

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** 認知症は大きく、アルツハイマー型、脳血管性、レビー小体型、前頭側頭型に分けることができる。近年、様々な薬剤の臨床治験が進められているが未だ成果が上がっていないのが現状であり、根本的な疾患の原因を突き止めてエビデンスに基づく創薬やドラッグリポジショニング、早期予知による発症の阻止を目指すことが必要となる。認知症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、開発に繋げる狙いがある。

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

老年病、特に認知症の患者数は全世界で増加の一途をたどっており、本邦においてもその患者数は500万人（2013年、厚生労働省研究班推計）に達する勢いである。大部分の認知症は、糖尿病や虚血性心疾患と同様に生活習慣病と捉えることができ、環境因子と遺伝因子が複雑に絡み合って発症すると考えられる。これまでの双子疫学研究による認知症、特に孤発性アルツハイマー病（Late-onset Alzheimer's disease; LOAD、本文中ではADと略す）の発症に与える遺伝因子の割合は58%~79%であることが証明されており、その大部分を遺伝因子が占めていることが明らかとなっている。したがって、遺伝因子群を同定し、分子マーカーとして使用することや、その役割を精査することから疾患の分子メカニズムが解明でき、エビデンスに基づく予防法や治療法の開発に大きく貢献できる。一方、これまでに欧米において認知症のゲノム解析が大規模に施行され20個程度の関連座位が同定されているが、日本人において再検証されたのはこの中のわずか4座位にとどまっている。このような違いは欧米人と日本人のゲノム構造の根本的な違いに大きく依存すると考えられ、日本人ゲノムに特化し大規模ゲノム解析が必要になる。また、欧米の解析で同定された疾患座位のどの遺伝子がどのように疾患に関わるかといった疾患の原因もほとんど明らかになっていないのが現状である。本研究では日本人、アジア人特有

のゲノムに特化したジェノタイピングアレイによるGWASと次世代シーケンサーを駆使した全RNA解析を統合することによる、疾患に直接的に関連した遺伝的バリエーション、遺伝子機能、遺伝子発現、スプライスバリエーションの同定、解析から真の疾患分子を同定および詳細な解析と共に、統計、機械学習的なアルゴリズムを用いた疾患の予知法や創薬のターゲットとなる分子パスウェイを探索することを目的とする（図1）。これまでに日本人を基盤として、大規模GWASおよびそれと同一サンプルを用いた大規模RNA配列解析等のオミクス解析を統合したスタディーデザインによる研究が行われた例は他になく、世界的に見ても独創的かつ画期的である。

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）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。7,132例のDNAについては東北メガバンク機構にて開発されたジャポニカアレイ（東芝への外注）によるジェノタイピングを施行し、東北メガバンク機構にて構築された日本人3,500人の全ゲノム配列を基に作製されたインピュテーションパネルを用いたインピュテーション解析を施行した。試験的なADを対象としたGWAS（AD 2,357例 vs CO 3,174例）の統計解析はplinkソフトウェアにより施行した。APOEジェノタイプについては、1000ゲノムデータ

（<http://www.internationalgenome.org/>）よりアフリカ人、アメリカ人、ヨーロッパ人、東アジア人（日本人を除く）、日本人rs429358、rs7412 SNP多型を取得し、それぞれのε4アレル頻度について、NCGGデータと比較した。イルミナ社アジアスクリーニングアレイを用いたジェノタイピングは200ngのDNAを用いて全

