

遺伝統計学による疾患病態解明



医療と技術

岡田 随象*

Department of Statistical Genetics,
Osaka University Graduate School of Medicine

Key Words : Statistical genetics, genome-wide association study, drug repositioning

はじめに

遺伝統計学 (statistical genetics) は、ヒトの遺伝情報と形質情報のつながりを統計学の観点から検討する学問である。古くは、メンデルの遺伝の法則まで歴史のある学問であり、疾患を中心とするヒト形質の遺伝的背景の解明に貢献を果たしてきた。以前は、実際に取得できるヒト集団の遺伝情報が、遺伝子マーカーおよびサンプル数の双方において限られていたため、シミュレーション解析や数理モデルに則った理論的な研究が主流であった。特に、集団遺伝学の分野におけるヒトゲノムの動態を評価する学問においては、20世紀において多くの発見が報告されている¹⁾。

21世紀に入り、ヒトゲノム配列の整備と共に、各人種集団におけるヒトゲノム全域での一塩基多型 (single nucleotide polymorphism; SNP) の詳細が明らかとなったことを受け、疾患感受性遺伝子変異をゲノムワイドにスクリーニングすることが可能となった。ヒトゲノム全域に位置する数十万~数百万カ所の SNP を、マイクロアレイ技術を用いて計測し、疾患罹患者と対象者との間で各 SNP の遺伝子変異頻度比較を行うゲノムワイド関連解析 (genome-wide association study; GWAS) が実現化したことにより、多くの疾患感受性遺伝子変異の同定へとつながった。現在、GWAS は国際共同研究コンソーシアムの主

導のもと、複数人種を対象に数万人~数十万という大規模で実施することが一般的となっている。既に1000以上のヒト形質の個人差に影響を与える遺伝子変異が、1万カ所以上同定されており、ヒト形質の遺伝的背景の解明が急速に進んでいる。SNP マイクロアレイのコストは、現在1サンプル辺り5000円程度まで下がっており、これは試料からのDNA抽出費用とほぼ同じである。以前のように、特定の疾患研究を目指して限られたサンプルのゲノムデータを構築する時代から、通常のサンプル収集作業の一環としてゲノムデータを作っておく時代へと変化が生じている。

近年、遺伝統計学に変革をもたらしたのが、次世代シーケンサー (next generation sequencing; NGS) の登場である。特に、大型次世代シーケンス機器の導入はヒトゲノム配列の解読のハイスループット化をもたらした。個人のヒトゲノムの全配列 (約30億塩基対) の解読を行う全ゲノムシーケンシング (whole-genome sequencing; WGS) は、現在1サンプルあたり10万円以下という低コストで実現している。次世代シーケンシングにより、集団内でごく少人数しか保有していない、低頻度の遺伝子変異 (レアバリエント) の直接的な観測が可能となったため、疾患ゲノム解析だけでなく集団遺伝学的な見地からも、新たな成果が得られている。なお、大型次世代シーケンシング機器は、高額な値段や頻繁の新規機種発売等の事情により、大学施設よりも企業が保有する例が多くなっている。ヒトゲノム研究の商業化が進んでいる例と解釈することもできる。

一方、ゲノム研究の低コスト化や大規模化はあくまで手段であり、最終目的ではない。ヒトゲノム研究の目的は疾患病態の解明や個別化医療の実現、ゲノム創薬の実現にある。しかし、大規模疾患ゲノム解析の成果を適切に解釈し、社会実装へとつながる



* Yukinori OKADA

1980年9月生まれ
東京大学大学院医学系研究科博士課程修了 (2007年)
現在、大阪大学 大学院医学系研究科遺伝統計学 教授 博士 (医学)
遺伝統計学
TEL : 06-6879-3971
E-mail : yokada@sg.med.osaka-u.ac.jp

具体的な方法論については、研究の遅れが指摘されている。筆者らは、遺伝統計学の活用がその解決に寄与できるとして、研究活動を進めている。

遺伝統計学による疾患病態の解明

大規模疾患ゲノム解析の成果を、遺伝統計学を用いて多彩な医学・生物学のデータベースと統合することに依り、疾患病態の解明につながると期待されている。近年注目を集めているのが、異なるサンプル集団由来のオミクス情報を、インフォマティクス技術を用いて統合する横断的オミクス解析 (trans-layer omics analysis) である。同一サンプル由来のオミクス情報を蓄積する従来の多層的オミクス解析 (multi-layer omics analysis) と比較して、オミクス情報層の数や質、対象サンプル数を飛躍的に上昇させることが可能となっている。また、

