Unraveling the genetic architecture of blood unfolded p-53 among non-demented elderlies: novel candidate genes for early Alzheimer's disease

Arash Yaghoobi[1] and Seyed Amir Malekpour[1*]

Abstract

Background Alzheimer's disease (AD) is a heritable neurodegenerative disease whose long asymptomatic phase makes the early diagnosis of it pivotal. Blood U-p53 has recently emerged as a superior predictive biomarker for AD in the early stages. We hypothesized that genetic variants associated with blood U-p53 could reveal novel loci and pathways involved in the early stages of AD.

Results We performed a blood U-p53 Genome-wide association study (GWAS) on 484 healthy and mild cognitively impaired subjects from the ADNI cohort using 612,843 Single nucleotide polymorphisms (SNPs). We performed a pathway analysis and prioritized candidate genes using an AD single-cell gene program. We fine-mapped the intergenic SNPs by leveraging a cell-type-specific enhancer-to-gene linking strategy using a brain single-cell multimodal dataset. We validated the candidate genes in an independent brain single-cell RNA-seq and the ADNI blood transcriptome datasets. The rs279686 between AASS and FEZF1 genes was the most significant SNP (*p*-value $= 4.82 \times 10^{-7}$). Suggestive pathways were related to the immune and nervous systems. Twenty-three candidate genes were prioritized at 27 suggestive loci. Fine-mapping of 5 intergenic loci yielded nine cell-specific candidate genes. Finally, 15 genes were validated in the independent single-cell RNA-seq dataset, and five were validated in the ADNI blood transcriptome dataset.

Conclusions We underlined the importance of performing a GWAS on an early-stage biomarker of AD and leveraging functional omics datasets for pinpointing causal genes in AD. Our study prioritized nine genes (SORCS1, KIF5C, TMEFF2, TMEM63C, HLA-E, ATAT1, TUBB, ARID1B, and RUNX1) strongly implicated in the early stages of AD.

Keywords Alzheimer's disease, U-p53, GWAS, SNP, Single-cell RNA-seq, Single-cell ATAC-seq

Introduction

Dementia has a prevalence of about 7.1% among Europeans [1]. The most common type of dementia (50–70% of cases) is Alzheimer's disease (AD) [2]. AD is a progressive neurodegenerative disease whose asymptomatic phase begins decades before the onset of symptoms [3, 4]. AD is a multifactorial disease, and several risk factors, such as cardiovascular diseases, oral infection, sleep disturbance, aging, genetic susceptibility, traumatic brain injuries, and air pollution, have been proposed for it [5]. Regarding the

*Correspondence:
Seyed Amir Malekpour
a.malekpour@ipm.ir
[1] School of Biological Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran 19395-5746, Iran

beginning of the pathological process years before the onset of AD symptoms, early diagnosis of people at high risk for AD is pivotal.

Although the exact pathophysiological mechanisms of AD are not fully understood, the Amyloid-beta (Aβ) hypothesis has been the dominant one in describing AD pathogenesis. According to this hypothesis, aging causes excessive accumulation of Aβ peptide in the brain which eventually leads to the accumulation of toxic brain plaques and the onset of symptoms of AD. Genetic studies have also shown that in rare cases of AD, there are mutations in genes related to Aβ metabolism, strengthening the Aβ hypothesis [6]. However, this hypothesis has been subjected to a closer examination in recent years. Indeed, it has been questioned to some extent because new drugs based on antibodies against Aβ have poorly improved the symptoms of patients with mild and moderate forms of AD [7–9]. Therefore, there may be other biological pathways involved in the early stages of AD.

Based on the observed drawbacks of the Aβ hypothesis, various models have been proposed to describe the underlying mechanisms of AD. One of the proposed mechanisms for AD is a defect in the antioxidant response system and the role of the p53 protein in this process [10]. In addition to playing a role in tumor suppression in the body, the p53 protein can prevent the progress of neurodegeneration through various pathways, such as synaptic function, neurite outgrowth, protection from oxidative agents, and axonal regeneration [11–14]. For the first time in 2002, there were reports of observing a conformationally changed structure of p53 protein in the peripheral cells of patients with AD [15]. This protein, called unfolded p53 (U-p53), was not seen in other diseases, such as cancer and other neurodegenerative diseases, and seemed specific to AD. In the following years, it was suggested that the oxidative agents, chronic subtoxic oxidative stress, and sublethal Aβ concentrations may increase the expression of the U-p53 protein [16, 17]. It is worth noting that brain is an essential source of oxidative agents' production since it consumes more than 20% of the body's oxygen [18–20]. The rise of oxidative substances in the body causes various alterations in post-translational modification, the tertiary structure, and physiological functions of the p53 protein [15]. Moreover, the Aβ peptide can also cause similar changes in the p53 structure [21–23]. Conversely, the p53 protein can also affect Aβ concentration [24]. Additionally, oxidative agents can also lead to the phosphorylation of tau protein, which is one of the pathological hallmarks of AD. Hence, pathways related to oxidative stress could play key roles in the early stages of neurodegeneration [25, 26].

In recent years, there has been a growing interest in developing blood biomarkers for the early diagnosis of AD [27]. Developing biomarkers for AD has two main advantages: First, aiding in early diagnosis of the disease; Second, helping to develop new drugs for AD. The second advantage is precious since the development of effective drugs for AD has slowed, and the available drugs mostly have little effect on improving or delaying the symptoms of this disease [28]. However, most AD biomarkers have not yet entered the official diagnostic protocols of the disease [29]. Recently, an antibody has been developed to identify the blood U-p53 protein, which had a promising performance in identifying people at high risk for AD in the asymptomatic and symptomatic stages of the disease [30]. In a recent study, the blood U-p53 protein has been proposed as a robust biomarker outperforming other well-known AD biomarkers, such as Aβ PET scan, for identifying people at risk of developing AD 6 years before the onset of symptoms (AUC > 98%) [14].

According to several studies, the heritability of AD is estimated between 60–80% [31]. So far, multiple genome-wide association studies (GWAS) have been performed on AD, which have led to the discovery of numerous single nucleotide polymorphisms (SNPs) and genes [32]. However, the exact mechanism of the disease remains unknown. GWAS studies focusing on AD as a phenotype encounter numerous challenges due to two primary reasons: First, the clinical diagnostic accuracy of AD is poor. Second, due to the influence of age on AD, it is recommended to utilize age-matched control groups in GWASs conducted on this disease. However, the feasibility of obtaining such control groups is often limited [33]. To overcome these challenges, an alternative approach involves conducting GWAS on early-stage biomarkers of AD and utilizing these biomarkers as endophenotypes. This approach allows for the identification of novel genes and SNPs associated with the biological pathways underlying the early stages of the disease [34]. Compared to traditional GWASs performed on binary phenotypes, this approach has several advantages: First, continuous quantitative phenotypes, such as U-p53 concentration, represent real laboratory measurements rather than subjective binary clinical diagnoses. Second, this approach enhances the statistical power of GWAS even with smaller sample sizes compared to studies focusing on binary phenotypes. Third, early-stage endophenotypes could capture slight differences in susceptibility to disease that could not be detected with clinical binary phenotypes. So far, several GWASs have been conducted on various AD biomarkers, such as Aβ, tau protein, phosphorylated tau protein, and sTREM concentrations of cerebrospinal fluid (CSF), which has led to the discovery of novel genes and pathways involved in AD [34, 35].

Considering the role of the U-p53 protein in the early stages of AD and the high performance of this biomarker

in predicting AD in the asymptomatic and early stages of this disease, we decided to conduct a GWAS on this biomarker as one of the AD endophenotypes to reveal novel genes and biological pathways involved in the early stages of AD. Our primary hypothesis was that the SNPs associated with blood U-p53 variation among older adults could shed light on novel loci involved in pathways related to the early stages of AD. By employing single-cell RNA-seq and multimodal datasets, as well as leveraging a single-cell enhancer-to-gene mapping strategy, we were able to establish connections between non-coding SNPs identified in our GWAS and their corresponding cell-specific target genes. This approach enabled us to gain deeper insights into the specific genes within brain cells that are causally implicated in the pathogenesis of AD.

Methods

Study design
We used the ADNI cohort for our research. In summary, ADNI is a cohort launched by Michael W. Weiner, MD, in 2003. The primary goal of this cohort is to gather various types of biomarkers, including body fluid biomarkers, neuroimaging, and genetic data, for developing tools for early diagnosis of AD. This cohort includes 3 phases: ADNI1, ADNIGO/2, and ADNI3. All samples have been collected using the relevant guidelines. A consent form has been obtained from all the cohort participants, and the regional ethics committees have also approved this cohort. The details of this cohort can be found on its website (http://adni.loni.usc.edu/). By personal communications, we had permission to access the database via the https://ida.loni.usc.edu/.

Participants
To carry out our GWAS, we first used the file "DIA-DEM_V2_08_30_22.csv" which contains a dataset of U-p53 concentration in the blood plasma of 584 subjects from the ADNI GO/2 phase. The genotypic data of the ADNI GO/2 phase participants are available in 2 distinct PLINK files ("ADNI_GO_2_OmniExpress.zip" and "ADNI_GO2_2nd.tar.gz files"). The first file has the genotypic data of 432 subjects and the second one consists of the genotypic data of 361 subjects. In the next step, we merged these two files using the PLINK 1.9 software (https://www.coggenomics.org/plink/) which yielded a single file containing genotypic data of 793 subjects. We also used the "ADNIMERGE.csv" file which contains demographic data of the participants of the ADNI 1, GO/2, and 3 phases.

