Research Article

Lower expression of *BIN1's* neuronal isoform in vulnerable excitatory neurons increases risk in Alzheimer's disease

Journal of Alzheimer's Disease Reports
Volume 9: 1–16
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/25424823241296018
journals.sagepub.com/home/alr

Rajesh Ranganathan, Siwen Li, Georgy Sapozhnikov, Shoutang Wang and You-Qiang Song

Abstract

Background: Neurons in Alzheimer's disease (AD) experience elevated DNA damage, with DNA repair sites enriched at enhancer regions of genes essential for neuronal survival. Excitatory neurons in the cortical superficial layers expressing *CUX2* and *RORB* (Cux2+/Rorb+), are selectively vulnerable in AD, but their relationship to single nucleotide polymorphisms (SNPs) in AD genome-wide association studies (GWAS) is unclear.

Objective: This study aimed to identify and characterize functional AD-GWAS SNPs using single-nucleus RNA sequencing data, focusing on selectively vulnerable neurons and DNA repair hotspots.

Methods: Filters were applied to identify candidate SNPs based on overlap with repair hotspots, RNA expression, transcription factor binding, AD association, and epigenetic significance. In vitro assays and analyses of large datasets from bulk RNA-seq ($n = 1894$), proteomics ($n = 400$), and single-nucleus RNA-seq ($n = 424$, 1.6 M cells) were conducted.

Results: *BIN1* SNP, rs78710909, met multiple criteria - located in an AD-GWAS locus, repair hotspot, and promoter region. rs78710909C exhibits 1.52× higher AD risk and 5.4× differential transcription factor binding. In vitro, rs78710909C shows greater enhancer activity and weaker p53 but stronger E2F1 binding. *BIN1*'s neuronal isoform is neuroprotective, but its AD expression is lower ($p < 0.01$). Moreover, only in AD and Cux2+/Rorb + neurons, rs78710909C is associated with a lower average *BIN1* neuronal isoform ratio ($p < 0.01$). The genes upregulated in neurons with lower neuronal isoform ratio were associated with the hallmarks of AD pathology.

Conclusions: In a disease-relevant mechanism, the *BIN1* SNP rs78710909C is associated with a lower ratio of *BIN1*'s neuronal isoform which increases the vulnerability of specific excitatory neurons in AD patients.

Keywords

alternative splicing, Alzheimer's disease, BIN1, DNA damage, DNA repair, genetic predisposition to disease, genome-wide association study, isoforms, neurons, single-cell RNA sequencing, vulnerability

Received: 19 August 2024; accepted: 6 October 2024

Introduction

Alzheimer's disease (AD) is an aging-related neurodegenerative disease defined by the accumulation of amyloid-β (Aβ) plaques and the intracellular neurofibrillary tangles of hyperphosphorylated tau in the brain.[1] The incidence of AD significantly increases with age, escalating from 5.3% in individuals aged 65–74 years to 34.6% in those over 85 years.[2] While early-onset familial AD (<65 years of age) can be attributed to mutations in genes like *PSEN1*, *PSEN2*, and *APP*, the more common late-onset sporadic AD appears to be a complex and multifactorial disease.[3]

Several AD genome-wide association studies (AD-GWAS) have identified genetic loci, beyond *APOE*, that are associated with AD risk, including *BIN1, CLU, PICALM, MS4A2, CR1,* and *ABCA7*.[4–6] However, variations

School of Biomedical Sciences, The University of Hong Kong, Hong Kong, China

Corresponding author:
You-Qiang Song, L3-63, Laboratory Block, 21 Sassoon Road, Hong Kong, China.
Email: songy@hku.hk

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (https://us.sagepub.com/en-us/nam/open-access-at-sage).

in gene expression across different brain cell types have limited the reliability of conclusions drawn from traditional bulk tissue analyses. For instance, the hypotheses regarding the role of the *BIN1* gene in AD has been revised based on findings that it is overexpressed in oligodendrocytes but reduced in neurons during disease progression.[7–9] Similarly, single-nucleus RNA sequencing (snRNA-seq) has revealed distinct microglial states and the importance of genes like *TREM2* in the neuroinflammatory response in AD brains.[10]

