

アルツハイマー型認知症のゲノム解析による 発症リスク遺伝子の同定と早期予防のための発症予測モデルの開発

国立長寿医療研究センター メディカルゲノムセンター
遺伝統計解析ユニット長 重水 大智

(共同研究者)

国立長寿医療研究センター メディカルゲノムセンター 研究員 秋山 真太郎

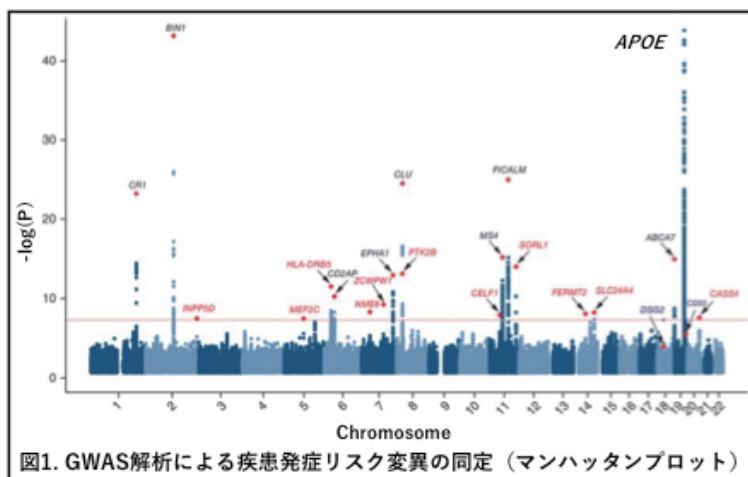
はじめに

認知症の高齢者は毎年増加傾向にあり、2025年にはその数は700万人に達すると言われている。しかしながら、その病態は未だ解明されておらず、根本的治療薬や予防法も十分に確立されていない。したがって、認知症の病態解明を含めた認知症の早期発見や診断法の確立が急務である。

アルツハイマー型認知症（AD）は家族歴ありの家族性、家族歴なしの孤発性、発症年齢65歳を境に早発性と遅発性に大別される。家族性ADは、単一遺伝子疾患としてこれまで疾患原因変異が報告されている。しかしながら、孤発性ADは、*APOE* 遺伝子の遺伝子多型が高頻度に認められること、アミロイドβ蛋白の脳への蓄積が認められること以外、ほとんど報告がない。また、*APOE* 遺伝子の遺伝子多型やアミロイドβ蛋白の脳への蓄積だけでは十分な疾患の発症機序の説明が難しく、未だ不明な点が多い複雑な多因子疾患と言われている。

近年、ヒトゲノム全体の一塩基多型（SNP）をジェノタイピングし、疾患群と健常者群を比較することで、疾患群に特異的な疾患発症リスク因子を網羅的に探索するゲノムワイド関連解析（Genome-Wide Association Study: GWAS）が行われ、様々なりスク因子が同定されている。ADにおいても、欧米のグループによる大規模検体（>70,000人）を用いたGWAS解析から*APOE* 以外の新規疾患発症リスク因子が11個同定されている⁽¹⁾（図1）。

また、次世代シーケンス（NGS）技術の目覚ましい発展により、ゲノムシーケンスが比較的安価に実装可能になり、DNA（バリエーション）レベルによるAD発症リスク因子の探索に加え、RNA（遺伝子発現）レベルによる発症リスク因子の探索も可能となってきている。それにより遺伝子レベ



ルでの新規疾患発症リスク因子もいくつか報告されている⁽²⁾(図2)。また、遺伝子発現の制御に大きく関わるmicroRNA(miRNA)が、ADを含む神経変性疾患に関連していることも報告されている⁽³⁾。

