

## 科学研究費助成事業 研究成果報告書

令和 4 年 6 月 13 日現在

機関番号：32206

研究種目：基盤研究(B) (一般)

研究期間：2019～2021

課題番号：19H03425

研究課題名(和文) 低頻度バリエーションに重点をおくアルツハイマー病の発症に関与する遺伝子の探索法開発

研究課題名(英文) Exploring rare variants associated with Alzheimer disease

研究代表者

辻 省次 (Tsuji, Shoji)

国際医療福祉大学・医学部・教授

研究者番号：70150612

交付決定額(研究期間全体)：(直接経費) 13,500,000円

研究成果の概要(和文)：サンプルサイズが小さくても、permutationにより関連解析の検出力を高めることをめざした。アルツハイマー病患者446例、認知機能正常な高齢健常者446例のexomeデータを用いて関連遺伝子の探索を行った。見出された候補遺伝子にはGC率が高い領域が含まれていたために、exon capture, library作成の最適化を行った。新規に、アルツハイマー病患者262例、認知機能正常高齢健常者260例を用いてexome解析を実施し、精度の高いデータを得た。当初見出した候補遺伝子については有意な関連は確認されなかったが、関連が示唆される複数の遺伝子が示され、permutationの意義を確認した。

研究成果の学術的意義や社会的意義

多因子疾患の発症に関わる遺伝子の探索は、ゲノムワイド関連解析 (genome-wide association study, GWAS) が行われる。関連解析では、個々の変異毎に、疾患群と対照群の間で検定を行うが、変異の数に対応した多重検定の補正が必要になり、有意水準としてのp値が極めて小さくなり、検出力を確保するために、巨大なサンプルサイズが必要となる。この課題を克服するために、permutationにより、有意水準としてのp値を定めることにより、検出力を高めることを検討し、この手法の有用性が高いことを見出し、多因子疾患の関連解析の新たな手法を提案した点に本研究の意義がある。

研究成果の概要(英文)：To improve the power of association studies for relatively small sample size, we employed permutation testing to determine the p-value to be used for association studies. Employing exome sequence data obtained from 446 cases of sporadic Alzheimer disease (AD) and 446 aged controls with preserved cognitive function, we identified a candidate gene associated with AD. Since the candidate gene contains GC-rich regions, we optimized the methods for exon capture and library construction. We newly conducted exome sequence analysis of 262 cases of AD and 260 aged controls with preserved cognitive function. We confirmed that sequence data with high quality were obtained with the improved procedures. Although significant association of the candidate gene was not replicated, we found several candidate genes possibly associated with AD. The current study confirmed the utility of permutation testing for exome association studies.

研究分野：分子遺伝学

キーワード：association study Alzheimer disease rare variants

## 1. 研究開始当初の背景

分子遺伝学研究の発展により、遺伝性神経疾患の病因遺伝子が多数見出されてきている。しかし、遺伝性に発症するのは全体の 5-10%程度で、大部分を占める孤発性神経疾患については、発症機構はまだ十分には明らかにされていない。孤発性疾患では、頻度の高い遺伝子多型が疾患発症に対して貢献しているという common disease-common variants hypothesis に基づきゲノムワイド関連解析 (GWAS)が行われてきたが、疾患発症に対する影響度の大きい遺伝子はほとんど見出されておらず、そのジレンマは、missing heritability と呼ばれ、分子遺伝学研究の大きな課題となっている。

研究代表者は、missing heritability を解き明かすことを目指し、孤発性の神経疾患の発症に対して影響度の大きい遺伝子の探索を行ってきており、*APOE*  $\epsilon$ 4 allele が孤発性早発型アルツハイマー病の発症リスクになっていること(1)、Gaucher 病の病因遺伝子である *GBA* の低頻度 variants がパーキンソン病発症のリスクを高めること(2, 3)、多系統萎縮症の発症のリスクも高めること(4)を報告してきた。多系統萎縮症については、多発家系の解析から *COQ2* 遺伝子が疾患感受性遺伝子であることを見出し(5)、coenzyme Q10 の補充による医師主導治験へと発展している。これまでに見出した variants はいずれも低頻度であり、common disease-multiple rare variants 仮説を実証するものである。これらの成果は genetics-driven approach, hypothesis-driven approach に基づいて発見したものであるが、genetics-driven approach, hypothesis-driven approach は有力であるものの、網羅的に疾患感受性遺伝子を同定するには、data-driven approach に基づく戦略が必要で、その実現には検定対象となる多数の variants に対する多重検定の問題を克服することが重要である。しかし、解析検体数 10,000 以上の規模の研究の実現は困難である。従って、解析検体数の規模が比較的小さくても、十分な検出力を確保する研究手法の開発と検証が必要であると考えられる。

