

統合失調症の新たな遺伝的メカニズムを解明

— 遺伝子産物の「切り出し」に影響する DNA 多型のビッグデータ解析から —

要旨

理化学研究所（理研）脳科学総合研究センター精神疾患動態研究チームの高田篤客員主幹研究員（現横浜市立大学大学院医学群遺伝学 講師）、加藤忠史チームリーダー、横浜市立大学大学院医学群遺伝学の松本直通教授らの共同研究チームは、遺伝子産物の「切り出し（スプライシング）」に影響する DNA 配列の個人差（DNA 多型）の網羅的な解析を行い、統合失調症^[1]の新たな遺伝的メカニズムを解明しました。

DNA から転写された RNA は通常、そのままタンパク質として翻訳されません。一次的な転写産物であるメッセンジャーRNA（mRNA）前駆体から翻訳される部分（エクソン）のみが切り出され、最終的にタンパク質の鋳型となる成熟 mRNA が形成されます。この切り出し過程は「スプライシング」と呼ばれます。また、一つの mRNA 前駆体から異なる組み合わせのエクソンを持つ成熟 mRNA が合成されることもあります。これは「選択的スプライシング」と呼ばれ、限られた数の遺伝子から多様な生命活動を生み出すメカニズムとして注目されています。

共同研究チームは CommonMind Consortium^[2]で公開されている 206 名分の死後脳の RNA シーケンスデータ^[3]を解析し、脳内の「選択的スプライシングを制御するゲノム領域（splicing quantitative trait loci: sQTL）」を網羅的に同定しました。さらに、これらの sQTL が過去のゲノムワイド関連解析（GWAS）^[4]でさまざまな疾患との関連が報告されているゲノム領域に多いこと、なかでも統合失調症関連領域で多く認められることを明らかにしました。さらに、統合失調症を対象とした大規模な GWAS のデータを解析し、脆弱 X 症候群^[5]の原因遺伝子と相互作用を示す *FXR1* やシナプス^[6]機能に関わる *SNAP91* などの選択的スプライシングを制御する sQTL が、特に統合失調症と強く関連することを明らかにしました。このことは、統合失調症と関連する一塩基多型（SNP）^[7]の少なくとも一部が sQTL として働き選択的スプライシングを制御することによって疾患リスクに寄与するという、新たな遺伝的メカニズムが存在することを示しています。

今後、これらの遺伝子の選択的スプライシングが脳神経系の機能や発達に及ぼす影響を検討することで、統合失調症の病態メカニズムの一端を解明できると期待できます。また、将来的には統合失調症に対する新たな治療法の開発に貢献する可能性もあります。なお、本研究で作成した sQTL のリストは、統合失調症のみならず、さまざまな疾患の遺伝的メカニズムの研究において有益なり

ソースとして利用できます。

本研究成果は、国際科学雑誌『*Nature Communications*』に掲載されます（2月27日付け）。

本研究は日本医療研究開発機構（AMED）「脳科学研究戦略推進プログラム」事業、日本学術振興会 科学研究費助成事業 若手研究（A）の支援を受けて行われました。

1. 背景

DNA から RNA が転写され、それを鋳型としてタンパク質が翻訳されるという情報伝達の流れは、分子生物学の中心教義（セントラルドグマ）として知られています。ヒトを含む真核生物では、転写された RNA がそのままタンパク質として翻訳されるわけではありません。一次的な転写産物であるメッセンジャー RNA（mRNA）前駆体から、翻訳される部分であるエクソン（ゲノムのうちタンパク質をコードしている部分）のみが切り出され、タンパク質の鋳型となる成熟 mRNA が形成されます。この切り出し過程は「スプライシング」と呼ばれます。また、一つの mRNA 前駆体から異なった組み合わせのエクソンを持つ成熟 mRNA が合成されることがあり、この機構は「選択的スプライシング」といいます。選択的スプライシングは高度な調節を受けており、しばしば特定の組織、細胞でのみ観察されます。特に霊長類の脳では、複雑な選択的スプライシングが行われていることが近年の研究で明らかになっています^{注 1,2)}。

