

厚生労働科学研究費補助金（認知症政策研究事業）

総括研究報告書

日本人認知症ゲノム解析を出発点としたオミックス-臨床情報統合解析による
疾患関連パスウェイの解析から診断、治療への応用に関する研究

研究代表者 尾崎 浩一

（国立長寿医療研究センター・メディカルゲノムセンター・臨床ゲノム解析推進部 部長）

研究要旨 認知症、特にアルツハイマー病は環境と遺伝因子が複雑に関与して発症することが知られているが（遺伝率60%~80%）、近年の疫学研究により遺伝因子の重要性が強く浮かび上がってきている。このような背景のもと、これまでに欧米において認知症のゲノムワイド関連解析（GWAS）が大規模に施行され疾患関連座位群が同定されているが、日本人において再現されたのはわずか数座位にとどまっており、ほぼ未解明であると言っても過言では無い。このギャップは欧米人と日本人のゲノム構造の違いに依存すると考えられ、日本人ゲノム構造に特化した解析が疾患の根本的な原因を探るには重要になるとともに、臨床情報等も加味してその機能的な側面を解明することがエビデンスに基づく診断、治療の開発に必要なことになる。本研究では認知症のサブタイプおよび様々な臨床情報に着目し、日本人、アジア人に特化した全ゲノムジェノタイピングプラットフォームによる大規模ゲノムワイド関連解析（GWAS）やそのメタ解析、全エクソーム、全ゲノム解析による網羅的な疾患感受性遺伝子の同定を出発点として、次世代シーケンサーを駆使した全RNA配列解析からの遺伝子発現、バリエーション（スプライシング等を含む）情報等といったオミックス解析情報を遺伝統計学的に統合することによる疾患の真の遺伝的バリエーション、ゲノム領域、関連機能の同定、解析を進める。さらに、これら真の疾患関連分子情報を用いた機械学習、人工知能等による *in silico*での疾患感受性分子生体内パスウェイの解明を進め、既存の薬剤のターゲットとなる分子パスウェイと相互比較することにより、ドラッグリポジショニング等の迅速な治療薬、予防薬の発見に繋がる解析を目指す。また、これらのオミックス情報と年齢、性別や血圧等の一般臨床データはもとより将来的にはMMSE等の認知機能の指標や磁気共鳴画像（MRI）やPET画像といった情報を含む臨床情報を統合して遺伝統計学的なアルゴリズムと機械学習、人工知能を駆使することにより正確な疾患発症予知、予測および正確な診断を目指したバイオマーカーの探索、開発が可能になる。

研究分担者

新飯田俊平 国立長寿医療研究センター
メディカルゲノムセンター
重水 大智 国立長寿医療研究センター
メディカルゲノムセンター
臨床ゲノム解析推進部

A. 研究目的

アルツハイマー型認知症（AD）については有効な治療法がなく、その原因についても不明なところが多い。ADは環境と遺伝因子が複雑に関与して発症することが知られているが、近年の疫学研究により遺伝因子の重要性（遺伝率60～80%）が強く浮かび上がってきている。このような背景のもと、これまでに欧米を中心としてADのゲノムワイド関連解析（GWAS）が大規模に施行され疾患関連座位群が同定されているが、日本人において同様のオッズ比を持って再現されたのはわずか数座位にとどまっており、日本人における大規模解析が必要となる。欧米人と日本人のゲノム構造は主成分分析などにより大きく異なることが知られており、日本人ゲノム構造に特化したゲノム解析が日本人における疾患の根本的な原因究明に重要になるとともに、臨床情報等も加味してその機能的な側面を解明することがエビデンスに基づく診断、治療の開発に必要な。本研究では国立長寿医療研究センターバイオバンク（NCGGバイオバンク）等によりリクルートされた日本人被験者のDNAサンプルを用いて、アジア人に特化した全ゲノムジェノタイプングプラットフォームによる大規模ゲノムワイド関連解析（GWAS）やそのメタ解析、全エクソーム、全ゲノム解析による網羅的な疾患感受性遺伝子の同定を出発点として、次世代シーケンサーを駆使した全RNA配列解析からの遺伝子発現、バリアン

