

iPS コホートと機械学習を用いたアルツハイマー病再構成

－CDiP テクノロジーによる無病社会に向けた孤発性高齢疾患の解説－

京都大学 iPS 細胞研究所の井上治久教授、近藤孝之特定拠点講師、理研バイオリソース研究センター iPS 創薬基盤開発チームの矢田祐一郎研究員、新潟大学脳研究所遺伝子機能解析学分野の原範和特任助教、池内健教授らの共同研究グループは、孤発性アルツハイマー病 (AD) の患者 102 人から樹立した iPS 細胞からなる iPS コホートを用いて、102 人分の大脳皮質神経細胞を作製し、複雑な孤発性 AD の病態を細胞種および病態ごとの表現型 (病的形質) に分解し、その背景の遺伝子データから AD の臨床リアルワールドデータを再構成する「Cellular dissection of polygenicity (CDiP) テクノロジー」を開発しました。

CDiP テクノロジーにより、細胞型・病態の表現型ごとに解析された遺伝的素因を基にした機械学習を通じて、AD コホート研究 (ADNI および J-ADNI) の臨床上の病態を再現することに成功しました。本研究成果は、孤発性 AD 病態の予測および回避への貢献が期待できます。

【本研究成果のポイント】

- 孤発性アルツハイマー病 (AD) ^{注1)} の患者 102 人から樹立した iPS コホート^{注2)} を用いて、複雑な孤発性 AD の病態を細胞種および病態ごとの表現型 (病的形質) に分解し、その背景の遺伝子データから AD の臨床リアルワールドデータを再構成する「Cellular dissection of polygenicity (CDiP) テクノロジー」を開発しました。
- iPS コホートを用いることで、iPS 細胞から分化誘導される細胞の種類と AD 病態の表現型 (病的形質) の組み合わせに特異的な背景の遺伝子のデータを探索することができるシステムとして cell GWAS という概念を構築し実施しました。
- 得られた遺伝子のデータと ADNI・J-ADNI ^{注3)} というアルツハイマー病総合診断体系実用化プロジェクトの臨床リアルワールドデータから、機械学習^{注4)} を用いて、孤発性 AD 病態発症に寄与するポリジーン^{注5)} の同定、個々人が AD 病態を発症するかどうか予測するアルゴリズム構築、AD の遺伝子治療の標的的同定といった社会実装へ向けた基盤実証を行いました。
- CDiP テクノロジーは、AD をはじめとする孤発性高齢疾患の病態を解説し、疾患の予知と回避を可能にすると期待できます。

1. 研究の背景

アルツハイマー病 (AD) は認知症を生じる神経変性疾患では最も患者が多く、2010 年時点で世界に 3000 万人いるとされています。超高齢社会の到来とともに AD 患者数は増え続け、根本的治療法がなければ、2030 年には 6000 万人に達し、2050 年には 1 億人を超えると推定されています。

近年の生物学的・遺伝学的研究の発展に伴い、多様な認知症、中でも AD の病態に関する分子生物学的な理解が進んできましたが、現時点では限られた対症療法しかありません。この背景には AD の病態が複雑であり、特に AD 患者の 95% を占める家族歴のない孤発性 AD については、病態の遺伝的な原因を探る有効なアプローチがないことが挙げられます。そこで共同研究グループは、孤発性 AD 患者の遺伝情報の特徴を明らかにし、それを基に臨床上の病態を再現することに挑みました。