また選択的スプライシングは、DNA 配列の個人差（DNA 多型）によって影響を受けることが知られており、「選択的スプライシングを制御するゲノム領域（splicing quantitative trait loci: sQTL）」の研究も徐々に進展しています。しかし、これまでに、最も複雑な選択的スプライシングが行われているヒトの脳由来のデータを用いて、大規模かつ網羅的に、高解像度の sQTL 解析を行った研究はありませんでした。

そこで、共同研究チームはヒトの死後脳から得られた転写産物データから、大規模な sQTL 解析を試みました。

注 1) GTEx Consortium. Human genomics. “The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans.” *Science* 348, 648–660. 2015.05

注 2) Merkin, J., Russell, C., Chen, P., and Burge, C.B. “Evolutionary dynamics of gene and isoform regulation in Mammalian tissues.” *Science* 338, 1593–1599. 2012.12

2. 研究手法と成果

共同研究チームはまず、CommonMind Consortium が公開しているヒト死後脳（前頭葉）サンプルの RNA シーケンス^[7]データのうち 206 名分を解析し、脳内で観察される選択的スプライシングを網羅的に同定しました。例えば、エクソン A、B、C の三つが並んでいるときに、ある成熟 mRNA にはエクソン B が含ま

れているのに別の成熟 mRNA には含まれていない、といった現象を検出しました。続いて、同じサンプルから得られた一塩基多型 (SNP) のデータと選択的スプライシングのデータを比較し、選択的スプライシングを制御するゲノム領域のマーカーとなる多型 (sQTL SNP) を同定しました。例えば、ある SNP の遺伝子型が A/A (アデニン/アデニン) だと特定のエクソンが成熟 mRNA に含まれるのに、遺伝子型が A/G (アデニン/グアニン) もしくは G/G (グアニン/グアニン) だと成熟 mRNA にエクソンが含まれない、といった多型が sQTL SNP です。統計学的な処理を行い、互いに独立と考えられる sQTL SNP を 1,539 個同定しました。

次に、さまざまなヒトの疾患を対象として行われたゲノムワイド関連解析 (GWAS) のデータを GWAS Catalog^[8] からダウンロードし、sQTL と疾患関連領域の関係について検討しました。まず、何らかのヒトの疾患との関連が報告されているゲノム領域に sQTL SNP が多いかどうかを、sQTL ではないと考えられる SNP と比較して検討したところ、疾患関連領域では有意に sQTL SNP が多いことが分かりました。続いて、個別の疾患と関連するゲノム領域に sQTL SNP が多いかどうかを解析したところ、統合失調症、多発性硬化症^[9] といった脳神経系の疾患と関連する領域で、特に sQTL SNP が多いことが分かりました (図 1)。

統計学的有意差 (P 値)^[10]、オッズ比^[10] の両方の観点から評価すると、特に統合失調症関連領域で sQTL SNP が多いという結果が得られました (図 1)。そのため、少なくとも一部の統合失調症関連領域では、sQTL が mRNA の選択的スプライシングに影響を与え翻訳されるタンパク質が変化することによって、疾患リスクに関わっていると考えられます。特に、ある SNP が統合失調症の有無と高い有意水準で関連し、かつその SNP が sQTL である場合には、その可能性が高まると想定されます。