Quality control and merging
To perform the GWAS, we conducted several stages of quality control on the samples with available genotypic information. We utilized ADNI GO/2 subjects genotyped by the Illumina HumanOmniExpress BeadChip in 2014. The genomic quality control steps were performed using the PLINK 1.9 software. First, we excluded participants with a missing genotype rate exceeding 5% and a heterozygosity rate deviating more than three standard deviations from the mean. Additionally, we assessed the concordance between participants' ascertained gender and the estimated inbreeding coefficient calculated by the PLINK software to identify potential gender mismatches. Next, we merged the output file with our phenotype file to extract subjects with both genotypic and phenotypic data. Subsequently, we excluded SNPs located on the sex chromosomes. SNPs with a minor allele frequency (MAF) less than 2%, a missing genotype rate exceeding 5%, and a Hardy–Weinberg equilibrium (HWE) *p*-value less than 10^{-6} were also excluded.

Phenotype
Blood samples of 593 participants from the ADNI cohort were analyzed using the AlzoSure® Predict method to measure the blood plasma concentration of U-p53 protein, from June 16 to July 12, 2022, in more than eight rounds. In summary, the AlzoSure® Predict method is a diagnostic laboratory method that uses LC–MS/MS or Ion Trap on protein-depleted plasma samples. This method is particularly relevant for individuals over 50 years old with a family history of AD, a documented genetic predisposition, and signs of mild cognitive impairments (MCI). It aims to predict the progression from MCI to AD. During each round of analysis, approximately 54–77 samples from ADNI subjects were included. Additionally, at least three quality control samples from the Australian Imaging, Biomarkers and Lifestyle (AIBL) biobank, previously analyzed and covering the full range of U-p53 protein expression during the early stages of AD progression, were included as well. Out of the 593 samples, nine samples were discarded from the analysis despite undergoing two rounds of analysis due to the inability to obtain results. Detailed information regarding the methodology of laboratory tests and quality control procedures can be found on the ADNI website.

Genome-wide association analysis
To account for population substructures as a potential confounding factor, we utilized the PLINK software to compute the principal components (PCs) of the genotypic data. The top 3 PCs were selected and incorporated as covariates in the analysis. This approach enabled us to control for genetic relatedness and address any potential biases introduced by population substructures [36]. To perform GWAS analysis, we used the linear regression

model with the "–assoc –linear" commands in the PLINK 1.9 software. Also, we used age and gender as two other covariates. To calculate the significant and suggestive *p*-value thresholds, we used the "–indep –pairwise" command in the PLINK software (window size = 50 Kb, Linkage disequilibrium (LD) $r^2 > 0.8$, and step size = 5) to calculate the number of independent SNPs. To evaluate the degree of genomic inflation for adjusting population substructures, we computed the λ-statistic. We also highlighted the observed versus expected *p*-values in the Q-Q plot using the qqman package in R. Power calculation was performed using the "genpwr" package in R. According to our calculation, our study has 80% power at the alpha = 2×10^{-7} (the significant threshold of our GWAS) to detect effect sizes of 0.14 and 0.32 when the MAF of SNPs is 0.50 and 0.05, respectively. Note that power is the probability of avoiding a type II error or false negative predictions, in which we are especially interested.

Pathway analysis

We used the Pathway Scoring Algorithm (Pascal) software developed by David Lamparter et al. to perform pathway analysis using our GWAS summary statistic. In short, Pascal is fed with the SNP *p*-values derived from the GWAS analysis. We selected the "sum" option which takes the average of association signals (*p*-values) of all SNPs around ± 50 Kb of each gene. The maximum number of SNPs considered for averaging the chi-squared statistics of each gene was set to 3000. We set the MAF cut-off value to 0.01. We also used the European population of the 1000 Genome Project phase 3 for calculating LD. The merge distance option was set to 1. More details about the Pascal methodology have been described elsewhere [37]. We explored the KEGG, Biocarta, and Reactome pathway databases, to identify the biological pathways that are significantly enriched with the SNP-associated genes.

Positional mapping of coding SNPs and step 1 fine mapping of intergenic SNPs using an AD single-cell gene program

Coding SNPs were simply annotated to their respective residing genes. For mapping intergenic SNPs, we employed a 2-step approach. At the step1, we first implemented a simple positional mapping approach by annotating intergenic SNPs to two nearest genes on the left and right sides of SNPs. Then, to produce a more AD-related comprehensive list of risk genes, we used an AD single-cell gene program constructed by Karthik A. Jagadeesh et.al. [38]. They used a single-nucleolus RNA-seq of the prefrontal cortex of 48 healthy and AD subjects from the ROS-MAP cohort omics project to construct this gene program. In summary, they computed a

gene-level nonparametric Wilcoxon's rank-sum statistic to identify the differentially expressed genes (DEGs), between cells from healthy and AD tissues for each cell type. Then, the gene weighting in the AD gene program was determined by transforming the gene *p*-values for each cell type. This transformation involved converting the *p*-values to X = −2log(*p*-value), which follows a χ2 distribution. The transformed values were then normalized to a grade between 0 and 1 using the min/max normalization formula g = (X − min(X))/(max(X) − min(X)). This process resulted in a relative weighting of genes in the AD gene program.

To associate our significant intergenic SNPs with nearby genes, we utilized the UCSC human genome build 38 and the MAGMA software (https://ctg.cncr.nl/software/magma). We annotated the SNPs to genes within a ± 500 Kb window. From the generated list of genes, we specifically selected those that overlapped with weights greater than 0.30 in any of the AD single-cell gene program for each cell type. These selected genes were then used for the subsequent fine mapping in step 2.

Exploring SNPs with high LD at intergenic genomic risk loci

In order to identify all the SNPs in LD with our suggestive and significant intergenic SNPs, we used the FUMA software (https://fuma.ctglab.nl/) and NIH LDlink webtool (https://ldlink.nih.gov/?tab=ldproxy). In FUMA, the SNP2GENE function was utilized to identify SNPs in LD with our suggestive and significant intergenic SNPs of interest. Additionally, the LDlink webtool was utilized to assess LD using the European population of the 1000 Genome Project phase 3 as the reference genome panel. A threshold of $r^2 > 0.6$ was chosen to determine high LD for SNPs, except for the major histocompatibility complex (MHC) region on chromosome 6, where a higher threshold of $r^2 > 0.9$ was adopted due to its complex LD structure.

Step 2 fine mapping of intergenic SNPs using an AD and healthy brain tissue multimodal dataset

In order to pinpoint potential risk genes more accurately, we sought to link our non-coding significant SNPs to their cell-type specific target genes using a brain tissue multimodal dataset. First, we downloaded a multimodal dataset isolated from post-mortem dorsolateral prefrontal cortex (DLPFC) tissues of seven AD and eight healthy controls from the GEO database (GSE214637). This dataset has simultaneously profiled the chromatin accessibility (single-cell assay for transposase-accessible chromatin with sequencing, or ATAC-seq) and gene expression (single-cell RNA-seq) from the same cells. More details about the samples of the original study can be found elsewhere [39]. We used the Seurat package to create an object containing

both single-cell ATAC-seq and RNA-seq sparse matrices. In the next step, we used a recently introduced Single-cell Enhancer Target gene mapping (SCENT) method to link our significant intergenic SNPs to their target genes. In summary, the SCENT method applies the Poisson regression to model the effect of chromatin accessibility on gene expression in order to identify potential enhanc-

we performed a linear regression analysis to examine the relationship between the concentration of U-p53 (dependent variable) and the expression of our candidate genes (independent variables), in blood. Age and gender were included as covariates in the analysis. The regression model can be represented as follows:

$$\text{U}-\text{p53 concentration} \sim \alpha + \beta 1 \times \text{candidate gene expression} + \beta 2 \times \text{age} + \beta 3 \times \text{sex} \tag{1}$$

ers and silencers for each gene in a cell-type specific manner. More details about the SCENT package can be found elsewhere [40]. Our primary hypothesis was that our significant non-coding SNPs or their proxy SNPs might be located in enhancers or silencers and regulate the expression of potential risk genes. So, we first explored the intersection of position of our significant SNPs or their proxies with ATAC-seq peaks using the UCSC human genome build 38 version. Then at each intersected region, we used the SCENT package to investigate the potential association of that ATAC-seq peak region signals with the expression level of candidate genes produced in the step 1 fine mapping. We also used log-normalized RNA counts, percentage of mitochondrial reads in each cell, age, and gender as covariates. The *p*-value less than 0.05 was chosen as the threshold indicating significant results.

Validating differential expression of final mapped genes between AD and healthy controls using an independent single-cell RNA-seq dataset

We obtained a single-cell RNA-seq dataset from an independent cohort with accession number GSE174367. The dataset consists of nuclei isolated from the prefrontal cortex of post-mortem brain tissues, including 11 late-stage AD samples and seven age-matched healthy control samples. The original study by Samuel Morabito et al. [41] provides complete details on this dataset. Briefly, the MAST package (v1.12.0) was applied to identify DEGs between AD and control samples in this single-cell RNA-seq dataset. The resulting DEG list was generated, with *p*-values adjusted using the Bonferroni correction method. Supplementary data 1 of Samuel Morabito et al. [41] also provides the DEG list between AD and control samples. We proceeded to identify the genes that were common between the DEGs from Samuel Morabito et al. [41] and our final mapped genes obtained from the previous steps.

