

## 要 旨

2001年にヒトゲノムドラフト配列が作成されたことを契機に全ゲノムの情報を用いたありふれた疾患（多因子疾患）の遺伝的関連解析（Genome Wide Association Study; GWAS）が開始され、2005年に遺伝情報の集団レベルのカタログである国際HapMap計画の第二次データが公開されてからは世界中で広く行われるようになった。これは単一の原因遺伝子変異を突き止めるメンデル遺伝性疾患の遺伝解析とは異なる潮流で、数十万から数千万におよぶゲノム上の個人間で異なる遺伝的配列のある部位（遺伝的変異部位）を利用して疾患と関係のある特徴的な遺伝的配列をスキャンする。そのような遺伝的変異の中でも代表的なものが一塩基多型（Single Nucleotide Polymorphism; SNP）である。GWASによって、アルツハイマー型痴呆<sup>1</sup>・関節リウマチ<sup>2</sup>などのありふれた疾患や、や赤血球数<sup>3</sup>のような誰でもばらつきを示すような数値についての遺伝因子に迫ることができるようになった。理化学研究所統合生命医科学研究センターは、その前身である遺伝子多型研究センターが世界で最初にGWASを行って<sup>4</sup>以降、常に世界の最先端と競いながらゲノム解析の発展に貢献してきた。

GWAS結果は頑健で、欧米系集団とアジア系集団という完全に独立なデータでも、極めて似通った解析結果をもたらす。また、ゲノムに対して、ゲノム関連を調べる対象を形質 (trait) と呼ぶが、GWASはこれまで疾患形質のみならず、さまざまな量の形質（例えば身長・体重や血液検査値など）、さらには遺伝子発現形質に関わる SNP (eQTL<sup>5</sup>) など、およそ生物においてなんらかの遺伝性を示すあらゆるデータについて、GWASはその関連 SNP を解き明かしてきた。このとき、数百万以上の遺伝的変異部位からなるゲノムデータの解析結果がこのように安定して強い再現性を示すことは驚きである。

その一方、GWASが検出した形質関連 SNP は、個々には効果が弱く、しかも1つの形質に関連する SNP が非常に多数（数千とも予想されている）あって、それらの生物学的意義は、まだまだ未解明な部分が多い。その中では、エピゲノムデータについて大規模データ解析を行った研究グループ（Roadmap Epigenomics Consortium）は、ゲノム上のエンハンサー領域を同定した後、GWASシグナルがそのエンハンサー領域に濃縮されて存在する（エンリッチしている）ことが示した<sup>6</sup>。この結果は、独立して行われた理化学研究所ライフサイエンス基盤技術センターのFANTOM研究においても確認されている<sup>7</sup>。

すなわち、GWASが検出していた SNP の効果は弱く、かつ1つの形質に対して百を超えるほどの多数の SNP が関連し、それらは遺伝子制御領域に生じていて、遺伝子発現量を変化させることで形質を変化させるのだろう。この生物学経路はさらに追うことができる。すなわち、GWASにより検出された SNP（ゲノム）による遺伝子制御領域（エピゲノム）の変化が、mRNA（トランスクリプトーム）を変化させ、タンパク質（プロテオーム）におそらく量的な影響を与え、その結果代謝物（メタボローム）を変化させて、そのような全体の変化の結果に環境要因も加わることで疾患が発症するのではないかと考えることができる。このような「オーム」のつく解析をまとめて「オミックス解析」と呼び、特定の種類の生体試料をすべて採取して全体として解析する研究手法のことを指すが、ありふれた形質に関連する SNP は個々の効果は弱くかつ数は多い（いろいろな遺伝子に渡って影響がある）わけなので、その下流のトランスクリプトーム、プロテオーム、メタボロームについても全体として変化を追っていく必要がある。そのような統合解析によってGWASの生物学的意義が解明され、あらたな疾患バイオマーカーの発見に結びつく可能性がある。



図1 ゲノムからオミックス解析を直線的に並べた概念図

われわれは理化学研究所統合生命医科学研究センターにおいて引き続き大規模 GWAS を行なっている<sup>8</sup>。<sup>9</sup>。これらは日本人集団において形質関連を示す DNA 配列のリストであると言える。さらに、白血球各分画について、RNA-seq 技術を用いたトランスクリプトーム解析を行っている<sup>10</sup>。

一方、現在京都大学においては、ながはま 0 次コホート研究の参加者による大規模なメタボロミクス解析を行っている。参加者は地域の一般集団からのボランティアであり、まずこのメタボロームデータは今後疾患の解析を行う際のリファレンスデータとなりうる。しかしそれだけではなく、このコホートにはゲノムデータもあるため、これを用いたメタボローム QTL (量的遺伝子座) 解析を行っている。

今回のセミナーでは、島津製作所との共同研究によるガスクロマトグラフ質量分析計 (GC-MS) 技術を用いた、ながはまコホートの数千人レベルのメタボロミクス研究を紹介する。その後、このメタボロームデータのゲノム解析について解説する。最後に理研の GWAS 結果と京都大学のメタボローム GWAS 解析結果を垂直統合することによって、どのような新たな地平線が見えてくるのかについて、現在までの最新の進行状況を説明する。

## 参考文献

1. IGAP consortium. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* 45, 1452-8 (2013).
2. Okada, Y. et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 506, 376-381 (2013).
3. Kamatani, Y. et al. Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nat. Genet.* 42, 210-5 (2010).
4. Ozaki, K. et al. Functional SNPs in the lymphotoxin- $\alpha$  gene that are associated with susceptibility to myocardial infarction. *Nat. Genet.* 32, 4-6 (2002).
5. The GTEx Consortium et al. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* (80-. ). 348, 648-60 (2015).
6. Farh, K. K.-H. et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 518, 337-43 (2015).
7. Andersson, R. et al. An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455-461 (2014).
8. Low, S.-K. et al. Identification of six new genetic loci associated with atrial fibrillation in the Japanese population. *Nat. Genet.* 49, 953-958 (2017).
9. Akiyama, M. et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nat. Genet.* 49, 1458-1467 (2017).
10. Ishigaki, K. et al. Polygenic burdens on cell-specific pathways underlie the risk of rheumatoid arthritis. *Nat. Genet.* 49, 1120-1125 (2017).