

## Review Article



# Genetics of Alzheimer's Disease

## OPEN ACCESS

Received: Nov 24, 2018

Accepted: Dec 6, 2018

### Correspondence to

Jong Hun Kim, MD, PhD

Department of Neurology, National Health Insurance Service Ilsan Hospital, 100 Ilsan-ro, Ilsandong-gu, Goyang 10444, Korea.

E-mail: jh7521@naver.com

© 2018 Korean Dementia Association

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

### ORCID iDs

Jong Hun Kim

<https://orcid.org/0000-0002-2594-1048>

### Conflict of Interest

The author has no financial conflicts of interest.

Jong Hun Kim

Department of Neurology, National Health Insurance Service Ilsan Hospital, Goyang, Korea

## ABSTRACT

Alzheimer's disease (AD) related genes have been elucidated by advanced genetic techniques. Familial autosomal dominant AD genes founded by linkage analyses are *APP*, *PSEN1*, *PSEN2*, *ABCA7*, and *SORL1*. Genome-wide association studies have found risk genes such as *ABCA7*, *BIN1*, *CASS4*, *CD33*, *CD2AP*, *CELF1*, *CLU*, *CR1*, *DSG2*, *EPHA1*, *FERMT2*, *HLA-DRB5*-*HLA-DRB1*, *INPP5D*, *MEF2C*, *MS4A6A*/*MS4A4E*, *NME8*, *PICALM*, *PTK2B*, *SLC24A4*, *SORL1*, and *ZCWPW1*. *ABCA7*, *SORL1*, *TREM2*, and *APOE* are proved to have high odds ratio (>2) in risk of AD using next generation sequencing studies. Thanks to the promising genetic techniques such as CRISPR-CAS9 and single-cell RNA sequencing opened a new era in genetics. CRISPR-CAS9 can directly link genetic knowledge to future treatment. Single-cell RNA sequencing are providing useful information on cell biology and pathogenesis of diverse diseases.

**Keywords:** Alzheimer's Disease; Genetics; CRISPR-Cas Systems; Single-Cell Analysis

## INTRODUCTION

The new genetics techniques that decode the human genome have revealed the risk genes for various diseases. Based on these, there have been many developments in diagnosis, prediction, and understanding of diseases. The genetic pathogenesis of Alzheimer's disease (AD) has also been revealed in the last decade. This review is about the introduction of the basic genetics concept and the AD genes revealed by each genetics technique, and promising genetics research techniques other than sequencing.

## ALLELE, MUTATION, SINGLE NUCLEOTIDE POLYMORPHISM (SNP), AND VARIANT

The "allele" refers to the mutation of a particular gene. They can exist as major or minor allele. A minor allele is an allele form that is less frequent in the population while a major allele is more frequent. Minor allele frequency (MAF) is a concept of frequency. MAF can be calculated by a following equation.

$$MAF = \frac{\text{heterozygote} + 2 \times \text{homozygote}}{2 \times \text{individual}}$$

The MAF should be understood as the frequency of the allele in the chromosome number, not the individual number. For an example, MAF 0.3 does not mean that 30% of the population have a minor allele. An individual has 2 chromosomes inherited from his mother and father, and homozygotes are considered when counting the number of alleles. The “SNP” means that 1 DNA nucleotide is replaced with another DNA nucleotide. Polymorphism can also mean that MAF is more than 1%. The “mutation” is less than 1% MAF and does not include the concept of DNA length. The “variant” covers all the concepts of allele, SNP, and mutation, which are described above. Allele, mutation, SNP, and variant are not strictly separated and used interchangeably. If the length of the changed nucleotides is more than 2, SNP cannot be used.

## COMMON VS. RARE ALLELE

In genetic study, MAF in the population are important, because clinical significances of mutations are different, and research methods are different according to MAF. Common alleles refer to alleles where their MAFs greater than 1% and can be studied using the genome-wide association study (GWAS). Rare alleles are those with less than 1% of MAF and are studied using next generation sequencing. The statistical power can be calculated by MAF and odds ratio (OR). The higher the MAF and the OR, the higher the statistical power.