ゲノム増幅を行い、断片化DNAをアレイにハイブリダイズした。iScanを用いて傾向を測定し、イルミナ社GenomeStudioソフトウェアを用いてクラスタリングを行った。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(TruSeq Stranded Total RNA Sample Preparation Kit; イルミナ社)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(ジーンウィズ株式会社、タカラバイオ株式会社)にてデータを得た。RNA発現量との発現 quantitative traits loci (eQTL)解析はplinkソフトウェアにより行った。全エクソーム解析は202例のAPOE ε4 ADリスクアレルを持たないAD患者由来DNAについてHiSeq2500(イルミナ社)を用いて配列決定をおこなった。エクソーム解析のパイプラインについては図2に示した。エクソーム解析により同定した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細胞にmyc tag配列を付加した正常及びバリエーションタンパクを強制発現し、myc tag抗体を用いた免疫染色により蛍光顕微鏡を用いて行った。また、炎症の中心的なメディエーターである Nuclear factor kappa B (NFκB)の活性に与えるバリエーションの影響については、安定的にルシフェラーゼ遺伝子を発現するHEK293細胞を構築し、この細胞に正常およびバリエーションタンパクを強制発現することにより行った。

(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、個人情報の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7,132例(サンプル内訳は図3に示した)についてジャポニカアレイによりジェノタイピングを行った。このジェノタイピングデータを用いてインピュテーション解析後、図4に示したクオリティーコントロール解析を施行した。試験的な解析として約775万SNPを用いたAD(2,357例)と対照コントロール(3,174例)のGWASを進めたところ、図5に示したようにAPOEバリエーションによるインフレーションが認められたが $\lambda_{gc}$ は1.00であり、集団の階層化は認められない結果であった。この解析でゲノムワイド有意性を示したローカスは染色体19番長腕のAPOE座位のみであり、既報のAPOEジェノタイプがこの集団においても最も強いADとの関連を示すことが再確認できた。また、ゲノムワイド有意性は獲得できなかったが、示唆的な統計値( $p < 10^{-5}$ )を示し、東アジア人にしかアレル頻度を持たないいくつかの疾患候補座位を同定することもできている。一方、近年欧米諸国のGWASにより同定された座位群(APOE座位を除く)でこの日本人集団においてADと明らかな関連を認められた座位は第8番染色体CLU座位、第11番染色体PICALM座位、第19番染色体ABCA7座位および第11番染色体のSORL1座位になる。図6にAPOE ε4

アレル頻度について、1000ゲノムデータ、NCGGデータを用いて比較した。NCGGのAD (NCGG AD) サンプルにおけるε4頻度が日本人コントロールJPT (1000ゲノムデータ、東京在住日本人) およびNCGGコントロール(NCGG CO)に比べて有意に高いことが認められるが、1000ゲノムデータより得られたアフリカ人のε4頻度がさらに高いことが印象的である。

遺伝的因子(ジェノタイプ)と遺伝子発現量(フェノタイプ)の関係について精査し、疾患発症に重要な遺伝子の絞り込み(発現 Quantitative Trait Loci; eQTL、図7)や分子機能を解明することを目的として、末梢血バフィーコート(主に白血球細胞)からのRNA抽出、全RNA配列解析について次世代シーケンサーを用いて進めてきている(1,000例~を目標)。これまでに約600検体について高品質RNAを抽出、全RNA配列解析ライブラリを作成し全配列解読を進めており、240例についてはパイロット的に発現解析、eQTL解析を行った。240サンプルを用いたeQTL解析では21,793遺伝子中1,393遺伝子がcis-eQTL(Permutation test; adjusted p-value < 0.05)を示した。これらの遺伝子群の中でADとコントロールで発現に差があり(10個)、かつGWASでのp値が0.05未満であったバリエーションが存在する遺伝子は2個存在しており、有力なAD関連候補遺伝子であると考えることができる。

GWASでは比較的頻度の高いコモンバリエーションを対象とした解析になるため、頻度の低い(アレル頻度 > 0.005)、いわゆるレアバリエーションの同定には向かない。そこで、本研究ではAPOEリスクバリエーションを持たないADサンプルを用いたエクソーム解析からのレアバリエーションの同定を試みた。その結果、図2に示すパイプラインによりアミノ酸の置換を伴う炎症関連分子内のバリエーションがADに強く関連することを同定した(オッズ比 6.1、 $P = \sim 10^{-5}$ )。