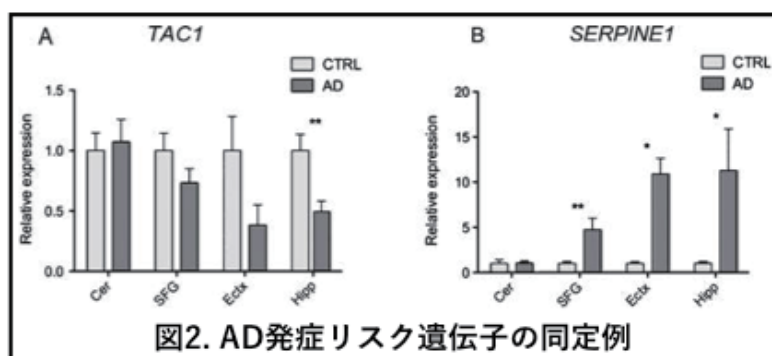


図2. AD発症リスク遺伝子の同定例

上述の疾患発症リスク因子の探索だけにとどまらず、それらを用いた疾患発症予測モデルの開発が2型糖尿病やリウマチをはじめとしたコモンディゼーズや癌に対して研究が行われはじめている⁽⁴⁻⁵⁾。しかしながら、ADに対しての研究はそれほど行われていない。そこで本研究では、AD発症者と健常者のmiRNAを網羅的に比較解析することで疾患発症リスク因子を同定し、そのリスク因子をもとにAD発症予測モデルの構築を行うことを目的とする。この研究はADの早期発見を可能にし、早期予防への貢献が期待される。

結 果

【解析に使用した検体】

本研究で使用した検体数は1,309例で、すべての検体の末梢血から血清を生成し、東レの3D-Geneマイクロアレイを用いて網羅的にmiRNAの発現量を測定した(2,562 miRNAs)。検体の統計値は、それぞれ疾患群と健常者群で平均年齢が79.2歳と71.7歳、男性の割合が30%と53%、APOE4アレルのカウント数の平均が0.5と0.19であった。全検体は発症予測モデル構築用(学習セット)と検証用(テストセット)に、年齢の偏りがないように1:1に分割した(表1)。

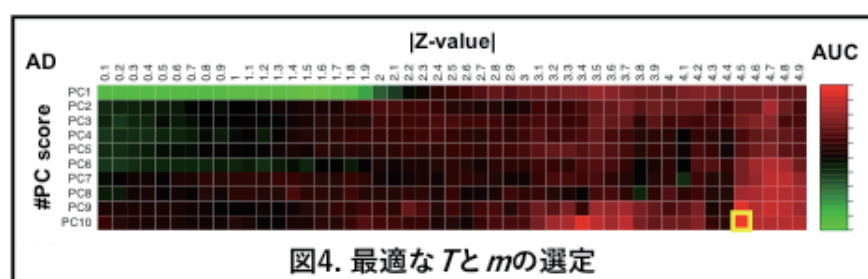
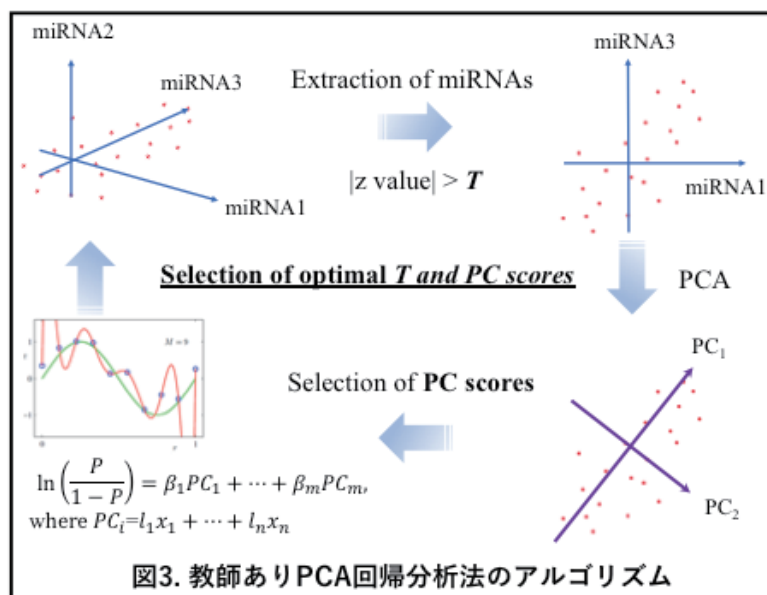
表1. 学習セットとテストセットの検体数と臨床情報の統計値

	学習セット				テストセット			
	症例数	年齢	性別 (M)	APOE	症例数	年齢	性別 (M)	APOE
疾患群	511	79.2	0.29	0.53	510	79.2	0.31	0.47
健常者群	144	71.7	0.49	0.22	144	71.8	0.56	0.15