During AD progression, certain neuronal populations are selectively vulnerable to accumulation of stress factors such as Aβ and phosphorylated tau, culminating in higher rates of apoptosis.[11] Cortical regions, particularly excitatory neurons and the entorhinal cortex-hippocampal pathway, are notably affected at early stages of the disease.[12–16] However, not all pathological tau bearing excitatory neurons are equally susceptible to death.[17] For instance, neurons near Aβ plaques suffer significantly more excitatory synaptic loss compared to those farther away.[18] snRNA-seq studies have identified *CUX2* as a marker of excitatory neurons in the superficial layers (L2-L4) of the neocortex[17,19–21] and neurons expressing both *CUX2* and *RORB* (Cux2+/Rorb+) as a selectively vulnerable subpopulation, with differential expression of genes related to AD risk factors, synaptic function, microtubule related pathways, neuroinflammation and biomarkers of neurodegeneration.[16,17,20,22] The relationship between selective vulnerability and SNPs identified in AD-GWAS remains to be fully elucidated.

Interestingly, adult neurons are subject to several sources of DNA damage but are required to have a lifespan that closely mirrors that of the organism. Therefore they rely heavily on DNA repair mechanisms for their survival and function.[23–25] A recent study using stem cell-derived neurons found that certain genomic "hotspots" are preferentially repaired, and these hotspots are enriched near genes critical for neuronal survival and in regulatory regions of the genome.[26] Other studies have corroborated that highly transcribed and critical genes are repaired preferentially in aging neurons.[27,28] Consequently, the accumulation of DNA damage with age and AD pathology may lead to the enrichment of DNA repair *hotspots* in AD-associated genes.

This study aims to integrate data on selectively vulnerable neuronal subtypes, AD genetic risk variants, and their associations with DNA repair hotspots to shed light on the underlying mechanisms of AD pathogenesis. By identifying causal variants, their target genes, the specific cell types involved, and the biological pathways linking genetic factors to disease progression, we hope to enhance our understanding of the molecular underpinnings of AD.

To achieve this objective, we have synthesized data from different sources, including functional and genomic annotation of AD-GWAS[4,5] SNPs (FUMA,[29] SNPnexus,[30] LDlink),[31] DNA repair *hotspots,*[26] gene regulatory elements (GeneHancer,[32] ANANASTRA),[33] transcription factor and RNA binding protein databases (RBPmap),[34] and whole genome sequencing, proteomics and snRNA-seq data (AD Knowledge Portal).[35–39]

Methods

Ethics approval

This study received IRB approval (UW 23-200).

Prioritization and annotation of candidate SNPs and genes

This study employed a multi-step approach to identify SNPs associated with both AD and DNA repair hotspots (Figure 1). The genomic coordinates (chromosome, range) of 61,178 DNA repair *hotspots* were obtained from "Table S1" of a previously published study.[26] We sourced AD-GWAS summary statistics from two different studies-11,632 SNPs from a seminal two-stage meta-analysis of 74.046 individuals that identified 20 genomic loci associated with AD (IGAP study),[5] and 3207 SNPs from an updated genome-wide meta-analysis and fine-mapping study (Finemap study).[4]

Using the "GenomicRanges" R package, we identified AD-GWAS SNPs that overlapped with DNA repair *hotspots*. We then functionally annotated these SNPs and identified protein-coding *cis*-genes using the FUMA[29] and SNP-Nexus[30] webservices (Supplemental Figure 1A).

To prioritize the genes, we created a "Combined Percentile Score" (CPS) based on the abundance of associated DNA repair hotspot peaks and the gene's RNA expression level obtained from "Table S3" of the Reid et al. study.[26] Genes were shortlisted if they met the following criteria: RNA expression > 2 TPM, repair hotspot peak percentile > 0.5, and CPS score > 0.25 (Supplemental Table 1). We obtained enhancer coordinates associated with these genes from the GeneALaCart webservice[32] and generated a list of prioritized genes, enhancers, and SNPs where DNA repair *hotspots* and SNPs overlapped with enhancer/promoter regions (Supplemental Table 2).

Next, we identified the SNPs and genes from Supplemental Table 2 that had significant association with increased AD risk or evidence of allele-specific differential transcription factor (TF) binding (Figure 1). These SNPs and genes, along with the Genomic Evolutionary Rate Profiling (GERP) score as reported in the Finemap study[4] and the 'functional significance' score from the DeepSea webservice[40] are summarized in Table 1. The functional significance score, with lower scores indicating higher significance, measures the significance of an SNP based on a geometric mean of the chromatin effects and evolutionary conservation metrics.