Validating association of the mapped genes with blood U-p53 at transcription level using ADNI whole blood transcriptomics

To validate some of the potential genes associated with the concentration of U-p53 and possibly early AD,

For each candidate gene, Eq. (1) was fitted to the U-p53 concentration and candidate gene expression data obtained from the ADNI dataset. To achieve this, we downloaded the "Microarray_Gene_Expression_Profile_Data.csv" file from the ADNI cohort which contains microarray transcriptome of 811 participants from the ADNI 1 and GO/2 phases. Briefly, the transcriptome profiling of 811 participants was performed at Bristol-Myers Squibb (BMS) laboratory using the peripheral blood samples. The Affymetrix Human Genome U219 Array (www.affymetrix.com) was used to determine the transcriptome profile. More details about the quality control steps and laboratory methods can be found on the ADNI website. When fitting Eq. (1) to the gene expressions from ADNI, the Wald statistic was calculated for the β1 as the coefficient of the candidate gene expression. The *p*-value, derived from the Wald statistic, was then used to assess the significance of the association, with a threshold of *p*-value less than 0.01 considered as indicative of significance.

The all steps performed in our study are graphically depicted in Fig. 1.

Results

Quality control and genome wide association results

Merging two genotypic files of the ADNIGO/2 phase yielded genotypic data of 793 subjects (425 men and 368 women) on 730,525 SNPs. Twenty participants were excluded due to a heterozygosity rate of more than three standard deviations from the mean. There was no discordance between the subject's ascertained gender information with gender estimation based on genotypes. Missing genotype rates of all participants were less than 5%. Among 773 remaining subjects, 484 were kept since they had both genotypic and blood U-p53 data. 79,777 SNPs with MAF of less than 2% were excluded. 16,483 SNPs were excluded due to a missing genotype rate of more than 5%. 255 SNPs with Hardy–Weinberg *p*-value less than 10^{-6} were also excluded. 21,167 SNPs on sex chromosomes were also excluded. Finally, the genotypic data of 612,843 SNPs among 484 participants passed all the quality control steps. The total genotypic rate was
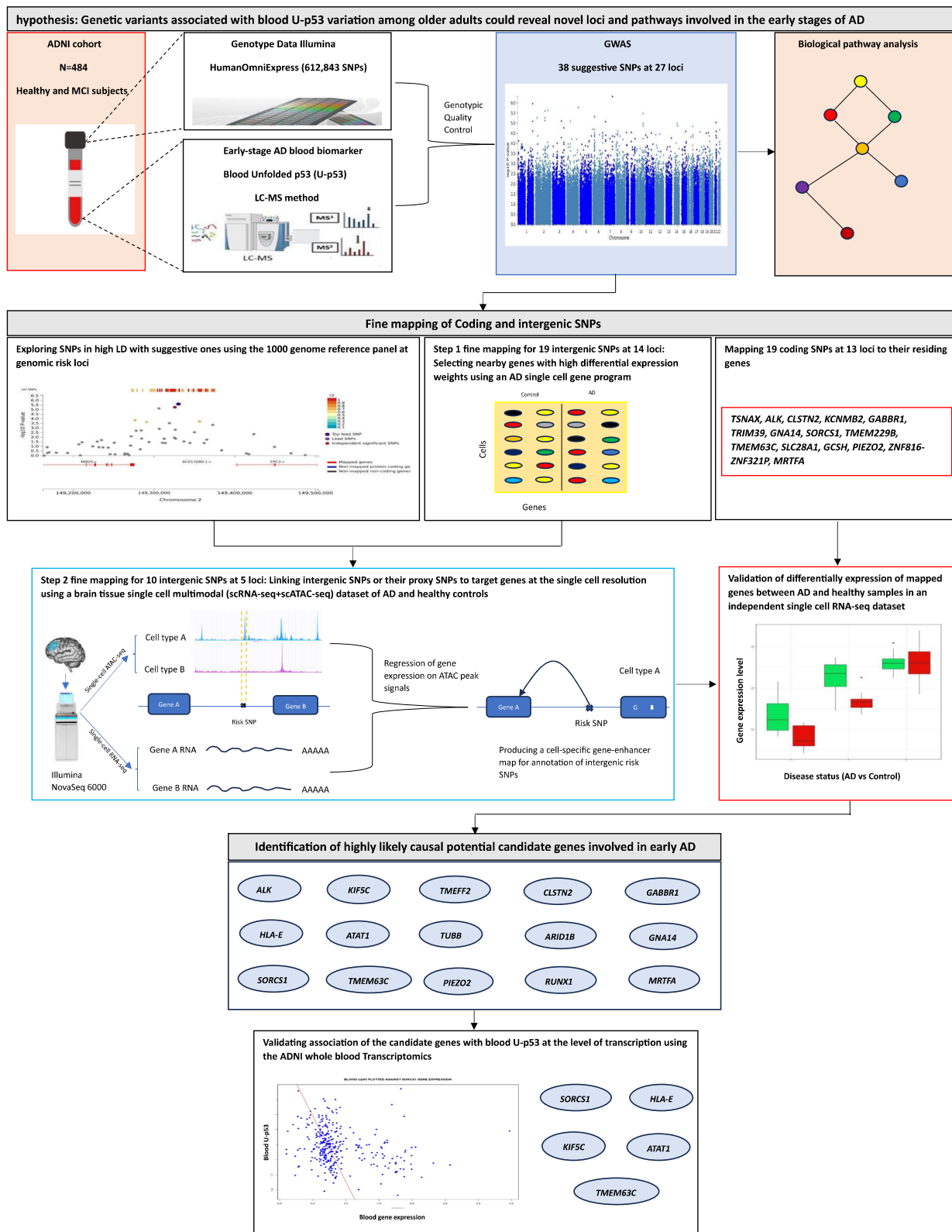
Fig. 1 Graphical depiction of all of the steps performed in the study. The figure shows the GWAS pipeline used in our study. The upper part of the figure shows the datasets, GWAS and pathway analyses. The middle part of the pipeline shows the steps for positional and functional fine-mapping of SNPs using brain single-cell RNA-seq and multimodal datasets. Lastly, the lower part of the figure shows the validated candidate genes in the independent single-cell RNA-seq and ADNI blood microarray transcriptome datasets

0.9986. Among the 484 participants of our study, there were 190 cognitively normal (CN) subjects (121 healthy subjects and 69 subjects with subjective memory complaints) and 363 patients with MCI (200 subjects with early mild cognitive impairments (EMCI) and 94 subjects with late mild cognitive impairments (LMCI)). The mean age of our study participants was 71.78 (±6.97). Finally, GWAS was performed on 484 subjects (260 male and 224 female). The λ-statistic indicating the degree of genomic inflation due to population structures was low in our analysis (λ=0.99184). The Manhattan and Q-Q plot of blood U-p53 GWAS are shown in Fig. 2. To check the robustness of our GWAS results, we performed linear mixed modeling to account for genetic relatedness using the Genome-wide Complex Trait Analysis (GCTA) software [42] (https://yanglab.westlake.edu.cn/software/gcta/#Overview), following the approach suggested by Junhao Wen et al. in their study [43]. The results obtained from the GCTA software align accurately with our final reported results obtained from the PLINK software. The GWAS summary statistics obtained from both software are provided in the Supplementary File 1.

We retained the independent SNPs by excluding 188,013 dependent SNPs (LD $R^2 > 0.8$ at 50 Kb window) from the total SNPs. Then, our GWAS's significant and suggestive thresholds were calculated based on the number of independent SNPs (424,830 SNPs). So, we chose 2×10^{-7} (~0.05/424,830) and 3×10^{-5} (~$200 \times 0.05/424,830$) as significant and suggestive *p*-value thresholds for our GWAS.

Among 612,843 SNPs, 38 were considered suggestive SNPs ($2 \times 10^{-7} < p\text{-value} < 3 \times 10^{-5}$). These 38 suggestive SNPs were located at 27 genomic risk loci. None of the SNPs reached the significant threshold level (*p*-value < 2×10^{-7}). The rs279686 had the smallest *p*-value and nearly reached the significant threshold (*p*-value = 4.82×10^{-7}). Among 38 suggestive SNPs, 19 were intergenic.

Pathway analysis results
The GWAS results provided genome-wide SNP *p*-values, which were subsequently used in Pascal package for gene scoring and pathway enrichment analysis. Pascal aggregates SNP signals within ±50 Kb of each gene and employs these per-gene aggregated signals for the pathway enrichment analysis. In brief, we examined the KEGG, Biocarta, and Reactome pathways to identify significant enrichments of genes associated with SNPs. For more information see Methods. Our pathway analysis revealed six suggestive biological pathways associated with U-p53, with *p*-values below 0.01, as plotted in Fig. 3. Top biological pathway was REACTOME_NFKB_ACTIVATION_THROUGH_FADD_RIP1_PATHWAY_MEDIATED_BY_CASPASE_8_AND10

(*p*-value = 7.16×10^{-4}). The other 5 suggestive pathways were REACTOME_NCAM_SIGNALING_FOR_NEURITE_OUT_GROWTH (*p*-value = 4.32×10^{-3}), BIOCARTA_IL17_PATHWAY (*p*-value = 5.47×10^{-3}), REACTOME_NCAM1_INTERACTIONS (*p*-value = 6.85×10^{-3}), BIOCARTA_PTEN_PATHWAY (*p*-value = 7.27×10^{-3}), and REACTOME_AXON_GUIDANCE (*p*-value = 9.46×10^{-3}).

