

「代謝調節機構解析に基づく細胞機能制御基盤技術」
平成 17 年度採択研究代表者

小田 吉哉

エーザイ(株)コアテクノロジー研究所オミクスグループ統括

定量的メタボロミクスとプロテオミクスの融合

1. 研究実施の概要

ゲノミクス、トランスクリプトミクス、プロテオミクスに引き続き、4番目のオミクスであるメタボロミクスも徐々に盛んになってきているが、要素技術の完成度は必ずしも高くない。よって現在一度に100種類以上の代謝物の同定ができる研究機関は一握りしかない。我々は分離分析の中でもルーチン性に優れた HPLC や GC を主体とした分析手法の確立、そして比較解析を可能にするために信頼性が高い定量分析法の開発、エンドユーザーの利便性を考えながら、種々の装置に対応可能で、かつ差分解析を実行できるソフトウェアの作成と配布を行ってきた。このような各技術基盤を開発しながら、その方向性の確認と新たな課題を探るために、診断を目的とした臨床試料分析や薬剤の作用機序推定を目的とした細胞内代謝物解析を行った。その結果一定の成果を得ると同時に代謝物の自動同定システムの必要性が浮き彫りになった。

2. 研究実施内容

1) 定量的メタボローム解析手法の確立

代謝物の解析では未知物質の同定作業の前に、違いがあるもの(正常と病態の違いなど)を見つけることが必要である。つまり定量法が非常に重要な位置を占める。現在の LC/MS 装置は再現性が高い結果を得やすくなりつつあるが、それは完成度が高い一般的なシステムである。生体内の微量成分を検出するためには、LC/MS の流速を毎分数マイクロリットルからナノリットルまで低流速化する必要がある。しかしナノ LC/MS は安定性が損なわれる傾向があることから信頼性の高い定量値を得ることは難しい。そこで我々はプロテオミクスで培ってきた網羅的内部標準物質すなわち安定同位体元素標識細胞を内部標準細胞として利用した。ここで哺乳類細胞が作り出す代謝物を安定同位体元素標識することは容易ではないが、大腸菌は最小培地で培養可能なため代謝

標識も簡単にできる。しかも主要な代謝物の多くは大腸菌からヒトまで共通である。我々は種々培養条件を吟味してから同位体標識した大腸菌を内部標準として用いることで精度の高い定量的メタボローム解析を実現した。これによって各種代謝酵素阻害による抗がん剤の作用機序解析が可能になった。

上記定量法は対象物質を数百種類以下にあらかじめ限定したメタボローム解析では非常に有用であるが、対象を限定せずに少しでも多くの代謝物を分析しようとする際は、内部標準として添加した大腸菌由来の代謝物が問題となる。なぜなら質量分析装置で検出されるピークが増えるため、もしくはナノLCカラムに負荷できるサンプル量が半減するため(大腸菌のサンプルも同時に負荷するため)、本来着目すべき代謝物を見逃してしまう危険性が高くなる。そもそも現在のメタボローム解析では未知代謝物の同定が難しいため、大腸菌由来か目的細胞(組織)由来か即座に判別できない。そこで内部標準を使わない方法を確立しようと、まず LC/MS の流速を下げ、感度を上げつつ、できるだけ安定した測定系の構築を行い、定量測定を行った。比較する試料を同じ日に処理して、同じ日に測定を行えば、許容される程度の再現性が得られたが、試料数が多くなって同一日以内に処理や測定が終わらない場合、病態と正常の違いよりも測定日の違いのほうが、差が大きくなることが判明した。そこで次に2種類の内部標準物質を用いてピーク面積の補正を行った結果、日間変動が小さくなり、定量解析に耐えうるデータを算出することが可能になった。この方法によって臨床試料を用いた疾患関連代謝物解析を行い、診断マーカー候補を得ることができた。

LC/MS を使った網羅的解析から変動幅が比較的大きい分子を見出す探索段階の場合、内部標準物質を使わない簡便な定量法も有効である。しかし先に述べたように質量分析特にナノLC/MS は再現性を犠牲にして感度を向上させていることからセミ定量としても精度が悪い。ところでこれまでに報告されているオミクスデータや我々が行ってきたデータから、例えば正常と病態の比較において変動する分子は一部であり、多くの分子は変動しない。このことは変動しない分子を疑似内部標準物質とすれば定量値を補正できることを意味している。課題は変動しない分子をどうやって見出すかという点である。ハウスキーピング分子でさえも変動する可能性があることから、まず LC/MS によって網羅的解析データを取得し、各々のピークの強度比を計算する。次に比較するサンプル間で、これらピーク強度比の比を計算する。観測された全てのピークの比を計算し、その値ごとに分布を調べ、もっとも分布が多いところが変動しない分子の集合体である(多くの分子は変動しないため)。これらを疑似内部標準物質とした。そして実際の定量値を算出する際には、目的のピークと HPLC の溶出時間が近い複数の疑似内部標準物質を選び、それぞれ定量値を計算して、それらの平均値を最終値とした。今回は自動解析ソフトの充実度と我々の開発した疑似内部標準法の確からしさの確認のため、既に種々の定量法が確立されているプロテオームデータを使って仮説の検証を行った。本法は質量分析一般に当てはまることから、自動ソフトウェアが充実さえすればメタボロミクスにも即座に利用できる。

