

小児期・移行期を含む包括的対応を要する希少難治性肝胆膵疾患の調査研究

ゲノムコホート・バイオバンクと希少疾患

研究分担者 栗山進一 東北大学災害科学国際研究所 教授

**研究要旨**

人工知能の一つである機械学習の手法を活用し、自閉スペクトラム症（ASD）が異種の疾患の集合体である可能性があることを世界で初めて発見した。ASDを分割して、「クラスターごとの患者群」と対照群で遺伝情報上の比較を行ったところ、「患者群全員」と対照群の比較では発見できなかった多くの有意な「違い」つまり疾患原因の候補が観察された。機械学習を用いて症例をより均質な集団にクラスタリングすることにより、それぞれの集団の特徴に応じた個人毎のアプローチ、個別化医療が可能となることが期待される。本研究はASDの遺伝的構造と病因を解明し、ASDの精密医療の開発を促進する手がかりを提供するものである。

**A. 研究目的**

ある疾患名でまとめられ潜在的に異質な集団を一塊として扱った「患者群全員」と対照群とのゲノムワイド関連解析（GWAS）は多くの疾患で実施されてきたが、自閉スペクトラム症（ASD）については有意な SNP はみられるものの、その効果は小さく、遺伝的感受性因子を特定したとはいえない状況であると指摘されてきた。しかし、遺伝的要因に関しては、一卵性双生児の ASD の一致率は 92%、二卵性双生児は 10%であることや、兄弟間の ASD の再発リスク比は 22 と高いことなどから、リスクの増加には遺伝的要因が強く寄与していることが示唆されていた。

近年ではデータセットに内在するパターンを明らかにできる人工知能技術である機械学習手法が開発されており、機械学習のクラスタリングアルゴリズムとそれに続く GWAS は、新規でより遺伝的に均質な集団、クラスターを明らかにするのではないかという仮説を立てこれを検証した。

**B. 研究方法**

米国のサイモンズ財団（Simons Simplex Collection）からのデータの分譲を受け、クラスター数 15 の k-means アルゴリズムを使用し、ASD が示す多様な症状とビタミン B6 処方歴を変数としてク

ラスタ分析を行った。597 例の ASD 症例と 370 例の対照のデータセットを使用し、従来の GWAS を実施した。2 番目のステップでは、クラスタリングの結果に基づいてケースを分割し、各サブグループとコントロール（クラスターベース GWAS）で GWAS を実施した。

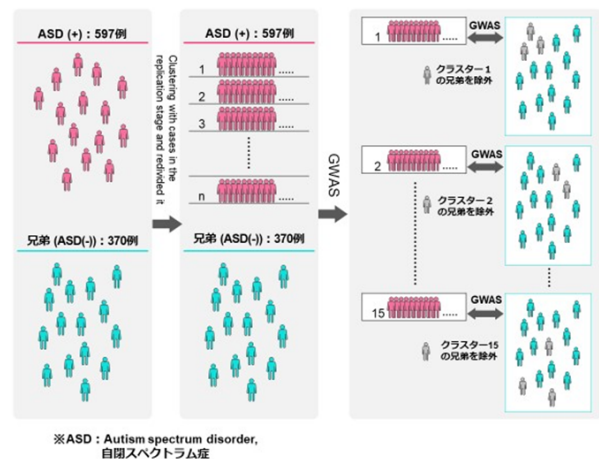


図 1 クラスターごとの患者群と対照群とで実施された GWAS の方法

**C. 研究結果**

従来の ASD という疾患名でまとめられた潜在的に異質な集団を一塊として扱った「患者群全員」と対照群を用いた GWAS 手法で行われた解析では、有意な関連は観察されなかった。クラスターに基づいた GWAS（クラスターベース GWAS）では、65 個の有

意な遺伝子座を特定した。これらの 65 の遺伝子座のうち、SRRM4 遺伝子内に位置する rs11064685 は、再現性を検証する解析でも症例と対照で有意に異なる分布を示し、ASD の原因のひとつである可能性が示唆された。