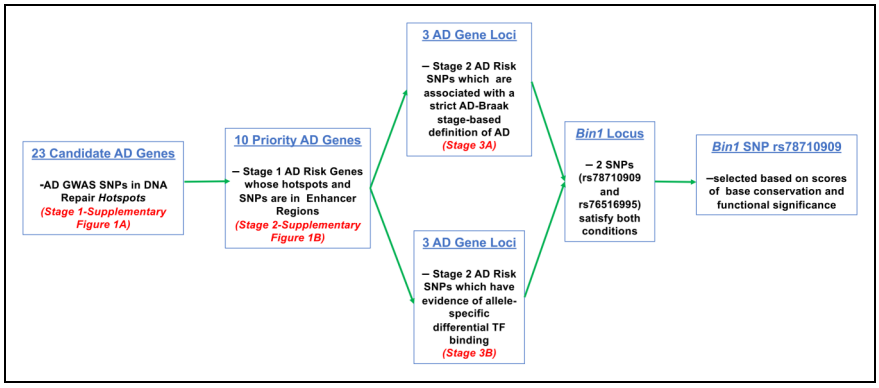

Figure 1. Flowchart to select the single SNP of *BIN1* rs78710909 as a highly prospective genomic variant for further analysis to explain possible causative relationship in AD. This SNP was chosen because of its genome-wide significance in AD-GWAS studies, location within a DNA repair hotspot and enhancer region, allele-specific differential TF binding (Figure 2), association analysis with strict definition of AD based on Braak-stage (Supplemental Table 4), and scores for GERP conservation score and functional significance (Table 1).

Allele-specific differential TF analysis

We characterized the SNPs from Supplemental Table 2 based on allele-specific differential transcription factor binding evidence. We chose to use the ANANASTRA webservice[33] (ANnotation and enrichment ANalysis of Allele-Specific TRAnscription factor binding at SNPs) because the database has been created through meta-analysis of more than 15,000 publicly available ChIP-Seq experimental data and is more robust compared to theoretical motif-analysis alone.

SNP association analysis with AD risk

Data sources. The genomic variant data used in this study was obtained from the Accelerating Medicines Partnership - Alzheimer's Disease (AMP-AD) consortium. This dataset harmonizes whole-genome sequencing (WGS) data from three major Alzheimer's disease cohorts: the Religious Orders Study and Memory and Aging Project (ROSMAP), Mount Sinai Brain Bank (MSBB), and the Mayo RNA-seq (Mayo) study.[35–38] The WGS data is available in VCF

Table 1. The *BIN1* SNP rs78710909 selected for further analysis based on converging lines of evidence among the SNPs from the Stage-3 in Figure 1. The *BIN1* SNP association with AD is based on: (A) genome wide significance of p-value of SNPs in the gene locus from the AD-GWAS Fine-mapping study. (B) Fisher's exact test with mid-p correction that indicates the allele 'C' of rs78710909 is associated with AD Risk. The source WGS data is from Synapse website for the consolidated ROSMAP, Mayo Clinic and MSBB studies. 675 Samples were selected (from 1894) for analysis based on Age > 60, *APOE* $\epsilon 4 = 0$ and AD defined based Braak stage only (AD>=4, Normal <=1). (C) 5.4× change between effect and reference allele in differential TF binding from the ANANASTRA database which is based on meta-analysis of more than 15,000 ChIP experiments (D) GERP conservation score (>+2 indicates a more conserved base) and DeepSea functional significance score (lower indicates more significant, based on geometric average of E-value across chromatin features and the evolutionary conservation scores. The data related to the chosen SNP rs78710909 is in bold type. In the color scheme below, "Green" indicates that the SNP passed the filter, while "Red" signifies that it did not.

