

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 5 月 17 日現在

機関番号：12601

研究種目：若手研究（B）

研究期間：2010 ～ 2011

課題番号：22791107

研究課題名（和文） 次世代シーケンサーを利用した自閉症候補遺伝子領域のターゲットリシーケンス

研究課題名（英文） Target resequencing of the candidate region of the susceptibility gene for autism by using the next-generation sequencer

研究代表者

栃木 衛（TOCHIGI MAMORU）

東京大学・医学部附属病院・講師

研究者番号：40456116

研究成果の概要（和文）：4 世代にわたり自閉症性障害を発症している家系を対象としたエクソーム解析の結果では、疾患と完全に共分離している複数の候補遺伝子を見出しており、新たな自閉症候補遺伝子の発見につながる可能性がある。また、自閉症性障害患者とその両親 163 組についてのコピー数多型解析では、500K 塩基対以上の大きさの新規 CNV が 7 組のサンプルで 8 つ認められており、そのうちの 1 つは 15 番染色体長腕 11.2 領域における重複であることが判明した。

研究成果の概要（英文）：As a result of the exome sequencing of a four-generation-extending autism spectrum disorder (ASD) family, several candidate genes, which completely co-segregate with the disease, were found and may lead to a new finding. In the copy number variation (CNV) analysis of 163 trios of ASD, eight de novo CNVs (> 500 kb) were found in seven trios, one of which was duplication in 15q11.2.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010 年度	1,800,000	540,000	2,340,000
2011 年度	1,200,000	360,000	1,560,000
年度			
年度			
年度			
総計	3,000,000	900,000	3,900,000

研究分野：医歯薬学

科研費の分科・細目：内科系臨床医学・精神神経科学

キーワード：自閉症、候補遺伝子、次世代シーケンサー、コピー数多型（CNV）

1. 研究開始当初の背景

自閉症は 90%以上の高い遺伝率を示すことから、多因子遺伝病の中でも特に強い遺伝要因の関与が示唆されている（Lintas and Persico, 2009）。しかし、染色体異常に伴って発症する場合や一部の症候群を除けば、現在のところはっきりした原因遺伝子の同定にまでは至っていない。その理由の 1 つは、

これまでの研究の多くが、一般集団中にも存在する比較的頻度の高い遺伝子多型が疾患の発症に関与するという Common Variant-Common Disease (CVCD) 仮説に基いて探索を進めていたことにあると考えられる。実際、最近報告された大規模ゲノムワイド関連研究の結果でも、最も有意な関連を示した多型のオッズ比は 1.19 に過ぎず、今後さらに感受性遺伝子の同定を進めようとする

れば少なくとも 1000 人単位のサンプル規模が必要であることが示された (Wang et al., 2009)。一方で、2003 年以降、Neurologin、SHANK3、Neurexin などのシナプス関連遺伝子との関連が相次いで指摘されてきたが、これらの遺伝子変異の頻度は 1% 程度かもしくはそれ以下と極めて稀なものである。また、マイクロアレイを用いてゲノムワイドに行われたコピー数多型 (CNV) 解析では、自閉症患者の 10% に de novo CNV が発見されたが、それらはそれぞれ異なるものであり、頻度としてはやはり極めて稀であることが示された (Sebat et al., 2008)。すなわち、遺伝子によってはその探索に CVCD 仮説ではなく、頻度は稀ではあるが疾患発症への寄与度の高い変異の存在を想定する Multiple Rare Variant-Common Disease (MRVCD) 仮説に基づいたアプローチも必要であると同時に、最近になって高度な多型性を示すことが明らかとなった CNV などの微細構造異常も含めた解析が重要であることが示唆されているものと考えられる。

報告者は、これまでに東京大学・東海大学・鳥取大学の 3 大学で結成された自閉症 DNA コンソーシアムに参加し、サンプルの拡充に寄与するとともに直接分子遺伝学的研究にも携わって 15 番染色体長腕 11-13 (15q11-q13) の領域を中心にその成果を発表してきた (Marui et al., 2004; Tochigi et al., 2007; Kato et al., 2008; Tochigi et al., 2008; Kato et al., 2008)。15q11-q13 はゲノムインプリンティングが起きる領域であり、遺伝子により母親由来・父親由来のどちらのアレルから発現されるかがメチル化などのエピジェネティックな機構により厳密に制御されている。これが欠失やメチル化異常などの原因によって異常をきたすと、Prader-Willi 症候群や Angelman 症候群が惹起されることはよく知られているが、これらの症候群が自閉症と共通する精神症状を示すこと、また自閉症罹患者でもこの領域で重複を中心とする構造異常が 1-3% にも上る高い頻度で認められることから、15q11-q13 は自閉症感受性遺伝子の候補領域としても注目されてきた。報告者らは、SNP を利用した関連研究の手法によりこの領域の解析を進め、その結果、父親由来アレルのみが発現する領域 (PED) に位置する SNURF-SNRPN とその下流の HBII-52、さらに母親由来アレルのみが発現する領域 (MED) に位置する ATP10C において有意な関連を見出した (Kato et al., 2008)。SNURF-SNRPN は同領域のインプリンティングの成立に必要な

