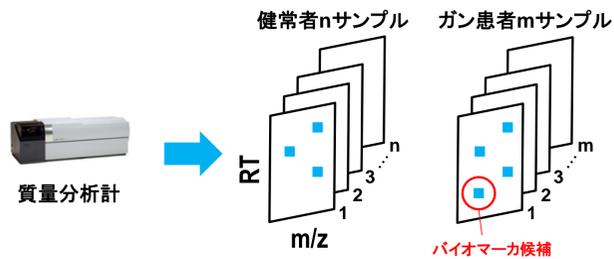


Fig. 1. Label free differential analysis using MS.

質量分析用解析ソフトウェアMass++には、ピークマトリクスを作成するための優れたピーク検出や前処理[1]、アライメント機能[2]が用意されている。さらに今回、バイオマーカー候補探索を目的とした統計解析、多変量解析、同定、ピーク値の箱ひげ図表示などの機能を開発したので紹介する。また、さまざまな手法を総合的、相補的に使用することの有用性についても報告する。

Mass++ は下記のWebサイトから無償でダウンロードできる。

<http://www.first-ms3d.jp/achievement/software/mass2>

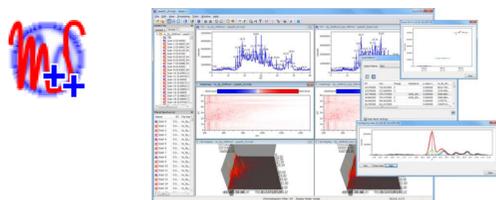


Fig. 2. Overview of Mass++.

### 2. Experimental Section

Mass++に実装した、差異解析のための統計解析・多変量解析機能の評価のため、下記サンプルデータ(LC-MALDIデータ)に対して解析を行った。

- <Sample>  
患者サンプルと健康人サンプルのモデルとして以下の混合サンプルを用いた。  
(1) "4 mix" group : 4種類の混合タンパク質の消化物 × 5  
(2) "5 mix" group : ((1)の4種類 + BSAの消化物) × 4  
→ 本解析ではBSA由来のペプチドが模擬のバイオマーカー候補である。

<Instrument>  
1D nanoflow LC(nanoLC)-AccuSpot-AXIMA system  
・Prominence nano-AccuSpot™ (島津)  
・AXIMA Performance™ (島津/Kratos, UK)



Fig. 3. AXIMA Performance.

### References

- [1] 「Mass++: 質量分析のピーク情報検出技術の開発」  
宇都宮ら. MSSJ 60th (今大会), 3P-09  
[2] 「疫病の診断、創薬研究に向けた3次元アライメントによるマーカー探索手法」  
青島ら. MSSJ 56th, 3P-41

### Acknowledgement

本研究は、日本学術振興会の最先端研究開発支援プログラムにより、助成を受けたものである。

### 3. Label-free Quantitation on Mass++

#### 3.1 Peak Matrix

Mass++を用いて作成されたピークマトリクスをFig. 4に示す。

RT	m/z	Substance	p value (u-...)	121017.4m...	121022.4m...
22.433332	1226.879883	HBB_BOVIL	0.391272	0.252371.7...	0.350607.8...
22.266666	1274.804932	HBB_BOVIL	0.177914	0.252371.7...	0.271551.2...
27.730000	1448.848511	HBB_BOVIL	0.177914	450085.09...	108752.1...
26.466666	1265.957886	HBB_BOVIL	0.177914	40208.57...	1317706.1...
24.544444	1357.881836	PIYGM_SABL	0.066195	886192.58...	345030.83...
28.900000	1053.698973	PIYGM_SABL	0.270348	3224873.2...	5059065.8...
29.333334	1442.814941	PIYGM_SABL	0.391272	724955.44...	661147.84...
29.400000	1355.837545	PIYGM_SABL	0.066195	817624.87...	704616.15...

Fig. 4. Peak Matrix.

ピーク位置、そのピーク位置におけるピーク値、同定結果、検定結果等をまとめた行列(マトリクス)。

#### 3.2 Create Peak Matrix

Mass++ではWizard形式によりピークマトリクスを作成する。

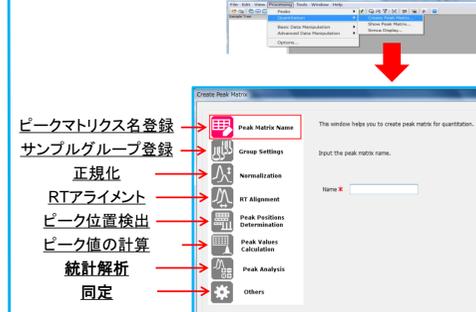


Fig. 5. Creating Peak Matrix using wizard.

Mass++には、各種の優れた一般的もしくは独自アルゴリズムのデータ処理機能が実装されている。  
○...独自アルゴリズム使用

- ピーク検出
    - ・MWD ○
    - ・GION ○
    - ・AB3D ○
    - ・mspeaks
  - アライメント
    - ・TIC
  - 正規化
    - ・内部標準
    - ・TIC
    - ・最大強度
  - ベースライン除去
    - ・BLT法 ○
    - ・線形ベースライン除去法
- etc. (詳細は[1][2])

### 3.3 Statistical Analysis

検定結果(p値)はピークマトリクスの「p value」列に表示される。

使える検定手法	RT	m/z	Substance	p value (u-...)	1210...
t検定 (Welchのt検定)	30.533332	1447.902466	ADH1_YEA...	0.902524	98671
ANOVA (分散分析)	26.433332	1640.121948	ALBU_BOVI...	0.010800	0.000
U検定 (Mann-WhitneyのU検定)	24.033333	1554.719604	ALBU_BOVI...	0.010800	0.000

U検定で、p値<0.05のピークの多くがBSA由来ペプチドと同定。

### 3.4 Identification

同定結果はピークマトリクスの「Substance」列に表示される。



Fig. 6. "Mascot MS/MS Ions Search" window.