							AD-GWAS Fine mapping[A]	Allelic Association with AD WGS Data (Minor vs. Ref Allele)[B]		Allele-specific Differential TF Binding (Minor vs. Ref Allele)[C]		Phylogenetic & Epigenetic Association[D]	
Gene Locus	Chr	SNP	Ref Allele	Minor Allele	LD r2	MAF	min p-value	min p	Odds Ratio (95% CI)	Log Fc	TF	GERP Score	DeepSea 'funsig'
BIN1	2	**rs78710909**	**G**	**C**	1.0	0.07	**2.4E-09**	**0.05**	**1.54 (0.98-2.40)**	**5.4**	**SP3**	**2.56**	**0.004**
		rs76516995	C	G								-7.59	0.016
CD2AP	6	rs1004173	C	T		0.19	6.2E-11	0.04	1.42 (1.01-2.00)			0.56	0.055
TSPAN14	10	rs1870140	A	G		0.06	2.9E-08	0.03	1.51 (1.04-2.20)			0.22	0.064
CLU	8	rs1532278	T	C	1.0	0.24	5.1E-25					3.04	0.016
		rs1532277	T	C								1.55	0.018
		rs1532276	T	C						5.6	RFX1	-0.85	0.059
CELF1	11	rs7947450	G	A	1.0	0.33	8.5E-11			2	MYCN	-0.21	0.017
		rs896817	C	T		0.32						-3.94	0.143

format from the Alzheimer's Disease Knowledge Portal (synapse file ID syn:11707420). Relevant covariate data, including diagnostic information and *APOE* genotypes, were sourced as CSV files for each cohort (MSBB: syn6101474, Mayo: syn23277389, ROSMAP: syn3191087).

Subsetting the data. Given the strong influence of the *APOE* ϵ4 allele on AD risk[41] and the late-onset nature of the disease,[2] our analysis focused on participants aged 60 years or older who did not carry any *APOE* ϵ4 alleles. Age was transformed to a classification variable of either 'above 84' or "below 84". Participants were considered as AD cases if they had a Braak stage score ≥4, and as cognitively normal controls if they had a Braak stage score ≤1. All other samples were excluded. This filtering process resulted in a final dataset of 675 participants (551 AD cases, 124 normal controls). Detailed covariate information for the selected samples is provided in Supplemental Table 3.

Genotypic association analysis. Association between SNPs and AD status (defined by Braak stage) was assessed using Fisher's exact test with mid-p correction ('stats' and 'epitools' packages in R) and logistic regression ('stats' package in R). Fisher's exact test with mid-p correction was used due to the low frequency of minor alleles for several SNPs. In cases where the frequency of the homozygous minor allele was very low (<3%), the data was transformed to allelic distribution, and the odds ratio for the effect allele versus the non-effect allele was calculated. This approach was necessary to mitigate the impact of sparse genotype data on the reported statistical significance. For SNPs in high linkage disequilibrium, such as *BIN1* rs78710909 and rs76516995, they were treated as a single locus for the association analysis. The genotypic association results were further corroborated using logistic regression adjusted for age (Supplemental Table 5).

Expression quantitative trait loci (eQTL) analysis. For samples with both WGS and RNA-sequencing data, the RNA-seq specimen IDs were cross-referenced with the WGS specimen IDs to map the gene-level and transcript-level RNA expression data (available from syn30821562 and syn3388564, respectively) with the genomic variant information. Similarly, for samples with both WGS and proteomic data, the proteomic specimen IDs were cross-referenced with the WGS specimen IDs to map the isoform-level protein expression data (available from syn17008935) with the genomic variants.

The differential expression of mRNA at the gene and transcript level, as well as protein expression, was assessed based on the SNP rs78710909 genotype and AD disease status (defined by Braak stage) using Student's t-test ('genefilter' package in R) and linear regression models ('stats' package in R). The mapping of mRNA transcripts to neuronal or ubiquitous isoforms is provided in Supplemental Table 6.

Luciferase and EMSA assays

Vector construction. To determine whether the alleles of the SNP rs78710909 (C/G) contribute to different enhancer activities, the indicated genomic sequence upstream of *BIN1* encompassing the corresponding rs78710909G was amplified from human genomic DNA of SH-SY5Y cells, and the *BIN1* enhancer sequence containing rs78710909C was synthesized by BGI Genomics (Shenzhen, China). The fragments were then inserted into the SacI and XhoI sites of pGL3-promoter vector (Promega, #E1761, Madison, WI, USA).

The expression vectors gab-p53 and gab-E2F1 were constructed by cloning the coding sequences of human p53 (1182bp) or E2F1 (1314bp) into the NheI and KpnI sites of pcDNA3.1 + vector (Addgene, #V790-20, Watertown, MA, USA). All constructs were confirmed by direct sequencing. All primer sequences used for cloning are listed in Supplemental Table 7.

