第51卷第2期 中国科学技术大学学报 Vol. 51, No. 2
2021年2月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Feb. 2021

Received: 2021-01-06; Revised: 2021-01-28 dio: 10.52396/JUST-2021-0008

SIPA1L2 as a risk factor implicated in Alzheimer's disease

WEN Canhong[1,2], LI Jiarui[1], CAI Jiahui[3], TAN Haizhu[1*]

1. School of Data Science, University of Science and Technology of China, Hefei 230027, China;
2. Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei 230026, China;
3. Department of Preventive Medicine, Shantou University Medical College, Shantou 515031, China

* Corresponding author. E-mail: linnanqia@126.com

Abstract: Alzheimer's disease (AD) is a common neurodegenerative disorder with high heritability. An increasing number of common variants have been found to be associated with AD, but these common variants can only explain a small proportion of the heritability. Theory and practice have shown that rare variants can explain the remaining heritability. We explored rare functional variants that altering susceptibility to AD among 600470 variants in 389 individuals (175 with AD and 214 with cognitively normal). Firstly, after imputing the missing genotypes on the Michigan imputation server, quality control and gene-based annotation were carried out. Secondly, the efficient resampling sequence kernel association test was performed on 311 annotated exonic variants. Finally, the underlying biological interpretations of the identified risk gene were predicted through several bioinformatics tools. The results showed that under the Bonferroni correction, the rare missense variant rs2275303 in SIPA1L2 gene was significantly associated with AD ($P=6.00E-04$), and its pathogenicity was verified by bioinformatics analysis. SIPA1L2 gene is expected to play an important role in the prevention, diagnosis, prognosis and treatment of AD.

Keywords: Alzheimer's disease; rare variants; association analysis; ER-SKAT

CLC number: Q811.4 **Document code**: A

1 Introduction

Alzheimer's disease (AD) is a chronic, incurable, common neurodegenerative disorder with high heritability (estimated heritability up to 80%)[1,2]. Linkage analysis revealed that the apolipoprotein E (APOE) gene was the first confirmed susceptibility gene for AD[3]. However, APOE can account for at most 50% of the total genetic effect in AD[4]. A large number of genome-wide association studies (GWASs) have successfully identified other genomic risk factors for AD. For example, two extensive GWAS, of 12000 probable AD cases and 18000 age-matched non-demented controls, revealed three new candidates for the genetic risk of late-onset or sporadic AD: CLU, CR1 and PICALM[5]. A staged association study and testing of suggestive loci revealed variants in ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP associated with AD[6]. Despite these discoveries, these risk common variants (minor allele frequency, MAF> 0.05) identified by GWAS account for less than half of the heritability to date[7,8]. The reasons might be the common variant hypothesis for GWAS, that allelic variants exist in more than 1% to 5% of the population, which is only partly implicated in usual diseases[9,10].

Hence, there are many explanations for the remaining unexplained heritability (also called missing heritability), such as rare variants ($0.005<MAF<0.01$) and structured variants, which are poorly detected by genotyping arrays[11]. Previous studies have revealed that rare variants are more likely to contain deleterious functional consequences to cause diseases[12]. Rare variants play an important role in complex human diseases. In a study of 14002 individuals and 202 genes, approximately 95% of exonic variants were rare variants and were present in only one or two individuals[13]. An increasing number of studies have shown that rare variants have large effect sizes in AD. For example, Nicolas et al. suggested that SORL1 is a major risk factor for the familial early-onset AD[14]. Rare variants in TREM2, SORL1, and ABCA7 contributed to AD in 1779 cases and 1273 controls[15]. ABCA7, a loss of the function rare variant, has been found related to AD.

How to detect rare variants in association studies is a challenge. The statistical power of single-variant association study is low for assessing the role of rare and

Citation: WEN Canhong, LI Jiarui, CAI Jiahui, et al. SIPA1L2 as a risk factor implicated in Alzheimer's disease. J. Univ. Sci. Tech. China, 2021, 51(2): 147-152.

low-frequency variants unless the sample size is huge. Multiple comparison adjustments are also required[16]. The main multiple-comparison problem is that type I error increases with each additional test. Hence, numerous statistical methods specifically designed to increase the power for the rare variant association analysis have been proposed. For example, the burden tests method assumes that all rare variants in the target region have effects on the phenotype in the same direction with the similar magnitude. The Sequence Kernel Association Test (SKAT), an extension of the C-alpha test, provides a robust test that is particularly powerful in the presence of protective and deleterious variants and null variants[17]. The Efficient Resampling-SKAT (ER-SKAT), calibrating single and gene-based rare variant association analyses, can improve computational efficiency over conventional resampling with low minor allele count (MAC) variant sets and can ameliorate the conservativeness of results in case-control studies[18].

