

**科学研究費助成事業 研究成果報告書**

平成 28 年 5 月 17 日現在

機関番号：17401

研究種目：挑戦的萌芽研究

研究期間：2014～2015

課題番号：26640092

研究課題名(和文) 溶出時間のデータベース化による超深度定量プロテオミクスの実現

研究課題名(英文) Ultra-deep quantitative proteomics with RT-normalized peptide database

## 研究代表者

大槻 純男(Ohtsuki, Sumio)

熊本大学・生命科学研究部・教授

研究者番号：60323036

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：本研究は、溶出時間の高精度な補正技術によるペプチド情報のデータベース化とシグナルとノイズの効率的な分離によって高深度定量プロテオミクス技術を開発することを目的とする。溶出時間補正およびシグナル・ノイズ分離のアルゴリズムを開発し、ヒト肝臓ミクロソームのトリプシン消化産物を用いて検証した。その結果、既存法と比較して開発法は精度よく偽陽性を除去し、濃度のタンパク質由来ペプチドを同定し比較できることが明らかとなった。また、異なる2施設間での補正精度を解析した結果、補正精度が低下するが、解析可能であることが明らかとなった。解析技術を用いて臨床検体を解析しネットワーク解析やバイオマーカー探索に応用した。

研究成果の概要(英文)：The purpose of the present project is to develop deep quantitative proteomics with peptide database by high-accurate retention time normalization and separation of signal to noise. The developed method for retention time normalization and false discovery rate was validated with trypsin-digested human liver microsomes. The developed method was able to eliminate false positives more efficiently and compare larger number of peptide than the standard method. The ability of retention time normalization was evaluated between two institutes, the accuracy of normalization was reduced, however the normalized retention time is possible to be applied to quantitative analysis by changing analysis settings. The developed method was applied to quantitative proteomic analyses for network analysis and biomarker discovery.

研究分野：プロテオミクス

キーワード：プロテオミクス 質量分析 溶出時間 データベース

### 1. 研究開始当初の背景

サンプル間のタンパク質発現量を網羅的に比較する定量プロテオミクスは生命科学、特にバイオマーカー研究、に必要な不可欠な技術である。定量プロテオミクスの最大の課題は「比較タンパク数を増加させ低発現タンパク質の網羅的比較を可能にする」ことである。これまでの定量プロテオミクスの多くがショットガンプロテオミクスを技術基盤としているため、上記の課題克服には、液体クロマトグラフィ(LC)と質量分析装置(MS)のハード面からの改良が主に行われている。しかし、ハードの改良は技術難易度を格段に上げ、さらに多くの投資を必要とする。

近年、サンプル中のMS/MS情報をすべて取得する新しいデータ取得法(MS/MS ALL)が複数報告された。この新しい取得法で得られたデータを各ペプチドの情報(質量情報、溶出時間)を用いてピークを同定することによって、ショットガンより高感度に、さらに再現性よく発現量を比較することが可能である。しかし、現在の技術では各ペプチドの情報を同じ試料の連続した計測で測定されたショットガン解析から取得している。従って、比較サイズは感度の低いショットガン解析によって制限され、新しいデータ取得法のポテンシャルを活用できていない。このボトルネックを解消する手段として、各ペプチドの情報を様々なサンプルから取得し、データベース化することで解決することを着想した

### 2. 研究の目的

本研究は、溶出時間の高精度な規格化と復元に技術によるペプチド情報のデータベース化とシグナルとノイズの効率的な分離によって、従来法と比較して比較規模と感度の上昇を実現する新しい超深度定量プロテオミクス技術を開発することを目的とする。定量プロテオミクスの比較規模拡大は重要な課題であるが、これまでは主にハード面での改良によって実現されてきた。本研究では、新しいアイデアを導入することによって性能の向上をソフトウェアによって実現する。本技術の実現によって技術難易度を上げることなくバイオマーカー探索をより深く実施することが可能となる。

### 3. 研究の方法

タンパク質試料は可溶化し、還元、アルキル化後、リジルエンドペプチダーゼ及びトリプシンによって消化した。消化産物を固相抽出によって脱塩し、ペプチド試料とした。ペプチド試料はnanoLC(C18カラム、75µm ID x 250 mm, 300 nL/min)によって分離し、ESI-Q-TOF(TripleTOF5600, SCIEX)によってデータを取得した。ペプチド同定はProteinPilot(SCIEX)を用いた。また、MS/MS ALLのデータを用いた、ピーク同定、ピーク面積計算はPeakView SWATH microapp