アルツハイマー病は、世界に先駆けて高齢化社会を迎えるわが国が克服すべき最大の課題である。アルツハイマー病の発症機構については、非常にまれにみられる家族性アルツハイマー病では、*APP*, *PSEN1*, *PSEN2* などの原因遺伝子が解明されているものの、大部分を占める孤発性アルツハイマー病については、これまで、ゲノムワイド関連解析を中心に研究がなされてきているが、疾患発症に対して大きな影響力のある疾患感受性遺伝子は、*APOE*  $\epsilon$ 4 allele を除いては見出されていない。一方、病理学的解析では、アミロイド $\beta$  (以下、 $A\beta$ ) を主成分とする老人斑の蓄積、それに引き続いて生じるリン酸化タウを主成分とする神経原線維変化の蓄積が知られているものの、このような凝集体が形成される最上流の原因は、いまだ明らかでない。これまでの臨床遺伝学的研究からは、孤発性アルツハイマー病において強い遺伝的要因が関与していることはよく知られている。しかしながら、missing heritability と呼ばれているように、存在するはずの強力な遺伝的要因が見出されていないジレンマを解決したいというのが本研究の核心をなす学術的「問い」である。

## 2. 研究の目的

本研究の目的は、孤発性アルツハイマー病の発症に対して、影響度の大きい遺伝的要因を解明すること、そのために、新たなゲノムインフォマティクス解析法を開発して検証することである。孤発性疾患の発症に関わる遺伝子の探索は、これまでゲノムワイド関連解析 (genome-

wide association study, GWAS) が行われてきたが, GWAS で見出される疾患感受性遺伝子のオッズ比は非常に小さく, 疾患の原因とその病態機序を解明するには至っていない. GWAS は一般集団における minor allele frequency (MAF) が 5%以上の頻度の高い一塩基多型 (single nucleotide polymorphisms, SNPs) を用いて, 患者集団と健常者集団における, ゲノム全域の SNPs の頻度の比較を行い (common variants association study), 統計学的に有意差がある SNPs を見出そうというものである. この考え方は, common disease-common variants hypothesis に基づくものであるが, 低頻度の variants は, 論理的にも common SNPs では検出が困難である. 疾患発症に対する影響度の大きい遺伝的要因を解明するためには, GWAS では検出が困難な低頻度の variants を網羅的に解析する必要があると考えられる (rare variants association study). しかしながら, 低頻度 variants になるに従い, 統計学的検出力が低下するというジレンマがある. すなわち, 第 2 種過誤をいかに避けるか (false negative の結果を避けて, 十分な検出力を確保する) という点が鍵となる. 第 2 種過誤を避けるためには, 1. 検出力を確保するために十分なサンプルサイズを確保する, あるいは, 2. サンプルサイズが小さくても, 仮説空間を減じることにより検出力を確保する (検証すべき variants の数を可能な限り減じる) という 2 つのアプローチが考えられる. 理想は, サンプルサイズを十分な検出力が得られる巨大なサイズにすることである. しかしながら, この規模を実現するには, 費用の点から, 個別研究としての実現性は極めて困難と考えられる. したがって, 本研究では, 仮説空間を適切に減じることにより, 比較的少ないサンプルサイズであっても第 2 種過誤を避けて, 疾患発症に対する影響度の高い遺伝子を効率よく見出す解析法を開発するという点に焦点を置いて研究を進める.