To identify the risk rare variants related to AD, we explored rare functional variants altering susceptibility to AD in 389 individuals with 600470 variants (175 individuals with AD and 214 cognitively normal (CN) controls) from the Alzheimer's disease neuroimaging initiative (ADNI) database (www.loni.ucla.edu/ADNI). After imputing missing genotypes on the Michigan imputation server, we performed quality control (QC) and gene-based annotation. Subsequently, we performed a gene-based association study of 311 selected annotated exonic variants in 254 genes by using ER-SKAT. The identified significant rare variant could explain some proportions of genetic heritability of AD.

2 Materials and methods

2.1 Participants

Data used in this study were downloaded from the ADNI-1 database. In total, 819 individuals were enrolled at the baseline, including 175 with AD, 398 with mild cognitive impairment (MCI), and 229 CN controls. The present study involved 175 individuals with AD and 214 CN controls with 600470 variants located on chromosomes 1-22 in PLINK data format. Information for the participants, including selection criteria, demographic data, and baseline assessments, are available on the ADNI website.

2.2 Statistical analysis

Genotype imputation can increase genome coverage and improve power in GWAS. It greatly boosts the number of single nucleotide polymorphisms (SNPs) tested in association studies and helps in discovering related loci[19]. Moreover, rare GWAS variants to be imputed are easier to be tagged than are common GWAS variants[20] and could increase the imputation quality for rare variants. In our study, pre-imputation QC was conducted before phasing and imputation. We eliminated SNPs with genotype call rate > 10%, 0.05 < MAF < 0.01 and Hardy-Weinberg equilibrium P < 1.00E−05. Then, we performed genotype imputation on the Michigan Imputation Server (Minimac4) with Haplotype Reference Consortium as reference panel (https://imputationserver.sph.umich.edu/index.html)[21]. The imputation quality metric Rsq (estimated R2, specific to each SNP) was considered an effective post-imputation filter. BCFTOOLS v1.9 (http://samtools.github.io/bcftools/bcftools.html) was used to filter imputed SNPs with R2 > 0.8. Post-imputation QC was performed again after genotype imputation.

Although the human exome represents less than 2% of the genome, it contains about 85% of the known disease-related variants[22]. The exon-transcribed sequence remains in mature RNA after the intron is removed by RNA splicing and can be expressed as a protein during protein biosynthesis[23]. To identify whether the variants in the exonic or splice region might cause protein coding changes and the amino acids that are affected by the mutations[24], we used ANNOVAR (http://annovar.openbioinformatics.org/en/latest/) for gene-based annotation. Therefore, only exonic and splicing (variant is within 2 bp of a splicing junction) variants were chosen from the results of ANNOVAR. The gene-based association study based on the selected exonic rare variants involved using the SKAT R-package v1.3.2. In this procedure, small-sample-adjusted SKAT null model was used for a binary phenotype correcting for the gender, APOE4 allele and age. Finally, significant SNP sets were obtained after controlling by family-wise error rate (FWER) = 0.05. We used the Minimum Achievable p-values (MAP), the lower limit of p-values for each variant set, to estimate the effective number of tests. MAP-adjusted and unadjusted quantile-quantile (Q-Q) plots were drawn.

2.3 Bioinformatics analysis

Several prediction tools were used to show the protein structure and verify the underlying biological interpretations of the identified gene. These bioinformatics databases and resources include NCBI (https://www.ncbi.nlm.nih.gov/gene/), genecards (https://www.genecards.org/), targetvalidation (https://www.targetvalidation.org/), varsome (https://varsome.com/), genepine (http://grch37.genepipe.ncgm.sinica.edu.tw/variowatch/), Noncode (http://www.noncode.org/), hosphosite (https://www.phosphosite.org/homeAction), gnomad (https://gnomad.broadinstitute.org/), UCSC Genome

Table 1. Demographics and clinical indicators of participants with Alzheimer's disease (AD) and cognitively normal (CN) controls.

category		AD	CN	p-value*
No. of participants, n(%)		175 (45%)	214 (55%)	
the gender, n(%)	M	93 (53%)	115 (54%)	0.9071
	F	82 (47%)	99 (46%)	
Age, years, mean (SD)		75.4 (7.4)	75.7 (4.9)	0.6620
APOE4 allele, n(%)	0	58 (33%)	156 (73%)	
	1	85 (49%)	53 (25%)	2.2E−16
	2	32 (18%)	5 (2%)	

[Note] * results of two independent-sample t tests for sex, age, and APOE4 allele.