Cell culture and luciferase reporter assay. Transformed human embryonic kidney (HEK293T) cells were grown in Dulbecco's Modified Eagle's Medium (Gibco, #12800017, Waltham, MA, USA) supplemented with 10% fetal bovine serum (Gibco, #10270106) and 100 units/mL penicillin-streptomycin (ThermoFisher Scientific, #15140122, Waltham, MA, USA) in a humidified environment of 37°C that contained 50 mL/L CO2. To evaluate the enhancer activities of the SNP, HEK293T cells were cotransfected in a 48-well plate with 100 ng of firefly luciferase reporter plasmid and 2 ng of pRL-CMV. To study the effect of p53 and E2F1 on *BIN1* enhancer activity, HEK293T cells grown in a 48-well plate were cotransfected with 2 ng of pRL-CMV, 100 ng of firefly luciferase reporter plasmid, and 200 ng of gab (empty vector), or gab-p53 (*p53*-expressing vector),[42] or gab-E2F1 (*E2F1*-expressing vector). Forty-eight hours after the transfections, the cell lysates were applied to luciferase assay using the dual-luciferase reporter system (Promega, #E1960, Madison, Wisconsin, USA).[43] The pRL-CMV vector expressing Renilla luciferase was used as the internal control to adjust for differences in the transfection and harvest efficiencies.

Electrophoretic mobility shift assay. Electrophoretic mobility shift assay (EMSA) was performed with the Chemiluminescent EMSA Kit (Beyotime, #GS009, Shanghai, China), according to the manufacturer's instructions. The following oligonucleotides that corresponded to the enhancer region of *BIN1* and covered the rs78710909C/G polymorphism were synthesized (underlined letters indicate polymorphism):

rs78710909C - 5′- TCCAGCTGCACCGCCCCCGGGCGGG GACTAATCCG –3′; rs78710909G - 5′- TCCAGCTGCAC CGCCCCGGGGCGGGGACTAATCCG –3′.

Additionally, a conserved DNA binding sequence of p53 was also synthesized:

5′-AGCTTAGACATGCCTAGACATGCCTA −3′.

Oligonucleotides were annealed and end-labeled with biotin using the EMSA Probe Biotin Labeling Kit (Beyotime, #GS008). Nuclear extracts from HEK293T cells were prepared as described. For the competition assay, a 100-fold molar excess of unlabeled oligonucleotide was added to the binding reaction mixture as a competitor before the addition of biotin-labeled probes. After incubation for 30 min at room temperature, the products were separated on pre-electrophoresed 5% non-denaturing polyacrylamide gel at 4°C and then transferred to a Nylon membrane (Beyotime, #FFN10). The membrane was subsequently cross-linked by UV light for 15 min and subjected to imaging using the ChemiDoc XRS + system (Bio-Rad, Hercules, CA, USA).

Single nucleus RNA-seq (snRNA-seq) data analysis

Data acquisition and processing. The fastq files for the snRNA-seq data generated from post-mortem human brains as part of the ROSMAP study were obtained from the AD Knowledge Portal (syn23650894). The relevant covariate information, including individual metadata, cell type annotations, and cell demultiplexing data, was accessed from the portal (syn3157322, syn21589957, syn31512863).

Transcript quantification and filtering. Transcripts were quantified from the fastq files using the Kallisto|bustools package.[44] The analysis focused specifically on transcripts corresponding to the *BIN1* gene. Cells with no expression of *BIN1* transcripts were excluded due to dropout events, leaving a final dataset of 418 individuals and approximately 770,000 viable cells with non-zero *BIN1* expression.

Isoform ratio analysis. For each cell, the ratio of the *BIN1* neuronal isoform to the total *BIN1* expression was calculated. This ratio was analyzed as both a continuous variable and a discrete variable, with cells classified as having high (≥0.45) or low (<0.45) *BIN1* neuronal isoform expression.

Cell type-specific analysis. Cell type annotations were obtained from the metadata provided in the AD Knowledge Portal, based on the annotations reported by Fujita et al.[20] Differential expression analysis of the *BIN1* isoform ratio was performed within specific cell types and subtypes, including *CUX2*(+) excitatory neurons in layers 2–4 (syn53694054) and *CUX2*(-) excitatory neurons (syn53694068). We also analyzed a subset of excitatory neurons with positive normalized expression (>=1) of both *RORB* and *CUX2*. Such Cux2+/Rorb + neurons have been identified as being selectively vulnerable in AD.[11,14–16,45]

Statistical analysis

The differential expression of the *BIN1* neuronal isoform ratio was evaluated using Student's t-test for the continuous variable and Fisher's Exact test for the discrete variable. In all the analyses, AD condition was defined by Braak-Stage alone (AD >=4, Normal <=1, or NotAD <4). Analyses were performed using the 'ggplot2', 'ggpubr', 'Seurat' and 'BPCells' packages in R.

