

# メタボロームデータに 影響を与える諸要因

Parameters that affect Metabolome Data

メタボロミクスは、他のオミクス技術と同様に同一試料から多くの標的物質レベルを一斉に測定できることが魅力である。これにより、少なくとも生物学的な実験（個体）間誤差や分析化学的な測定間誤差を最小化できるという長所を持ちあわせている（多くの化学種を個別の方法で測定する場合、手法間の振れやより多くの試料を必要とするだろう）。また、高価ではあるが分解能の高い測定機器を用いることで前処理プロセスを簡略化することも、誤差要因を低減させる効果がある。しかし、これまで行われてきた少数物質を対象としたターゲット解析と同様に、試料自体がもつ不確定さに測定値が大きく影響される。また、その性格上、すべての代謝物質に最適な条件で測定を行うことは不可能であり、メタボロミクス特有の不確定さ（試料調製から測定までの工程で生じる分析化学的不確かさ）も存在する。ここでは、メタボロミクスデータに影響を与える要因について述べ、その回避法についてもできる限り紹介する。

## 6.1 生物学的要因

### 6.1.1 個体差（個人差）

生物はたとえ同じ種であっても遺伝的に個体差が見られ、表現型に違いが生じる。また、双子やクローン動物のようにゲノムが同一であったとしても、DNAのメチル化パターンや

染色体の形態(ヒストンの修飾)の違い、環境要因などにより個体間差が生じる。工業製品における個体差の把握は、その製品に求められる基本性能を評価することで行われ、基準を満たさない製品は排除される。すなわち、合目的性が最優先される。生物学実験においても、合目的性(想定される実験結果を検証するという目的)を優先する場合は、目的を達成するために最小限の基準を設けることがあるが、ときに恣意的であるとの非難を受けるので、注意が必要である。個体差を最小化するために一般に用いられる指標としては、性別、年齢(年齢)、体重を挙げることができる。ヒトの場合は、体重の代わりにボディマスインデックス(BMI)が適用され、人種を考慮する場合もある。生物学においては、これらの指標を用いて個体群を一致させることは一般的であり、メタボロミクスにおいても推奨される(当然、これらの指標を説明変数とする場合はこの限りではない)。ただし、先入観のない解析がメタボロミクスの長所でもあることから、これ以上の指標を導入する場合は、合理的な説明が必要であろう。また、医薬品や特定食品の投与量は、単位体重あたり一定量とするのが一般的である。

### 6.1.2 概日リズムと日内変動

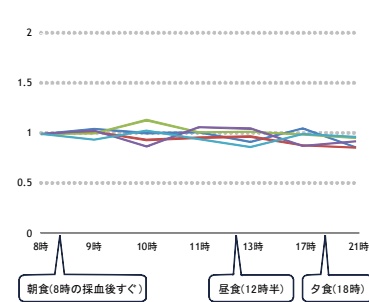


図 6.1 クレアチニンの日内変動

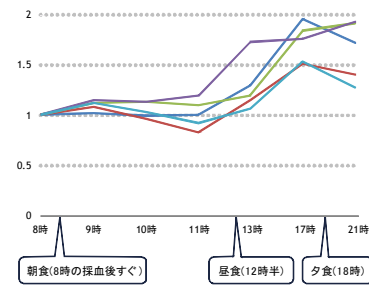


図 6.2 プロリンの日内変動

生物は約 24 時間周期で生理機能を変動させる能力を持っており、概日(サーカディアン)リズムや体内時計などと呼ばれている。シアノバクテリア(*Synechococcus elongates*)では、3種類の時計遺伝子と呼ばれる遺伝子群により制御されており、光の照射などの外的刺激(または社会的要因)により補正されている。この作用により、血圧は日中高く、睡眠時に低下する。体温も朝方は低く、夕方にかけて高くなっていく。代謝もこの概日リズムに支配されていることが明らかになっており、実験動物からの試料調製時刻や被験者からの採血、採尿時刻の設定には注意を要する。理化学研究所の上田のグループは、マウス血漿を定時採取してメタボローム解析を行い、濃度変化が概日リズムに支配されている 42 種類の物質を明らかにした<sup>132</sup>。これらの代謝物質には、リゾフォスファチジルコリン類やアミノ酸類、尿素回路中間体であるシト