Browser on Human (https://genome.ucsc.edu/), and Online Mendelian Inheritance in man (OMIM) (https://omim.org/entry/).

3 Results

3.1 Basic demographic features

Table 1 describes the characteristics of the AD and CN groups. 93 (53.1%) participants were male in the AD group, and 115 (54%) in the CN group. The average age in AD and CN groups was 75.4 (SD = 7.4) and 75.7 (SD = 4.9). Approximately 67% and 27% of the AD and CN groups carried at least one APOE4 allele.

3.2 Significant gene identified by ER-SKAT

After pre-imputation QC, 8,121 variants and 389 individuals were retained. We chose an Rsq threshold of 0.8 for SNPs with 0.005 < MAF < 0.01 after genotype imputation, which resulted in a total of 389 participants and 28511 variants passing post-imputation. In the annotation step, a total of 311 exonic SNPs in 254 gene sets were reserved for the statistical analysis, and ER-SKAT was used for analysis of 311 exonic variants in 254 gene sets with 389 samples. The MAP was determined for each gene. The effective number of tests for Bonferroni correction was 46 (alpha level 0.05) and the number of resampling was 1000. We identified the significant variant using the Bonferroni correction, that is, the p-value of the found variant rs2275303 in the SIPA1L2 gene is 6.00E−04, which is smaller than 0.05/46. Figure 1 shows MAP-adjusted and unadjusted quantile-quantile (Q-Q) plot of ER-SKAT-hybrid p-values from analysis of all exonic variants.

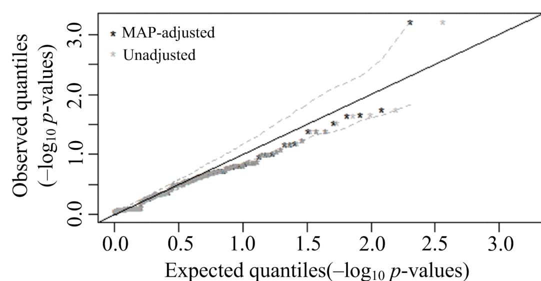

Figure 1. The x-axis is the MAP-adjusted or unadjusted expected quantile of $-\log_{10}$ p-values, and the y-axis is quantiles of $-\log_{10}$ p-values. Observed p-values are plotted against the MAP-adjusted expected quantiles (black dots) and unadjusted expected quantiles (gray dots). The dashed line represents a 95% confidence interval based on 1000 random draws from the MAP-based mixture distribution.

3.3 Protein structure and functions of SIPA1L2 derived from bioinformatics analysis

The protein primary structure of the SIPA1L2 gene is shown in Figure 2. The accession number for SIPA1l2 protein is NP_065859.3 (for human SIPA1l2). Structural analysis revealed a conserved RapGAP domain followed by a PDZ signaling domain. The PDZ domain may be responsible for specific protein-protein interactions, as most PDZ domains bind C-terminal polypeptides and bind internals (non-C-terminal). SPAR_C region is the C-terminal domain of SPAR protein.

We also found that the identified variant in SIPA1L2 is a missense mutation. The Gene Ontology annotations related to SIPA1L2 included GTPase activator activity. SIPA1L2 is suggested to be most abundant in granule cells of the dentate gyrus and cerebellum and show RapGAP activity for the small GTPases Rap1 and Rap2. The gene is involved in two related pathways in the Ras signaling pathway including the Ras and Rap1 signaling pathways (one KEGG pathway is the Rap1 signaling pathway, which belongs to Ras signaling pathway). In mammalian models, some evidence has implicated components of the Ras signaling pathway in the aging and metabolic

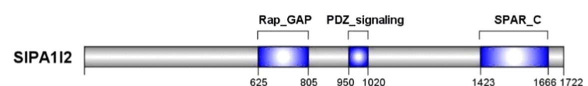

Figure 2. The protein primary structure of the SIPA1L2 gene.