### 3. 研究の方法

本研究で提案する解析方法は, 4 つのステップで構成される. 本研究では, 解析データとして, 既にこれまでの研究で得ている, 孤発性アルツハイマー病患者 446 例, 認知機能正常な高齢健常者 446 例の全エクソンシーケンスのデータを用いる. 第一ステップとして, 近年になり充実してきている, 一般集団における variants の頻度情報 (国内 (ToMMo), 海外 (gnomAD), およびインハウスの control データベース) を用いて, 一般集団において見出される variants について, 一定の頻度以上の variants を検証対象から除去するようにする. 第二ステップにおいて, 上記の条件で variants を除去した上で, 残る variants について遺伝子ごとにデータを統合する (遺伝子レベルで, variants を collapse させる). 第三のステップとして, 遺伝子ごとに, 孤発性アルツハイマー病患者群, 認知機能正常な高齢健常者群において, 低頻度 variants を有する症例数について Fisher の正確検定を行う. 第四のステップにおいて, 多重検定の補正に関して, 患者群と健常者群のサンプルをランダムに並べ替える (permutation) ことによって, false discovery rate が 5%未満となる最大の p 値をカットオフ値として定めるようにする. これらの方法で, 孤発性アルツハイマー病に有意に関連すると判定された遺伝子については, 検証に用いた患者, 健常者群とは別に, 準備が完了している孤発性アルツハイマー病患者集団, 認知機能正常な高齢健常者集団を用いた validation study を行う.

### 4. 研究成果

アルツハイマー病を対象とし, exome 解析が完了している孤発性アルツハイマー病患者 446

例，認知機能正常な高齢健常者 446 例のゲノムデータを用いて，minor allele frequency やその他の条件を検討することにより仮説空間を適切に減じ，孤発性アルツハイマー病に有意に関連すると判定される遺伝子を見出した．当初見出した遺伝子は GC 率の高い領域を多く含むため，すでに実施した各 446 例のデータにおいて coverage が不十分な領域が少なからず認められることが判明した．さらに，相同性の高い pseudogene が存在する事が判明し，pseudogene 由来の塩基配列を区別した解析ができるかどうかについても検討が必要となった．GC 率の高い領域で coverage が不十分となる現象については，複数の異なるサンプル調整試薬を用いて作成したライブラリについて NovaSeq を用いて解析を実施し，サンプル調整試薬を変更することにより，GC 率の高い領域における coverage の低下をごくわずかにまで抑えることができることを確認した．また，pseudogene の問題については，read length をこれまでの 75bp x 2 paired-end から，100bp x 2 paired-end と read 長を長く取ることにより，pseudogene 由来の配列を区別することができることを確認した．

上記の最適化された条件で，実際に validation data set での exome 解析を開始し，実データでの検討により，標的遺伝子のコーディングエクソンに存在する全塩基について，genotype quality および read depth が十分な値で保たれており，さらに pseudogene 由来の配列情報を区別した想定通りの良好なデータが得られていることが確認できた．exon capture, library 作成方法について検討を行い，課題を克服できる exome 解析の諸条件を確立した．これらの課題を克服できる最適化された条件により，実際に validation data set での exome 解析を開始し，想定された通りのクオリティの高いデータが得られていることを確認した．validation については，孤発性アルツハイマー病患者 262 例，認知機能正常な高齢健常者 260 例を用いて検討を進めた．解析の完了したデータについて，permutation testing によるシミュレーションを行い，allele frequency, CADD value を考慮に入れて，有意水準としての p value を定め，関連解析を実施した．当初見出していた候補遺伝子については有意な関連は確認されなかったが，関連が示唆される遺伝子が見出された．

本研究では，permutation testing により，一般的に行われる多重検定の補正と比較して，有意水準としての p 値を緩和することができ，小さいサンプルサイズであっても検出力を確保できることを示すことができた．今後，様々な疾患の発症に関連する遺伝子の探索研究に応用できると期待される．