ト（スプライシング等を含む）情報等といったオミックス解析情報を遺伝統計学的に統合することによる疾患の真の遺伝的バリエーション、ゲノム領域、関連機能の同定、解析を進める。

B. 研究方法

国立長寿医療研究センター・メディカルゲノムセンター・バイオバンク（NCGGバイオバンク）、新潟大学脳研究所、理化学研究所およびバイオバンクジャパンによりリクルートされた認知症及びコントロールサンプルを用いて解析を行った。約22,700例のDNAについてはillumina社のアジアスクリーニングアレイ（ASA）およびジャポニカアレイによる全ゲノムジェノタイプングを施行し、ASAデータについては理化学研究所にて構築された日本人約7,000人の全ゲノム配列等によるインピュテーション解析を施行した。今回GWASを施行したアルツハイマー病（LOAD）3,962例、コントロール4,074例の内、NCGGサンプルについてはジャポニカアレイにより全ゲノムジェノタイプングを行い、東北メディカルメガバンクにおいて3,500人の全ゲノムリファレンス配列により構築されたパネルを用いてインピュテーション解析を行った。新潟データ（Affymetix GeneChip 6.0アレイ）については1000ゲノムデータを用いてインピュテーションを行った。ゲノムワイド関連解析（GWAS）は plinkソフトウェアにより施行した。本GWASにより同定した新規有意座位（ $P < 5 \times 10^{-6}$ ）についてはマルチプレックスPCR-インベーター法を用いてLOAD 1,216例、コントロール2,446例を用いた再検証解析を施行した。トランスエスニックGWASは日本人GWAS統計値と欧米人（IGAPステージ1データ；AD 21,982サンプル、コントロール41,944サンプル）統計値についてメタ解析を行った。ポリジェニックリスクスコア（PRS）

の構築に用いるバリエーションの選定はLD plunningおよびP値による閾値法を使用した。全RNA解析についてはNCGGバイオバンクのバフィーコートより高純度のRNAを抽出し、全RNA配列解析用ライブラリ作製キットを用いて、高精度のRNAライブラリを構築した。全RNA配列解析については外注にてデータを得た。遺伝子発現による好中球を含む12免疫細胞のサンプルでの割合の算出はCIBERSOTプログラムを使用した。SHARPIN遺伝子におけるターゲット再配列解析をLOAD 240例、コントロール 240についてサンガーシーケンス法により行った。同定したSHARPINバリエーションの機能解析として、正常、バリエーションタンパクでの細胞内局在の違いはHEK293細胞正常及びバリエーションタンパクを強制発現し免疫染色により蛍光顕微鏡により行った。

(倫理面への配慮)

本研究は、「人を対象とする医学系研究に関する倫理指針」および「ヒトゲノム・遺伝子解析研究に関する倫理指針」に則り、国立研究開発法人国立長寿医療研究センター倫理・利益相反委員会の承認を得て施行されている。すべての検体において書面による同意を取得していると共に、研究対象者個人の尊厳と人権の尊重、個人情報保護等について倫理的観点から十分に配慮しながら研究を遂行している。研究参加者のプライバシーを尊重し、結果については秘密を厳守し、研究の結果得られるいかなる情報も研究目的以外に使用されることは行わない。また、ゲノム情報を保存するサーバ等の記憶媒体に個人を特定できるような情報を一緒に格納していない。国立研究開発法人国立長寿医療研究センターの定める「保有する個人情報の保護に関する規程」に基づき、個人情報保護管理者が厳格に守秘する。