Positional mapping of coding SNPs and step 1 fine mapping results
First positional mapping of coding SNPs was performed based on the human genome hg38 version, which resulted in mapping 15 genes. In the next step, positional mapping and functional gene mapping of intergenic SNPs using the AD single-cell gene program yielded 36 genes in various brain cell types. The results of all suggestive SNPs and their step 1 mapped genes are summarized in Table 1.

Step 2 fine mapping results
None of our suggestive intergenic SNPs were located in ATAC peaks, but using LD information at the suggestive intergenic loci, we could identify seven loci out of 14 suggestive loci that contain at least one SNP that is both in LD with the suggestive SNPs ($r^2 > 0.6$) and is located in ATAC peaks of the multimodal dataset used in our analysis. At each seven loci, we investigated the association of each identified ATAC peak with the expression level of the candidate genes obtained from step 1 fine mapping, at each eight brain cell types using the SCENT method. At five loci out of the seven, we could map the suggestive noncoding SNPs to their putative target risk genes in five brain cell types. The fine-mapped genes and their target cell-type resulted from this step were: *EPC2* (*p*-value = 0.01 in Microglial cells and 3.62×10^{-6} in Oligodendrocyte progenitor cells (OPCs)), *KIF5C* (*p*-value = 8.94×10^{-3} in OPCs), *TMEFF2* (*p*-value = 1.33×10^{-11} in excitatory neurons), *ARID1B* (*p*-value = 0.002 in excitatory neurons), *HLA-E* (*p*-value = 0.03 in Microglial cells), *ATAT1* (*p*-value = 0.01 in excitatory neurons), *TRIM26* (*p*-value = 0.02 in excitatory neurons), *TUBB* (*p*-value = 0.001 in OPCs, 0.01 in Astrocytes, and 0.02 in inhibitory neurons), and *RUNX1* (*p*-value = 0.002 and 0.03 in Microglial cells). At loci 7 and 16, expression of no candidate genes was significantly associated with the ATAC peak signal of that region. The results of the significant genes with their related cell types are summarized in Table 2.

Validating differential expression of final mapped genes between AD and healthy controls in an independent single-cell RNA-seq dataset results
There were 15 genes common between our final 23 mapped genes and the DEG list obtained from the Morabito et al. [41] dataset. These 15 genes included: *ALK*,

A.

B.
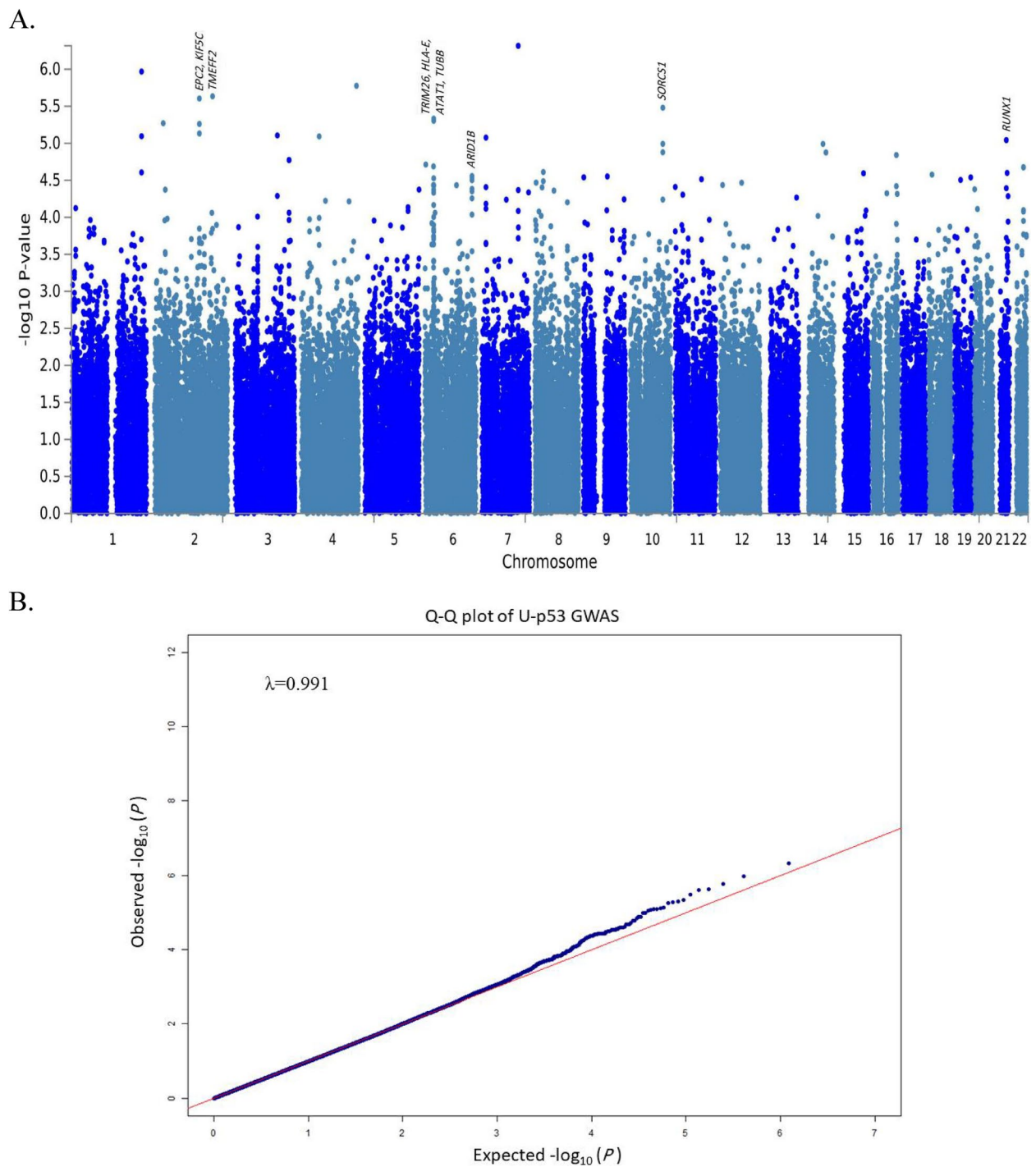
Fig. 2 Manhattan and Q-Q plot of the blood U-p53 GWAS. **A** Manhattan plot of the blood U-p53 GWAS. Only six suggestive loci with strong evidence of involvement in AD pathogenesis in our analyses are shown. **B** Q-Q plot of the blood U-p53 GWAS results showing low genomic inflation due to the population substructures ($\lambda = 0.99184$)

KIF5C, TMEFF2, CLSTN2, GABBR1, HLA-E, ATAT1, TUBB, ARID1B, GNA14, SORCS1, TMEM63C, PIEZO2, RUNX1, and *MRTFA.*

Validating association of the mapped genes with blood U-p53 at the transcription level in ADNI whole blood transcriptome dataset results

By merging ADNI whole blood microarray transcriptome

Fig. 3 Bar plot for the −log10 of the *p*-value of selected pathways that are enriched with SNP-associated genes. A −log10(*p*-value) greater than 2 corresponds to pathways that are significantly enriched, with corresponding *p*-values less than 0.01, as determined using the Pascal package
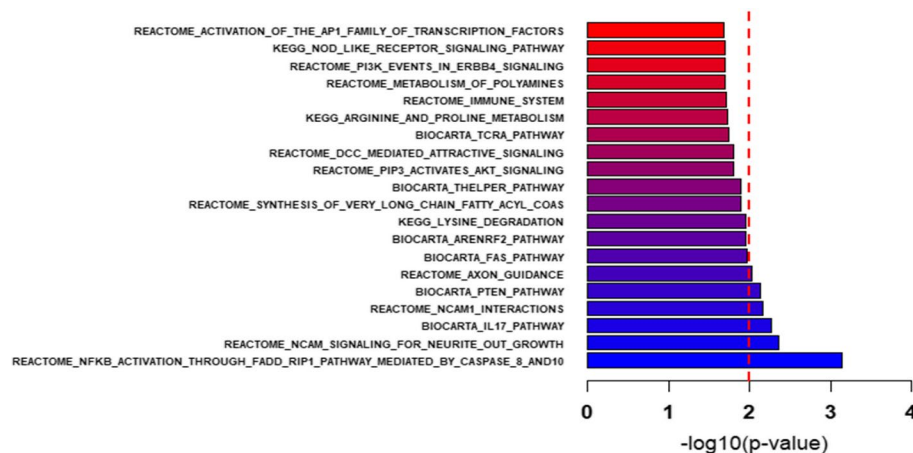
dataset with the U-p53 data, 322 subjects had both blood transcriptome and the blood U-p53 data. All probes of the 15 genes associated with blood U-p53 obtained from previous mapping steps were derived from ADNI whole blood transcriptome microarray dataset to investigate their possible association of expression level with blood U-p53. Among these genes, five genes were replicated at the blood transcript level. These 5 genes were *SORCS1* (*p*-value = 2.89×10^{-5}), *HLA-E* (*p*-value = 4.37×10^{-4}), *ATAT1* (*p*-value = 2.20×10^{-3}), *KIF5C* (*p*-value = 2.38×10^{-3}), and *TMEM63C* (*p*-value = 9.93×10^{-3}).