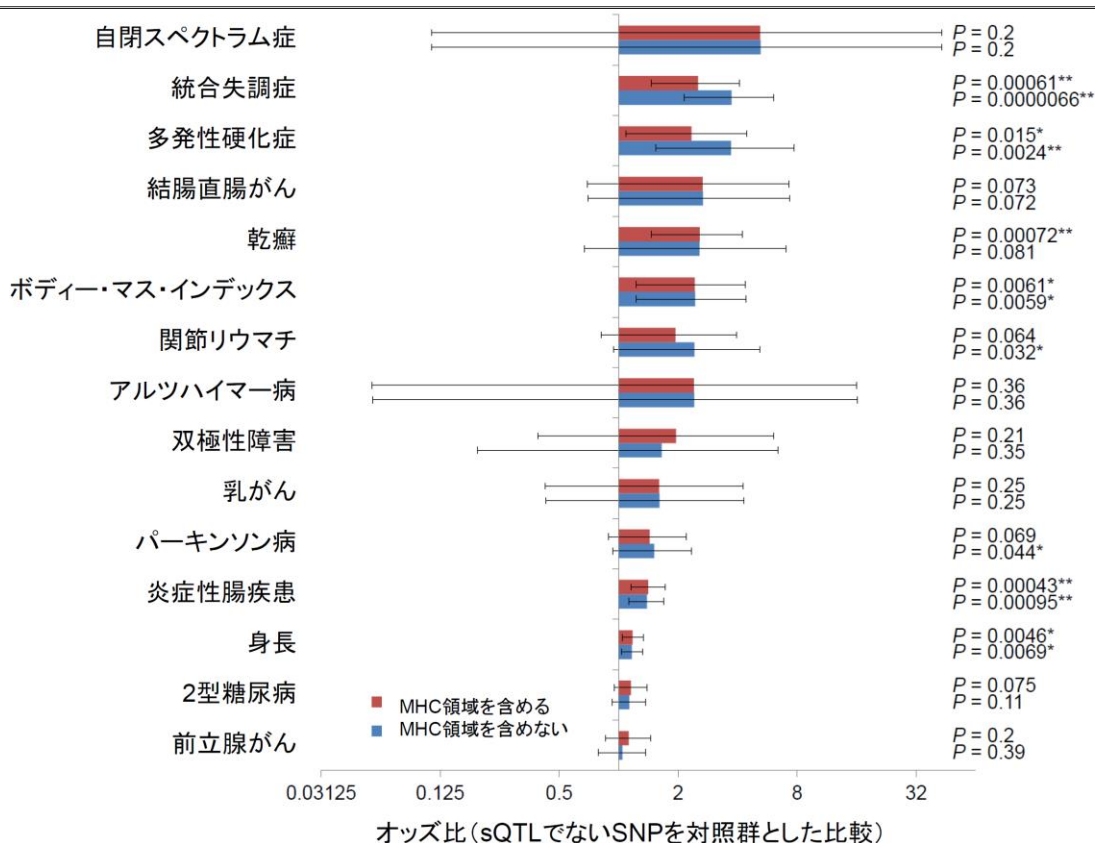


図 1 GWAS データを用いた各疾患の関連ゲノム領域における sQTL SNP の局在に関する検討

主要組織適合遺伝子複合体 (MHC、下記参照) のデータを含めた解析 (赤のバー) と MHC を含めない解析 (青のバー) を行った。横軸はオッズ比 (右に行くほど、疾患関連領域に sQTL SNP が多い) を示しており、MHC を含めない解析でオッズ比が高い順に上から並べた。なお、自閉スペクトラム症でオッズ比が高いのに統計学的に有意な関連でないのは、自閉スペクトラム症を対象として行われた GWAS の規模が小さいためである。逆に、炎症性腸疾患で影響がごく小さいのに統計学的な関連が認められているのは、大規模な GWAS が行われているためである。

主要組織適合遺伝子複合体とは、免疫反応に必要な遺伝子を多く含むゲノム領域。ヒトゲノムの中で最も多様性が高い領域の一つであり、さまざまな疾患と関連している一方、その解析は複雑であるため、結果にバイアスを生じる原因となることがある。そのため、本研究では MHC を含めた解析と含めない解析の両方を行った。MHC は major histocompatibility complex の略。