C. 研究結果

認知症を含むDNA検体約22,700例についてASAによるジェノタイプデータおよび1,000例のバフィーコートからの全RNA配列の蓄積を進めている(表1)。これまでに得たデータ(NCGGジャポニカデータ、新潟データ)を用いてインピュテーション後、クオリティコントロール解析を施行し、日本人計8,036例を用いたAD GWAS(図1、2)およびその再検証解析を進めた(計11,698例)。このGWASで新規に示唆的な有意性($P < 5 \times 10^{-6}$)を示す座位を7個発見し再検証解析を行った。第4番染色体上のFAM47E遺伝子内に存在するバリエーションはGWAS有意性($P < 10^{-8}$)を示し、オッズ比0.65と比較的強く関連することが示された(表2)。このバリエーションは日本人で頻度が高く、民族特異的疾患感受性バリエーションであると考えられた。このバリエーションはアミノ酸変化を伴うものではなかったため、遺伝子発現に与える影響(遺伝子発現質的量的座位; eQTL)について公的なデータベースを確認したところ、このバリエーションが副腎皮質や脳でFAM47EやSCARB2遺伝子の発現を上昇することが判明した(図3)。また、示唆的な有意性を示す新たな2バリエーションをBANK1遺伝子およびLINK00899遺伝子上に同定した。トランスエスニックメタGWASを通した解析(図4)では新たに第6番染色体上OR2B2遺伝子上のバリエーションがGWAS有意性を持ってLOADに関連(オッズ比1.34、表3)することを発見するとともに、他7座位(MTSS1L、CLEC3B、EFL1、FAM155A、NTM、C1S、TSPAN14)の示唆的な有意性をもってLOADに関連することを発見した。日本人GWASを用いたPRSの算出も試みた。サンプルの2/3を探索セット、1/3を検証コホートとして使用し、最適のPRSを選別し、最終的には年齢、性別、APOEジェノタイプを加味することにより

ROC曲線下面積 (AUC) を算出したところ、新規座位 *FAM47E* も含めたAUCは0.736であり、それを抜くことにより0.734となった (図5)。この*FAM47E*のPRSへの寄与は統計学的に有意であり、疾患発症に関連することが裏図けられた。

末梢血バフィーコート (主に白血球細胞) から全RNAを抽出し、全配列解析を次世代シーケンサーにより進めてきた。これまでに約1,000例について高品質RNAを抽出、全配列解読を進めている (表1)。610例について約22,000種類の遺伝子発現解析を通して好中球の細胞数がLOADで増加することを発見している (図6)。また、ADとコントロールにおける網羅的な遺伝子発現差異の解析 (図5) からADに関連した機能的モジュールとハブ遺伝子群を同定している (図7)。さらにこれらの情報を用いた機械学習を施行し軽度認知障害からADへの移行を予測できるAD症予測モデルをAUC 0.727の正確性をもって構築した (図8)。

一方、SHARPIN 遺伝子中のレアバリエントがLOADと強い遺伝的な関連 (OR=4.6、 $P < 0.0001$) があり、このバリエント (G186R) がタンパク局在を変化させることにより免疫炎症の活性を低下させることがLOADに關係する機能の一部であることを示してきた。このSHARPINについてさらなる疾患関連バリエントを同定しその機能解明から疾患との関連をさらに明らかとすることを目的として、SHARPIN 遺伝子領域の再遺伝子配列解析をLOAD 240例、コントロール 240について行った。その結果、LOADやコントロールのみにしか認められない低頻度バリエントをSHARPIN 遺伝子上に同定した。これらのバリエントについてさらにサンプル数を増加して関連解析を施行したところADと有意に関連するアミノ酸変化を伴うバリエントを同定

している (表4)。このバリエントについてHEK293細胞に強制発現し、細胞内局在を観察したところLOAD感受性G186Rと同様にドット状のタンパク局在を示すことが観察された (図9)。また、G186Rと同様に免疫系の重要な転写因子NFkBの活性に与える影響を見たところ、この変異によってNFkBの活性がG186R同様に低下することが判明した (図9)。これらの結果は複数のバリエントがコードするSHARPINの機能変化がLOADの発症に関与することを裏付けている。