ENCODE (Encyclopedia of DNA Elements) プロジェクトや Roadmap Epigenomics プロジェクトなどの国際コンソーシアムにより、多数の細胞組織におけるエンハンサーやプロモーター領域などのエピゲノム情報のカタログが構築されたことは、ヒト形質における組織特異性の網羅的な検討を可能とした。

我々は、日本人集団 17 万人を対象に、60 の量的形質に対するゲノムワイド関連解析を実施した^{3,4)}。対象となった量的形質には、身長・肥満といった身体計測値、赤血球・白血球・血小板の分画といった血液検査測定値、脂質・肝機能・腎機能・血糖値・電解質・炎症マーカーといった生化学検査測定値、血圧や心エコーなどの生理検査測定値における個人差が含まれている。1500 以上の関連遺伝子変異が同定され、そのうち半数以上は新規の報告となった (図1)。

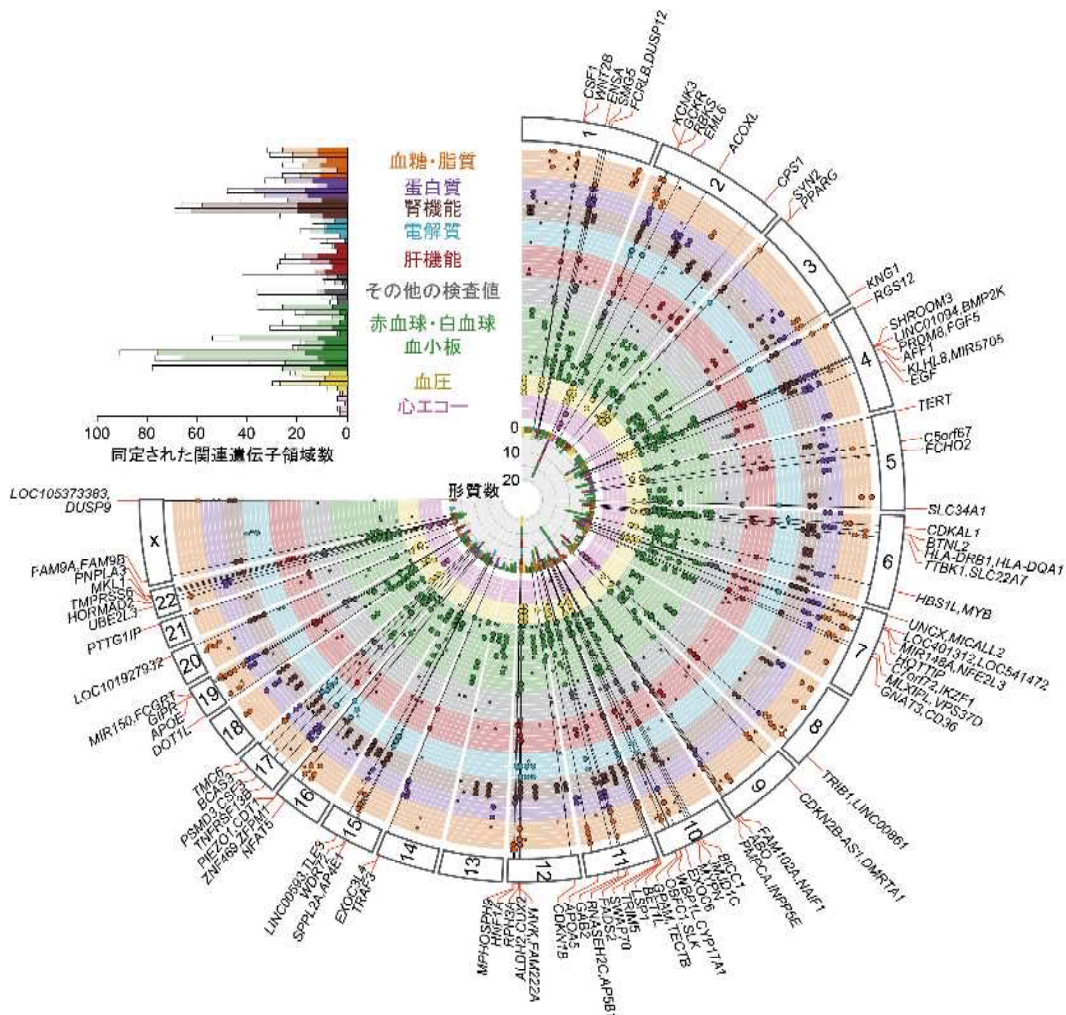


図1：日本人集団 17 万人を対象とした量的形質のゲノムワイド関連解析結果。文献4より改変して引用。

更に、40の疾患におけるゲノムワイド関連解析の結果を追加し、100のヒト形質のゲノムワイド関連解析を対象とした横断的オミクス解析を実施した。米国ハーバード大学で開発された遺伝統計解析手法であるLDSC (linkage disequilibrium score regression) を使用して、各疾患における細胞組織の特異性や、疾患同士の遺伝的背景の相関関係を網羅的に検討した。その結果、身長や腎機能の個人差に影響を与える遺伝子変異は、各々、骨系統組織や腎臓組織のエンハンサー領域に特異的に分布していることなど、既存の生物学的知識に合致した結果が得られた。免疫系に対して抑制的に働く制御性T細胞においては、最も強固な関連を示したのはバセドウ氏病であった。このように、疾患ゲノム情報とエピゲノム情報のみを使用しているにもかかわらず、既存の生物学的知識を再構築することに成功した点は、大規模ゲノム・エピゲノム解析の成果を活用した横断的オミクス解析の可能性を示す結果と考えられた。

ヒトゲノム情報を活用した日本人集団の歴史の解明