さらに、Psychiatric Genomics Consortium^[11]による 15 万人以上を解析した統合失調症の GWAS データを用いて、ゲノムワイドな有意水準^[12]で統合失調症と関連する sQTL SNP を探索しました。その結果、四つのゲノム領域 (染色体 3 番に二つ、6 番と 14 番の一つずつ) にある *NEK4*、*FXR1*、*SNAP91*、*APOPT1* という遺伝子の選択的スプライシングに関連する sQTL SNP が、統合失調症リスクとも強く関連することが分かりました (図 2)。それぞれの機能としては、*NEK4* は DNA 損傷の修復などに関わるタンパク質を作る遺伝子であり、脳で多く発現しています。*FXR1* から作られるタンパク質は、脆弱 X 症候群の原因遺伝子 (*FMR1*) から作られるタンパク質と相互作用を示すことが知られています。*SNAP91*からはシナプスを介した情報伝達に重要なタンパク質が作られ、*APOPT1*はミトコンドリアに局在し、細胞死などに関連するタンパク質を作ります。

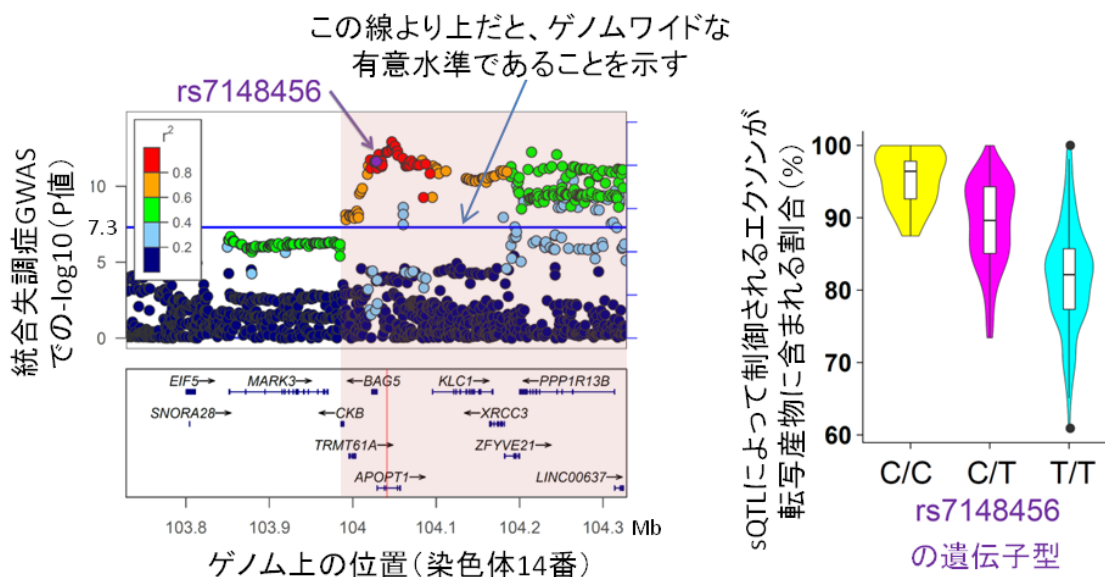


図 2 統合失調症と強く関連する染色体 14 番にある *APOPT1* 遺伝子の sQTL SNP

rs7148456 という ID の SNP は、ゲノムワイドな有意水準 (p 値が 5×10^{-8} 未満、 $-\log_{10}(5 \times 10^{-8}) \approx 7.3$ 以上) で統合失調症と関連しており、さらに本研究で sQTL SNP として同定されている。左図上の一つ一つの丸は各 SNP に対応しており、rs7148456 との連鎖 (近接した SNP が同時に子に伝達されること) の程度 (r^2) によって色分けされている。左図右半分の範囲 (背景がピンクの部分) でゲノムワイドな有意水準の関連 SNP が認められた。右図は、rs7148456 の遺伝子型が、*APOPT1* 遺伝子の選択的スプライシングに及ぼす影響を示している。遺伝子型が C/C (シトシン/シトシン) だと、選択的スプライシングを受けるエクソンがほとんどの転写産物に含まれるのに対して、T/T (チミン/チミン) では平均で 8 割ほどしか含まれない。