II. 研究の概要

1) iPS 細胞を用いた孤発性 AD の病的形質の定量化

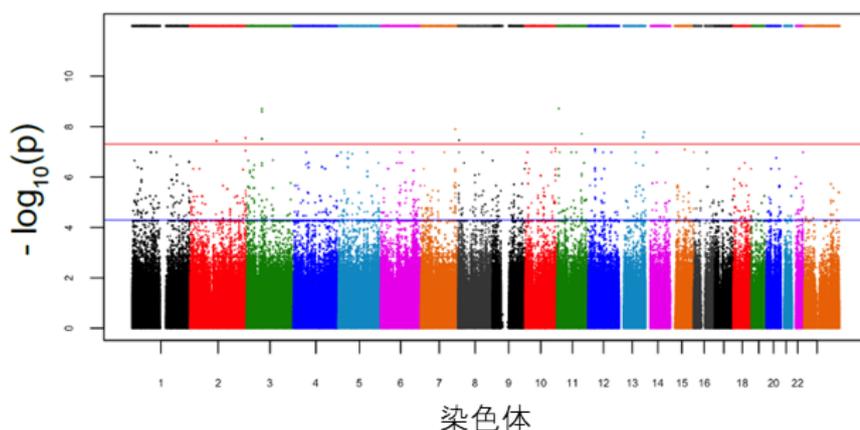
孤発性 AD 患者 102 人の末梢血細胞から iPS 細胞を樹立し、研究に用いる「iPS コホート」を構築しました。この iPS コホートを脳皮質の神経細胞へ分化誘導しました。得られた神経細胞において、AD の病態表現型の一つである老人斑^{注1)}の主要構成成分「アミロイド β ($A\beta$)^{注6)}」の産生量を測定したところ、個人の $A\beta$ 産生動態は一様ではなく、同じ孤発性 AD 患者の中でも多様な表現型があることを確認しました。

2) iPS コホートを用いた cell GWAS

次に、孤発性 AD 患者の病態の多様性をもたらす遺伝的背景の探索を行いました。これまで、孤発性 AD の遺伝的背景を調べるために、AD の発症有無や発症年齢などの臨床情報をもとにしたゲノムワイド関連解析 (GWAS)^{注7)} が試みられてきました。この臨床情報に基づいた GWAS (clinical GWAS) から、AD との関連が示唆される一塩基多型 (SNP)^{注8)} や疾患感受性領域 (遺伝子座)^{注9)} のデータが得られました。しかし、AD の病態は複雑で、生活習慣や依存症などが影響し、さらにさまざまな細胞種や病態表現型が関連して発症するため、細胞種あるいは病態表現型ごとに遺伝的背景を調べることはできませんでした。AD では細胞種ごとに異なる病態表現型が現れることから、病態を細胞レベルで観察しながら、直接的に細胞種や病態と結び付く遺伝情報を得ることが重要であると考えました。

そこで、iPS コホートから得られた AD 病態指標の一つである $A\beta$ 42/40 比率^{注6)} を表現型として、神経細胞レベルでの GWAS (cell GWAS) を行い、 $A\beta$ 42/40 比率に関連する遺伝子座を探索した結果、24 個の遺伝子座を同定しました (図 1)。このうち 5 個の遺伝子は $A\beta$ 産生に関連することが報告されており、8 個の遺伝子は clinical GWAS で AD に関連することが示唆されていました (図 1)。