ルリンやオルニチン、4- グアニジノ酪酸などが含まれていた。これらの代謝物質は採血時間により大きく濃度が変動しているため、安定なデータが得られる採血時間を設定する必要がある。ヒトの場合、血液中のアミノ酸は午後から夜中にかけて高値となる傾向がある (図 6.2) ので、午前中に採血する方が安定なデータが得られるが、代謝物質により変動時間は一定しないので、できるだけ採血時刻をそろえるように心がけたい。

### 6.1.3 細胞集団の不均一性

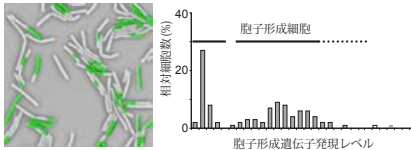


図 6.3 野生型枯草菌の胞子形成

色付きの細胞が胞子形成を開始しているもの (左)。このように同じ条件下でも応答する細胞としない細胞が混在することがある。

メタボロミクスでは、通常数十 mg の細胞から代謝物質を抽出し、測定に供する。これは、真性細菌では  $10^9$  細胞、高等動物細胞では  $10^7$  細胞程度である。しかし、これだけの数の均一な細胞を集めるのは容易ではない。とくに動物から得られる組織試料は、微細構造を考慮すると均一な細胞塊を得ることはほとんど不可能である。組織内の微細構造まで考慮したメタボロミクス

は、まだ今後の技術革新を待たねばならない。手術や内視鏡により採取されたがん部位も、実際は正常細胞とがん細胞がモザイク状に混ざり合っており、純粋にがん細胞だけを集めることは難しい。レーザーマイクロダイセクション (LMD) 法によりがん細胞部位だけを切り取ることも可能であるが、レーザーにより局所的に高熱がかかることで代謝物質組成への影響が懸念され、またメタボロミクスに供するのに十分な量の試料を集めるのも容易ではない。そこで、検体を採取する際に、組織切片の病理解析により、どの程度のがん細胞が試料に含まれているかを把握しておくことが肝要である。

実験室において培養された細胞や微生物も、不均一である場合がある。培養した細胞に薬剤を暴露する場合は、細胞の応答は比較的均一に起きる。しかし、栄養飢餓や低酸素、温度変化など細胞が自然界でも直面し得る環境変化においては、細胞が適応機構として不均一な応答をすることがある。例えば、ある種の微生物は栄養飢餓により抗生物質を産生するが、産生細胞のみが同時に耐性機構を発現しており、周囲の非産生集団は仲間が放出する抗生物質によって溶菌する。これは、栄養飢餓状態において集団の一部が溶けることで、飢餓応答した細胞ができるだけ長く増殖する (もしくは飢餓耐性細胞に分化する) 栄養分を与えるためと考えられている<sup>133</sup>。そして、このような不均一性の発現は、分化に関与する遺伝子の発現もしくは代謝反応が正のフィードバック制御\*を受けていることによりカタストロ

\* シグナル伝達や代謝反応において、産物が上流の反応を正に制御する (活性化する) フィードバック様式。初期入力の高確率的な振れによって、各細胞 (細胞) の応答が二峰性 (オン、オフの二極応答) となる。

フィー\*を起こすことで説明される。このように、細胞が本来の機能として不均一応答性をもつ系では、遺伝子操作により正のフィードバック制御を抑制することで均一な応答集団を作り出すことができ、得られるメタボロミクスデータも大きく異なる<sup>48</sup>。

#### 6.1.4 採血と液性成分の分離

近年、血液のメタボローム解析により疾病バイオマーカーを探索する試みが多くなされている。血液は臨床の現場においてよく制御された状態で採取できるため、最も信頼性の高い試料であると言える。血液の液性成分の分離法としては、血清が最もよく用いられるが、用途に応じて各種手法によって調製された血漿を用いる場合もある。血清は、採血後一定時間血液を静置することで凝固反応を誘発し、遠心分離によって血餅成分を沈殿させた上清である。一般にシリコン製の血清分離剤入の真空採血管が用いられるが、若干の溶出物が認められる。