*In vitro* 解析の結果、このバリエーションは炎症に中心的な役割を果たす転写因子であるNFκBの活性を制御することが判明した。

#### D. 考察

ジャポニカアレイによる7,132例のジェノタイプング、インピュテーションにより得られたバリエーション数は775万個であり、他の標準的なジェノタイプングアレイと同等の結果と考えられる。ADとコントロールを用いたGWASでは、我々の使用したサンプルに集団の階層化がほとんど認められず、擬陽性は少ないと考えられたが、ゲノムワイド有意性を超えた染色体座位はAPOE座位のみであり、現段階では新規の疾患感受性バリエーション群の同定には至っていない。今後さらなるサンプルの増加、欧米人を含む既報GWASとのメタ解析により新たな疾患感受性座位が同定できるものと考えている。日本人(NCGG)約7,000例におけるAPOEのリスクアレル頻度については、ε4アレルがADで非常に高く、既報通り疾患との関連が非常に強く示唆される。一般アフリカ人におけるε4アレル頻度の高さには目を引くものがあるが、人種的に他のプロテクティブな遺伝的因子あるいは環境因子が存在するのかもしれない。全RNA解析についてはバフィーコートからのRNA抽出後のRIN値(RNA Integrity Number; RNAの品質の指標)は概ね7を超えており全RNAライブラリ調整には適しており、600例についてライブラリを作成、配列解析を行い、問題なくリファレンスの配列にマッピングできることが確認できている。240例の試験的なeQTLおよび発現差異検定の結果いくつかの遺伝子がADに関連する可能性を得ており今後の検証が期待される。エクソーム解析からはこれまでに同定されていないレアバリエーションを同定し、そのバリエーションが炎症、免疫系の制御に関連している可能性が浮かび上がっている。海外でADと関連するとして同定された *TREM2*

遺伝子内のレアバリエーションについても同様に脳における炎症、免疫系での制御がADに関連すると報告されていることから、今回同定した分子の詳細な解析からADの病態解明の一助になる可能性が大きい。

#### E. 結論

日本人に特化したジェノタイピングアレイによる日本人認知症を含む約7,000例のジェノタイピングを施行し、ADのGWASを試みた結果、APOE座位のみがゲノムワイド有意性を示したが、示唆的な統計値 ( $p < 10^{-5}$ ) を示す東アジア人にしかアレル頻度を持たないいくつかの疾患候補座位も同定しており、今後のサンプル数増加による解析やメタ解析による検証が期待される。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、分子機能の解明へと発展させることができると共に、全ゲノムやエクソームシーケンス解析による疾患に強く寄与するレアバリエーションの同定と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探索、さらにはドラッグリポジショニングを含む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 ε4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. Shigemizu D, Akiyama S, Asanomi Y, Boroevich KA, Sharma A, Tsunoda T, Matsukuma K, Ichikawa M, Sudo H, Takizawa S, Sakurai T, Ochiya T, Ozaki K, Niida S. Risk prediction models for dementia constructed by supervised principal component analysis using miRNA expression data. *Communications biology*. 2019,2, 77, doi: 10.1038/s42003-019-0324-7.
2. Saji N, Niida S, Murotani K, Hisada T, Tsuduki T, Sugimoto T, Kimura A, Toba K, Sakurai T. Analysis of the relationship

between the gut microbiome and dementia: a cross-sectional study conducted in Japan. *Scientific Reports* 2019, 9, 1008, doi: 10.1038/s41598-018-38218-7.

3. Shigemizu D, Miya F, Akiyama S, Okuda S, Boroevich K, Fujimoto A, Nakagawa H, Ozaki K, Niida S, Kanemura Y, Okamoto N, Saitoh S, Kato M, Yamasaki M, Matsunaga T, Mutai H, Kosaki K, Tsunoda T. IMSindel: An accurate intermediate-size indel detection tool incorporating de novo assembly and gapped global-local alignment with split read analysis. *Scientific Reports*. 2018, 8(1), 5608, doi: 10.1038/s41598-018-23978-z.