(図 1) iPS コホートから樹立した脳皮質神経細胞の $A\beta$ 42/40 比率を対象とする cell GWAS



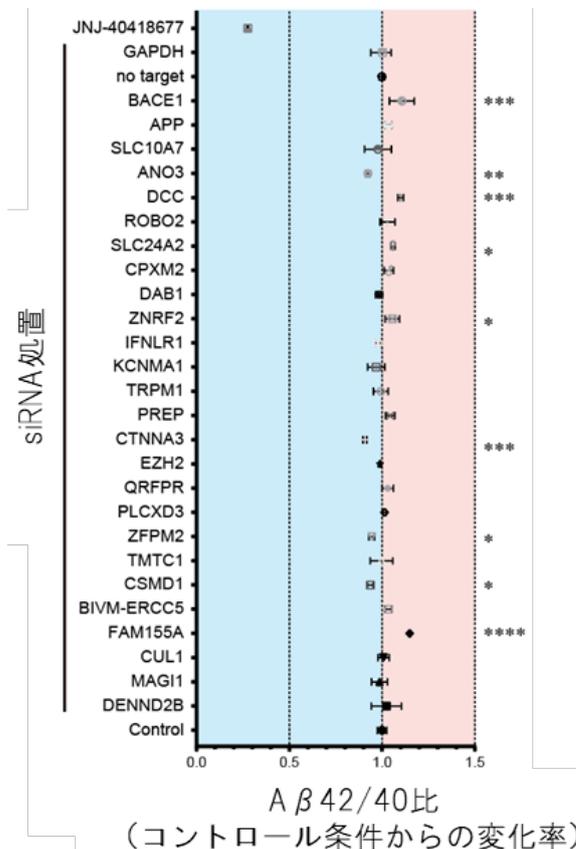
解析対象となった遺伝子座の p 値のマンハッタンプロット。赤線が $p = 5 \times 10^{-8}$ のゲノムワイドな有意水準を示す。24 個の遺伝子座がゲノムワイドな有意水準を満たし、そのうち 11 個が新しく同定された。

従って、cell GWAS は大規模な患者群を必要とする clinical GWAS の結果と一部重複しており、神経細胞での A β 42/40 比率と関連する遺伝子を適切に抽出していると考えられます。また、残りの 11 個は新しい AD 関連遺伝子として同定されました。これらの遺伝子がこれまで見いだせなかったのは、おそらく clinical GWAS だけでは多様な交絡因子^{注10)} がノイズとなったためだと考えられます。

3) cell GWAS による治療標的遺伝子の同定

見いだされた遺伝子がどのように A β 産生動態に影響するかを調べるために、siRNA^{注11)} を用いて、cell GWAS で見いだした 24 個の遺伝子を患者 iPS 細胞由来の脳皮質神経細胞においてそれぞれノックダウン^{注12)} した際の A β 42/40 比率を定量しました。すると、8 個の遺伝子をノックダウンしたときに A β 42/40 比率が変動したことから、A β の凝集から蓄積への経路が制御可能となり、今後、これらの 8 遺伝子が治療の標的となる可能性が示されました (図 2)。

(図 2) siRNA を用いた A β 42/40 比率変動の結果



cell GWAS で同定された 24 個の遺伝子と A β の産生経路に関わる 4 個の遺伝子の計 28 個を siRNA によりそれぞれノックダウンすると、8 個の遺伝子で対照群と比べて A β 42/40 比率が変化した。なお、一番上の JNJ-40418677 は A β 42/40 比率を変化させる陽性対照群となる化合物である。*は統計的に有意な変化であることを示す。

4) cell GWAS によるレアバリエントの発見

さらに、これら 24 個の遺伝子が AD の発症に直接影響するかどうかを検証するために、日本人 AD 患者集団を対象とした J-ADNI 研究 (アルツハイマー病の画像診断を用いたコホート研究) において取得されたデータをエクソーム解析^{注13)} したところ、cell GWAS から見いだした KCNMA1 遺伝子に AD のリスクとなるレアバリエント^{注14)} が存在することが分かりました。さらに、J-ADNI

とは別のコホートを対象にした米国の ADNI 研究のエクソーム解析においても、KCNMA1 遺伝子に AD のリスクとなる同じレアバリエントが見いだされたことから、このレアバリエントは異なる地域で普遍的であることがわかりました。

これらの結果から、cell GWAS で得られた AD の遺伝的背景の理解により、臨床では均質な AD 患者群の中から遺伝情報に基づいて特定の集団を抽出でき、AD のより詳細な区分にもつながることが示されました。