3. 今後の期待

本研究で得られた、ヒト疾患関連ゲノム領域では sQTL SNP が有意に多いという所見は、これらの疾患と関連する SNP の少なくとも一部が、sQTL として働き選択的スプライシングを制御することによって疾患リスクに寄与するという、新たな遺伝的メカニズムが存在することを示しています。なかでも、脳組織サンプル由来のデータを用いて本研究で同定した「脳 sQTL」は、統合失調症の遺伝的リスクをよく説明すると考えられます。

また通常 GWAS では疾患関連座位がある程度幅広いゲノム領域として同定され、多くの場合その中には複数の遺伝子が存在します。例えば図 2 では左図の中央から右半分の範囲 (背景がピンクの部分) でゲノムワイドな有意水準の関連 SNP を認めています。その中だけでも 9 個の遺伝子 (*CKB*, *TRMT61A*, *BAG5*, *APOPT1*, *KLC1*, *XRCC3*, *ZFYVE21*, *PPPIR13B*, *LINC00637*) があり、どの遺伝子の機能が統合失調症と関わっているかを特定するのは容易ではありません。しかし今回の sQTL のデータを用いると、それら複数の遺伝子の中で“*APOPT1* が特に有望な疾患感受性遺伝子の候補であり、その選択的スプライシングの異常が病態に関わっている可能性がある”という仮説を構築することができます。

こういった仮説をさらに検証するためには、sQTL を導入した細胞や動物を用いた実験や、患者由来のサンプルの解析などを行う必要がありますが、今回の

sQTL のデータリソースを用いることによって、統合失調症のみならずさまざまな多因子疾患において、有望な候補遺伝子を絞り込むことができると考えられます。

また今回の研究は、CommonMind Consortium の RNA シーケンスデータ、Psychiatric Genomics Consortium の統合失調症 GWAS データ、GWAS Catalog に登録されているその他の疾患の GWAS データなど、大規模データを複合的に解析し、生物学的に重要な知見を引き出したという点でも意義深いと考えられます。

今後、こうしたビッグデータの複合的解析をさらに推進することによって、精神神経疾患を含む多因子疾患の遺伝的メカニズムの解明につながると期待できます。

4. 論文情報

<タイトル>

Genome-wide identification of splicing QTLs in the human brain and their enrichment among schizophrenia-associated loci

<著者名>

Atsushi Takata, Naomichi Matsumoto and Tadafumi Kato

<雑誌>

Nature Communications

<DOI>

10.1038/NCOMMS14519

5. 補足説明

[1] 統合失調症

統合失調症は幻覚や妄想、意欲の低下、感情の平板化などを主要な症状とする精神疾患。発症後には社会的機能の低下を伴うことが多い。生涯罹患率は人口の約 1%と報告されている。

[2] CommonMind Consortium

疾患に対する専門知識と、脳組織サンプルおよびデータを提供するための官民パートナーシップで、マウント・シナイ医科大学、ペンシルベニア大学、ピッツバーグ大学、武田薬品工業株式会社、エフ・ホフマン・ラ・ロシュ社などが参加している。

詳細は <https://www.synapse.org/#!/Synapse:syn2759792/wiki/69613>

[3] RNA シーケンス

RNA から逆転写した cDNA (RNA と相補的な DNA) を、次世代シーケンサーで解析する実験手法。RNA の定量のみならず、新規転写産物の検出や、選択的スプライシングイベントの定量、新規同定を行うことができる。

[4] ゲノムワイド関連解析 (GWAS)

ゲノム中の数十万から数百万の SNP を網羅的に調べ上げ、疾患の有無や、身長・体重などの形質と関連するゲノム領域を同定する研究手法。GWAS は genome-wide association study の略。

