研究ノート

メタボローム解析が明らかにするヒトの代謝環境

東北メディカル・メガバンク機構＆未来型医療創成センター
小柴生造　(koshiba@megabank.tohoku.ac.jp)

１．はじめに

人の体は様々な分子で構成されており、タンパク質のような大きな分子から酢酸のような小さな分子まで多種多様である。これらの分子は互いに独立ではなく、小さな分子から大きな分子が合成される一方、逆に大きな分子が分解されて小さな分子の集合になる。このような体の中でおこっている各種反応は代謝とよばれており、生体内の分子のうち、相対的に分子量の小さい分子（分子量<1,000~1,500）は特に代謝物と呼ばれている。ヒトは毎日の食事で摂取する炭水化物やタンパク質、脂質などを消化（分解）し、代謝物として腸から吸収し、栄養源にするとともに、不要になったものを代謝物まで分解して外部に放出する。

人の体の中には数千〜数万種類の代謝物が存在するといわれているが、どの分子がどの程度存在するのか不明な点が多い。特に代謝物は遺伝要因と環境要因（食事、運動、喫煙など各種生活習慣等）双方の影響を受けて変動するため、同じ分子でもヒトによって血液中の濃度に個人差がある。このため個人の代謝環境の違いを明らかにすることは個人の健康状態の重要な指標となる。代表的な例が血糖や尿酸といった健康診断で利用されている代謝物であるが、その他にも様々な代謝物（アミノ酸、有機酸、糖、脂質等）が個人の健康状態を反映するマーカーとして利用され始めている。

代謝物を健康の指標にするためには、まず健康な人の集団において血液中にどのような代謝物がどの程度存在するのか明らかにする必要がある。東北メディカル・メガバンク機構では一般住民15万人を対象とした大規模な健康調査を行っており、その際に収集した血液や尿などを分析することで代謝物を含む各種分子の種類や量を明らかにしてきた。このように生体内に存在する様々な分子を網羅的に解析する手法はオミックス解析とよばれる。本原稿では当機構でこれまでに行ってきたオミックス解析のうち、とくにメタボローム解析と呼ばれる代謝物解析について報告する。

２．メタボローム解析の方法と日本人多層オミックス参照パネル

代謝物の解析は一般に核磁気共鳴(NMR)法と質量分析(MS)法の２種類の方法が使用されており、当機構ではこの２種類の方法で主にヒトの血液中の代謝物を解析している。NMR 法の利点としては定量性が高いのと、長期に安定して測定できるため万単位の検体を解析・比較することができる点にある。一方 MS 法に比べると感度が低いため比較的濃度の高い代謝物を対象としている。一方 MS 法の利点としては感度が高いため濃度の低い代謝物も測定することができる点にあるが、定量性が低いのと測定時間が長いのが課題である。このため当機構では両方の方法を目的に合わせて使い分けて解析を行っている。具体的にはまず NMR を使用して多数の検体をハイスループットに解析するともに、その中から一部の検体を選択して MS で詳細な解析を行うことで、多くの対象者の血液中の代謝物の特徴を解析するとともに、ヒトの血液中の代謝物を網羅的に解析することを目指している[1]。

我々はこれまでに 2.5 万人以上の一般住民の血液中の代謝物のメタボローム解析を実施し、その

成果を「日本人多層オミックス参照パネル（Japanese Multi Omics Reference Panel (jMorp)）」として無償で一般に公開している（https://jmorp.megabank.tohoku.ac.jp/）。このデータベースは代謝物のみならず、ゲノムやタンパク質などヒトの様々な構成要素について当機構の成果を元に統計情報を公開しているデータベースで、日本人の標準的な遺伝子多型情報（塩基配列の違い）や代謝物の濃度分布情報を知ることができる（図１）。