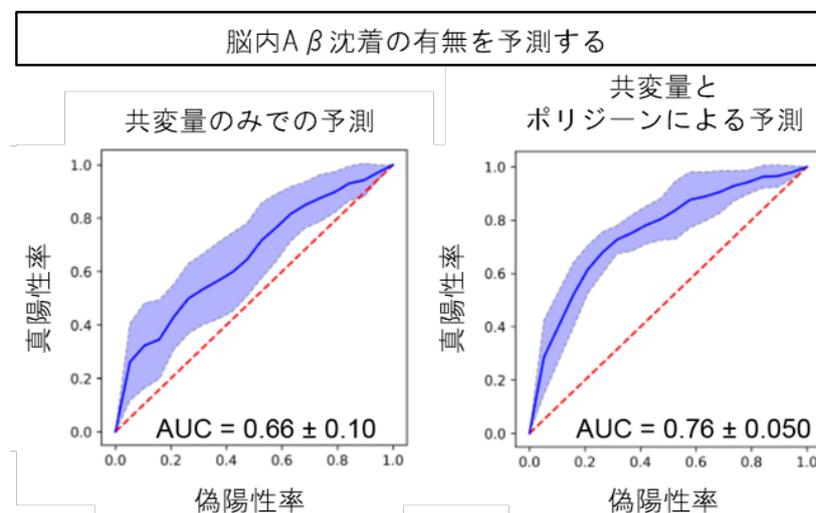
5) cell GWAS から得たポリジーンデータセットによる AD 発症予測

同定されたレアバリエントにより、単独の遺伝子情報から大きな AD 発症リスクを予測できませんが、このレアバリエントを持つ患者の数は 100 人に 1 人以下と限定的です。また、AD は複数の遺伝子（ポリジーン）の複合的な作用によって発症すると考えられています。

そこで、cell GWAS で得られた大脳皮質神経細胞の $A\beta$ 42/40 比率に関連する 496 個のポリジーンデータセットを用い、個人の AD 発症リスクの予測に取り組みました。予測対象のデータとして、米国の ADNI 研究から取得した実臨床リアルワールドデータを用いました。

496 人の ADNI 研究参加者の SNP 情報を基に、それぞれの参加者のポリジーンの遺伝型^{注15)}を定めました。このポリジーン情報を用いた機械学習により、 $A\beta$ の脳内沈着を検出する PET 検査^{注16)}の結果を予測しました。すると、年齢、性別、 $A\beta$ の脳内蓄積に関わっているとされる APOE4 遺伝型^{注15)}の情報だけを用いたときに比べて、それらにポリジーン情報を加えると PET 検査の予測精度が向上することがわかりました（図 3）。この結果から、cell GWAS で見いだしたポリジーンは実臨床における $A\beta$ 蓄積に関与することが示唆され、cell GWAS の結果が脳内 $A\beta$ 沈着の有無の予測に貢献することが示されました。