血漿は、採血時に抗凝固剤を添加することで凝固反応を抑制し、速やかに遠心分離することで血球成分を沈殿させた上清である。臨床検査では、目的に応じていくつかの抗凝固剤を使い分ける(表 6.1)。メタボロミクスでは一般に血清および血漿が用いられるが、抗凝固剤の選択は注意が必要である。ヘパリンは生体由来であるが故に代謝物質が混入するので推奨されない。NaF(フッ化ナトリウム)は解糖系酵素であるエノラーゼを抑制することで、正確な血糖値を測定するために用いられる。しかし一般の検査ではあまり用いられない上に、メタボロームプロファイルも異なる(図 6.4)。EDTA およびクエン酸は、血中のカルシウムをキレートすることで抗凝固作用を発揮する。実際、EDTA 血漿とクエン酸血漿はメタボロームプロファイルも似た傾向を示す(図 6.4)。金属イオンはときにメタボロームの定量データに影響することがあるので、これらの抗凝固剤はメタボローム解析に適していると考えられる。全血は、ゲノムやトランスクリプトーム解析で用いられるが、これは遺伝子もしくは遺伝子転写産物が主に血球成分に含まれていることが理由である。メタボロミクスにおいても、血球を含めた解析を行いたい場合は全血を用いる。採血の際は、真空採血管ホルダ付き針での採血でも翼状針を用いたシリンジ採血でもかまわない。ただし、溶血を避けるため、できるだけ太い採血針(22 ゲージ程度)が好ましい。採血後は真空採血管を転倒させて混和する

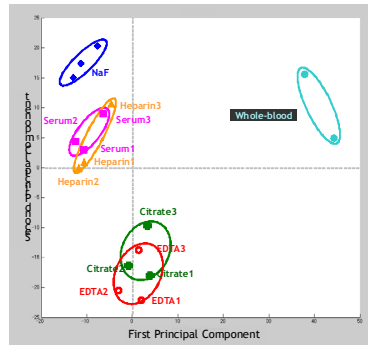


図 6.4 前処理方法とメタボロームプロファイル

\* フィールズ賞を受賞したフランスの数学者ルネ・トムが提唱した、力学系の分岐理論。不連続な応答(スイッチング)を説明する理論として各分野に応用されている。

表 6.1 臨床検査で用いられる血液試料

血液試料	処理方法	用途
血清	静置による凝固後、遠心分離により得る	酵素法による検査 たんぱく質解析
ヘパリン血漿	ヘパリン添加後に遠心分離により得る	化学検査
EDTA 血漿	EDTA-Na もしくは EDTA-K 添加後に遠心分離により得る	メタボローム 内分泌検査
クエン酸血漿	クエン酸添加後に遠心分離により得る	メタボローム 凝固検査
NaF 血漿	NaF(フッ化ナトリウム) 添加後に遠心分離により得る	血糖値の測定
全血	抗凝固剤添加後そのまま保存	全血由来成分の評価 ゲノム、トランスクリプトーム 血球計数

が、溶血を最小限とするために、ゆっくりと数回転倒させる。また、採血量が少ない(2ml程度)場合は転倒混和での血液落下ストロークが長くなることで溶血することがあるので、採血量は 5ml 以上としたい。遠心分離条件は、真空採血管の推奨値を採用するのが理想であるが、遠心分離機の回転数が上げられない場合は、少なくとも 800Xg の平均遠心力を確保し、15 分以上行うことが推奨される。血漿(血清)分離前に冷蔵庫で真空採血管を一昼夜放置することは、メタボロームの濃度変化を免れない。基本的に通常通りの採血、血漿(血清)分離操作で構わないが、以上のことを確認することを推奨している。

### 6.1.5 臓器や病巣組織の採取

メタボローム解析では、手術によって採取された病巣部位を用いることがあるが、試料は速やかにサンプルチューブに入れて液体窒素に浸けることが推奨される。試料の切除から

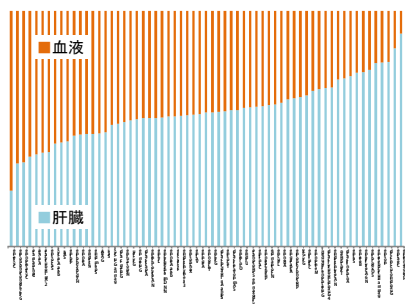


図 6.5 肝臓中の残留血液の影響

保存までの時間を記録しておく、解析の際に参考になることが多い。実験動物の臓器試料の採取においては、できれば頸椎脱臼による安楽死を行い、麻酔剤を用いないことが望ましい。臓器により相当量の血液の混入が想定される場合は、臓器灌流もしくは全身灌流を行う場合がある。とくに肝臓の場合、放血後であっても残留血液は全肝臓重量の 30% 以上となることがあり、血液成分の混入を無視できない場合が多い。