[5] 脆弱 X 症候群

X 染色体上の *FMR1* 遺伝子の異常が原因となる遺伝性疾患で、知的障害や特徴的な顔貌などの症状を示す。

[6] シナプス

神経細胞同士の情報伝達に関わる構造。情報を伝える細胞と伝えられる細胞の間には約 20 ナノメートル (1 億分の 2 メートル) のすき間がある。情報を伝える細胞はこのすき間に神経伝達物質を放出し、伝えられる細胞側の神経伝達物質受容体がそれを受け取ることにより神経情報が伝わる。

[7] 一塩基多型 (SNP)

DNA 配列の個人差のうち、標準的な配列と 1 塩基だけ異なっており、かつ集団内で 1% 以上の頻度で観察されるもの。1 人あたり 2 コピーのゲノムを持つので、50 人中 1 人が、片方のゲノムで SNP を持っているとならば頻度 1% となる。SNP は single nucleotide polymorphism の略。

[8] GWAS Catalog

さまざまな疾患、形質を対象とした GWAS のデータがまとめられたカタログ。米国の国立ヒトゲノム研究所 (National Human Genome Research Institute) と英国の欧州バイオインフォマティクス研究所 (European Bioinformatics Institute) によって運営されている。詳細は <https://www.ebi.ac.uk/gwas/>

[9] 多発性硬化症

脱髄疾患の一種で、脳や脊髄、視神経などに神経症状が出て、しびれや麻痺などを引き起こす。治癒と再発を繰り返すことが知られており、日本では特定疾患に認定されている難病。遺伝要因、環境要因の双方が発症に関与していると考えられるが、その原因は解明されていない。

[10] P 値、オッズ比

P 値は、偶然にそのようなことが起こる確率のことで統計学的有意差を示す指標。数値が低いほど有意水準が高いことを表す。オッズ比は、ある事象の起こりやすさについて二つの群で比較したときの違いを示す統計学的尺度の一つ。本研究では、疾患関連領域に含まれるか否かという事象について、sQTL SNP と sQTL ではない SNP の 2 群を比較してオッズ比を算出している。

[11] Psychiatric Genomics Consortium

精神疾患のゲノム解析データを共有し、大規模研究を行うための国際コンソーシアム。詳細は <https://www.med.unc.edu/pgc>

[12] ゲノムワイドな有意水準

GWAS では数十万から数百万の SNP を一度に調べるため、偶然疾患などに関連しているように見える SNP が多数検出される（偽陽性の問題）。そのため、「P 値が 5×10^{-8} 未満」という高い有意水準で関連を認めた SNP のみを真の陽性としており、この P 値の基準をゲノムワイドな有意水準と呼ぶ。

6. 発表者・機関窓口

<発表者> ※研究内容については発表者にお問い合わせ下さい
理化学研究所 脳科学総合研究センター 精神疾患動態研究チーム
客員主幹研究員 高田 篤（たかた あつし）
（現横浜市立大学大学院医学群 遺伝学 講師）
チームリーダー 加藤 忠史（かとう ただふみ）
TEL：045-787-2606（高田）、048-467-6949（加藤）
FAX：045-786-5219（高田）、048-467-6947（加藤）
E-mail：atakata@brain.riken.jp、もしくは、atakata@yokohama-cu.ac.jp（高田）
kato@brain.riken.jp（加藤）

横浜市立大学大学院医学群 遺伝学
教授 松本 直通（まつもと なおみち）
TEL：045-787-2606 FAX：045-786-5219
E-mail：naomat@yokohama-cu.ac.jp

<機関窓口>
理化学研究所 広報室 報道担当
TEL：048-467-9272 FAX：048-462-4715
E-mail：ex-press@riken.jp

<事業に関するお問い合わせ>
日本医療研究開発機構 戦略推進部 脳と心の研究課
Tel：03-6870-2222 Fax：03-6870-2244
E-mail: brain-pm@amed.go.jp