D. 考察

現在までに認知症関連検体を含む約22,700検体についてASAを用いた日本人に特化したジェノタイピングを施行している。これは認知症関連としてASAで解析した本邦最大のデータとなっている。約11,700例を用いた日本人に特化したGWASおよび再検証解析からはGWAS有意性を示す新規LOAD感受性座位を同定しており最も有意なバリエントは脳における遺伝子発現などに影響を与えることも判明している。さらに欧米人LOAD GWASデータとのトランスエスニックメタ解析からは新たに民族間共通の染色体座位がGWAS有意性を持って同定されており、着実に遺伝因子群の発見に至っている。解析の大規模化によりさらに新たな日本人特有や民族間共通のAD感受性座位が同定できADの全容解明に寄与できる。RNA-seqを用いたオミックス解析については、現在までに998例について配列決定を完了した。610サンプルのRNA-seqデータを用いた解析よりLOADにおける好中球の上昇、LOAD関連ハブ遺伝子群の同定からこれらの情報を用いた機械学習を施行しLOAD発症予測モデルを構築した (前向きコホートにおけるAUC 0.727)。これらは更なるコホートによる検証を経て臨床応用が可能になる。また、日本人において初めて我々が同定したLOAD

感受性遺伝子である *SHARPIN* については、さらに *LOAD* に強い影響を与える可能性のレアバリエントが発見されており、複数のバリエントが疾患感受性である可能性が浮上してきた。今後の解析では *SHARPIN* をターゲットにした創薬の可能性も十分考慮できる。

E. 結論

民族に特化した *LOAD* のゲノムワイド関連解析を通して新たに11個の疾患感受性染色体領域 ($P < 1 \times 10^{-5}$ 、表5) を同定した。これらはこれまでの欧米における大規模解析で見つかったものではなく、日本人に特異的あるいは日本人で影響が強い *LOAD* 感受性座位であると考えられた。今後の日本人サンプル数増加による解析や欧米人とのトランスエスニックメタ解析によるさらに大規模な解析を施行することにより疾患の全容解明に繋げることができると考えられる。全RNA配列解析及びその統合解析においても同様に大規模化することによる真の疾患感受性パスウェイの同定、分子機能の解明、それらの情報を用いた機械学習などから正確度の高い疾患予測モデルの構築へと発展させることができる。またここで同定した *SHARPIN* をはじめとして疾患に強く寄与するレアバリエントの全ゲノムやエクソームシーケンス解析によるさらなる同定やこれらの統合解析から認知症の予知診断に有用なポリジェニックリスクスコアの構築、ドラッグリポジショニングなど創薬へと発展させることが可能になる。

F. 健康危険情報

本研究集団では *APOE* e4 アレルを一つ持つ場合、AD に対するオッズ比は2.57程度となる。95%信頼区間は2.21~2.99。

G. 研究発表

1. 論文発表

1. *Translational Psychiatry* 11, 151. doi: 10.1038/s41398-021-01272-3. Ethnic and trans-ethnic genome-wide association studies identify new loci

- influencing Japanese Alzheimer's disease risk. 2021. [Shigemizu D](#), Mitsumori R, Akiyama S, Miyashita A, Morizono T, Higaki S, Asanomi Y, Hara N, Tamiya G, Kinoshita K, Ikeuchi T, [Niida S](#), [Ozaki K](#).
2. *Alzheimer's Research & Therapy* 12: 145. doi: 10.1186/s13195-020-00716-0. Prognosis prediction model for conversion from mild cognitive impairment to Alzheimer's disease created by integrative analysis of multi-omics data. 2020. [Shigemizu D](#), Akiyama S, Higaki S, Sugimoto T, Sakurai T, Boroevich KA, Sharma A, Tsunoda T, Ochiya T, [Niida S](#), [Ozaki K](#).
3. *Journal of Human Genetics* Published Online 26 October. doi: 10.1038/s10038-020-00864-z. Association of an IGHV3-66 gene variant with Kawasaki disease. 2020. Johnson T, Mashimo Y, Wu JY, Yoon D, Hata A, Kubo M, Takahashi A, Tsunoda T, [Ozaki K](#), Tanaka T, Ito K, Suzuki H, Hamada H, Kobayashi T, Hara T, Chen CH, Lee YC, Liu YM, Chang LC, Chang CP, Hong YM, Jang G, Yun S, Yu J, Lee KY, Kim JJ, Park T, Lee JK, Chen YT, Onouchi Y.
4. *Nature Genetics* 52: 1169-1177 doi: 10.1038/s41588-020-0705-3. Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. 2020. Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, Matsunaga H, Ieki H, [Ozaki K](#), Onouchi Y, et al.
5. *Nature Genetics* 52 (7): 669-679. doi: 10.1038/s41588-020-0640-3. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. 2020. Ishigaki K, Akiyama M, Kanai M, Takahashi A, Kawakami E, Sugishita H, Sakaue S, Matoba N, Low SK, Okada Y, Terao C, Amariuta T, Gazal S, Kochi Y, Horikoshi M, Suzuki K, Ito K, Koyama S, [Ozaki K](#), Niida S. et al.
6. *Circulation: Genomics and Precision Medicine* 13 (3), e002670. doi: 10.1161/CIRCGEN.119.002670. Transethnic meta-analysis of genome-wide association studies identifies three new loci for coronary artery disease. 2020.