## GWAS

GWAS is based on a technology that can detect hundreds of thousands of mutations at once using DNA chips. DNA chips can detect predetermined site mutations. Therefore, GWAS is conducted for the common alleles. If a DNA chip is designed for rare alleles, only small information can be obtained, because most of the results will be monomorphic. During meiosis, chromosomes perform homologous recombination which gives more diversity to the offspring. The nucleotides closely located to each other tend to move together after the homologous recombination. These highly related DNA fragments are called linkage disequilibrium blocks. Therefore, mutations that show significant association with diseases in GWAS results may be the cause of some of those disease, but they may be detected as markers associated with causal mutations. After performing GWAS, it is necessary to perform sequencing around the significant markers to find the most significant mutations. Before and after the GWAS, it is necessary to perform several quality control tasks to minimize false associations. Finally, significant results are selected from the multiple test corrections. The variants with  $p$  value  $<5 \times 10^{-8}$  ( $0.05/10^6$ ) are recognized as the significant variants in the GWAS.

GWAS was first published in 2005 and *CFH* was found to play an important role in age-related macular pigmentosa.<sup>1</sup> GWAS associated with AD was also performed to identify several risk genes. The largest studies to date have been performed by Lambert et al.,<sup>2</sup> in 2013, in which 74,046 people of International Genomics of Alzheimer's Project (IGAP) were involved. No further large-scale study has been performed since the study. The authors identified 20 genes (*ABCA7*, *BIN1*, *CASS4*, *CD33*, *CD2AP*, *CELF1*, *CLU*, *CR1*, *DSG2*, *EPHA1*, *FERMT2*, *HLA-DRB5-HLA-DRB1*, *INPP5D*, *MEF2C*, *MS4A6A/MS4A4E*, *NME8*, *PICALM*, *PTK2B*, *SLC24A4*, *SORL1*, and *ZCWPWI*) associated with AD.

## NEXT-GENERATION SEQUENCING (NGS)

NGS is a technique that can dramatically increase the efficiency compared to the existing sequencing techniques and can decipher individual's genome quickly. Because NGS can read DNA, it can study both common and rare alleles. However, it is still more expensive than GWAS and is mainly used to study rare alleles. In 2013, *TREM2* was elucidated as an AD risk gene using NGS.<sup>3</sup> In addition, the A673V mutation of the *APP* gene, which is known as a gene of autosomal dominant familial AD, was elucidated as a protective mutation of AD. Persons with this mutation had a low production of amyloid  $\beta_{1-42}$  and subjects with *APOE*  $\epsilon 4$  homozygote were able to survive without dementia if they had A673V mutation in *APP*. The result showed the hope of anti-amyloid treatment.<sup>4</sup>

The studies targeting rare alleles have low statistical power and try to overcome such problems using gene-based analysis. The frequency can be increased by adding the rare alleles that occurs in the boundary of genes. Another advantage of the gene-based analysis is that the threshold of multiple test corrections can only achieve  $p$  value =  $0.05/20,000$  (number of genes in human genome) =  $2.5 \times 10^{-6}$ . However, the disadvantage is that the mutations used in the assay are mixed with risk, protective, and neutral mutations, which may offset the results.

The Kari Staffanson's team of Icelandic deCode genetics, which discovered *TREM2* and a protective mutation in *APP*, uncovered a new powerful AD risk gene with interesting approach. In 2010, Dickson et al.,<sup>5</sup> suggested a synthetic association hypothesis in order to explain the low heritability of GWAS genes. The various rare alleles may be synthetically associated with the common variants that are significant in GWAS and are the real cause of the diseases. The Kari Staffanson's team tested the cryptic association theory and raised the threshold for multiple corrections by targeting only AD genes found in the large-scale GWAS of IGAP.<sup>2</sup> In addition, they performed gene-based analysis using only loss-of-function variations in order to overcome the disadvantage of gene-based analysis, and found that *ABCA7* is an AD risk gene with a high OR.<sup>6</sup>

## FAMILIAL & EARLY-ONSET AD (EOAD) GENES

*APP*, *PSEN1*, and *PSEN2* are well-known autosomal dominant familial AD genes. These genes are found in about 70% of families with more than 1 EOAD, and account for only 20% of sporadic EOAD when the sporadic EOAD are younger than 50 years of age.<sup>7</sup> *ABCA7* and *SORL1* can cause autosomal dominant AD depending on the types of mutations.<sup>6-9</sup> *TREM2* is also known to be associated with EOAD.<sup>10</sup>