regulation[25]. Cellular senescence is a feature of age-associated diseases, such as AD and osteoarthritis[26]. Oncogenic mutations in Ras result in its hyperactivation to trigger senescence[27]. From OMIM, SIPA1L2, a 1514-amino acid protein, shares significant similarity with SIPA1L1, and SIPA1L2 is moderately expressed in all adult and fetal tissues and specific adult brain regions. Target validation showed that on microarray analysis of six brain areas from AD patients and normal individuals, RNA expression was decreased in the posterior cingulate cortex of AD patients.

Alias for SIPA1L2 Gene is Spine-associated RapGAP 2 (SPAR2). SPAR2 is a novel GTPase activating protein (GAP) for the small GTPase Rap that shows significant sequence homology to SPAR. SPAR is a synaptic RapGAP, which has been reported to regulate the spine morphology in hippocampal neurons[28]. Overexpression of Rpph1 increased the density of dendritic spines in primary cultured hippocampal pyramidal neurons, while the knockdown of Rpph1 had the reverse effect, which may represent a compensatory mechanism at the early stage of the AD pathogenesis[29]. Like SPAR, SPAR2 interacts with the recently described synaptic scaffolding protein ProSAP interacting Protein (ProSAPiP), which in turn binds to the PDZ domain of ProSAP/Shank post-synaptic density protein. SPAR2 transcripts were mainly expressed in cerebellar and hippocampal granule cells. In a recent study, SPAR2 was identified as a brain-derived neurotrophic factor (BDNF) responder gene that is upregulated during cerebellar granulosa cell development[30]. In neurons, BDNF is transported in amphisomes, which signal locally at presynaptic boutons during retrograde transport to the soma. In hippocampal neurons, TrkB-signaling endosomes are actually amphisomes with local signaling capacity in the context of presynaptic plasticity during retrograde transport. The autophagosomal protein LC3 regulates RapGAP activity of SIPA1L2 and controls the retrograde transport and local signaling of TrkB[31]. Hence, SIPA1L2 is a potential rare risk variant for AD. After a series of bioinformatics analysis, the functions of SIPA1L2 gene in biomedicine were obtained, see Table 2.

4 Discussion

Although genotype imputation better enables the imputation of genotypes of rare variants and low-frequency variants from existing GWAS datasets, we analyzed rare variants with only $0.005 < MAF < 0.01$ in our study. There may not be sufficient shared reference individuals to predict those variants, which are not directly genotyped data in GWAS datasets. The imputation accuracy for variants decreases with decreasing MAF. Too-low imputation accuracy for very rare variants is one obstacle for the whole-genome rare-variant analysis. In addition, gene-based annotation is important for rare-variant GWAS and interpreting the final results. In our rare-variant study, we focused on exome regions because annotating the functional consequences of non-exome variants is still not developed. Also, the whole-genome rare-variants analysis is challenging because annotating the non-exome variants is difficult.

An increasing number of rare-variant association methods have been built depending on the underlying genetic architectures of complex traits. We used the ER-SKAT method to analyze the rare-variant association with AD based on integrating genomic and demographic data in our study. ER-SKAT aggregates variants within a specified region without considering the direction of the effect for individual variants, and it successfully solves the problem of linkage disequilibrium between genetic variations. Therefore, ER-SKAT is more suitable to detect associations in which both risk and protective variants or numerous non-causal variants are present. Lee et al. proposed an efficient resampling method based on ER-SKAT to calibrate single and gene-based rare-variant association analyses in case-control studies. This method can improve computational efficiency > 1000-fold and significantly shorten the calculation time over conventional resampling for low minor-allele-count variant sets. It ameliorates the conservativeness of results by using the mid-p-value and the estimated minimum achievable p-value for each test. By applying this method, we obtained the calibrated QQ plots and the number of effective tests to judge the significance and identified the significant gene SIPA1L2.

Table 2. Functions of SIPA1L2 gene.

Chromosome	SNP	Gene	Function
1	rs2275303	SIPA1L2(SPAR2)	1. Mutation in SIPA1L2 is a Missense mutation. 2. Gene Ontology annotations of SIPA1L2 gene include GTPase activator activity. 3. SIPA1L2 is a member of the SIPA1L family of neuronal RapGAPs. 4. SIPA1L2 is related with Ras signaling pathway. 5. SPAR has been reported to regulate spine morphology in hippocampal neurons. 6. SPAR2 is identified as a BDNF responder gene.