Matsunaga H, Akiyama M, Takahashi A, Nomura S, Ozaki K, Onouchi Y, Suna S, Ogishima S, Yamamoto M, Satoh M, Sasaki M, Yamaji T, Iwasaki M, Tsugane S, Tanaka K, Naito M, Akai W, Tanaka H, Sakata Y, Morita H, Matsuda K, Murakami Y, Akazawa H, Kubo M, Kamatani Y, Komuro I, Ito K.

Mitsumori R, Niida S, Ozaki K, The Japan Society of Human Genetics, 2020/11/18、国内(Web).

2. 学会発表

1. 櫻井 孝、尾崎浩一 大脳白質病の臨床的とリスク 第 39 回日本認知症学会学術集会 シンポジウム 19 (名古屋、2020 年 11 月 27 日)
2. アミロイドカスケード仮説 ; アルツハイマー病の 原因遺伝子, 感受性遺伝子からの考察 宮下 哲典, 原 範和, 春日健作, 菊地 正隆, 尾崎 浩一, 新飯田俊平, 池内 健 第 39 回日本認知症学会、シンポジウム、2020/11/26~28、国内 (Web)
3. 日本人コホート JGSCAD, J-ADNI, NCGG, ToMMo に おけるバリエーション解析 : *APOE* 宮下 哲典, 原 範和 , 春日 健作, Liu Lixin, 樋口 陽, Zhu Bin, 月江 珠緒, 長谷川舞衣, Adyfitrah Yusran, 石黒 敬信, 村上 涼太, 菊地 正隆, 中谷 明弘, 尾崎 浩一, 新飯田俊平, 赤澤 宏平, 桑野 良三, 岩坪 威, 池内 健 第 39 回日本認知症学会、シンポジウム、2020/11/26~28、国内 (Web)
4. Kokubo M, Ozaki K, Nakanishi H, Ohta H, Shimizu A, Arai H, Sakurai T. Genome-Wide Association Study identifies two novel chromosome loci associated with Cerebral White Matter Hyperintensities volume in Japanese population. Alzheimer's Association International Conference® 2020, 2020/7/28 (Web).
5. Kikuchi M, Miyashita A, Hara N, Shigemizu D, Ozaki K, Niida S, Ikeuchi T, Nakaya A. Polygenic analysis of late-onset Alzheimer's disease in a Japanese population. 2020 Alzheimer's Association International Conference, 2020/7/27 (Web).
6. Identification of loci influencing four types of cognitive impairment、ポスター、Mitsumori R, Asanomi Y, Shigemizu D, Akiyama S, Morizono T, Niida S, Ozaki K、The Japan Society of Human Genetics、2020/11/18、国内 (Web)。
7. Identification of the novel rare variants in SHARPIN associated with the risk of late-onset Alzheimer's disease、口頭、Asanomi Y, Shigemizu D,

表 1 全ゲノムジェノタイピングおよび全 RNA 配列解析数

(AD ; アルツハイマー病、MCI ; 軽度認知障害、VaD ; 血管性認知症、FTD ; 前頭側頭型認知症、DLB ; レビー小体型認知、NPH ; 正常圧水頭症、PD ; パーキンソン病、CN ; 認知機能正常)、その他 ; 主観的認知障害 (SCI)、大脳皮質基底核変性症、病型不明の認知症など