##### 2. 学会発表

1. 認知症のリキッドバイオプシーを旨とした血清マイクロRNAバイオマーカーの探索, ポスター, 浅海裕也, 重水大智, 茅野光範, 松熊佳奈, 市川真紀子, 須藤裕子, 滝澤聡子, 尾崎浩一, 新飯田俊平, 第10回日本RNAi研究会/第5回日本細胞外小胞学会JSEV, 2018/8/30, 国内.
2. The construction of risk prediction models for dementia with supervised principal component analysis using miRNA expression data, ポスター, Shigemizu D, Akiyama S, Asanomi Y, Boroevich KA, Sharma A, Tsunoda T, Matsukuma K, Ichikawa M, Sudo H, Takizawa S, Sakurai T, Ozaki K, Niida S, 第10回日本RNAi研究会/第5回日本細胞外小胞学会JSEV, 2018/8/30, 国内.
3. 遅発性アルツハイマー病新規リスクレアバリエーション候補の関連解析, 口頭, 浅海裕也, 重水大智, 光森理紗, 森大気, 新飯田俊平, 尾崎浩一, 第63回日本人類遺伝学会, 2018/10/12, 国内.
4. 日本人における遅発性アルツハイマー病患者由来末梢血単核細胞の発現量の形質遺伝子座解析, ポスター, 森大気, 重水大智, 秋山真太郎, 光森理紗, 浅海裕也, 新飯田俊平, 尾崎浩一, 第63回日本人類遺伝学会, 2018/10/12, 国内.
5. 日本人における4種の認知症病型のゲノムワイド関連解析, ポスター, 光森理紗, 浅海裕也, 重水大智, 森大気, 秋山真太郎, 新飯田俊平, 尾崎浩一, 第63回日本人類遺伝学会, 2018/10/12, 国内.
6. The construction of risk prediction models for dementia with supervised principal component analysis using miRNA

expression data, 口頭, Shigemizu D, Akiyama S, Asanomi Y, Boroevich KA, Sharma A, Tsunoda T, Matsukuma K, Ichikawa M, Sudo H, Takizawa S, Sakurai T, Ozaki K, Niida S, 第 63 回日本人類遺伝学会, 2018/10/12, 国内.

7. 認知症の血中マイクロ RNA マーカー探索, ホスター, 櫻井 孝, 重水大智, 浅海裕也, 佐治直樹, 尾崎浩一, 松熊佳奈, 市川真紀子, 須藤裕子, 近藤哲司, 滝澤聡子, 新飯田俊平, 第 37 回日本認知症学会学術集会, 2018/10/12, 国内.

8. 認知臨床ゲノム情報データベース構築に関する開発研究, ホスター, 池内 健, 新飯田俊平, 佐々木貴史, 尾崎浩一, 新井康通, 中谷明弘, 柿田明美, 鈴木一詩, 齋藤祐子, 村山繁雄, 橋詰良夫, 寺田整司, 吉田真理, 嶋田裕之, 三村 將, 岡野栄之, 岩坪 威, 秋山治彦, 森啓, 第 37 回日本認知症学会学術集会, 2018/10/12, 国内.

9. The construction of risk prediction models for dementia with supervised principal component analysis using miRNA expression data, ホスター, Shigemizu D, Akiyama S, Asanomi Y, Boroevich KA, Sharma A, Tsunoda T, Matsukuma K, Ichikawa M, Sudo H, Takizawa S, Sakurai T, Ozaki K, Niida S, The American Society of Human Genetics 2018, 2018/10/19, 国外.

10. Trans-ethnic meta-analysis of genome-wide association studies for coronary artery disease reveals genetic differences between Japanese and Europeans. ホスター, Matsunaga H, Koyama K, Ozaki K, Ito T. The American Society of Human Genetics 2018 (San Diego), 2018/10/19, 国外.