残留血液の影響は代謝物質により異なるが、灌流を行わなかった場合、アミノ酸測定データでは 50% 程度が血液由来となる (図 6.5)。臓器細胞のメタボロームデータを得たい場合は灌流を施すことが推奨されるが、灌流による栄養飢餓および酸素欠乏の影響がデータ解釈に強く影響する場合はその限りではない。薬剤投与の影響を試験する場合など、適切なコントロールを設定できる場合は、灌流を施した方が良好な結果が得られることが多い。

## 6.2 化学的要因

### 6.2.1 代謝物質の抽出プロセス

メタボローム解析を行うためには、生体試料から代謝物質を抽出する必要がある。抽出操作に求められる要件として、1) 可能な限り速やかに細胞をクエンチし、代謝反応を停止させる、2) 代謝物質の損失を可能な限り最小限に止めるような抽出法を用いる、3) 用いる分析法を妨害する混入物を可能な限り最小限に止めることが挙げられる。メタボロミクスでは少なくとも数百種の代謝物質を一斉に分析するため、すべての代謝物質に最適な手法は存在せず、最大公約数的手法を採用せざるを得ない。しかし、現在用いられている分析手法はある特定の代謝物質のもつ性質に基づいているため、それらの代謝物質に適した手法が開発されている。一般に細胞のクエンチングにはメタノールなど極性の高い有機溶媒が用いられ、そこに適当量の水を加えることで比較的満足できる抽出効率を実現している。しかし、フリーのチオール基 (SH 基) をもつ代謝物質など、容易に酸化される代謝物質もあり、これらを正確に定量するためには、修飾反応を行う場合もある。また、CE-MS によるメタボローム解析では、水溶性代謝物質に最適な抽出法を用いるため、長鎖脂肪酸 (おおそ炭素数 8 以上) の場合は十分な抽出効率を得られない。また、試料中のタンパク質の混在は分析に悪影響を及ぼすため、有機溶媒による沈殿法や限外ろ過による除去を行う。アミノ酸など十分な水溶性をもつ分子の場合、有機溶媒共存下で除タンパク質処理を行うことで、タンパク質への吸着ロスもほぼ解決できている。注目している代謝物質の化学的性質が他と異なる場合は、それらに適した抽出法を選択することが肝要である。

### 6.2.2 分離分析における諸要因

メタボローム解析では、多種類の代謝物質を一斉に分離分析するのが原則であるため、分離法の性能が定量性に直接影響を与える。生体由来の試料の場合は、適切な前処理を行っていてもマトリクス効果の影響を少なからず受けるので、標準物質 (純粋な化学物質の混合物) による性能データだけでは参考にならないことが多い。用いる分離手法の性能について

は、生体試料を用いた場合のシグナルの濃度直線性を十分に考慮し、データの評価を行わなければならない。一般に、LC-MSは理論段数が低いため分離が不十分であることがあり、また注入試料量が多い場合があるのでイオンサプレッション(12 ページ脚注)による定量性の低下が起りやすい。GC-MSは分離の理論段数は高いが、誘導体化を行っている場合は定量性に影響することがあるので、別途検証することが望ましい。CE-MSの場合は、理論段数が高く分離は十分であるが、極端に電荷が高い(もしくは低い)物質は分離性能が十分でないことがあるので、移動時間が短い(もしくは長い)場合は注意が必要である。これらの評価については文献<sup>134</sup>に詳しいので参照されたい。

各代謝物質が高い定量性をもって測定されているかどうかを判断する指標として、標準添加法<sup>135</sup>によるシグナル強度の濃度直線性の検証や添加回収率<sup>136</sup>の検証がある。これらを多数の代謝物質を一斉分析するメタボロミクスに適用するのは容易ではなく、また試料ごとにマトリクス状態が異なるため、実用的ではない。そこで、測定する試料をすべて等量ずつ混合した測定を行い、全試料データの平均値とどの程度一致するかで評価する方法がある<sup>135</sup>。