以上、平成19年度に我々は LC/MS を使ったメタボローム解析のために3つの定量法を開発した。

2) LC/MS 解析用ソフトウェアの開発

質量分析装置から得られるデータを解析するためには通常、装置開発会社が装置とともに供給しているソフトウェアを使って解析を行っている。クロマトグラフィーなど分離手法と組み合わせずに得られた質量スペクトルデータであれば、 m/z と強度との二次元情報のみであり、薬物動態研究における定量分析では測定したい分子をあらかじめ想定し、その分子だけのデータを取得することからデータ量も極端に少ない。質量分析装置に付随しているソフトウェアはこのような型にはまった使い方に合わせて開発しており、実際質量分析ユーザーの大多数をしめている。プロテオミクスそしてメタボロミクスは知名度こそ高まったが、質量分析装置メーカーからすればユーザー数は少なく、しかも解析法が一般化していないため、ユーザーによって、あるいは目的によって様々なデータ解析が行われている。そのうえオミクス解析では複数の異なった質量分析装置を使っている研究室も珍しくないが、異なる装置メーカーから得られたデータを別のメーカーのソフトウェアで解析することは通常できない。このようにオミクス現場のニーズに合わせたソフトウェアの開発は商業的には極めて難しい状況にある。我々はオミクス解析に必要な情報を研究者のニーズに合わせて引き出せるユニバーサルなソフトウェア、Mass++を開発してきた。平成19年度は対応できる機種を増やしたことで現在は質量分析主要4メーカー（Applied Biosystems、Thermo Fisher、Waters、島津）のデータ解析が可能となり（一部バージョンによっては未対応）、LC/MS や GC/MS のデータにおいて（原理的には CE/MS のデータも）、二つの測定データ（例えば正常と病態モデルの2種類）の差分解析を可能にし、三次元表示もより使いやすくした。Mass++はプラグイン方式なので各ユーザーが必要な機能を追加したり、不要な機能を削除して動作性を速くしたりすることができる。CREST ホームページからダウンロードして無償かつ自由に利用できるようになっている。

3) 実試料によるメタボローム測定

ヒト由来のガン細胞を使ってリン酸基を含む極性代謝物に絞った分析法を開発し、これらを精度よく一斉定量することで抗がん剤の作用機序解明に役立っている。

患者さんの脳髄液や血漿試料のメタボローム解析を行って診断マーカーの探索を行っている。

これら詳細については特許性の有無を確認後、特許申請や論文発表を行ってから順次公開していく。

3. 研究実施体制

(1) 「小田」グループ

① 研究分担グループ長: 小田 吉哉 (エーザイ(株)、統括課長)

② 研究項目

定量的プロテオミクスとメタボロミクスの融合

4. 研究成果の発表等

(1) 論文発表(原著論文)

1. “Pseudo Internal Standard Approach for Label-Free Quantitative Proteomics.” T. Tabata, T. Sato, J. Kuromitsu, and Y. Oda, Anal Chem. 79, 8440-8445 (2007).

(2) その他

①その他の成果発表

本プロジェクトにて開発した質量分析用ユニバーサルソフトウェア Mass++を無償提供中

<http://www.cellmetabo.jst.go.jp/ja/topics/ver.html>

LC/MSによるオミクスデータのセミ定量法 PIS がニュースとして2カ所で取り上げられました。

No-fuss quantitative proteomics with pseudo-internal standards

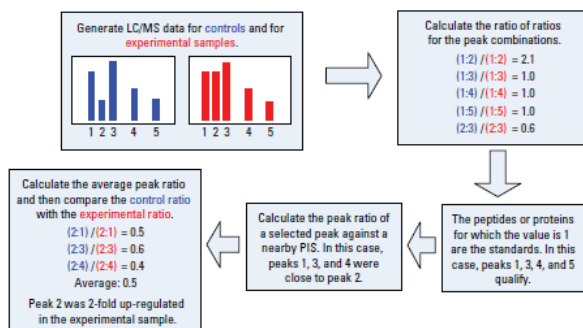
MS may be the defining tool of proteomics, but it is far from quantitative. Differences in sample preparation, instrument calibration, ionization efficiencies, and matrix interferences mean that ion intensity alone cannot be used to gauge the abundance of any particular peptide or protein. Researchers have developed several strategies to circumvent this problem, including the use of stable-isotope-labeled “spike-in” standards, differential metabolic labeling, and molecular tagging. Yet, these approaches are cumbersome, are expensive, or vary in efficiency from sample to sample.