解析プラットフォーム	AD	DLB	FTLD	MCI	VaD	NPH	PD	CN	その他	総数
SNP array	6,407	119	25	1,171	70	128	0	13,508	1,352	22,780
Transcriptome (mRNA)	356	-	-	400	-	-	-	242	-	998

図 1 GWAS ストラテジー

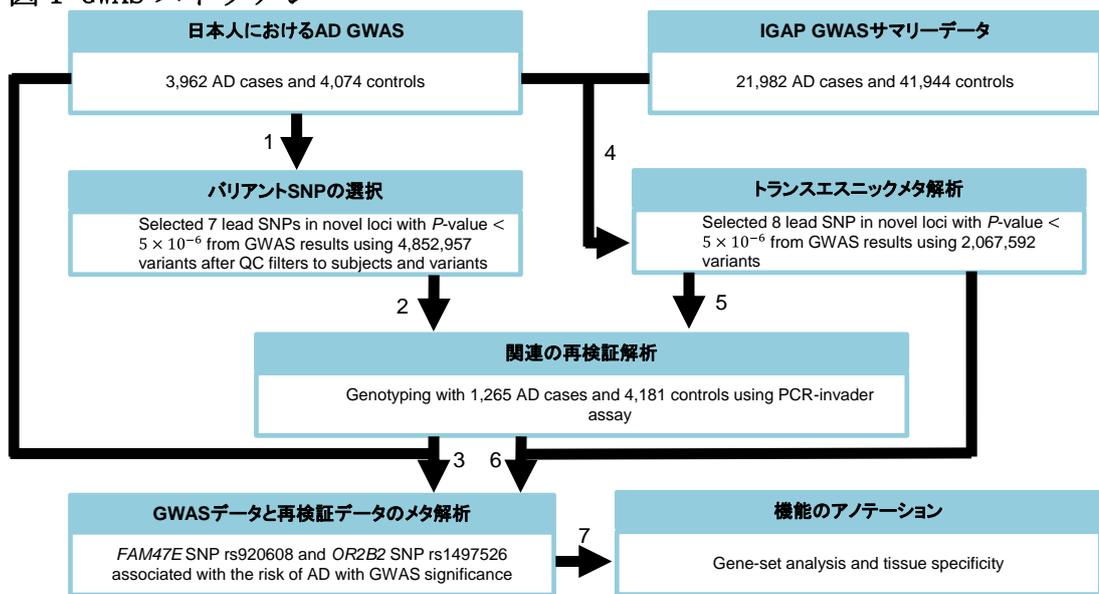


図 2 日本人 GWAS による新規アルツハイマー感受性座位 (FAM47E) の同定

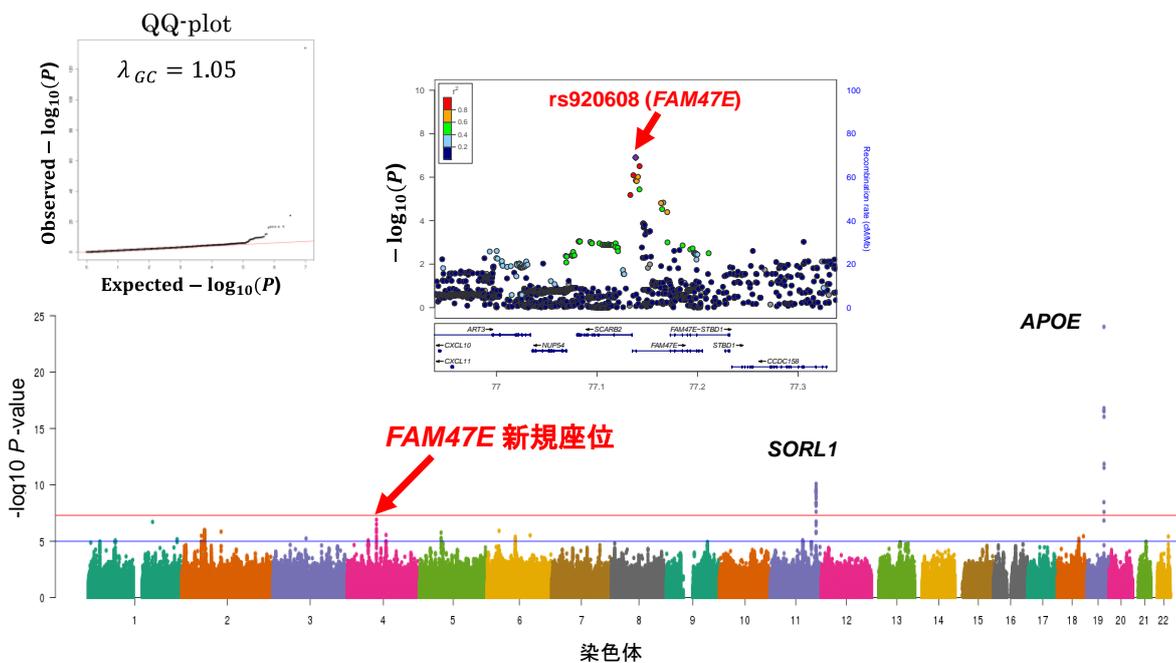


表 2 日本人 LOAD GWAS で上位を示したローカスの再検証とメタ解析

マーカーID	アレル	近傍遺伝子	ステージ (N)	サンプル数		アレル頻度		オッズ比	95% 信頼区間	P
				疾患	対照	疾患	対照			
rs920608	C/A	FAM47E	GWAS (8,030)	3,959	4,071	0.030	0.046	0.63	0.53-0.75	1.25×10^{-7}
			再検証 (3,662)	1,216	2,446	0.037	0.046	0.74	0.57-0.97	2.72×10^{-2}
			メタ解析 (11,692)	5,175	6,517	0.032	0.046	0.65	0.57-0.75	5.34×10^{-9}