領域 (Imprinting Center) とオーバーラップしており、コードされるタンパクはインプリンティングの制御に関与していると考えられている。HBII-52 は SNURF-SNRPN と一括して転写される non-coding RNA で、核小体低分子 RNA (snoRNA) に分類され、その機能はセロトニン 2C 受容体の RNA 編集制御であることが明らかにされている (Kishore and Stamm, 2006)。これらの発現が脳に限局されていること、また、自閉症とセロトニン機能異常の関連が示唆されていることを考えると報告者らの得た結果は極めて興味深いものであり、さらに詳しくこの領域を解析する意義があると考えられた。そこで、次に報告者は MRVCD 仮説からのアプローチの試みとして、アレイ CGH の方法により同サンプルを解析し、同領域における微細構造異常の有無を検討した。その結果、複数のサンプルでシグナル低下領域が認められ、その中には HBII-52 領域における異常を示すサンプルも存在した。CVCD 仮説・MRVCD 仮説双方からのアプローチによる結果から同じ領域の関与が示唆され、かつ機能的にも疾患の病態に関与している可能性が十分考えられることから、報告者は HBII-52 を中心とする領域に焦点を当ててさらに解析を進めた。具体的には、アレイ CGH によって示されたシグナル低下部位の確定、および他のサンプルも含めた微細構造異常や変異のさらなる探索を目標としていたが、HBII-52 は類似の配列が 55kb にわたって 47 コピー繰り返されていることから構造的に従来型シーケンサーや定量 PCR による解析が通用せず、困難を強いられていた。そのような中で、発展の目覚ましい次世代シーケンサーの技術を利用すれば、同領域を効率的かつ詳細に解析できる可能性があることに気づき、本研究を着想した。

2. 研究の目的

本研究は、自閉症罹患者を対象として、次世代シーケンサーを用いたターゲットリシーケンスによる 15 番染色体長腕 11-13 領域の解析を行い、自閉症と関連する遺伝子変異・微細構造異常の探索を行うものである。特に、繰り返し構造などの要因により、従来の方法論では困難であった HBII-52 領域の解析が可能かどうかを確認し、可能な場合はアレイ CGH により同領域での異常が認められているサンプルに適用してシーケンズレベルでのデータを得ることを目的とする。また、可能な範囲で他のサンプルにも応用し、新規の変異・微細構造異常の同定を試みると

ともに、将来的に多検体でリシーケンスを実施するための技術的な蓄積を図る。

3. 研究の方法

(1) ターゲットリシーケンスの方法論的検討

利用可能なターゲットリシーケンスの技術のうち、本計画の目的に対してどの方法が最も妥当であるかの検証を行う。具体的には、ゲノム濃縮の方法として、circularizing probeやRNA probeを用いた solution phase のものと、DNA マイクロアレイを利用した solid phase のものがあり (Summerer, in press)、さらに次世代シーケンサーとしても Roche 454 GS FLX sequencing、Illumina Genome Analyzer sequencing、Applied Biosystems SOLiD sequencing と複数の選択肢があるため、これらのうちどれが最適な方法であるかの検証を行う。これは、それぞれ異なる原理であることから、得られるデータには特徴があり、また実施に要するコスト・手間の点でも違いがあると考えられるためである。特に、類似の配列が繰り返されている HBII-52 の領域が十分な精度を持って解読されるかどうかは本研究計画における大きな関心の1つであるため、優先的な検討事項とする。

なお、ターゲットリシーケンスの技術によっても HBII-52 領域の解読が困難である可能性は否定できないが、過去の文献 (Harismendy et al., 2009) を参照すると純粋な repeating element である SINE などの領域でも解読できることが示されているので、少なくとも全くデータが得られないという事態は考えにくい。最終的に解像度や精度の点で問題が残る場合は、従来型シーケンサーや定量 PCR などの方法によるデータの補完も考慮する。

(2) 自閉症サンプルでのリシーケンス実施

既に収集を終え、利用可能となっている自閉症罹患者の DNA のうち、これまでに報告者がアレイ CGH で HBII-52 領域にシグナル低下を認めているサンプルを中心にリシーケンスを実施する。

(3) 得られたリシーケンスデータの解析

既存のプログラムによりアセンブリを行い、シーケンスデータを判読可能な状態とする。データをこれまでに得ている所見とも比較しつつ、構造異常・変異などの確定を行う。必要に応じ、従来型シーケンサーや定

量 PCR などによる結果の確認を行う。

(4) 臨床データとの関連・遺伝様式の解析

構造異常・変異などが確定されれば、既に得ている臨床データとの関連について解析を行い、臨床症状や病態に与えている影響について考察を行う。また、両親・兄弟などの DNA サンプルが利用可能であるため、これらのサンプルについても解析を行い、遺伝様式を確定させる。