#### <引用文献>

1. Okuizumi K, Onodera O, Tanaka H, Kobayashi H, Tsuji S, Takahashi H, Oyanagi K, Seki K, Tanaka M, Naruse S, Miyatake T, Mizusawa H and Kanazawa I. ApoE-ε4 and early-onset Alzheimer's. *Nature Genet* 7:10-11, 1994
2. Mitsui J, Mizuta I, Toyoda A, Ashida R, Takahashi Y, Goto J, Fukuda Y, Date H, Iwata A, Yamamoto M, Hattori N, Murata M, Toda T and Tsuji S. Mutations for Gaucher disease confer a high susceptibility to Parkinson disease. *Arch Neurol* 66:571-6, 2009.
3. Sidransky E, Aasly JO, Aharon-Peretz J, Annesi G, Barbosa ER, Bar-Shira A, Berg D, Bras J, Brice A, Chen C-M, Clark ON, Condroyer C, De Marco EV, Dürr A, Eblan MJ, Fahn S, Farrer M, Fung H-C, Gan-Or Z, Gasser T, Gershoni-Baruch R, Giladi N, Griffith A, Gurevich T, Januario C, Kropp R, Lang AE, Lee-Chen G-J, Lesage S, Marder K, Mata IF, Mirelman A, Mitsui J, Mizuta I, Nalls MA, Nicoletti G, Oliveira C, Ottman R, Orr-Urtreger A, Pereira LV, Quattrone A, Rogaeva E, Rolfs A, Rosenbaum H, Rozenberg R, Samii A, Samaddar T, Schulte C, Sharma M, Singleton A, Spitz M, Tan E-K, Tayebi

- N, Toda T, Troiano A, Tsuji S, Wittstock M, Wolfsberg TG, Wu Y-R, Zabetian CP, Zhao Y and Ziegler SG. International multi-center analysis of glucocerebrosidase mutations in Parkinson disease. *New Engl J Med* 361:1651-1661, 2009
4. Mitsui J, Matsukawa T, Sasaki H, Yabe I, Matsushima M, Dürr A, Brice A, Takashima H, Kikuchi A, Aoki M, Ishiura H, Yasuda T, Date H, Ahsan B, Iwata A, Goto J, Ichikawa Y, Nakahara Y, Momose Y, Takahashi Y, Hara K, Kakita A, Yamada M, Takahashi H, Onodera O, Nishizawa M, Watanabe H, Ito M, Sobue G, Ishikawa K, Mizusawa H, Kanai K, Hattori T, Kuwabara S, Arai K, Koyano S, Kuroiwa Y, Hasegawa K, Yuasa T, Yasui K, Nakashima K, Ito H, Izumi Y, Kaji R, Kato T, Kusunoki S, Osaki Y, Horiuchi M, Kondo T, Murayama S, Hattori N, Yamamoto M, Murata M, Satake W, Toda T, Filla A, Klockgether T, Wüllner U, Nicholson G, Gilman S, Tanner CM, Kukull WA, Stern MB, Lee VM-Y, Trojanowski JQ, Masliah E, Low PA, Sandroni P, Ozelius LJ, Foroud T and Tsuji S. Variants associated with Gaucher disease in multiple system atrophy. *Ann Clin Transl Neurol.* 2: 417–426, 2015. DOI: 10.1002/acn3.185
  5. Mitsui J, Matsukawa T, Ishiura H, Fukuda Y, Ichikawa Y, Date H, Ahsan B, Nakahara Y, Momose Y, Takahashi Y, Iwata A, Goto J, Yamamoto Y, Komata M, Shirahige K, Hara K, Kakita A, Yamada M, Takahashi H, Onodera O, Nishizawa M, Takashima H, Kuwano R, Watanabe H, Ito M, Sobue G, Soma H, Yabe I, Sasaki H, Aoki M, Ishikawa K, Mizusawa H, Kanai K, Hattori T, Kuwabara S, Arai K, Koyano S, Kuroiwa Y, Hasegawa K, Yuasa T, Yasui K, Nakashima K, Ito H, Izumi Y, Kaji R, Kato T, Kusunoki S, Osaki Y, Horiuchi M, Kondo T, Murayama S, Hattori N, Yamamoto M, Murata M, Satake W, Toda T, Dürr A, Brice A, Filla A, Klockgether T, Wüllner U, Nicholson G, Gilman S, Shults CW, Tanner CM, Kukull WA, Lee V M-Y, Masliah E, Low PA, Sandroni P, Trojanowski JQ, Ozelius L, Foroud T and Tsuji S. Mutations of *COQ2* in familial and sporadic multiple system atrophy. *New Engl J Med* 369:233-44, 2013

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	池内 健  (Ikeuchi Takeshi)  (20372469)	新潟大学・脳研究所・教授    (13101)	
研究分担者	田中 真生  (Tanaka Masaki)  (30774252)	国際医療福祉大学・医学部・講師    (32206)	
研究分担者	石浦 浩之  (Ishiura Hiroyuki)  (40632849)	東京大学・医学部附属病院・講師    (12601)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関