Discussion

To our knowledge, this is the first GWAS on blood U-p53 as one of AD's novel, specific and sensitive biomarkers. Using a single-cell multimodal and two single-cell RNA-seq datasets for fine mapping of risk SNPs, we prioritized *KIF5C*, *EPC2*, *TMEFF2*, *HLA-E*, *ATAT1*, *TUBB*, *ARID1B*, and *RUNX1* as potential risk genes with strong evidence of involvement in the early stages of AD. We also suggested the *SORCS1* gene as the other important gene involved in AD, with robust evidence at genotypic and blood transcript levels. Aβ processing (*SORCS1* and *TMEFF2*) and axonal transportation (*KIF5C*, *TUBB*, *ATAT1*) were two potential common biological pathways among the identified genes. A graphical illustration of results of multiple analyses sources across three loci are shown in Fig. 4.

We identified six suggestive biological pathways associated with blood U-p53 and possibly involved in AD. Two of them (NFKB activation through FADD and RIP1 pathway mediated by CASPASE 8 and 10 and IL17 pathway) highlight the importance of immune system involvement in AD. Indeed, the NFKB is an inflammatory

transcription factor that has important roles in AD pathogenesis [44]. Interestingly, another two suggestive pathways were related to the nervous system (NCAM signaling for neurite outgrowth and Axon guidance). We identified three SNPs at locus 18, residing in the *SORCS1* gene. Additionally, there was a strong and inverse association between the *SORCS1* gene blood expression and blood U-p53 (*p*-value = 2.89×10^{-5}). *SORCS1* belongs to the Sortilin family of vacuolar protein sorting-10 (Vps10) domain-containing proteins. The *SORL1* gene, one of the other genes of this family, is one of the most famous and replicated loci in multiple AD GWASs to date [45]. Wei Xu et al. suggested that rs10884402 and rs950809 in intron 1 of the *SORCS1* gene were associated with late-onset AD in the Chinese Han population [46]. It has been shown that the *SORCS1* gene alters the Amyloid beta precursor protein (APP) processing [47]. It also has been shown that overexpression of the *SORCS1* gene in cultured cells lowers Aβ generation [48]. Recently, Alfred Kihoon Lee et al. proposed an important role for *SORCS1* in rescuing Aβ oligomers-induced Neurexin dysfunction and synaptic pathology [49].

The rs4853647 SNP at locus four was significantly associated with blood U-p53 in our study. This SNP resides between *CAVIN2* and *TMEFF2* genes. By using enhancer-to-gene linking strategy, it was revealed that this region was strongly associated with the *TMEFF2* gene expression in excitatory neurons (*p*-value = 1.33×10^{-11}). *TMEFF2* (Transmembrane Protein With EGF Like And Two Follistatin Like Domains 2) gene belongs to the tomoregulin family of transmembrane proteins. Based on the Protein atlas database, it is highly expressed in brain and prostate tissues. Although, to date, this locus has never been reported to be associated with AD in previous GWASs, Hyun-Seok Hong

Table 1 Blood U-p53 GWAS summary statistics of suggestive SNPs and their potential candidate genes

Locus number	SNP	Ref. allele	Alt. allele	Chr	Frequency	*p*-value	Position	Candidate genes based on position (Intragenic SNPs)	Candidate genes based on AD single-cell gene program (intergenic SNPs)
1	rs6673974	A	G	1	0.04	2.47×10^{-5}	EGLN1-TSNAX	-	EGLN1
1	rs7519365	A	G	1	0.04	1.075×10^{-6}	TSNAX	TSNAX	-
1	rs11122309	T	C	1	0.04	8.031×10^{-6}	TSNAX-DISC1	-	EGLN1
2	rs10171094	C	T	2	0.03	5.365×10^{-6}	ALK	ALK	-
3	rs3936439	A	G	2	0.33	7.373×10^{-6}	MBD5-EPC2	-	EPC2, KIF5C, MBD5
3	rs16828910	T	G	2	0.32	5.48×10^{-6}	MBD5-EPC2	-	EPC2, KIF5C, MBD5
3	rs921243	T	G	2	0.33	2.481×10^{-6}	MBD5-EPC2	-	EPC2, KIF5C, MBD5
4	rs4853647	G	A	2	0.09	2.317×10^{-6}	CAVIN2-TMEFF2	-	TMEFF2
5	rs4462908	C	T	3	0.05	7.825×10^{-6}	CLSTN2	CLSTN2	-
6	rs4624543	G	A	3	0.42	1.69×10^{-5}	KCNMB2	KCNMB2	-
6	rs9844737	G	A	3	0.42	1.69×10^{-5}	KCNMB2	KCNMB2	-
7	rs1482699	T	C	4	0.06	8.079×10^{-6}	IGFBP7-ADGRL3	-	ADGRL3
8	rs10520544	C	A	4	0.04	1.674×10^{-6}	DCTD-WWC2	-	TENM3
9	rs9405684	T	C	6	0.13	1.94×10^{-5}	FAM50B-PRPF4B	-	-
10	rs3025642	C	T	6	0.05	2.05×10^{-5}	GABBR1	GABBR1	-
11	rs9393999	A	G	6	0.14	2.97×10^{-5}	TRIM26-TRIM39	-	-
11	rs12174823	T	C	6	0.12	4.958×10^{-6}	TRIM26-TRIM39	-	-
11	rs2240058	A	G	6	0.12	4.653×10^{-6}	TRIM39	TRIM39	-
12	rs2444783	A	G	6	0.14	2.89×10^{-5}	NOX3-ARID1B	-	ARID1B
12	rs2250759	C	T	6	0.13	2.77×10^{-5}	NOX3-ARID1B	-	ARID1B
13	rs2159808	C	A	7	0.30	8.384×10^{-6}	DGKB-AGMO	-	DGKB
14	rs279686	A	G	7	0.27	4.82×10^{-7}	AASS-FEZF1	-	PTPRZ1
15	rs4732990	C	T	8	0.15	2.44×10^{-5}	DUSP4-SARAF	-	SARAF
16	rs1158989	A	G	9	0.27	2.88×10^{-5}	SMARCA2-VLDLR	-	SMARCA2, VLDLR, KCNV2
17	rs7860549	A	G	9	0.07	2.80×10^{-5}	GNA14	GNA14	-
18	rs7071222	A	G	10	0.11	1.32×10^{-5}	SORCS1	SORCS1	-
18	rs2245123	T	C	10	0.09	1.02×10^{-5}	SORCS1	SORCS1	-
18	rs17121635	A	G	10	0.10	3.307×10^{-6}	SORCS1	SORCS1	-
19	rs17184650	C	T	14	0.11	1.02×10^{-5}	TMEM229B	TMEM229B	-
20	rs1508299	T	G	14	0.02	1.33×10^{-5}	TMEM63C	TMEM63C	-
21	rs7166440	T	C	15	0.04	2.54×10^{-5}	SLC28A1	SLC28A1	-
22	rs12444974	G	A	16	0.07	$1.44 \times 10-5$	GCSH	GCSH	-
23	rs196956	A	G	18	0.25	$2.65 \times 10-5$	PIEZO2	PIEZO2	-
24	rs8109900	C	T	19	0.21	$2.89 \times 10-5$	ZNF816-ZNF321P	ZNF816-ZNF321P	-
25	rs2835119	A	G	21	0.02	9.048×10^{-6}	RUNX1-SETD4	-	RUNX1, SETD4, CBR1
26	rs3787909	G	A	21	0.03	2.51×10^{-5}	ERG-ETS2	-	ETS2
27	rs6001930	T	C	22	0.09	2.11×10^{-5}	MRTFA	MRTFA	-
27	rs6001931	G	A	22	0.09	2.11×10^{-5}	MRTFA	MRTFA	-

et al., for the first time, showed that the TMEFF2 protein is related to Aβ metabolism and toxicity and could protect neurons from Aβ-induced toxicity [50]. This example highlights the power of an endophenotype-based GWAS in identifying novel loci previously reported to be related to the disease mechanisms in experimental studies.

At locus 3, there were three suggestive intergenic SNPs between the *MBD5* and *EPC2* genes. Interestingly, in a 2011 GWAS on AD CSF biomarkers among ADNI cohort subjects, the rs4499362 SNP between *EPC2* and *KIF5C* genes at this locus was associated with CSF total tau protein [51]. The authors proposed the EPC2 gene

Table 2 SCENT package Summary statistics of blood U-p53 candidate genes at intergenic risk loci