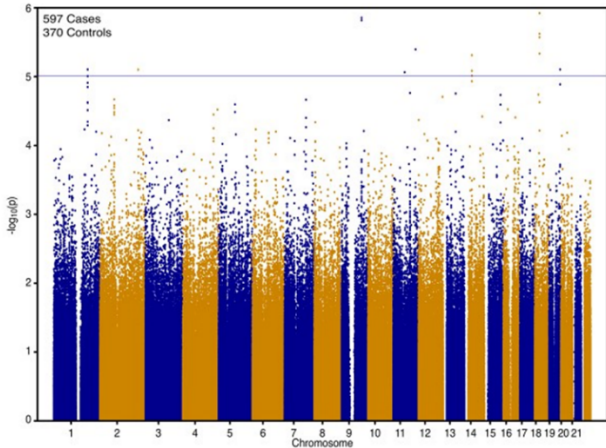


図 2 患者群全員と対照群とで実施された GWAS のマンハッタンプロット

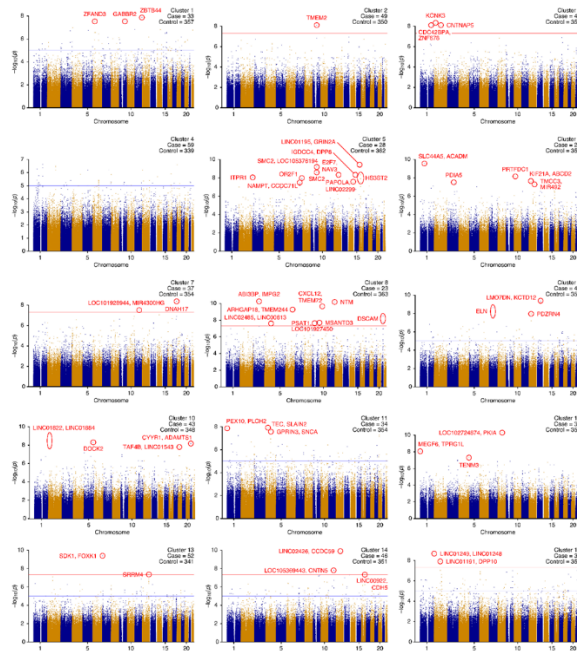


図 3 クラスターごとの患者群と対照群とで実施された GWAS のマンハッタンプロット

#### D. 考察

本研究結果は、クラスタリングが比較的均一な病因をもつサブグループを首尾よく特定できる可能性を示唆している。本研究は、データセットが複数の異種のサブグループで構成されている場合、均質

な個体の数があるかに少ないサブグループでさえ、影響が大きく統計学的に有意な遺伝的要因を検出できることを示した。

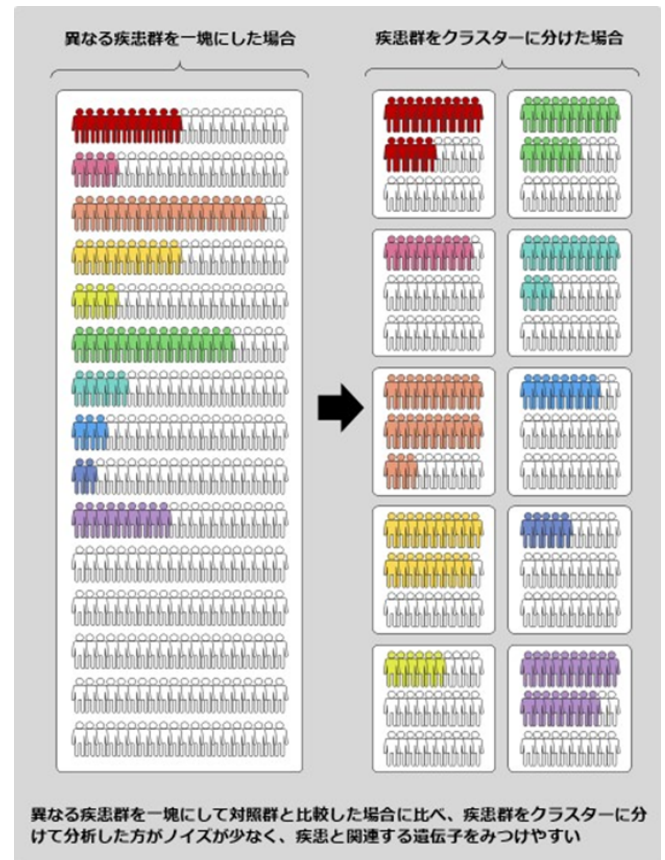


図 4 仮想上の患者群全員とクラスターごとの患者群の比較: クラスターに分けた方がノイズが少ない

これまでの多くの GWAS 研究は、サンプルサイズを大きくすればするほど多くの有意な遺伝子座を特定できることを示してきたが、得られた結果ではオッズ比がせいぜい 1.5 程度と低く、創薬のターゲットとするなど疾患対策を立てるためには不十分な発見であることが指摘されてきた。

本研究のデータは、広く定義された疾患名よりも ASD のより均一なサブグループを識別する手段としてのクラスターベース GWAS の重要性を示している。

本研究結果は、ASD の遺伝的構造と病因を解明し、ASD の精密医療の開発を促進する手がかりを提供するものである。

#### E. 結論

本研究結果から機械学習を用い症例をより均

質な集団にクラスタリングすることにより、それぞれの集団の特徴に応じた ASD の個別化医療が可能となることが期待される。

機械学習の手法を用いてその疾患の特徴を表していると思われる表現型に関する変数を使いサブグループを見つけ出すことで、多くの疾患の個別化医療が大いに進展することが期待される。

## G. 研究発表

査読付論文

タイトル: Clustering by phenotype and genome-wide association study in autism

著者: Akira Narita, Masato Nagai, Satoshi Mizuno, Soichi Ogishima, Gen Tamiya, Masao Ueki, Rieko Sakurai, Satoshi Makino, Taku Obara, Mami Ishikuro, Chizuru Yamanaka, Hiroko Matsubara, Yasutaka Kuniyoshi, Keiko Murakami, Fumihiko Ueno, Aoi Noda, Tomoko Kobayashi, Mika Kobayashi, Takuma Usuzaki, Hisashi Ohseto, Atsushi Hozawa, Masahiro Kikuya, Hirohito Metoki, Shigeo Kure, Shinichi Kuriyama

掲載誌: Translational Psychiatry

掲載日: 2020年8月17日

DOI: 10.1038/s41398-020-00951-x

URL: <https://rdcu.be/b6jMh>

## H. 知的財産権の出願・登録状況

1. 特許取得  
なし
2. 実用新案登録  
なし
3. その他  
なし