11. Genetic analysis for late-onset Alzheimer's disease in Japanese population, 口頭, Ozaki K, Asian Forum on Alzheimer's and Dementia 2018, 2018/11/23, 国外.

12. 認知症の血液 miRNA マーカー開発, 口頭, 新飯田俊平, 第 6 回 JMAC シンポジウム, 2019/1/24, 国内.

13. The construction of risk prediction models for dementia with supervised principal component analysis using miRNA expression data, ホスター, Shigemizu D, Akiyama S, Asanomi Y, Boroevich KA, Sharma A, Tsunoda T, Matsukuma K, Ichikawa M, Sudo H, Takizawa S, Sakurai T, Ozaki K, Niida S, 第 6 回 JMAC シンポジ

ウム, 2019/1/24, 国内.

14. Exploration of serum microRNA biomarkers for dementia-risk prediction, ホスター, 浅海裕也, 重水大智, 茅野光範, 松熊佳奈, 市川真紀子, 須藤裕子, 滝澤聡子, 櫻井孝, 尾崎浩一, 新飯田俊平, 第 6 回 JMAC シンポジウム, 2019/1/24, 国内.

図1 大規模ゲノム・オミックス情報の収集と解析

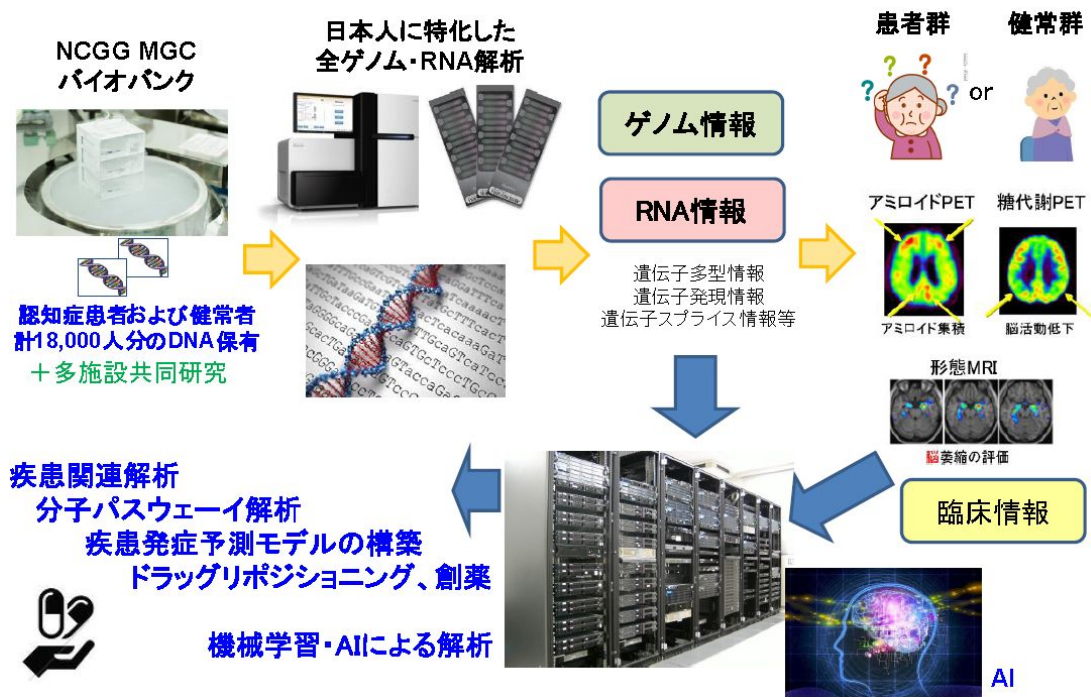


図2 エクソーム解析による LOAD 関連レアバリアントの同定パイプライン

(CADD ; Combined Annotation Dependent Depletion = 変異の有害性の検定法、VCF ; 変異情報ファイル、NC subjects ; 認知機能正常サンプル)