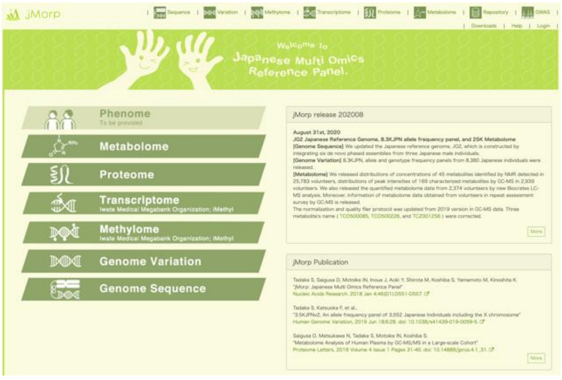

図１：日本人多層オミックス参照パネルのホームページ。代謝物の他、ゲノムやプロテオーム等の情報も公開。

現在公開している代謝物情報としては、2.5 万人の血液中の 45 種類の代謝物（アミノ酸、有機酸、糖等）の濃度分布（NMR 法）の他に、2 千人分についてはより詳細に 421 種類の代謝物（主に脂質など）の濃度分布（LC-MS/MS 法）等を公開している（図２）。代謝物の濃度には性差があるため男女別に色分けして公開している。また各代謝物について年齢層毎の分布情報も公開している（図３）。代謝物は加齢によって変動することが多いため、各年齢層毎に標準的な分布情報を提供している。

また代謝物間の相関情報も提供している。代謝物は様々な化学反応により相互に関連しており、体内で複雑なネットワークを形成している。このため特定の代謝物と相関する代謝物を調べることは、体内の代謝環境を知る上で非常に重要な情報となる。このため各代謝物について相関する代謝物をリスト化し、表とネットワーク図の２形式で表示している。

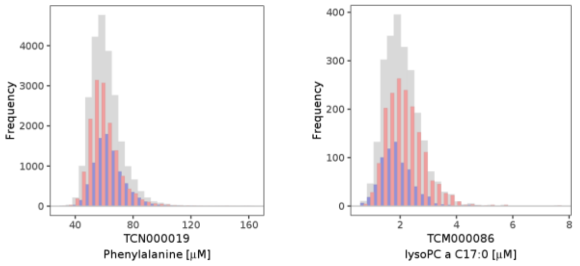

図２：代謝物の濃度分布（横軸：濃度（uM)、縦軸：人数）。（左）NMR で測定したフェニルアラニン（アミノ酸）、（右）MS で測定した LysoPC（脂質）

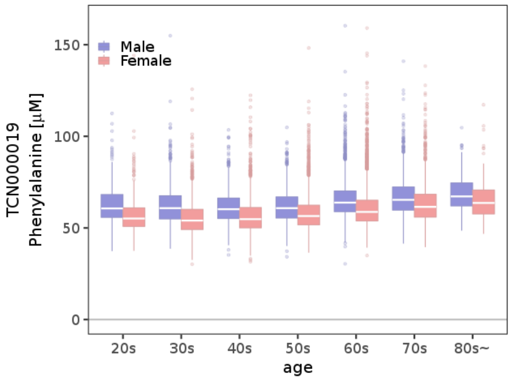

図３：年齢層毎の代謝物（フェニルアラニン）の濃度分布の例（横軸：20 代から 80 代までの各年齢層。縦軸：濃度）。

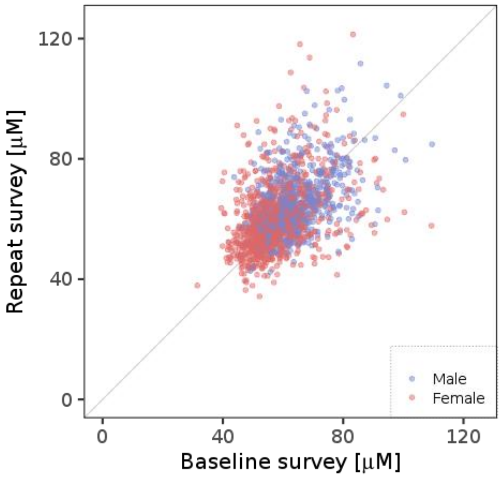

図４：代謝物（フェニルアラニン）の経時変化（横軸：最初の調査時。縦軸：４年後の調査時（追跡調査））。

一方、当機構では健康調査を継続して実施ており、最初の健康調査参加者の４年後の血液試料を現在収集中である。そこで同一人物の４年間の前

後の血液中の代謝物を解析し、４年前と現在の代謝物濃度の違いを明らかにし公開している（図４）。興味深いことに代謝物によって変動の傾向は異なり、４年間で大きく変動する対象者がある程度存在する代謝物もあれば、ほとんど変動せず濃度が維持されるような代謝物も存在している。後で述べるように変動が少ない代謝物の場合、個人間の濃度の違いは、遺伝的要因が強く関与している場合があることが判明している。遺伝子との関連が明らかになった代謝物については、遺伝子多型の違い（塩基の違い）による代謝物の分布の違いについても公開している。