## 6.3 データプロセッシングにおける諸要因

### 6.3.1 データ標準化

定量データの標準化とは、得られた代謝物質量をどのような単位で比較するかという問題に他ならない。例えば、培養細胞の場合は単位細胞数あたり、臓器は単位重量あたり、血液は単位体積あたりの代謝物質量で評価するのが一般的である。しかし、正常細胞とがん細胞では形態が異なるであろうし、肝肥大を起こした場合に正常な肝臓とどのように比較すればよいのであろうか。これらの問いに答える統一的な見解は得られていないのが現状であるが、それを決めることこそ研究者の資質と経験が問われるとも言える。メタボロミクス以外の実験データがある場合は、当然そのデータの標準化法に一致させるべきであり、メタボロミクス特有の考え方がある訳ではない。実際の標準化法としては、臓器試料の場合上記のもの他に単位タンパク質量あたりなどが用いられる。ただし、重量に比べタンパク質量が標準化の基準となる必然性を説明することは困難なことが多く、かつタンパク質量の測定はメタボローム解析以上に大きな誤差を伴う。このような場合、実験者の経験や考え方が最も重要である。細胞や臓器に依存せず、常に一定量存在する代謝物質は見つかっておらず、明

\* 濃度既知の標準物質溶液を試料に添加し、特定物質シグナル強度の濃度直線性を検定する手法。

\*\* 濃度既知の標準物質溶液を試料に添加して測定を行い、標準物質溶液で得られるシグナル強度分の増加が認められるかを検証する。通常両者の比率で表現され、理想的な定量性が得られている場合は100%となり、分離支持体への吸着などが起きている場合は低い値となる。

確な回答はない。ただし、尿の場合はこれまでと同様に単位クレアチニンあたりで標準化することでの確な比較ができる場合が多い。

### 6.3.2 統計処理 (検定)

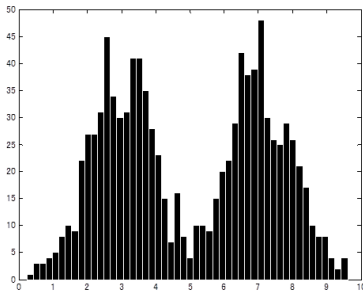


図 6.6 二峰性分布の例

実験を複数回行うもしくは検体数が複数の場合、統計処理を行ってデータを評価することとなる。しかし、平均値と標準偏差の算出においては、母集団が正規分布に従わない場合は正確な検定ができない。しかし、一般に正規分布していることを仮定できる場合は限られている。このような場合、平均値の代わりに中央値を用いる方が合理的である場合がある。また、極端に標準誤差が大きい場合は、定量値が定量限界に近く、正確な値が得られていない場合がある。

このような場合、データの絶対値やSN比などを考慮して測定精度を評価し、信頼できる数値を把握する必要がある。また、データが正規分布しておらず、多峰性の分布をしていることもある\*(図 6.6)。検体数が多いときは正規性の検定を行うことが必要である。データの分布において正規性を仮定できない場合は、ノンパラメトリックな検定を行うことが可能である(この場合も多峰性データでは意味をなさない)。検定におけるp値は、有意水準5%で評価する場合が多いが、より厳密に行う場合はp値の補正を適切に行う必要がある。最も厳密な補正方法はボンフェローニの補正である。これは、n個の独立した仮説を検定する際は、それぞれの有意水準を1/nとすることであり、メタボローム解析の場合は同時に測定した代謝物質数nにp値を乗じることで補正するのが正しい\*\*。

\* 身近な例では、国語や社会の試験結果はおおよそ正規分布するので平均値や標準偏差がそれ相当の意味をもつ。一方で数学や物理学の試験では、極端に点数の高い群と点数の低い群に分かれるので、平均値は最頻値から大きく外れ、平均値が意味をなさなくなり、かつ標準偏差は極端に大きくなる。

\*\* 例えば、100代謝物質を一斉分析して、それらのうちある代謝物質の検定結果が $p=0.005$ であったとする。ボンフェローニの補正を行うと、その物質の有意水準は $p=0.005 \times 100=0.5$ となり、一般的な有意水準5%( $p=0.05$ )より大きいため、有意とは言えない。