（図 3） cell GWAS で見いだした遺伝子群と機械学習による AD 病態の予測



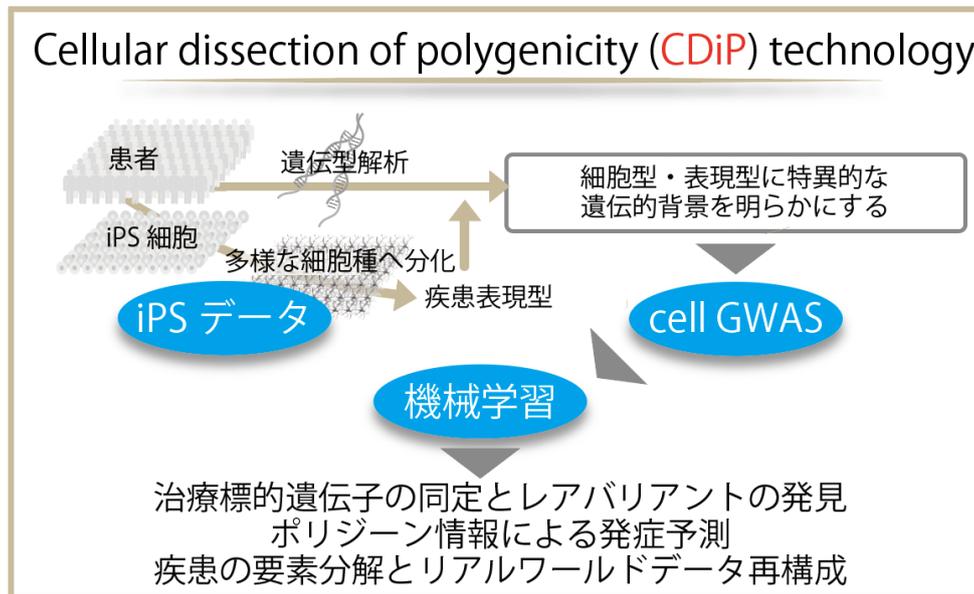
年齢・性別・APOE4 遺伝型の情報（共変量）を用いて ANDI 研究参加者の脳内 $A\beta$ の蓄積状況を予測した場合（左）に比べて、年齢・性別・APOE4 遺伝型の情報と cell GWAS で得られたポリジーン情報を用いた予測（右）は精度が上昇した。AUC は統計学用語で 0~1 の値をとる。1 に近いほど判別能が高いことを示す。

6) CDiP テクノロジーによる病態の再構成

以上のように、iPS 細胞から大脳皮質神経細胞を作製し、複雑な AD 病態を細胞種および病態ごとの表現型（病的形質）に分解し、その背景の遺伝子データから AD の実際の病態を再構成する新技术を「Cellular dissection of polygenicity (CDiP) テクノロジー」と命名しました（図 4）。

そしてこの CDiP テクノロジーを用いることで、孤発性 AD 病態発症に寄与するポリジーンを同定し、個々人が AD 病態を発症するかどうか予測するアルゴリズム構築、AD の遺伝子治療の標的的同定といった社会実装へ向けた基盤実証を行いました。

(図4) CDiP テクノロジーの概要



iPS コホートから抽出した「iPS データ」を用いることで、iPS 細胞から分化誘導される細胞の種類と AD 病態の表現型 (病的形質) の組み合わせに特異的な背景の遺伝子データを探索できるシステムとして「cell GWAS」を開発した。さらに、得られた遺伝子のデータと、ADNI、J-ADNI というアルツハイマー病総合診断体系実用化プロジェクトの臨床データから、「機械学習」を用いて、孤発性 AD 発症に関わるポリジーン同定の同定、個々人が AD を発症するかどうか予測するアルゴリズムの構築、AD の遺伝子治療の標的的同定といった社会実装へ向けた基盤実証を行った。CDiP テクノロジーは、AD をはじめとする孤発性高齢疾患の病態を解析し、疾患の予測と回避を可能にする。

今回、主に孤発性 AD 患者の脳皮質神経細胞を用いて Aβ に関する遺伝子情報を特定し、治療標的の同定、レアバリアントの発見、病態発症の予測を行いました。しかし、脳には神経細胞だけではなく、グリア細胞^{注17)}であるアストロサイトやミクログリアなども存在しており、AD の複雑な病態はそれぞれの細胞種が示す多様な病的形質の総和として形成されます。CDiP テクノロジーは、細胞種と病的形質の組み合わせごとにポリジーンデータおよびリスク予測情報を提供し、組み合わせ数を増やすことで AD 病態をより正確に予測できると考えられます。

III. 今後の展開

今後、CDiP テクノロジーにより、AD の脳皮質神経細胞のみならず、認知症の多種多様な病態をさまざまな細胞種で解析することにより、複雑な孤発性高齢疾患を分解・再構成し、個人の病態発症予測の社会実装が可能になると期待できます。そして、病気が生じていない無病期から個々人に必要な回避法を提案できる無病社会が期待されます。