Locus num	Significant Genes	Suggestive SNPs	Proxy SNPs	LD R^2	ATAC peak region	SCENT p-value	Brain cell types
3	EPC2	rs3936439	rs11687034	0.77	Chr2:148,614,153–148,614,910	3.62×10^{-6}	OPC
3	EPC2	rs3936439	rs11687034	0.77	Chr2:148,614,153–148,614,910	0.01	Microglia
3	KIF5C	rs3936439	rs11687034	0.77	Chr2:148,614,153–148,614,910	8.94×10^{-3}	OPC
4	TMEFF2	rs4853647	rs72912394, rs72912395	0.68	Chr2:191,864,620–191865537	1.33×10^{-11}	Excitatory
11	HLA-E	rs12174823	rs9391807, rs9391806	0.91	Chr6:30,259,185–30260006	0.03	Microglia
11	ATAT1	rs12174823	rs9391807, rs9391806	0.91	Chr6:30,259,185–30260006	0.01	Excitatory
11	TUBB	rs12174823	rs9391807, rs9391806	0.91	Chr6:30,259,185–30260006	0.02	Inhibitory
11	TUBB	rs12174823	rs9391807, rs9391806	0.91	Chr6:30,259,185–30260006	0.01	Astrocyte
11	TUBB	rs12174823	rs9391807, rs9391806	0.91	Chr6:30,259,185–30260006	0.001	OPC
11	TRIM26	rs12174823	rs9391807, rs9391806	0.91	Chr6:30,259,185–30260006	0.02	Excitatory
12	ARID1B	rs2250759	rs2603438	0.87	Chr6:157,507,177–156507907	0.002	Excitatory
25	RUNX1	rs2835119	rs73365738	0.96	Chr21:35,720,219–35721068	0.002	Microglia
25	RUNX1	rs2835119	rs12481979	0.81	Chr21:35,696,977–35,697,804	0.03	Microglia

as a novel gene involved in AD pathogenesis. This locus contains three genes (*MBD5*, *EPC2*, *KIF5C*), all highly expressed in the brain, making fine-mapping of this locus more challenging. Using the enhancer-to-gene linking approach, we could map the suggestive region to *KIF5C* and EPC2 genes in OPC and *EPC2* gene in microglial cells. However, unlike the *KIF5C* gene, the *EPC2* gene was not differentially expressed between AD and healthy controls in an independent RNA-seq dataset. Additionally, our analysis revealed that there was a significant direct association between blood expression of the *KIF5C* gene and blood U-p53. So, these two genes may be involved in AD pathogenesis in a cell-specific manner. Although the exact function of the *EPC2* (enhancer of polycomb homolog 2) gene has not been fully characterized, it may relate to AD pathogenesis through the regulation of epigenetic mechanisms and chromatin remodeling [51]. The *KIF5C* (Kinesin Family Member 5C) is required for anterograde axonal transportation. Axonal transportation defects play important roles in AD pathophysiology [52].

We identified three SNPs along with many SNPs which had subthreshold *p*-values at locus 11. We could fine-map

this region to 4 genes across various brain cell types; 3 (*HLA-E*, *ATAT1*, *TUBB*) were differentially expressed between AD and healthy cells. Interestingly, *ATAT1* and *TUBB* are involved in axonal transportation. *HLA-E* is involved in immune pathways. Interestingly, the blood expression of both *ATAT1* and *HLA-E* genes were also significantly associated with blood U-p53. This example highlights the challenges of fine mapping of GWAS SNPs and complexity of potential shared enhancers regulating multiple genes in different cell types in some of the risk loci identified in GWASs. Future studies based on CRISPR technology may unravel this complexity and pinpoint the true causal genes.

At locus 12, we identified two SNPs between *NOX3* and *ARID1B* genes. In a study in 2014, V K Ramanan et al. showed that the rs938448 SNP at this locus was associated with amyloid deposition in the brain [53]. Although at first glance, the *NOX3* gene seems to be the potential causal gene due to its involvement in oxidative agents' production, our single-cell enhancer-to-gene linking analysis revealed that the region is significantly associated with the expression of the *ARID1B* gene in excitatory neurons. *ARID1B* (AT-Rich Interaction Domain 1B) gene

(See figure on next page.)

Fig. 4 Graphical illustration of GWAS results at multiple levels across three loci with strong evidence of their involvement in AD pathogenesis. **A** The locus number 4 between *CAVIN2* and *TMEFF2* genes showing two SNPs (rs72912394, rs72912395) with LD $r^2 = 0.68$ with the suggestive SNP (rs4853647) located within ATAC peak regions across endothelial, excitatory neuron, and pericyte cells. Our analysis revealed that this regulatory intergenic region is a potential silencer of the *TMEFF2* gene in excitatory neuron cells. **B** The locus number 18 shows the suggestive SNP (rs17121635) within the *SORCS1* gene. The left bottom picture shows the risk effect of the alternative allele G of this polymorphism by increasing the U-p53 concentration in blood. The right bottom picture shows the significant inverse association between the blood *SORCS1* gene expression and the blood U-p53. **C** The locus number 25 showing two SNPs (rs73365738, rs12481979) with LD $r^2 = 0.96$ and LD $r^2 = 0.81$ with the suggestive SNP (rs2835119) located within two distinct ATAC peak regions in the microglia cells. Our analysis revealed that both intergenic regions regulate the expression of the *RUNX1* gene in microglia cells

Fig. 4 (See legend on previous page.)
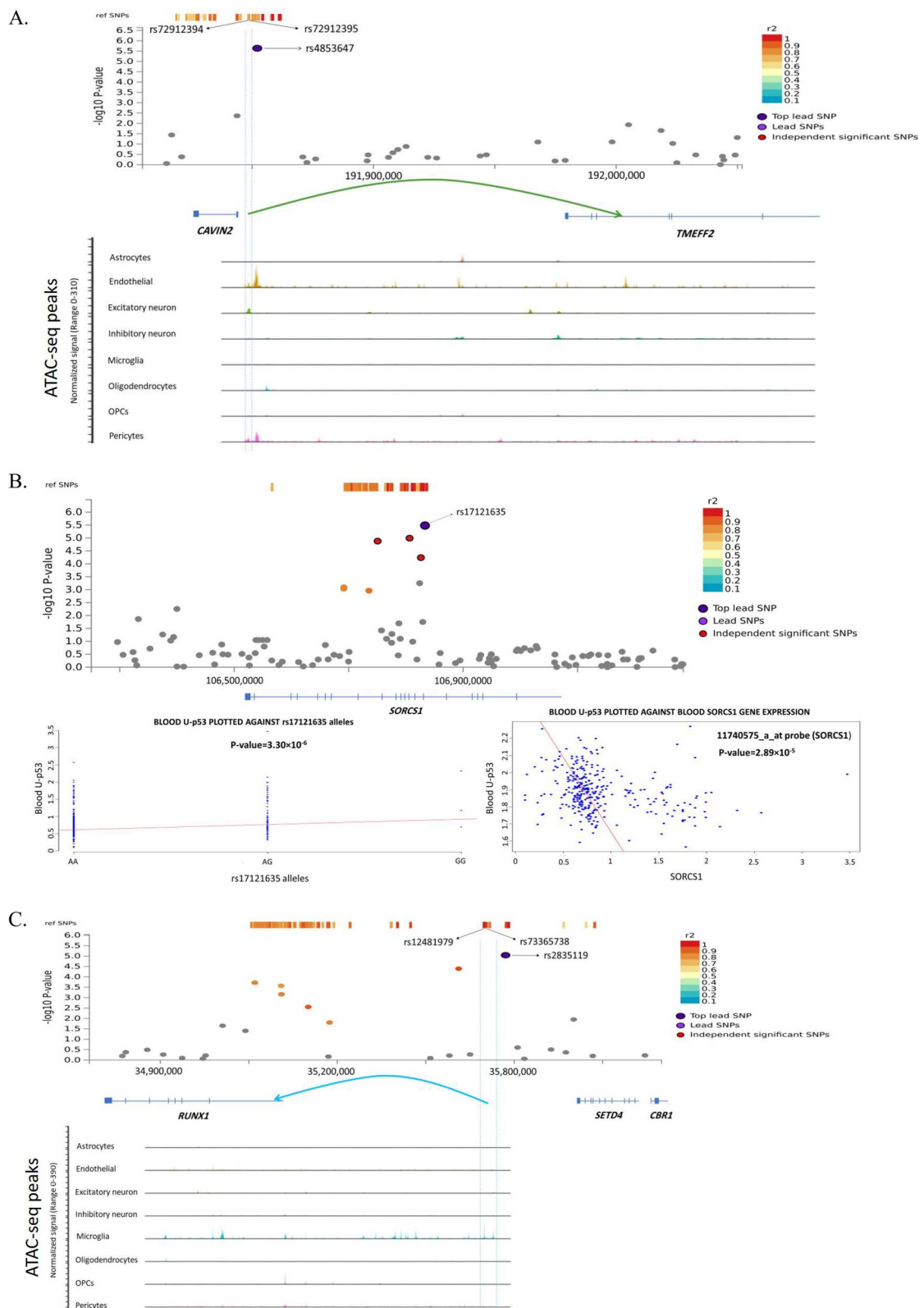
is a subunit of the mammalian SWI/SNF complex. It has an important role in chromatin remodeling. Although the exact mechanisms of this gene in AD pathogenesis are unclear, *ARID1B* mutations have been reported as monogenic causes of autism spectrum disorder (ASD) and intellectual disabilities. It has been shown that *ARID1B* haploinsufficiency can impair inhibitory synaptic function and increase excitation to inhibition (E/I) balance in the brain [54]. This example highlights the challenges of finding target genes of intergenic SNPs based solely on potential disease-related biological mechanisms.

At locus 25, the rs2835119 SNP, located between *RUNX1* and *SETD4* gene, was associated with blood U-p53. Analyzing two different open chromatin regions harboring two SNPs with high LD with the rs2835119 revealed that both regions were potential enhancers and silencers of the *RUNX1* gene in microglial cells. Ashok Patel et al. showed that variants (rs4816501) in *RUNX1* were associated with AD among patients with Down syndrome [55]. The *RUNX1* (RUNX Family Transcription Factor 1) gene encodes a transcription factor regulating the expression of numerous genes involved in normal hematopoiesis. It has been shown that *RUNX1* is involved in the regulation of TLR1/2 and TLR4 signaling pathways and inflammatory cytokine production [56]. The relevance of this gene in AD pathogenesis may be through immune system pathways. Future experimental studies are needed to uncover this.