4. 研究成果

平成 22 年度は主として 1) 次世代シーケンサーを利用したシーケンスデータの蓄積と、2) これまでに報告者が得ている、15 番染色体長腕 11-13 領域 (15q11-q13) に微細構造異常を示すサンプルのアレイ CGH データの確認を中心に実施した。1) については、実施費用が高額となることも考え、仮に 15q11-q13 領域の解析が困難であった場合でもデータが無駄にならないよう、4 世代にわたり自閉症性障害を発症している家系を対象としてエクソーム解析を実施し、得られたデータを当研究課題にも用いることとした。具体的には末梢血由来の DNA を用いてショットガンシーケンスライブラリーを作成した上で、Agilent 社製 Sure Select ヒト全エクソームキットを用いてキャプチャーを行い、さらにそれを Illumina 社製 Genome AnalyzerII により pair-end 75-bp の条件下にて解析した。塩基の決定には Illumina pipeline 1.5 を使い、得られたシーケンスは CLC Genomic Wrokbench ソフトウェアを用いてマッピングを行った。一方 2) については、定量 PCR を用いた確認が困難であることが判明したため、Affymetrix 社製 Affy SNP6.0 を用いた解析による確認を試みた。具体的には、当該サンプルをも含む自閉症性障害患者とその両親 163 組を用い、得られたデータについて PennCNV による HMM モデルを用いて CNV の判定を行った。Affymetrix 社提供による APT ツールにより標準化を行うと共に、サンプルや SNP を対象にデータのクリーニングを行って残った 136 組のデータのうち、500K 塩基対以上の大きさの de novo CNV が 7 組のサンプルで 8 つ認められており、その中の複数が 15q11-q13 領域に位置していることが判明した。

平成 23 年度は、1) 前年度次世代シーケンサーを用いて行った、4 世代にわたり自閉症性障害を発症している家系を対象としたエクソーム解析のデータ解析と、2) やはり前

年度自閉症性障害患者とその両親 163 組について得た Affymetrix 社製 Affy SNP6.0 のデータに関するコピー数多型 (CNV) 解析を引き続き実施した。1) について解析を進めた結果、発端者で 11,082 個の変異を見出し、うち 5,560 個が非同義置換あるいは挿入/欠失を伴うものであった。この中から、疾患と完全に共分離している候補遺伝子として、ストップコドン変異 3 個、フレームシフト変異 3 個を選び出した。現在のところ、当初の解析対象であった 15q11-q13 領域と直接結びつく知見は得られていないが、サンガーシーケンスによる確認や機能解析により新たな自閉症候補遺伝子の発見につながる可能性のある結果を得ている。2) では、健常対照者 487 名分のデータを新たに追加し、Birdsuite、QuantiSNP、PennCNV の 3 つのソフトウェアを用いることにより、CNV としてより信頼性の高い判定を得ることを目指した。その結果、サンプルや SNP を対象にデータのクリーニングを行って残った 136 組のデータのうち、500K 塩基対以上の大きさの de novo CNV が 7 組のサンプルで 8 つ認められており、そのうちの 1 つは 15q11.2 領域における重複であることが判明した。やはり当初の目的であった 15q11-q13 に微細構造異常を示すサンプルのアレイ CGH データの確認には至らなかったが、これ以外にも NLGN4X、CNTN4 遺伝子などに影響を与える CNV の存在や、健常対照群との比較において疾患群で有意に CNV が多く存在していることなどを見出した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計 3 件)

- ① 桑原斉、島田隆史、文東美紀、垣内千尋、岩本和也、栃木衛、佐々木司、金生由紀子、笠井清登. エクソンシーケンスを用いた自閉症スペクトラム障害家系の遺伝解析. 第 33 回日本生物学的精神医学会、2011 年 5 月 21 日-22 日、東京
- ② Kuwabara H, Shimada T, Bundo M, Kakiuchi C, Iwamoto K, Tochigi M, Sasaki T, Kano Y, Kasai K. Analysis of genetic inheritance in autism spectrum disorders by exome sequencing. International meeting for autism research 10th annual meeting, May 12th-14th, 2011, San Diego

- ③ Kuwabara H, Shimada T, Bundo M, Kakiuchi C, Iwamoto K, Tochigi M, Sasaki T, Kano Y, Kasai K. Analysis of genetic inheritance in autism spectrum disorders by whole exome sequencing. American Society of Human Genetics 60th annual meeting, Nov 2nd-6th, 2010, Washington DC

[その他]

ホームページ等

http://npsy.umin.jp/her_group.html

6. 研究組織

(1) 研究代表者

栃木 衛 (TOCHIGI MAMORU)

東京大学・医学部附属病院・講師

研究者番号：40456116