大規模ヒトゲノム情報を活用することで、各人種集団の歴史を紐解くことが可能となる。例えば、鎌状赤血球症の原因となる、ヘモグロビン遺伝子の変異は、マラリアが発生する地域の住民において高頻度に保有されている。これは、同遺伝子変異がマラリア感染への耐性を持つため、遺伝子変異を有する個体が結果として優先的に選択されてきた、という歴史を示している。同様に、チベットに住む住民では、高所であるための低酸素環境に適応した遺伝子変異を有する、欧米人は身長を高くする遺伝子変異が増えているなど、各人種のゲノム情報を紐解くことで、人種特異的な選択圧を知ることができる。

我々は、日本人集団における選択圧解析を実施するため、日本人集団2000名における全ゲノムシーケンス情報を用いた選択圧のゲノムワイドなスクリーニングを実施した⁵⁾。特に、全ゲノムシーケンス情報を用いたことで、集団中で非常に低い頻度を持つ遺伝子変異の検討が可能になり、過去数千年という「近い過去」における選択圧の検討が可能になった。これは、従来のSNPマイクロアレイを用いた選択圧解析が過去数万年という「遠い過去」を対象としていたことと対照的であり、日本人集団が

日本列島に居住してきてからの生活環境における選択圧の検討が初めて可能となった。

実際の選択圧の解析には、全ゲノムシーケンス情報を対象とした最新の解析手法である、singleton density score (SDS) 法を採用した。SDS法の適用により、日本人集団において有意な選択圧を示す4カ所の遺伝子領域を複数同定することに成功した (*ADH1B* 遺伝子、*MHC* 領域、*ALDH2* 遺伝子、*SERHL2* 遺伝子)。さらに、日本人集団を対象とした既存のゲノムワイド関連解析によって同定されたヒト形質関連遺伝子変異における選択圧を定量化することで、どの形質に対して選択圧がはたらいていたかを検討した。その結果、日本人集団において最も強い選択圧の影響を受けていたのは、飲酒量などのアルコール代謝関連形質および、脂質・血糖値・尿酸値・電解質・タンパク質など、栄養代謝に関わる形質であることが明らかとなった (図2)。これは、日本人集団において、他の人種集団とは異なる選択圧が働いていたことを示す結果と考えられた。一方、過去の形質や生活環境を直接観測することはできな