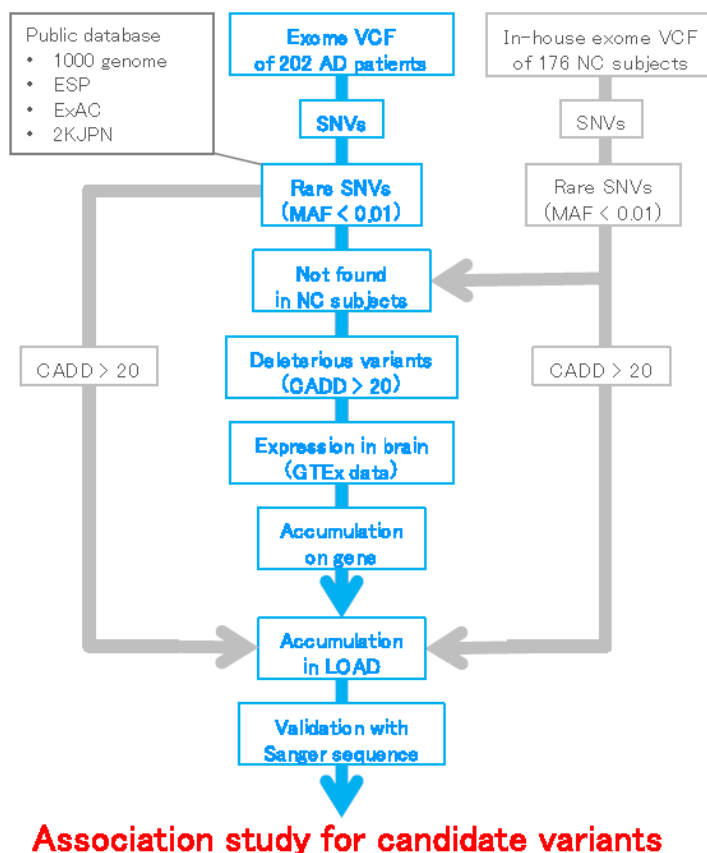


図3 ジェノタイピングサンプルの分類

(MCI ; 軽度認知障害、VaD ; 血管性認知症、FTD ; 前頭側頭型認知症、DLB ; レビー小体型認知症)

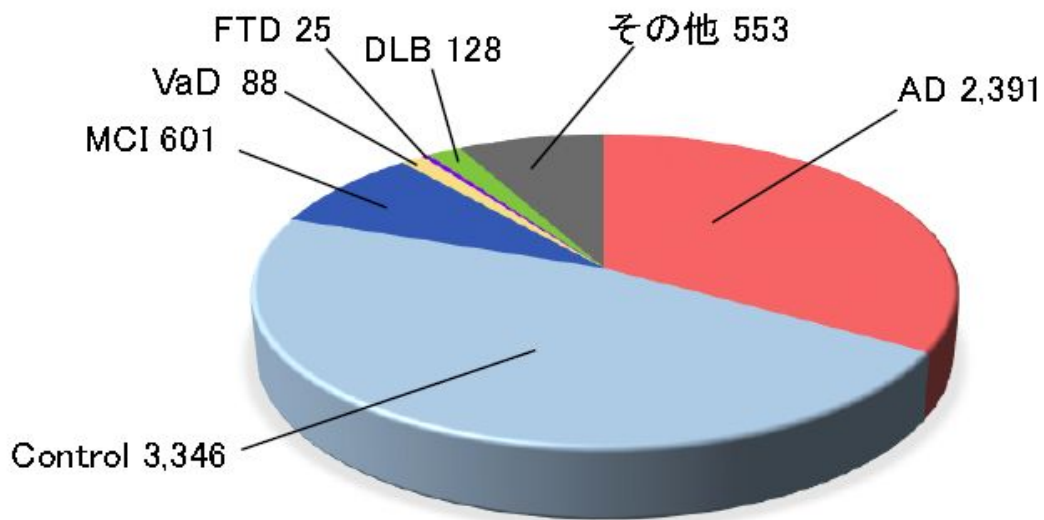


図4 クオリティーコントロール (QC)