IV. 研究成果の公表

これらの研究成果は、2022年2月18日午前1時(日本時間)に英国科学雑誌『Nature Aging』でオンライン掲載されました。

論文タイトル: “Dissection of the polygenic architecture of neuronal A β production using a large sample of individual iPSC lines derived from Alzheimer’s disease patients”

著者: Takayuki Kondo^{1,2,3}, Norikazu Hara⁴, Satoshi Koyama⁵, Yuichiro Yada^{2,3}, Kayoko Tsukita^{2,3}, Ayako Nagahashi^{1,2}, Takeshi Ikeuchi⁴, Kenji Ishii⁶, Takashi Asada⁷, Tetsuaki Arai⁷, Ryo Yamada⁵, Japanese Alzheimer’s Disease Neuroimaging Initiative (J-ADNI), Alzheimer’s Disease Neuroimaging Initiative (ADNI), Haruhisa Inoue^{1,2,3,*}

著者の所属機関:

1. 理化学研究所革新知能統合研究センター (AIP) iPSC細胞連携医学的リスク回避チーム
2. 理化学研究所バイオリソース研究センター (BRC) iPSC創薬基盤開発チーム
3. 京都大学 iPSC細胞研究所
4. 新潟大学脳研究所遺伝子機能解析学分野
5. 京都大学医学部統計遺伝学分野
6. 東京都健康長寿医療センター神経画像研究チーム
7. 筑波大学医学部精神神経科
8. 京都大学医学部附属病院 先端医療研究開発機構 (iACT)

* 責任著者

doi: 10.1038/s43587-021-00158-9

V. 本研究への支援

本研究は、下記機関より支援を受けて実施されました。

- 日本医療研究開発機構 (AMED) 再生医療実現拠点ネットワークプログラム (iPSC細胞研究中核拠点)
- 日本学術振興会科学研究費補助金 (17K16121, 18K18452, 20K16599, 21H02807)
- 京都大学医学部附属病院先端医療研究開発機構 iACT
- 鈴木謙三記念医科学応用研究財団
- 上原記念生命科学財団、日本医療研究開発機構 (AMED) (JP20dk0207045)

VI. 用語解説

注1) アルツハイマー病 (AD)、老人斑

ドイツの精神科医アロイス・アルツハイマー博士により1905年に報告された、進行性の記憶障害を伴う認知症。老齢期に発病し、徐々に記憶力の低下が進行し、意思疎通ができなくなる。その病理的特徴としては、脳内に老人斑と呼ばれるタンパク質の沈着が見られ、この老人斑の主成分がアミロイド β (A β) というペプチドであることから、A β の過剰な蓄積がアルツハイマー病の発症に深く関わっていると考えられてきた。

注2) コホート

コホートは人々の集団を意味し、コホート研究はこの集団を対象とした研究のこと。臨床コホート研究では、特定のライフイベントの有無や疾患の罹患群と健常群などを比較しながら、疾患の背景を探索する。本研究では、AD 患者コホート研究の参加者から iPS 細胞を樹立し、「iPS コホート」を構築した。

注3) ADNI、J-ADNI

ADNI は 2003 年に米国で開始されたアルツハイマー病の画像診断を用いた先導的研究のこと。AD の治療薬の治験を行うにあたって、進行スピードを正確に評価するなどの薬効の評価が困難であることから、MRI や PET 検査などの画像診断法、脳脊髄液などの体液のバイオマーカー測定などを併用し、AD の病理学的変化を正確に診断した上で、その進行過程を正確に評価することを目的としている。2008 年より日本でも ADNI (J-ADNI) が開始され、進行経過の精密な追跡評価が行われた。ADNI は Alzheimer's Disease Neuroimaging Initiative の略。

注4) 機械学習

データを分析する方法の一つで、コンピューターに明確な指示を与えなくても、大量のデータをもとに自律的に反復学習する技術。データの品質が学習結果として得られる予測などの精度に重要である。