最後に当機構では妊娠時の代謝の変化についても情報を提供している（図５）。一般に女性は妊娠に伴い体内の代謝環境が大きく変化する。本データベースでは、妊娠中及び出産後の各代謝物の濃度分布の違いについて公開している。妊娠に伴う代謝変化は妊娠高血圧症候群など妊娠特有の疾患や低出生体重児など新生児の健康を研究する上で重要な知見を与えている。

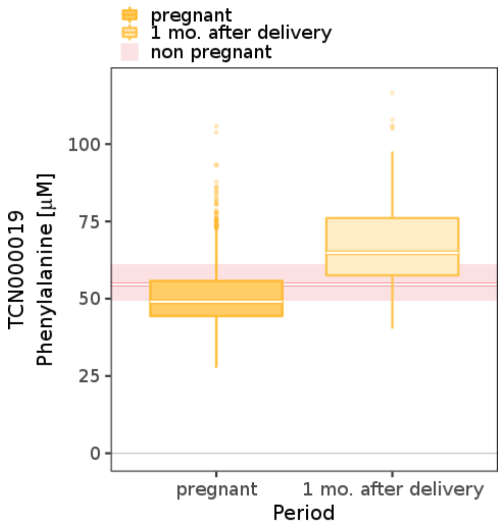

図５：妊娠中（左）と産後（右）における代謝物（フェニルアラニン）の濃度分布の違いの例（縦軸：濃度）。

このように当機構のデータベースでは日本人集団における標準的な代謝環境の特徴について情報を公開し、広く内外の研究者に利用されている。また来年度からは研究開始後１０年後の追跡調査が始まる予定で１０年経過後の代謝環境の変化についても今後して公開するとともに、１０年の間に疾患を発症した対象者の発症前後の血液を重点的に解析することで、疾患を発症する前、もしくは発症初期に早期に診断することができる早期診断マーカーや予防マーカーの開発を本格的に進める予定である。

３．ゲノムと代謝物の関係

体内の代謝は環境要因（生活習慣等）のみならず遺伝要因の影響を受けている。その影響の効果は代謝物によって異なり、関係する遺伝要因の種類も異なっている。一般にゲノム（遺伝子の総体）と表現型（身長のような身体的特徴や、疾患発症の有無、さらに代謝環境のような分子レベルの個体差など）との関連を調べる方法として genome wide association study (GWAS)という統計解析の手法が行われているが、とくに代謝物とゲノムの関連を調べる方法は metabolome genome wide association study (MGWAS)と呼ばれている。

当機構では最初のMGWAS解析として約500人を対象としてゲノムと代謝物の関連を解析し、５種類の代謝物についてそれぞれ影響を与える遺伝子の多型（塩基の違い）を同定した[3]。つまり５種類の遺伝子の中のそれぞれ特定の場所の塩基配列が個人によって異なることが、血液中の５種類の代謝物の濃度にそれぞれ影響を与えていることが判明した。遺伝子の多型にはその遺伝子がコードするタンパク質のアミノ酸配列を変化させるものと変化がないものと２種類存在するが、同定した５つの遺伝子多型はいずれもコードするタンパク質のアミノ酸を変化させるものであった。このことはアミノ酸の変化により各タンパク質の機能が変化したこと、そのことにより血液中の代謝物の濃度が変化したことを示している。実際５種類のタンパク質のうち３種類は対応する代謝物に直接働いて分解する機能を持っており、残りの２種類についても対応する代謝物の代謝経路上で働くタンパク質であった。前述したように代謝物は化学反応により別の代謝物とつながり複雑なネットワークを形成しており、遺伝子多型による代謝への影響はネットワークを介して関係する代謝物群に