## POLYGENIC RISK SCORES (PRSS)

The heritability of AD is about 70%, but only 30% of the AD heritability can be explained by the known genes.<sup>11</sup> The phenomenon that cannot account for heritability by known genes was common in complex genetic diseases/traits. Purcell et al., published excellent results in 2009 through schizophrenia GWAS.<sup>12</sup> They showed that a large number of mutations with a  $p$  value  $< 0.5$  were highly associated with schizophrenia and bipolar disorder. Mutations with small

effect sizes showed additive effects on diseases and explained the expected heritability of the complex genetic disease/trait.<sup>13</sup>

A PRS study of AD was first published in 2015<sup>14</sup> where the IGAP study, a large-scale GWAS for AD,<sup>2</sup> was divided into discovery set and replication set. PRS was constructed by adding the  $\beta$  values of the mutations with  $p$  value  $<0.5$  which are selected in discovery set. The PRS showed high association with AD and had predictive values in replication set.

## POPULAR GENETICS TECHNIQUES

### CRISPR-CAS9

The fascinating thing about genetics is that the outcome is crucial to finding the pathogenesis of the disease. In AD, the discovery of *PSEN1*, *PSEN2*, *APP*, *APOE*, and *TREM2* played a crucial role in making pathogenesis of AD. In addition to finding pathogenesis, CRISPR-CAS9 opened the possibility of applying genetic research results directly to therapy. CRISPR-CAS9 is a third-generation gene scissors, which has the advantage of being highly accurate and efficient compared to the previous generations gene scissors. Also, it can be manufactured within a day in a laboratory.<sup>15</sup>

Currently, the most popular CRISPR-CAS9 is made from *Streptococcus pyogenes*. Within CRISPR-CAS9, there is a guide RNA, which makes CRISPR-CAS9 find a specific site in a genome. The guide RNA of *Streptococcus pyogenes* is 20 base pair long and researchers can modify the guide RNA sequence in order to change the action site of CRISPR-CAS9. The primary role of CRISPR-CAS9 is cutting the DNA. After cutting DNA, hosts can restore the defect and cause mutations in 2 ways. First, the non-homologous end joining (NHEJ) causes 0–10 base pair deletion when it links the breaking DNA sequence. If the repair occurs without deletion, the CRISPR-CAS9 cut it again because of its high efficiency, so only the deletion of 1–10 base pairs can be observed. Because 3 DNA bases encode 1 amino acid, gene knock-out due to frameshift occurs at 60% when targeting the exon. Second, homology-directed repair (HDR) can cause mutations. If external DNA oligonucleotides that have the same shape around the truncated part and contain another nucleotide sequence segment in the middle is inserted, a new nucleotide sequence segment will be inserted when the truncated part is restored. In the cells, NHEJ are more frequent than HDR, which makes intended mutagenesis difficult. Recently, researches are solving this problem.<sup>16</sup> In addition, off-target effects need to be solved in order to apply CRISPR-CAS9 to clinical setting.

### Single-cell RNA sequencing

The genomes of all cells in a person are the same, however, the expressed proteins differ according to tissues or cells. Therefore, in order to observe actual events in a cell, DNA studies are not helpful, rather, transcribed RNA are the helpful ones. RNA can be reversely transcribed into DNA and be sequenced. RNA can be quantified and the kinds of proteins produced in a cell found by NGS. In recent years, a technique called multiple displacement amplification (MDA) has been developed that amplifies very small amounts of RNA using bacteriophage DNA polymerase.<sup>17</sup> This MDA made possible single-cell RNA sequencing. The advantage of single-cell RNA sequencing is that possibilities can be controlled and the results obtained are more refined. Single-cell RNA sequencing is widely used in a variety of researches. Scientists are trying to make human cell atlas by analyzing the RNA profiles of all cells in all tissues using single-cell RNA sequencing.<sup>18</sup>

## CONCLUSIONS

Many researches have been carried out on AD, but most of the known knowledge to date is based mostly on genetics findings. These are the powerful aspects of genetics. Based on genetic knowledge for AD, scientists are increasing therapeutic targets.<sup>16,19-22</sup>