Among intragenic SNPs that their relevant genes were differentially expressed between AD and healthy cells, products of some genes like *GABBR1* and *CLSTN2* are located in synapses. Among the other genes, it has been shown that overexpression of the *MRTFA* gene reduces the accumulation of Aβ peptide [57]. Heterozygous deletion of the other gene, *GNA14*, has been reported in an early-onset AD patient [58]. By studying the 3xTg-AD and tauC3 mouse AD models, Jisu Park et al. proposed that the *ALK* gene has a crucial role in tau-mediated neurodegeneration [59].

At the proteomics level, we investigated whether there are significant differences in the protein levels of our proposed genes between the brain tissues of AD patients and control subjects. To achieve this, we utilized the NeuroPro database (https://neuropro.biomedical.hosting/) [60], which has compiled a comprehensive list of differentially expressed proteins by conducting a meta-analysis of 38 brain proteomic studies across different stages of AD. Among the 15 proposed genes in our final list, protein levels of six genes were differentially expressed between brain tissues of AD and control subjects. Interestingly, the protein levels of TMEFF2 and TMEM63C were upregulated in the frontal cortex of patients with preclinical stages of AD compared to the controls [61]. This suggests that TMEFF2 and TMEM63C may serve as novel target genes for early-stage intervention and should be further investigated in future studies, as they have received less attention in the context of AD. The other four proteins, including the SORCS1, CLSTN2, TUBB, and GABBR1, were differentially expressed between patients with AD and healthy controls.

Although using a quantitative phenotype in our GWAS increased the power of our study, due to their high cost and laboratory challenges, GWASs on endophenotypes always suffer from relatively small sample sizes compared to traditional GWASs [45]. Future studies are needed to increase the sample sizes for various AD biomarkers, such as U-p53. However, our study highlights the power of leveraging two approaches in AD GWASs: First, it showed using a powerful early-stage biomarker of AD as an endophenotype in AD GWASs can lead to identifying novel loci and genes involved in the early stages of AD. Second, it showed that integration of GWAS results with single-cell datasets and leveraging powerful enhancer-to-gene linking strategies can pinpoint the true causal genes in a cell-specific manner.

Conclusions

In conclusion, through the integration of GWAS results on U-p53 as a reliable early-stage AD biomarker with brain single-cell RNA-seq and multimodal datasets, we have identified several significant genes associated with blood U-p53 and potentially with AD, particularly among non-demented subjects. These genes include SORCS1, TMEFF2, TMEM63C, KIF5C, HLA-E, TUBB1, ATAT1, ARID1B, and RUNX1. We believe that these proposed candidate genes hold substantial value for further exploration in future experimental studies, both in terms of their functional and structural aspects, due to their potential roles in the early stages of AD. For instance, genes like TMEFF2 and TMEM63C have received comparatively less attention in the context of AD, but there are some evidences indicating that their protein expression levels are altered in the preclinical stages of the disease. This highlights the importance of investigating these genes further. Furthermore, our results could provide valuable insights for the design of future drug targets.

Given the limited availability of cohorts where blood U-p53 has been measured, conducting GWASs on this biomarker across additional cohorts in the future would be feasible and could offer novel insights into the pathogenesis of AD. Using continuous endophenotypes like U-p53 in neurodegenerative diseases produces various benefits compared to traditional binary trait GWASs, such as increasing the statistical power of GWASs, reducing the disease heterogeneity which is often observed in complex diseases, and providing specific biological context for the associated genes [62].

Supplementary Information

Supplementary Material 1.

Availability of data and materials
The DIADEM_V2_08_30_22.csv, ADNI_GO_2_OmniExpress.zip, ADNI_GO2_2nd.tar.gz, Microarray Gene Expression Profile Data.csv, and ADNIMERGE. csv files from the ADNI cohort are available on the https://ida.loni.usc.edu/ web portal via a personal account. The AD gene program file from the Karthik A. Jagadeesh et al. study [38] can be found on https://alkesgroup.broadinstitute.org/LDSCORE/Jagadeesh_Dey_sclinker/gene_scores/gene_scores/disease_progression_programs/Alzheimers/. The brain single-cell multimodal dataset is available on the GEO website via the accession code GSE214637. The brain single-cell RNA-seq dataset used for validation of the candidate genes is found on the GEO website via the accession code GSE174367. Other datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Bacigalupo I, Mayer F, Lacorte E, Di Pucchio A, Marzolini F, Canevelli M, et al. A Systematic Review and Meta-Analysis on the Prevalence of Dementia in Europe: Estimates from the Highest-Quality Studies Adopting the DSM IV Diagnostic Criteria. J Alzheimers Dis. 2018;66(4):1471–81.
2. Winblad B, Amouyel P, Andrieu S, Ballard C, Brayne C, Brodaty H, et al. Defeating Alzheimer's disease and other dementias: a priority for European science and society. Lancet Neurol. 2016;15(5):455–532.
3. Aisen PS, Cummings J, Jack CR Jr, Morris JC, Sperling R, Frölich L, et al. On the path to 2025: understanding the Alzheimer's disease continuum. Alzheimers Res Ther. 2017;9(1):60.
4. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011;7(3):270–9.
5. Zhang XX, Tian Y, Wang ZT, Ma YH, Tan L, Yu JT. The Epidemiology of Alzheimer's Disease Modifiable Risk Factors and Prevention. J Prev Alzheimers Dis. 2021;8(3):313–21.
6. Kametani F, Hasegawa M. Reconsideration of Amyloid Hypothesis and Tau Hypothesis in Alzheimer's Disease. Front Neurosci. 2018;12:25.
7. Buttini M, Masliah E, Barbour R, Grajeda H, Motter R, Johnson-Wood K, et al. Beta-amyloid immunotherapy prevents synaptic degeneration in a mouse model of Alzheimer's disease. J Neurosci. 2005;25(40):9096–101.
8. Doody RS, Thomas RG, Farlow M, Iwatsubo T, Vellas B, Joffe S, et al. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. N Engl J Med. 2014;370(4):311–21.
9. Salloway S, Sperling R, Fox NC, Blennow K, Klunk W, Raskind M, et al. Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. N Engl J Med. 2014;370(4):322–33.
10. Clark JS, Kayed R, Abate G, Uberti D, Kinnon P, Piccirella S. Post-translational Modifications of the p53 Protein and the Impact in Alzheimer's Disease: A Review of the Literature. Front Aging Neurosci. 2022;14:835288.
11. Merlo P, Frost B, Peng S, Yang YJ, Park PJ, Feany M. p53 prevents neurodegeneration by regulating synaptic genes. Proc Natl Acad Sci U S A. 2014;111(50):18055–60.
12. Di Giovanni S, Knights CD, Rao M, Yakovlev A, Beers J, Catania J, et al. The tumor suppressor protein p53 is required for neurite outgrowth and axon regeneration. Embo j. 2006;25(17):4084–96.
13. Silva AR, Santos AC, Farfel JM, Grinberg LT, Ferretti RE, Campos AH, et al. Repair of oxidative DNA damage, cell-cycle regulation and neuronal death may influence the clinical manifestation of Alzheimer's disease. PLoS ONE. 2014;9(6):e99897.
14. Piccirella S, Van Neste L, Fowler C, Masters CL, Fripp J, Doecke JD, et al. A Conformational Variant of p53 (U-p53AZ) as Blood-Based Biomarker for the Prediction of the Onset of Symptomatic Alzheimer's Disease. J Prev Alzheimers Dis. 2022;9(3):469–79.
15. Uberti D, Carsana T, Bernardi E, Rodella L, Grigolato P, Lanni C, et al. Selective impairment of p53-mediated cell death in fibroblasts from sporadic Alzheimer's disease patients. J Cell Sci. 2002;115(Pt 15):3131–8.
16. French PW. Unfolded p53 in non-neuronal cells supports bacterial etiology of Alzheimer's disease. Neural Regen Res. 2022;17(12):2619–22.
17. Abate G, Frisoni GB, Bourdon JC, Piccirella S, Memo M, Uberti D. The pleiotropic role of p53 in functional/dysfunctional neurons: focus on pathogenesis and diagnosis of Alzheimer's disease. Alzheimers Res Ther. 2020;12(1):160.
18. Shulman RG, Rothman DL, Behar KL, Hyder F. Energetic basis of brain activity: implications for neuroimaging. Trends Neurosci. 2004;27(8):489–95.
19. Chong ZZ, Li F, Maiese K. Oxidative stress in the brain: novel cellular targets that govern survival during neurodegenerative disease. Prog Neurobiol. 2005;75(3):207–46.
20. De Felice FG, Velasco PT, Lambert MP, Viola K, Fernandez SJ, Ferreira ST, et al. Abeta oligomers induce neuronal oxidative stress through an N-methyl-D-aspartate receptor-dependent mechanism that is blocked by the Alzheimer drug memantine. J Biol Chem. 2007;282(15):11590–601.
21. Lanni C, Nardinocchi L, Puca R, Stanga S, Uberti D, Memo M, et al. Homeodomain interacting protein kinase 2: a target for Alzheimer's beta amyloid leading to misfolded p53 and inappropriate cell survival. PLoS ONE. 2010;5(4):e10171.
22. Lanni C, Necchi D, Pinto A, Buoso E, Buizza L, Memo M, et al. Zyxin is a novel target for β-amyloid peptide: characterization of its role in Alzheimer's pathogenesis. J Neurochem. 2013;125(5):790–9.
23. Uberti D, Cenini G, Olivari L, Ferrari-Toninelli G, Porrello E, Cecchi C, et al. Over-expression of amyloid precursor protein in HEK cells alters p53 conformational state and protects against doxorubicin. J Neurochem. 2007;103(1):322–33.
24. Singh AK, Pati U. CHIP stabilizes amyloid precursor protein via proteasomal degradation and p53-mediated trans-repression of β-secretase. Aging Cell. 2015;14(4):595–604.
25. Liu Z, Li T, Li P, Wei N, Zhao Z, Liang H, et al. The Ambiguous Relationship of Oxidative Stress, Tau Hyperphosphorylation, and Autophagy Dysfunction in Alzheimer's Disease. Oxid Med Cell Longev. 2015;2015:352723.
26. Zahid S, Khan R, Oellerich M, Ahmed N, Asif AR. Differential S-nitrosylation of proteins in Alzheimer's disease. Neuroscience. 2014;256:126–36.
27. Frisoni GB, Boccardi M, Barkhof F, Blennow K, Cappa S, Chiotis K, et al. Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. Lancet Neurol. 2017;16(8):661–76.
28. Cummings J. The Role of Biomarkers in Alzheimer's Disease Drug Development. Adv Exp Med Biol. 2019;1118:29–61.

29. Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimers Dement. 2018;14(4):535–62.

30. Abate G, Vezzoli M, Polito L, Guaita A, Albani D, Marizzoni M, et al. A conformation variant of p53 combined with machine learning identifies Alzheimer disease in preclinical and prodromal stages. J Pers Med. 2020;11(1):14. https://doi.org/10.3390/jpm11010014.

31. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry. 2006;63(2):168–74.

32. Lake J, Warly Solsberg C, Kim JJ, Acosta-Uribe J, Makarious MB, Li Z, et al. Multi-ancestry meta-analysis and fine-mapping in Alzheimer's disease. Mol Psychiatry. 2023;28(7):3121–32. https://doi.org/10.1038/s41380-023-02089-w.

33. Escott-Price V, Hardy J. Genome-wide association studies for Alzheimer's disease: bigger is not always better. Brain Commun. 2022;4(3):fcac125.

34. Jansen IE, van der Lee SJ, Gomez-Fonseca D, de Rojas I, Dalmasso MC, Grenier-Boley B, et al. Genome-wide meta-analysis for Alzheimer's disease cerebrospinal fluid biomarkers. Acta Neuropathol. 2022;144(5):821–42.

35. Liu C, Yu J. Genome-Wide Association Studies for Cerebrospinal Fluid Soluble TREM2 in Alzheimer's Disease. Front Aging Neurosci. 2019;11:297.

36. Oatman SR, Reddy JS, Quicksall Z, Carrasquillo MM, Wang X, Liu CC, et al. Genome-wide association study of brain biochemical phenotypes reveals distinct genetic architecture of Alzheimer's disease related proteins. Mol Neurodegener. 2023;18(1):2.

37. Lamparter D, Marbach D, Rueedi R, Kutalik Z, Bergmann S. Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. PLoS Comput Biol. 2016;12(1):e1004714.

38. Jagadeesh KA, Dey KK, Montoro DT, Mohan R, Gazal S, Engreitz JM, et al. Identifying disease-critical cell types and cellular processes by integrating single-cell RNA-sequencing and human genetics. Nat Genet. 2022;54(10):1479–92.

39. Anderson AG, Rogers BB, Loupe JM, Rodriguez-Nunez I, Roberts SC, White LM, et al. Single nucleus multiomics identifies ZEB1 and MAFB as candidate regulators of Alzheimer's disease-specific cis-regulatory elements. Cell Genom. 2023;3(3):100263.

40. Sakaue S, Weinand K, Dey KK, Jagadeesh K, Kanai M, Watts GFM, et al. Tissue-specific enhancer-gene maps from multimodal single-cell data identify causal disease alleles. medRxiv. 2022:2022.10.27.22281574.

41. Morabito S, Miyoshi E, Michael N, Shahin S, Martini AC, Head E, et al. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. Nat Genet. 2021;53(8):1143–55.

42. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88(1):76–82.

43. Wen J, Zhao B, Yang Z, Erus G, Skampardoni I, Mamourian E, et al. The Genetic Architecture of Multimodal Human Brain Age. bioRxiv. 2023:2023.04.13.536818.

44. Sun E, Motolani A, Campos L, Lu T. The pivotal role of NF-kB in the pathogenesis and therapeutics of Alzheimer's disease. Int J Mol Sci. 2022;23(16):8972. https://doi.org/10.3390/ijms23168972.

45. Andrews SJ, Renton AE, Fulton-Howard B, Podlesny-Drabiniok A, Marcora E, Goate AM. The complex genetic architecture of Alzheimer's disease: novel insights and future directions. EBioMedicine. 2023;90:104511.

46. Xu W, Xu J, Wang Y, Tang H, Deng Y, Ren R, et al. The genetic variation of SORCS1 is associated with late-onset Alzheimer's disease in Chinese Han population. PLoS ONE. 2013;8(5):e63621.

47. Reitz C, Tokuhiro S, Clark LN, Conrad C, Vonsattel JP, Hazrati LN, et al. SORCS1 alters amyloid precursor protein processing and variants may increase Alzheimer's disease risk. Ann Neurol. 2011;69(1):47–64.

48. Lane RF, Raines SM, Steele JW, Ehrlich ME, Lah JA, Small SA, et al. Diabetes-associated SorCS1 regulates Alzheimer's amyloid-beta metabolism: evidence for involvement of SorL1 and the retromer complex. J Neurosci. 2010;30(39):13110–5.

49. Lee AK, Yi N, Khaled H, Feller B, Takahashi H. SorCS1 inhibits amyloid-β binding to neurexin and rescues amyloid-β-induced synaptic pathology. Life Sci Alliance. 2023;6(4):e202201681. https://doi.org/10.26508/lsa.202201681.

50. Hong HS, Maezawa I, Petrlova J, Zhao XY, John CV, Jin LW. Tomoregulin (TMEFF2) binds alzheimer's disease amyloid-β (Aβ) oligomer and AβPP and protects neurons from Aβ-induced toxicity. J Alzheimers Dis. 2015;48(3):731–43.

51. Kim S, Swaminathan S, Shen L, Risacher SL, Nho K, Foroud T, et al. Genome-wide association study of CSF biomarkers Abeta1-42, t-tau, and p-tau181p in the ADNI cohort. Neurology. 2011;76(1):69–79.

52. Wang ZX, Tan L, Yu JT. Axonal transport defects in Alzheimer's disease. Mol Neurobiol. 2015;51(3):1309–21.

53. Ramanan VK, Risacher SL, Nho K, Kim S, Swaminathan S, Shen L, et al. APOE and BCHE as modulators of cerebral amyloid deposition: a florbetapir PET genome-wide association study. Mol Psychiatry. 2014;19(3):351–7.

54. Moffat JJ, Smith AL, Jung EM, Ka M, Kim WY. Neurobiology of ARID1B haploinsufficiency related to neurodevelopmental and psychiatric disorders. Mol Psychiatry. 2022;27(1):476–89.

55. Patel A, Rees SD, Kelly MA, Bain SC, Barnett AH, Thalitaya D, et al. Association of variants within APOE, SORL1, RUNX1, BACE1 and ALDH18A1 with dementia in Alzheimer's disease in subjects with Down syndrome. Neurosci Lett. 2011;487(2):144–8.

56. Bellissimo DC, Chen CH, Zhu Q, Bagga S, Lee CT, He B, et al. Runx1 negatively regulates inflammatory cytokine production by neutrophils in response to Toll-like receptor signaling. Blood Adv. 2020;4(6):1145–58.

57. Zhang W, Yang Y, Xiang Z, Cheng J, Yu Z, Wang W, et al. MRTF-A-mediated protection against amyloid-β-induced neuronal injury correlates with restoring autophagy via miR-1273g-3p/mTOR axis in Alzheimer models. Aging (Albany NY). 2022;14(10):4305–25.

58. Lazarczyk MJ, Haller S, Savioz A, Gimelli S, Bena F, Giannakopoulos P. Heterozygous Deletion of Chorein Exons 70–73 and GNA14 Exons 3–7 in a Brazilian Patient Presenting With Probable Tau-Negative Early-Onset Alzheimer Disease. Alzheimer Dis Assoc Disord. 2017;31(1):82–5.

59. Park J, Choi H, Kim YD, Kim SH, Kim Y, Gwon Y, et al. Aberrant role of ALK in tau proteinopathy through autophagosomal dysregulation. Mol Psychiatry. 2021;26(10):5542–56.

60. Askenazi M, Kavanagh T, Pires G, Ueberheide B, Wisniewski T, Drummond E. Compilation of reported protein changes in the brain in Alzheimer's disease. Nat Commun. 2023;14(1):4466.

61. Johnson ECB, Carter EK, Dammer EB, Duong DM, Gerasimov ES, Liu Y, et al. Large-scale deep multi-layer analysis of Alzheimer's disease brain reveals strong proteomic disease-related changes not observed at the RNA level. Nat Neurosci. 2022;25(2):213–25.

62. Farias FHG, Benitez BA, Cruchaga C. Quantitative endophenotypes as an alternative approach to understanding genetic risk in neurodegenerative diseases. Neurobiol Dis. 2021;151:105247.

Publisher's Note