図 3 eQTL マッピングによる FAM47E 座位の FAM47E および近傍遺伝子発現に与える影響

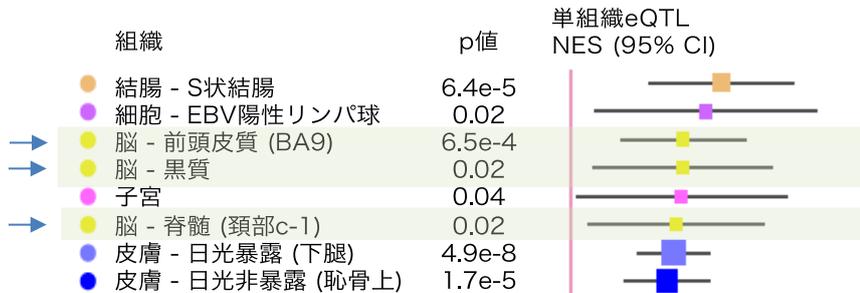


図 4 日本人および既報欧米人 (IGAP) データを用いたトランスエスニック GWAS

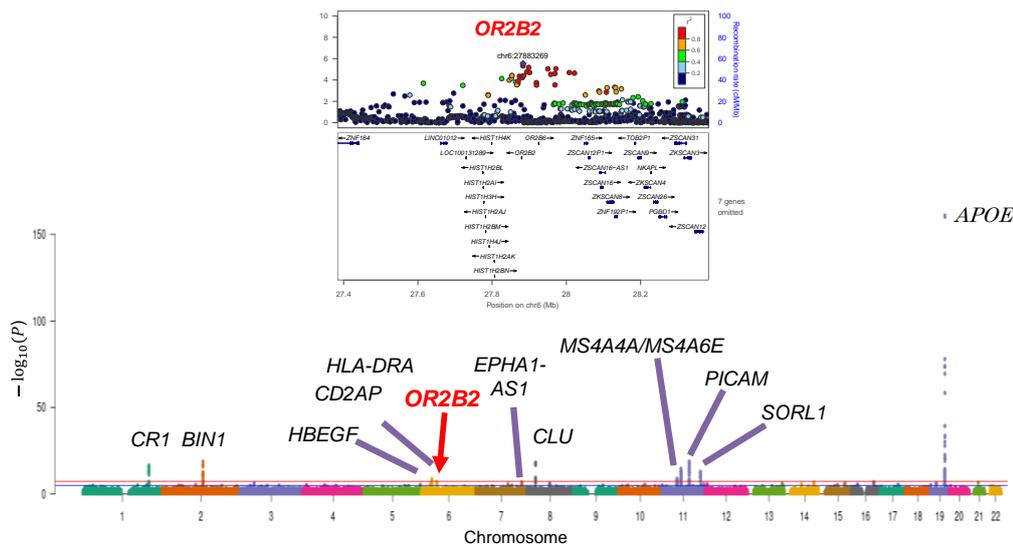


表 3 民族集団間 GWAS で上位を示したローカスの再検証とメタ解析

民族集団間メタGWAS				再検証解析				統合メタ解析			
遺伝子	マーカー ID	P値	代理 SNPs	オッズ比 (IGAP, GWAS)	サンプル数		アレル頻度		オッズ比	P値	メタ解析P値
					疾患	対照	疾患	対照			
OR2B2	rs1497525	2.96×10^{-6}	rs1497526	1.12, 1.18	1,156	2,385	0.15	0.12	1.34	1.98×10^{-4}	2.14×10^{-8}

民族集団間メタ解析; 日本人: 3962 vs 4074、欧米人: 21982 vs 41944
 再検証解析; 日本人: 1156 vs 2421
 (IGAP; International Genomics of Alzheimer's Project)

表 4 SHARPIN 候補バリエントの関連解析結果

Phase	No. of samples		No. of variants Hetero (Homo)		MAF		OR	95%CI	P