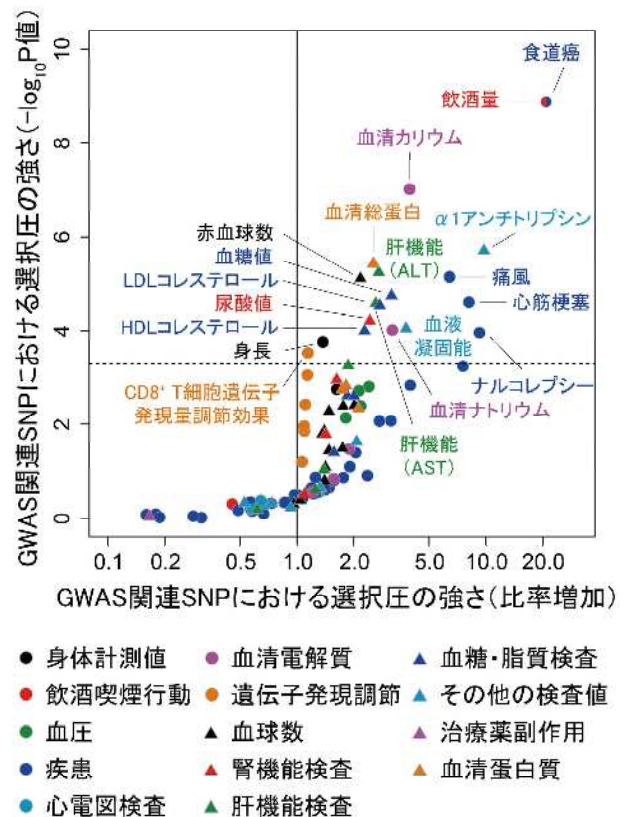


図2：日本人集団において疾患や臨床検査値に影響を与える遺伝的変異における選択圧の強さ。文献5より改変して引用。

いため、何故アルコール代謝や栄養代謝に関わる遺伝子変異が、日本人集団において選択圧の対象となっていたかは、諸説あるものの依然として不明である⁶⁾。

おわりに

大規模化を続けるヒトゲノム・エピゲノム情報を遺伝統計学の解析手法に基づき活用することで、疾患病態の解明や集団の歴史の解明につながる事が示されつつある。本稿では割愛させて頂いたが、個別化医療やゲノム創薬に向けた試みも始まっている^{1,7,8)}。

本邦においては、ゲノムデータ解析や遺伝統計学の専門家が不足していることが、かねてより指摘されている。当教室では、該当分野に関わる若手研究者の人材育成を目的に、夏期短期セミナーとして「遺伝統計学・夏の学校@大阪大学」を開催している(図3)⁹⁾。

参考文献

- 1) 岡田 随象, 遺伝統計学とヒトゲノム研究, 医学のあゆみ, 266巻5号
- 2) GWAS catalog, <https://www.ebi.ac.uk/gwas/>
- 3) Akiyama M et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nat Genet* 2018;49:1458-1467.
- 4) Kanai M et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet* 2018;50:390-400.
- 5) Okada Y et al. Deep whole-genome sequencing reveals recent selection signatures linked to evolution and disease risk of Japanese. *Nat Commun* 2018;9:1631.
- 6) 岡田 随象, 遺伝統計学における選択圧解析とがんゲノム進化解析 実験医学増刊 2018;36:236-240
- 7) Okada Y et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014;506:376-381.
- 8) Okada Y et al. Construction of a population-specific HLA imputation reference panel and its application to Graves' disease risk in Japanese. *Nat Genet* 2015; 47:798-802.
- 9) 遺伝統計学・夏の学校@大阪大学, http://www.sg.med.osaka-u.ac.jp/school_2017.html

**遺伝統計学
夏の学校
@大阪大学** 参加費：無料

夏の大阪で、ゲノム解析の世界に触れてみませんか？

日時：平成30年8月25日(土)～8月27日(月)
 場所：大阪大学吹田キャンパス 大学院医学系研究科
 バイオメディカル教育研究棟9F 遺伝統計学教室
(夏の学校では宿泊・交通の手配は行わず、参加者自身にお任せしています)

対象：大学学部生以上 (大学院生・勤務医・社会人の方も歓迎です)
 形式：講義・ゲノムデータ解析演習・特別セミナー
(演習に使用するPCおよび解析用ゲノムデータは、夏の学校で用意いたします)

定員：35名 (応募者多数の場合、所属や学年を考慮して調整させて頂く場合があります)

時間帯	一日目 (8/25)	二日目 (8/26)	三日目 (8/27)
午前 (9:00~12:00)	遺伝学入門 Linux入門	プログラミング入門	ゲノムデータ 解析演習①
午後 (13:00~16:30)	統計学入門 統計ソフトウェア入門	ゲノムデータ 解析演習②	特別セミナー

特別セミナー講師陣
 石川 俊平 先生 (東京医科歯科大学 難治疾患研究所)
 伊藤 薫 先生 (理化学研究所 生命科学研究所センター)
 白石 友一 先生 (国立がん研究センター 研究所)
(演習演習内容および特別セミナーの内容は、変更となる場合があります)

URL : http://www.sg.med.osaka-u.ac.jp/school_2018.html
 申込み : office@sg.med.osaka-u.ac.jp (担当：岡田 随象)

共催：医学系研究科バイオインフォマティクス・イニシアティブ、新学術領域研究・システム連携次元

図3：遺伝統計学・夏の学校@大阪大学