## REFERENCES

1. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005;308:385-389.  
[PUBMED](#) | [CROSSREF](#)
2. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* 2013;45:1452-1458.  
[PUBMED](#) | [CROSSREF](#)
3. Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. *N Engl J Med* 2013;368:107-116.  
[PUBMED](#) | [CROSSREF](#)
4. Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson PV, Bjornsson S, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature* 2012;488:96-99.  
[PUBMED](#) | [CROSSREF](#)
5. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. *PLoS Biol* 2010;8:e1000294.  
[PUBMED](#) | [CROSSREF](#)
6. Steinberg S, Stefansson H, Jonsson T, Johannsdottir H, Ingason A, Helgason H, et al. Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. *Nat Genet* 2015;47:445-447.  
[PUBMED](#) | [CROSSREF](#)
7. Nicolas G, Wallon D, Charbonnier C, Quenez O, Rousseau S, Richard AC, et al. Screening of dementia genes by whole-exome sequencing in early-onset Alzheimer disease: input and lessons. *Eur J Hum Genet* 2016;24:710-716.  
[PUBMED](#) | [CROSSREF](#)
8. Cuyvers E, De Roeck A, Van den Bossche T, Van Cauwenberghe C, Bettens K, Vermeulen S, et al. Mutations in ABCA7 in a Belgian cohort of Alzheimer's disease patients: a targeted resequencing study. *Lancet Neurol* 2015;14:814-822.  
[PUBMED](#) | [CROSSREF](#)
9. Pottier C, Hannequin D, Coutant S, Rovelet-Lecrux A, Wallon D, Rousseau S, et al. High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. *Mol Psychiatry* 2012;17:875-879.  
[PUBMED](#) | [CROSSREF](#)
10. Slattery CF, Beck JA, Harper L, Adamson G, Abdi Z, Uphill J, et al. R47H TREM2 variant increases risk of typical early-onset Alzheimer's disease but not of prion or frontotemporal dementia. *Alzheimers Dement* 2014;10:602-608.e4.
11. Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet* 2012;13:537-551.  
[PUBMED](#) | [CROSSREF](#)
12. International Schizophrenia Consortium; Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009;460:748-752.  
[PUBMED](#)
13. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell* 2017;169:1177-1186.  
[PUBMED](#) | [CROSSREF](#)
14. Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, et al. Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain* 2015;138:3673-3684.  
[PUBMED](#) | [CROSSREF](#)
15. Shalem O, Sanjana NE, Zhang F. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet* 2015;16:299-311.  
[PUBMED](#) | [CROSSREF](#)

16. Ma H, Marti-Gutierrez N, Park SW, Wu J, Lee Y, Suzuki K, et al. Correction of a pathogenic gene mutation in human embryos. *Nature* 2017;548:413-419.  
[PUBMED](#) | [CROSSREF](#)
17. Lasken RS, McLean JS. Recent advances in genomic DNA sequencing of microbial species from single cells. *Nat Rev Genet* 2014;15:577-584.  
[PUBMED](#) | [CROSSREF](#)
18. Rozenblatt-Rosen O, Stubbington MJ, Regev A, Teichmann SA. The human cell atlas: from vision to reality. *Nature* 2017;550:451-453.  
[PUBMED](#) | [CROSSREF](#)
19. Shi Y, Yamada K, Liddelow SA, Smith ST, Zhao L, Luo W, et al. ApoE4 markedly exacerbates tau-mediated neurodegeneration in a mouse model of tauopathy. *Nature* 2017;549:523-527.  
[PUBMED](#) | [CROSSREF](#)
20. Ulland TK, Song WM, Huang SC, Ulrich JD, Sergushichev A, Beatty WL, et al. TREM2 maintains microglial metabolic fitness in Alzheimer's disease. *Cell* 2017;170:649-663.e13.  
[PUBMED](#) | [CROSSREF](#)
21. Huang YA, Zhou B, Wernig M, Südhof TC. ApoE2, ApoE3, and ApoE4 differentially stimulate APP transcription and A $\beta$  secretion. *Cell* 2017;168:427-441.e21.  
[PUBMED](#) | [CROSSREF](#)
22. Wang Y, Cella M, Mallinson K, Ulrich JD, Young KL, Robinette ML, et al. TREM2 lipid sensing sustains the microglial response in an Alzheimer's disease model. *Cell* 2015;160:1061-1071.  
[PUBMED](#) | [CROSSREF](#)