	Cases	Controls	Cases	Controls	Cases	Controls			
1 st cohort	2,722	9,348	73 (1)	179	0.014	0.0096	1.45	1.10-1.90	0.0075
2 nd cohort	2,321	2,636	57	42	0.012	0.0080	1.55	1.04-2.31	0.031
Combined*	5,043	11,984	130 (1)	221	0.013	0.0092	1.48	1.18-1.85	0.00075

MAF; minor allele frequency, OR; odds ratio, CI; confidence interval

*; P value was calculated by Mantel-Haenszel test.

図 9 新規 SHARPIN リスクバリエントの解析

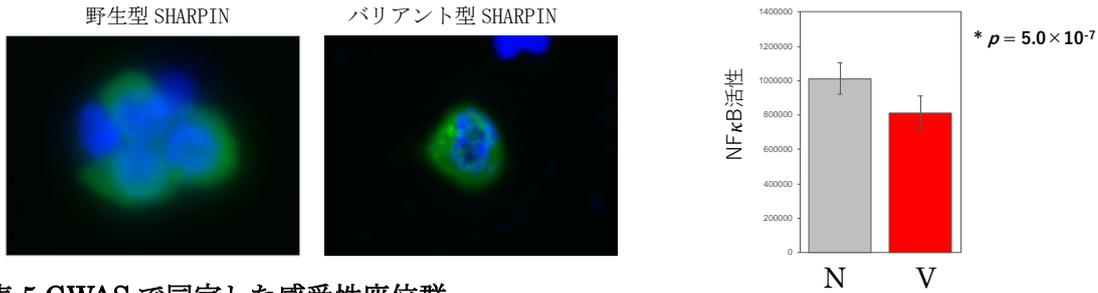


表 5 GWAS で同定した感受性座位群

	GWAS 有意性 ($P < 5 \times 10^{-8}$) を示した新規座位	$P < 1 \times 10^{-5}$ の有意性を示した新規座位
日本人の GWAS 解析	<i>FAM47E</i>	<i>BANK1, LINC00899</i>
民族集団横断型メタ解析	<i>OR2B2</i>	<i>MTSS1L, CLEC3B, EFL1, FAM155A, NTM, CIS, TSPAN14</i>