注5) ポリジーン

単一遺伝子によるメンデル型遺伝では説明できない形質発現において、遺伝要因を構成する遺伝子群。ヒトの身長や植物の多収穫性、孤発性アルツハイマー病の発症などがポリジーンにより規定されると考えられている。

注6) アミロイド β ($A\beta$)、 $A\beta 42/40$ 比率

アミロイド β は、家族性アルツハイマー病の原因遺伝子である APP がコードするアミロイド β 前駆体タンパク質から、 β -セクレターゼや γ -セクレターゼなどの酵素による切断の結果産生される 40 アミノ酸程度のペプチドのこと。 $A\beta 42/40$ 比率は、42 個のアミノ酸からなる $A\beta 42$ と 40 個のアミノ酸からなる $A\beta 40$ の比率のこと。より凝集性の高い $A\beta 42$ の存在比率である $A\beta 42/40$ 比率が高いと、AD のリスクにつながる。

注7) ゲノムワイド関連解析 (GWAS)

特定の遺伝的変異を疾患や身体的特性に関連付けるための遺伝学的アプローチ。疾患や身体的特性に関連する遺伝子マーカーから、どのように疾患や身体的特性が生じるのかを理解し、予防および治療戦略を開発できる。GWAS は genome-wide association study の略。

注8) 一塩基多型 (SNP)

ゲノム塩基配列中の一塩基の変異には多様性が見られ、集団内で 1%以上の頻度で見られる場合を SNP と定義する。ほとんどの SNP は健康や発達に影響を与えないが、そのうちいくつかは人間の健康において非常に重要であることが証明されている。SNP は single nucleotide polymorphism の略。

注 9) 疾患感受性領域 (遺伝子座)

疾患の発症に関連している染色体上の領域のこと。

注 10) 交絡因子

調べようとする因子以外の因子で、病気の発生に影響を与える事象。AD の発症において遺伝的背景を調べる場合は、性別、生活習慣、併存疾患、服薬歴などが交絡因子として想定され、遺伝的背景の解明の障害となる。

注 11) siRNA

二本鎖 RNA で実験的に RNA 干渉を起こして標的の遺伝子発現を抑制し、標的遺伝子の役割を知るためなどに用いられる。

注 12) ノックダウン

特定の遺伝子の量を減少させる操作。本操作により標的の遺伝子発現量が低下し表現型が生じた場合、その標的の遺伝子の働きや表現型との関連性を明らかにすることができる。

注 13) エクソーム解析

ヒトゲノムにおける全てのタンパク質コード領域 (エクソーム) の塩基配列を決定すること。アルツハイマー病など、多様な遺伝的背景が病態にとって重要と想定されている疾患の研究や臨床診断に適用される。

注 14) レアバリエント

SNP より稀少で、集団における頻度が 1%以下の遺伝子変異。GWAS 解析と相補的に用いられ、頻度は低いものの病態の遺伝学的背景を構成することが明らかになりつつある。

注 15) 遺伝型、APOE4 遺伝型

遺伝型とはある遺伝子の塩基配列パターンのこと。APOE4 遺伝型は、 $A\beta$ の脳内蓄積に大きく関わっているとされている APOE 遺伝子のタイプ。遺伝子のタイプは ϵ (イプシロン) 2、 ϵ 3、 ϵ 4 が二つ一組で 6 パターンの遺伝型を構成しており、 ϵ 4 の場合、すなわち APOE4 遺伝型を持つと $A\beta$ の脳内蓄積が多くなることが知られている。

注 16) PET 検査

体内の目的の物質に集積する放射性薬剤を体内に投与し、その分布を特殊なカメラでとらえて画像化する検査。

注17) グリア細胞

脳の神経細胞の生存や発達のための脳内環境の維持と代謝的支援を行っている細胞。脳に分布する主なグリア細胞は、アストロサイト、オリゴデンドロサイトおよびミクログリアの3種に分類される。

参照文献：京都大学 iPS 細胞研究所 (CiRA) プレスリリース