【疾患発症予測モデルの構築】

教師ありPCAロジスティック回帰法を用いて、疾患発症予測モデルの構築を行った(図3)。この予測モデル構築には、miRNAの選定に必要な閾値 T とPCA解析から得られるPCスコアの数 m 、を決める必要がある。この2つのパラメータの組み合わせを網羅的に調べ

ることで、最適なモデル構築に必要な組み合わせを同定した(グリッドサーチ)。その組み合わせの評価には、学習セットを用いた10-fold cross validationを用いた。その結果、 $T=4.5$, $m=10$ の組み合わせによって構築された78個のmiRNAからなるAD発症予測モデルが、AUC 0.877と高い予測能を示した(図4)。この予測モデルは、独立なテストセットでその予測能の再現性を検証したが、学習セットの場合と同様、高い予測能が確認された(AUC=0.874)。



【コホートデータによる検証】

上述の予測モデルの構築、検証は後ろ向きデータによるもので、つまり疾患群と健常者群を正確に分類できたことになる(Classification)。したがって、構築したAD発症予測モデルが前向きデータ(コホートデータ)においても適応可能かどうかの検証を行った(Prediction)。検証実験には軽度認知症(MCI) 32症例の追跡データを用いた。少なくとも6ヶ月以上の予後が観察され、32症例中10症例がADに移行したデータである。このデータを用いてMCIのADへの移行を予測できるかを調べた。表2がその一致度を示した結果である。ADに移行した検体すべてを、本研究で構築したAD発症予測モデルは移行すると予測し(感度1.0)、また予測モデルが移行しないと予測した4症例も移行していない結果となった(陰性的中率1.0)。一方で、予測モデルが移行すると予測した28症例中18症例は移行していない症例であり(特異度0.18)、偽陽性が高い結果となった(表2)。

表2. コホートデータによる予測モデルの性能評価

		コホートデータ		
		MCI からAD	MCI からMCI	計
予測モデルの予測	MCI からAD	10	18	28
	MCI からMCI	0	4	4
計		10	22	32

考 察

本研究では、新たに78個の疾患発症リスク因子 (miRNA) を同定し、それらを用いた疾患発症予測モデルは高精度な予測に成功した。しかしながら、コホートデータによる検証では、偽陽性が高いことが確認された。これはMCIのデータからADへの移行を予測したため、多くの検体がADへ移行すると予測したことが示唆される。現在データ収集中のため未確認であるが、健常者からADへ移行した検体による再検証が必要である。また、その他のさまざまな機械学習、例えば、ランダムフォレスト、サポートベクターマシン、ナイーブベイズ、回帰モデル (LASSO、Elastic net) や深層学習による予測モデルとも比較検証が必要かもしれない。将来的には、miRNAのデータだけでなく、バリエーションデータ、遺伝子発現データ等のオミクスデータによる予測モデルの構築がさらなる予測能の向上に貢献すると期待している。

要 約

網羅的なmiRNAのデータに、教師ありPCAロジスティック回帰法を応用することで、78個の疾患発症リスク因子を同定し、そのリスク因子によるアルツハイマー型認知症を高精度に予測する疾患発症予測モデルの開発を行った (AUC 0.87)。今後、バリエーションデータ、遺伝子発現データ等を組み合わせたオミクス解析により、さらなる予測能の向上を期待する。

文 献

1. Lambert, JC. et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet.* 2013 Dec;45 (12) :1452 – 1458
2. Magistri, M. et al. Transcriptomics Profiling of Alzheimer's Disease Reveal Neurovascular Defects, Altered Amyloid-beta Homeostasis, and Deregulated Expression of Long Noncoding RNAs. *J Alzheimers Dis.* 2015;48 (3) :647 – 665
3. Cogswell JP, et al. Identification of miRNA changes in Alzheimer's disease brain and CSF yields putative biomarkers and insights into disease pathways. *J Alzheimers Dis.* 2008 May;14 (1) :27– 41.
4. Shigemizu, D. et al. The prediction models for postoperative overall survival and disease-free survival in patients with breast cancer. *Cancer Med* 2017 Jul;6 (7) :1627 – 1638.
5. Shigemizu, D. et al. The Construction of Risk Prediction Models Using GWAS Data and Its Application to a Type 2 Diabetes Prospective Cohort. *PLoS One* 2014 9 (3) , e92549.