In the November 15 issue of *Analytical Chemistry* (pp 8440–8445), Yoshiya Oda of the Laboratory of Core Technology at Eisai Co., Ltd. (Japan), and colleagues describe a novel label-free approach to quantitative proteomics that addresses these shortcomings. The technique uses pseudo-internal standards (PIS), identified on the fly during data analysis, to provide relative quantification data; this is similar to the approach used in molecular biology techniques such as northern blotting.

“The unique feature of this is the internal standards are not chosen beforehand or spiked in,” Oda says. “Instead, the internal standards are chosen from the data set, going on the assumption that the majority of proteins do not change when stimuli are applied to cells.”

In northern blotting, relative gene expression levels are quantified by computing the ratio of abundance of a target RNA (i.e., the RNA of interest) versus that of a “housekeeping” RNA, the levels of which should remain constant across experimental conditions.

Similarly, in Oda’s approach, LC/MS spectra collected from two samples (say, control and drug-treated cells) are analyzed in a pairwise fashion both within



PIS-based quantitation in proteomic MS.

and across samples to identify peaks for which the intensities relative to other peaks do and do not change. Suppose that the two samples each produce five peaks. Each pairwise ratio of peak intensities (1:2; 1:3, etc.) in the control sample is compared with the same ratio in the experimental sample. If the ratio of ratios (e.g., 1:2 in the control/1:2 in the sample) equals 1—that is, the relative abundances don’t change—the peaks are designated PIS.

These PIS are then used to compute the abundances of peaks of interest, which do change across samples. Oda notes that the retention times of the standard and target peptide must be within 1 minute of each other, because the farther apart two peaks are, the more they will differ in mobile-phase composition, interferences, and so on. “Similar environments as much as possible [are] important for normalization of mass spec data,” he says.

To test the approach, Oda and his team used both PIS and ion intensity data to compare the abundances of individual brain peptides from normal and transgenic mice. In four experiments performed on the same day, the precision (coefficient of variance) of PIS-determined values was 6.9–8.0%, compared with 8.0–14.3% by ion intensity. Precision suffered in experiments performed across days, however, with values of

58–70% for PIS compared with 88–97% for ion intensity data; precision with stable-isotope-based techniques, by comparison, was typically <20%.

Analysis of samples that were spiked with known quantities of human serum albumin showed that PIS quantitation is more linear than the ion intensity approach, especially when multiple standards per target peak are used and

at lower target protein concentrations. That, says Oda, is probably because lower-abundance proteins are more prone to ion suppression effects.

Importantly, PIS-based quantitation was largely in agreement with data calculated with isobaric tagging for relative and absolute quantitation (iTRAQ) reagents: 72% of PIS values were within 1.5-fold of iTRAQ values. In contrast, 44% and 49% of ion intensity and emPAI (exponentially modified protein abundance index) values, respectively, were within 1.5-fold of iTRAQ values.

PIS has been used extensively at Eisai for the discovery of Alzheimer’s disease biomarkers, says Oda. Yet, he adds, it isn’t the best choice for those who need more accurate or precise data (e.g., to support an application to the U.S. Food and Drug Administration) or for those who are interested in very large changes—say, 10-fold or more.

Nevertheless, Oda continues to refine his process, for instance, by using peak-smoothing algorithms and higher-resolution instruments to resolve and quantify overlapping peaks.

And, he notes, the approach can be applied to MALDI mass spectra and to other classes of molecules (such as metabolites), too, as long as the abundances of most molecules remain constant across conditions. ■

—Jeffrey M. Perkel

Pseudo-internal standards for label-free quantitation

Stable-isotope labeling is time-consuming and expensive, and ion-intensity data are not always reliable, so Yoshiya Oda and co-workers at Esai Co. (Japan) and Japan Science and Technology Corp. developed an alternative method in which the levels of pseudo-internal standards (PIS) are analyzed. This label-free method requires no additional reagents and is simple to perform.

Just as unchanging expression levels of housekeeping genes are used for normalization in northern-blotting procedures for RNA analysis, proteins whose levels do not change under different conditions or among samples

in proteomics experiments are used as surrogate standards in the PIS method. The researchers got the idea when they noticed that the levels of most proteins are the same within a given quantitative proteomics experiment.

With the new quantitation method, the relative peak intensities of analytes of interest are compared with those of PIS. These standards are determined during a proteomics MS experiment and are represented by peaks that are present in a 1:1 ratio between samples. For analyses, PIS within a 1 minute retention time interval of a target peak are chosen for comparisons.

Oda and co-workers tested the method on mouse brain extracts. The precision

of the PIS strategy was better than that of ion-intensity-based quantitation techniques but was not as good as that of stable-isotope techniques. To determine the accuracy of the approach, the researchers compared PIS and ion-intensity data with data derived from isobaric tagging for relative and absolute quantitation (iTRAQ) measurements. Accuracies of unknown samples are difficult to assess, so in this experiment, iTRAQ results were taken as the correct answer. PIS data were more similar to the iTRAQ data than those from ion-intensity measurements were. The investigators say the method also could be applied to metabolomics analyses. (*Anal. Chem.* 2007, 79, 8440–8445)