However, our research has some limitations. First, the identified gene has not been verified with subsequent molecular genetics experiments. Second, the number of samples in the ADNI-1 dataset is limited for the rare-variant association research.

In conclusion, we found that SIPA1L2 is significantly associated with AD. Our research has demonstrated that rare variants may disrupt the gene function of AD as well. Hence, determining the gene function of rare variants with full-genome resequencing is a promising area for the AD genomic research.

5 Declarations

Availability of data and material

The ADNI was initiated in 2003 by the U. S. National Institute on Aging, the U. S. National Institute of Biomedical Imaging and Bioengineering, the U. S. Food and Drug Administration (FDA), private pharmaceutical companies, and nonprofit organizations as a public-private partnership, led by the principal investigator Michael W. Weiner, MD. The primary goal of ADNI is to test whether serial MRI, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of the mild cognitive impairment and early Alzheimer's diseases.

Code availability

Our codes and software lists are available for free at https://github.com/ruijiali/SIPA1L2.

Acknowledgments

The work is supported by the National Natural Science Foundation of China (11801540), the Natural Science Foundation of Anhui (BJ2040170017), the Fundamental Research Funds for the Central Universities (WK2040000016) and the Science and Technology Program of Guangzhou (202002030129); the Science and Technology Planning Project of Guangdong Province (2017A010101030), the National Key Research and Development Program of China (2018YFC1315400) and the Third Medical Technology Projects of Shantou City in 2018 (ShanFuKe【2018】128-08).

Conflict of interest

The authors declare no conflict of interest.

Author information

WEN Canhong received the BS degree in 2009 from Hefei University of Technology and the PhD degree in 2015 from Sun Yat-sen University. Currently, she is an associate professor in the School of Data Science, and the Department of Statistics and Finance in the School of Management at the University of Science and Technology of China. Her main research interests include statistical learning, biostatistics, and image genetics.

TAN Haizhu (correspongding author) received the BS degree in 1999 from Shantou University of Computer Technology and the PhD degree in 2014 from Sun Yat-sen University. Currently, she is an associate professor in the Department of Preventive Medicine, Shantou University Medical College, China. Her main research interests include biostatistics, bioinformatics and image genetics.

References

[1] Verheijen J, Sleegers K. Understanding Alzheimer disease at the interface between genetics and transcriptomics. Trends in Genetics, 2018: 34(6): 434-447.

[2] Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nature Genetics, 2009, 41(10): 1088-1093.

[3] Pericak-Vance M A, Bebout J L, Gaskell P C, et al. Linkage studies in familial Alzheimer disease: Evidence for chromosome 19 linkage. American Journal of Human Genetics, 1991, 48(6): 1034-1050.

[4] Farrer L A, Cupples L A, Haines J L, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: A meta-analysis. Jama, 1997, 278(16): 1349-1356.

[5] Kok E H, Luoto T, Haikonen S, et al. CLU, CR1 and PICALM genes associate with Alzheimer's-related senile plaques. Alzheimer's Research & Therapy, 2011, 3(2): 12.

[6] Hollingworth P, Harold D, Sims R, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nature Genetics, 2011, 43(5): 429-435.

[7] Zuk O, Schaffner S F, Samocha K, et al. Searching for missing heritability: Designing rare variant association studies. Proceedings of the National Academy of Sciences, 2014, 111(4): E455-E464.

[8] Cruchaga C, Chakraverty S, Mayo K, et al. Rare variants in APP, PSEN1 and PSEN2 increase risk for AD in late-onset Alzheimer's disease families. PLOS ONE, 2012, 7(2): e31039.

[9] Reich D E, Lander E S. On the allelic spectrum of human disease. Trends in Genetics, 2001, 17(9): 502-510.

[10] Pritchard J K. Are rare variants responsible for susceptibility to complex diseases? The American Journal of Human Genetics, 2001, 69(1): 124-137.

[11] Manolio T A, Collins F S, Visscher P M, et al. Finding the missing heritability of complex diseases. Nature, 2009, 461(7265): 747-753.