(SCIEX)によって行った。独自アルゴリズムによる計算及びデータベース構築はExcel, Access, SQL server (Microsoft) 及び Origin (Origin lab)を用いた。

### 4. 研究成果

本研究は、溶出時間の高精度な規格化と復元の技術によるペプチド情報のデータベース化とシグナルとノイズの効率的な分離によって高感度・高深度定量プロテオミクス技術を開発することを目的とする。

偽陽性の除去には False Discovery Rate(FDR)が汎用される。そこで巨大ペプチドデータベースを用いた解析に最適化された FDR 計算法の開発を行った。溶出時間およびシグナル・ノイズ分離の FDR アルゴリズムを検証するために、ヒト肝臓マイクロソームのトリプシン消化産物に 200 種類のペプチド混合液の濃度を変えて添加し、同一のサンプルを MS/MS ALL 法及びショットガン法によって解析した。得られた解析データをすでに構築した初期技術によって検証した結果、データベースを用いた方法は従来のショットガンと比較してより低濃度までペプチドを同定できることが明らかとなった。FDR については、市販ソフトウェアによる FDR と比較して開発法による FDR が濃度既知のペプチドのシグナル・ノイズ分離を正確に行えることを確認した。また、同一ヒト肝臓マイクロソームを 2 ヶ月の間をおいて解析したデータを既存の方法と開発方法で比較解析した結果、開発方法では比較ペプチド数が 1.7 倍に増加しただけでなく、2 倍以内の誤差のペプチドが 90% から 93% に増加した(図 1)。

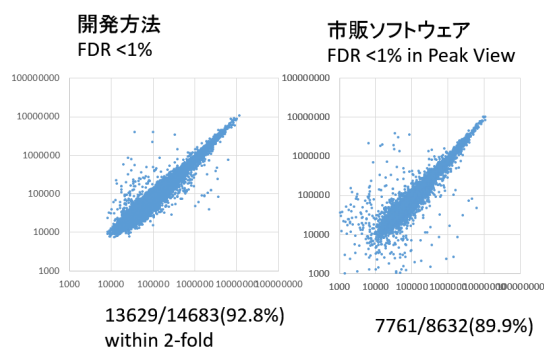


図1. 同一検体を定量解析した際の開発法と既存法の比較

さらに、市販ヒト血清を試料として非分画及び等電点による 12 分画のサンプルを MS/MS ALL 法及びショットガン法によって解析した。非分画血清のショットガン解析と比較して MS/MS ALL 法とデータベースを用いた手法ではペプチド数で 1.4 倍、タンパク質数で 1.8 倍同定数が増加した。さらに、MS/MS ALL 法とデータベースを用いた手法で同定されたペプチド及びタン

パク質のそれぞれ 89%、72% は非分画もしくは分画試料のショットガン解析で同定されていた。

開発技術を用い、異なる 2 施設において同一のデータベースを用いて解析を行った (図 2)。その結果、同一施設では 95% のペプチドについて誤差が 0.66 分以内であった。一方、データベース構築と異なる施設のデータの場合は、95% のペプチドについて誤差が 2.7 分以内であった。この結果から、データベース構築とデータ解析は同一施設で実施した方が精度の高い形跡が行えることが明らかとなった。しかし、異なる施設であってもピーク同定の溶出時間幅を広げることによってデータベースを利用できる。

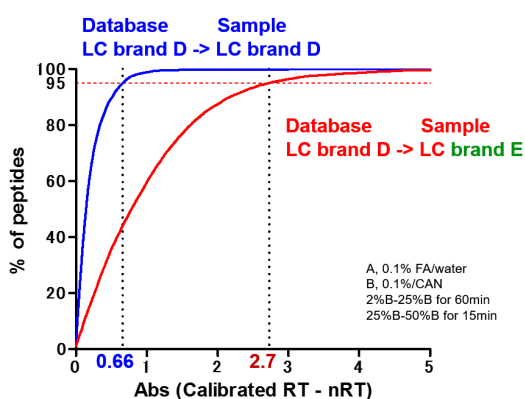


図 2. 同一施設間 (青) と異なる施設間 (赤) の補正誤差の比較

溶出時間補正はこれまでに複数の方法が報告されている。本研究では内部標準ペプチドを添加しない補正基準ペプチドについて開発方法 (オーバーラップシフト法) と共通ペプチド法 (CiRT 法)、また補正方法について開発方法と線形補間及び線形最小二乗法を比較解析した。開発方法と他 3 法について 100 解析のデータからデータベースを構築し比較した結果、開発方法であるオーバーラップシフト法が最も誤差の少ない溶出時間補正が可能であることが明らかとなった。