Mass++上から、直接データベース検索にかけることが可能。

同定結果は内部データベースに保存されており、ユーザは適時に同定結果の詳細を確認できる。

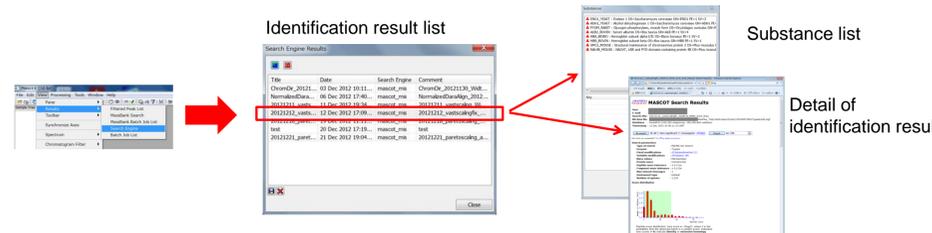


Fig. 7. Confirmation of identification result from "Search Engine Results" window.

### 3.4 Multivariate Analysis

■SIMCA-Plug-in 「SIMCA™」はUmetrics社が開発・販売している多変量解析ソフトウェア。SIMCA解析結果の読み込み・表示が可能なMass++プラグインを開発。

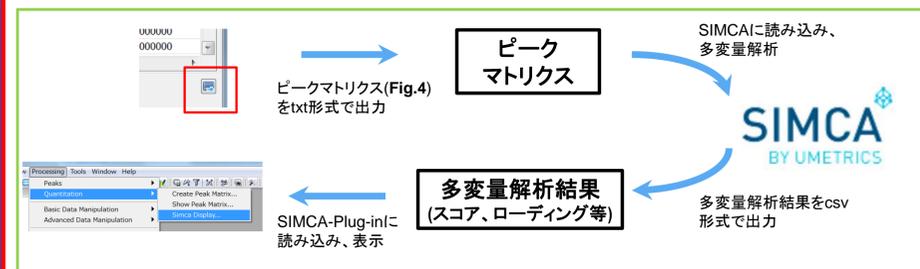


Fig. 8. Read and display of multivariate analytical result using SIMCA-Plug-in.

#### ■ピークマトリクスとの連携

- ・多変量解析結果のグラフは"Display SIMCA Result" window上にプロットされる。
- ・詳細を確認したいデータ点をダブルクリックすると、ピークマトリクスにおける該当のピーク行が最上部に移動し、強調表示される。

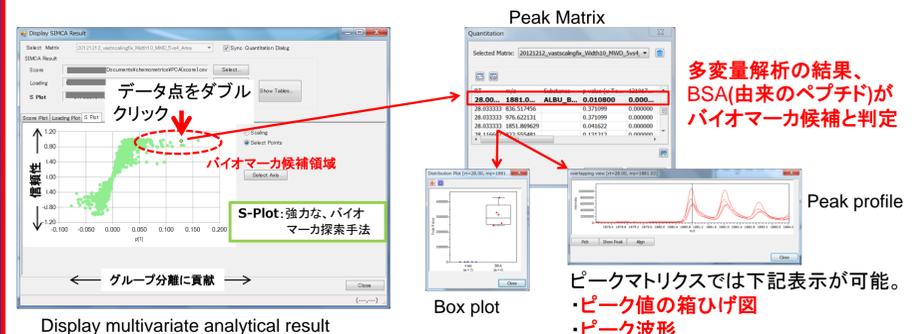


Fig. 9. Cooperation of SIMCA-Plug-in and Peak Matrix.

#### ■連携の利点

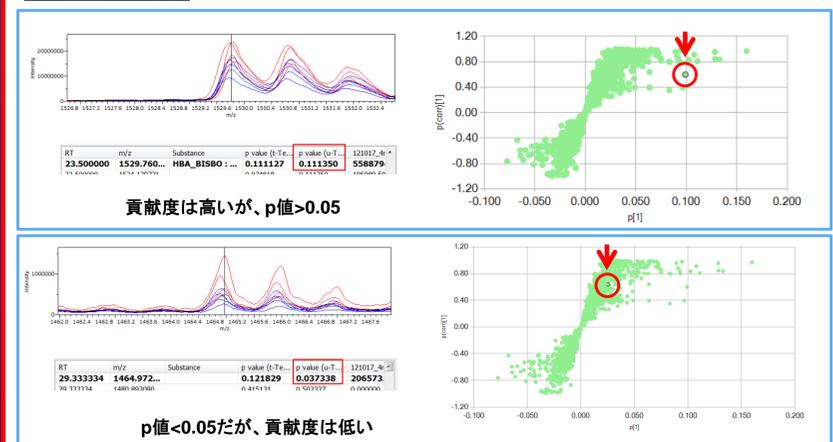


Fig. 10. Cooperation of multivariate and statistical analysis results.

→ 統計解析と多変量解析を相補的に用いた上で、ピーク波形も確認できる。

### 4. Conclusion

- (1) Mass++に、差異解析のための下記機能を開発した。
  - ・統計解析
  - ・多変量解析
  - ・同定
  - ・ピーク値の箱ひげ図
  - ・ピーク波形表示
- (2) LC-MALDIデータに対して(1)を用いて解析した結果、(模擬の)バイオマーカー候補を特定できた。
- (3) (1)の機能を連携させることで、統計解析と多変量解析を相補的に用いた上で、ピーク波形も確認できる。
- (4) バイオマーカー候補探索の為に、(3)を積極的に用いるべき。