- ✓ Imputation
  1. Info score > 0.4
  2. Hard call probability > 0.9
- ✓ Exclusion criteria for sample QC
  1. Sex inconsistency
  2. Inbreeding coefficient  $|F| > 0.1$
  3. Genotype missingness > 0.05
  4. Kinship coefficient  $\theta > 0.2$
  5. PCA using 1KG
- ✓ Exclusion criteria for marker QC
  1. Call rate < 0.98
  2. Minor Allele Frequency (MAF) < 0.001
  3. Hardy-Weinberg Equilibrium (HWE) in controls <  $1 \times 10^{-3}$
  4. Strand orientation (SNPflip)



図5 AD vs control による Quantile-Quantile plot

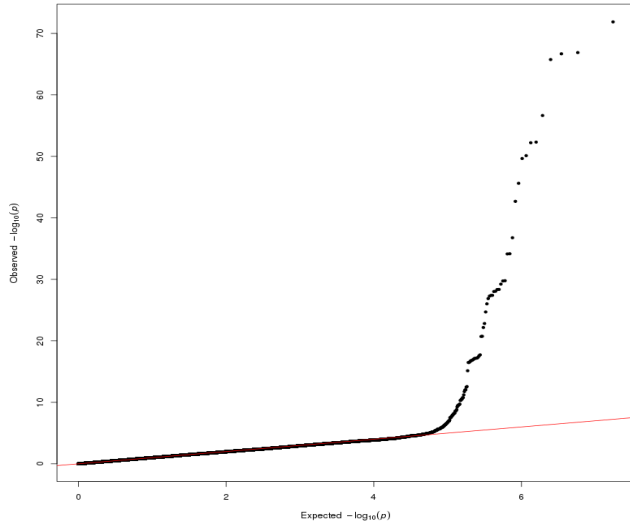


図6 APOE ε2, ε3, ε4 の各人種および日本人 AD、control における頻度  
 (NCGG\_AD、CO 以外は 1000 ゲノムデータベースより抽出)  
 (Hap\_JPT; 東京在住日本人、EAS; 東アジア人、EUR; ヨーロッパ人、AMR; アメリカ人、  
 AFR; アフリカ人)

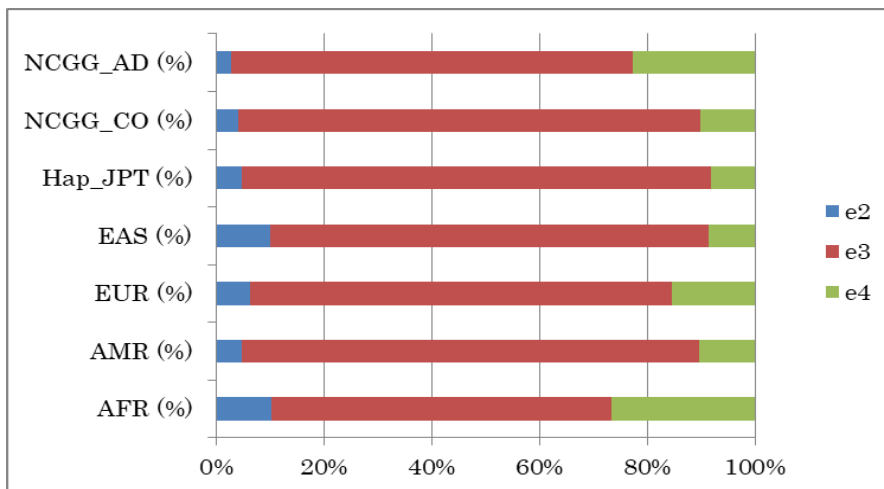


図7 発現量に関連する遺伝的バリエーションの同定 (eQTL 解析)