も影響を与えていることを示唆している。

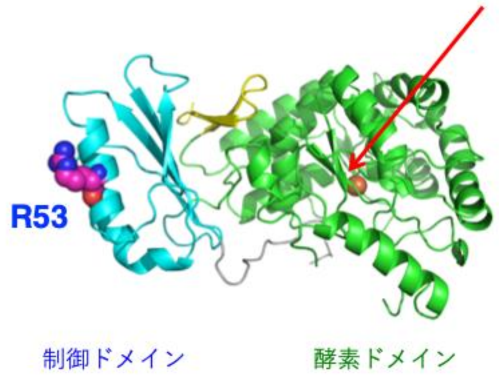

図６：タンパク質（phenylalanine hydroxylase）の立体構造。フェニルアラニンを代謝する酵素で、赤矢印の位置が活性中心。アミノ酸変異（R53）は酵素ドメインではなく、制御ドメイン上に位置する。

一方遺伝子多型によるアミノ酸変異がタンパク質に与える影響を立体構造を元に検討したところ、いずれの変異もタンパク質の立体構造上、化学反応を引き起こす活性中心から離れた位置、もしくは活性を持つ領域（酵素ドメイン）とは別に存在しその活性を制御する領域（制御ドメイン）に位置していることが判明した（図６）。このことはこれらの変異がタンパク質の機能を喪失させるような大きな影響を与えるわけではないが、ある程度の影響を与える比較的マイルドな効果を持つ変異であることを示唆している。GWAS 解析は解析対象である人の集団中で比較的多くの人に見られる遺伝子多型を同定する手法であり、今回同定された５種類の遺伝子多型はいずれも１割程度以上の人に見られるものであった。このように比較的頻度が高く見られる多型はその効果がマイルドであることが多い。一方大きな影響を与えるような多型はヒトの生存に関わるためその頻度は低いことが多い（数パーセント以下）。以上のことは、人類史の中で比較的マイルドな効果を持つ多型はその効果が弱いが故に消えずに残り、ヒトの集団の中で代謝などの表現型の個人差を生み出す原因となっていることを示唆している。そしてそのようなマイルドな多型がもたらすアミノ酸変異はタンパク質の活性中心から離れた（影響の小さい）位置に存在すことを示唆している。このことは逆に効果が不明な遺伝子多型の効果を推定する上でタンパク質の立体構造が非常に役に立つことを示している。

当機構では引き続き MGWAS 解析を進めておりさらに多くの種類の代謝物について遺伝子多型の影響を受けていることを明らかにしてる（投稿中）。

４．まとめ

NMR 法は様々な研究開発分野で利用されているが、医学研究においては MRI（磁気共鳴画像）法が広く臨床や研究に使われる一方、NMR 法は余り普及が進んでいない。しかし長期に安定して定量性の高いデータを取得することができるメリットを生かして、当機構のように大規模な健康調査を行っている世界各地の研究機関や病院で利用が増えている。当機構では今後も NMR メタボローム解析によるヒト集団の代謝環境の解明を進める一方、タンパク質の立体構造解析等も進めることで、生命現象を分子レベルで解明することを目指して研究を進めていく。

謝辞

本研究に際し東北メディカル・メガバンク機構の皆様には多大なご協力をいただきました。この場を借りて厚く御礼申し上げます。また当機構のコホート調査にご協力いただきました宮城・岩手両県の住民の皆様にも深く感謝いたします。

本研究は国立研究開発法人日本医療研究開発機構および文部科学省による「東北メディカル・メガバンク計画」、「ゲノム医療実現推進プラットフォーム事業先端ゲノム研究開発（GRIFIN）」、先端研究基盤共用促進事業、各種科研費の支援の元に行われています。

参考文献

[1] S. Koshiba, et al., Genes Cells, 23(2018) 406-417.

[2] S. Tadaka, et al., Nucleic Acids Res., 46(2018) D551-D557

[3] S. Koshiba, et al., Sci. Rep., 6(2016) 31463.