[12] Coventry A, Bull-Otterson L M, Liu X, et al. Deep resequencing reveals excess rare recent variants consistent with explosive population growth. Nature Communications, 2010, 1(1): 1-6.

[13] Nelson M R, Wegmann D, Ehm M G, et al. An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. Science, 2012, 337(6090): 100-104.

[14] Nicolas G, Charbonnier C, Wallon D, et al. SORL1 rare variants: A major risk factor for familial early-onset Alzheimer's disease. Molecular Psychiatry, 2016, 21(6): 831-836.

[15] Bellenguez C, Charbonnier C, Grenier-Boley B, et al.

Contribution to Alzheimer's disease risk of rare variants in TREM2, SORL1, and ABCA7 in 1779 cases and 1273 controls. Neurobiology of Aging, 2017, 59:220. e1-220. e9.

[16] Rice T K, Schork N J, Rao D C. Methods for handling multiple testing. Advances in Genetics, 2008, 60(4): 293-308.

[17] Seunggeun L, Wu M C, Lin X. Optimal tests for rare variant effects in sequencing association studies. Biostatistics, 2012, 13(4): 762-775.

[18] Seunggeun L, Christian F, Sehee K, et al. An efficient resampling method for calibrating single and gene-based rare variant association analysis in case - control studies. Biostatistics, 2016, 17(1): 1-15.

[19] Das S, Forer L, Sidore C, et al. Next-generation genotype imputation service and methods. Nature Genetics, 2016, 48(10): 1284-1287.

[20] Thakker R V, Whyte M P, Eisman J A, et al. Genetics of Bone Biology and Skeletal Disease. Salt Lake City, UT: Academic Press, 2017.

[21] McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nature Genetics, 2016, 48(10): 1279-1283.

[22] Dijk E V, Auger H, Jaszczyszyn Y, et al. Ten years of next-generation sequencing technology. Trends in Genetics, 2014, 30(9): 418-426.

[23] Vrieze S I, Malone S M, Pankratz N, et al. Genetic associations of nonsynonymous exonic variants with psychophysiological endophenotypes. Psychophysiology, 2014, 51(12): 1300-1308.

[24] Wang K, Li M Y, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research, 2010, 38(16): e164.

[25] Spilker C, Sanhueza G, Böckers T M, et al. SPAR2, a novel SPAR-related protein with GAP activity for Rap1 and Rap2. Journal of Neurochemistry, 2008, 104(1): 187-201.

[26] Cathy S. Ras signaling in aging and metabolic regulation. Nutrition and Healthy Aging, 2017, 4(3): 195-205.

[27] Naylor R M, Baker D J, Deursen V. Senescent cells: A novel therapeutic target for aging and age-related diseases. Clinical Pharmacology & Therapeutics, 2013, 93(1): 105-116.

[28] Dimauro T, David G. Ras-induced senescence and its physiological relevance in cancer. Current Cancer Drug Targets, 2010, 10(8): 869-876.

SIPA1L2:阿尔茨海默症的风险基因

温灿红[1,2],李瑞佳[1],蔡嘉辉[3],谭海珠[3*]

1. 中国科学技术大学大数据学院,安徽合肥 230027;

2. 中国科学技术大学管理学院统计与金融系,安徽合肥 230026;

3. 汕头大学医学院预防医学教研室,广东汕头 515031

*通讯作者. E-mail:linnanqia@126.com

摘要: 阿尔茨海默症(AD)是一种常见的具有高度遗传性的神经退行性疾病.越来越多的常见变异被发现与AD相关,但这些常见变异只能解释遗传力的一小部分.理论和实践表明罕见变异可解释剩余的遗传力.我们在389个个体(175个AD和214个认知正常)的600470个变异中探索能够改变AD易感性的罕见功能性变异.首先,在密歇根填补服务器上对缺失的基因型进行填补后,进行了质量控制和基于基因的注释.其次,对311个注释的外显子变异进行有效的重采样序列核关联测试.最后,通过几种生信工具对鉴定出的风险基因的潜在生物学解释进行了预测.结果显示,在Bonferroni校正下,SIPA1L2基因中的罕见错义变异rs2275303与AD显著相关(P=6.00E-04),并通过生信分析对其致病性进行了验证.SIPA1L2基因有望在AD的预防、诊断、预后和治疗中发挥重要作用.

关键词: 阿尔茨海默症;罕见变异;关联分析;有效重抽样序列核关联测试