100 解析のショットガンデータをデータベース化することによって UniProt データベースに登録されている 68,288 分子の内、21,673 分子 (31.7%) のタンパク質を解析できるペプチドのデータを蓄積した。さらに、公開されているペプチドデータベース (SWATH atlas) のデータを同様に溶出時間補正できるように変換し、融合することによって 31,751 分子 (46.5%) のタンパク質を解析できるペプチドのデータベースを構築できた。ただし、公開データベースを利用した場合、上記の異なる施設間での補正に相当し、溶出時間の補正誤差は 2 分以上に大きくなったため定量精度は低下する (図 3)。

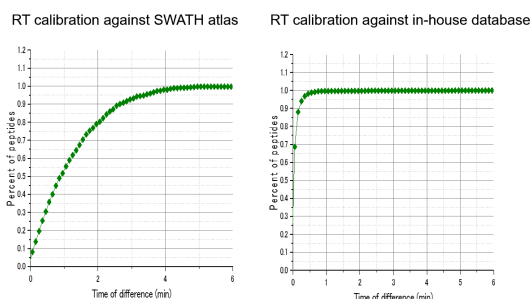


図 3. 公開データベース SWATHatlas (左) 及び in-house データベース (右) を用いた時の補正誤差

これらの構築データベースを用いて統合失調症患者の前頭葉プロテオームの解析、脳腫瘍患者の髄液プロテオームの解析、がん患者血液のプロテオーム解析を実施し、ネットワーク解析やバイオマーカー探索に応用した。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

### 〔雑誌論文〕(計 1 件)

K. Nakamura, M. Hirayama-Kurogi, S. Ito, T. Kuno, T. Yoneyama, W. Obuchi, T. Terasaki, S. Sumio: large-scale multiplex absolute protein quantification of drug-metabolizing enzymes and transporters in human intestine, liver and kidney microsomes by SWATH-MS: comparison with MRM/SRM and HR-MRM/PRM. *Proteomics*. Accepted for publication.

### 〔学会発表〕(計 8 件)

大槻純男: MS/MS ALL (SWATH) によって進化する定量プロテオミクス、日本プロテオーム学会 2014 年会、2014 年 7 月 17 日、筑波

大槻純男: "多"がつくと大変な標的定量プロテオミクス:一緒に楽をしましょう!、第 87 回日本生化学会大会、2014 年 10 月 16 日、京都

S. Ohtsuki: Ion library accumulation by RT calibration. ABSCIEX users meeting, Baltimore, USA, 15 June, 2014

大槻純男: 定量的標的プロテオミクスを活用したがんバイオマーカー研究、第 22 回 HAB 研究機構学術年会、2015 年 6 月 26-27 日、東京

大槻純男: 質量分析でタンパク質を測ると何が見えるか、第 40 回西日本薬剤学研究会、2015 年 8 月 21-22 日、九重

大槻純男: 基礎と臨床をつなぐ次世代プロテオミクス、第 1 回川島カンファレンス、2015 年 11 月 14?15 日、岐阜

大槻純男: 生命科学・薬物動態研究者による定量プロテオミクス開発と応用、

BMAS2015 SCIEX ランチョンセミナー、  
2015年8月22日、長崎  
Sumio Ohtsuki: Targeted proteomic  
approach to cancer biomarkers and  
diagnosis. Japna-US Workshop for Early  
Cancer Research: Biomarker Discovery for  
Early Cancer Detection. Tokyo, Japan, 7-8  
Mar, 2016.

〔図書〕(計0件)

〔産業財産権〕

出願状況(計1件)

名称：質量分析におけるペプチドピークの同  
定・定量のためのデータベース作成方法  
発明者：大槻純男  
権利者：国立大学法人熊本大学  
種類：特許  
番号：PCT/JP2014/003600  
出願年月日：2014年7月7日  
国内外の別：国際(PCT)

取得状況(計0件)

〔その他〕

ホームページ等

<http://www.ohtsuki-lab.jp/>

## 6. 研究組織

### (1) 研究代表者

大槻 純男 (OHTSUKI, Sumio)

熊本大学・大学院生命科学研究部・教授

研究者番号：60323036