Results

Identification of BIN1 SNP rs78710909 as a candidate AD risk variant

Using a multi-step filtering approach, we identified 149 SNPs and 119 associated genes with potential for AD risk (Supplemental Figure 1A). From this set, we prioritized 10 genes, 14 enhancers, and 23 SNPs based on overlap between regulatory regions of the genes and DNA repair hotspots (Supplemental Figure 1B).

Four SNPs, including rs78710909 and rs76516995 in the *BIN1* gene, showed significant allele-specific differential binding with known transcription factors based on data from the ANANASTRA database (Table 1, Figure 2). Specifically, the *BIN1* SNPs rs78710909 and rs76516995 demonstrated more than 5-fold change in allele-specific transcription factor binding.

BIN1 SNP rs78710909 is associated with higher AD risk. Age emerged as the most significant risk factor for AD in *APOE* ϵ4 non-carriers. For individuals above 84 years of age, there is a 7.9x-fold differential higher risk of AD compared to younger people (Supplemental Table 4A).

The *BIN1* SNP rs78710909 is in a significant AD-GWAS locus (Supplemental Table 2) and was identified in a fine-mapping study.[4] Since the homozygous risk allele genotype 'CC' is infrequent (only 2.8% of the samples), we focused on the AD risk of the 'C' allele compared to the reference allele 'G' after excluding the homozygous 'CC' genotype. The 'C' allele of the BIN1 SNP rs78710909 increased AD risk by 1.52 times (95% CI:0.98–2.38). Based on association analysis of the candidate SNPs, the 'T' allele of the *CD2AP* SNP rs1004173 and the 'A' allele of *TSPAN14* SNP rs1870140 are also associated with higher AD risk (Table 1). The results from the association analysis were concordant with the regression analysis of the SNPs of *BIN1* and *TSPAN14* with AD risk after adjusting for Age, however, the *CD2AP* SNP rs1004173 is not significant after adjusting for Age (Supplemental Table 5).

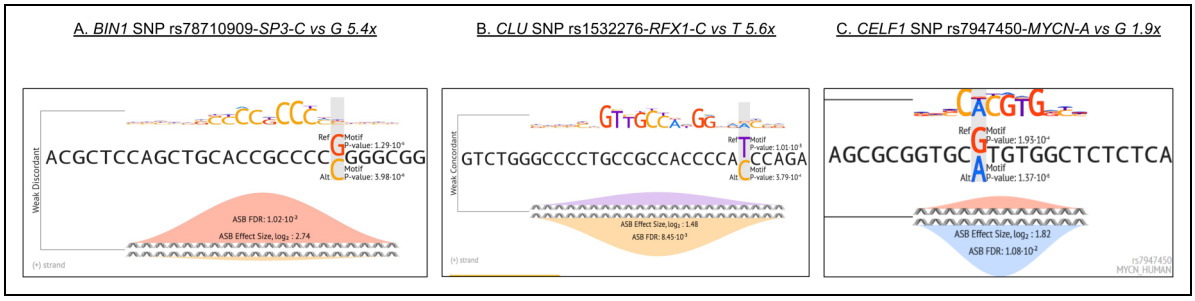

Figure 2. Four SNPs in the *BIN1*, *CLU*, and *CELF1* AD-GWAS loci showed allele-specific differential binding with transcription factors. The analysis of differential binding is based on data available from the ANANASTRA webservice- a meta-analysis of TF binding data from more than 15,000 ChIP experiments. Diff_Binding_fc is the difference between the effect size between the reference and effect alleles in ChIP experiments; Motif_log2_fc is the differential binding based on theoretical motif site binding analysis (A) rs78710909, hg38 coordinates: 2:127107346; Diff_binding_fc = 5.4; Motif_log2_fc = 1.7 (B) rs1532276, hg38 coordinates: 8:27608640; Diff_binding_fc = 5.6; Motif_log2_fc = 1.4 (C) rs7947450, hg38 coordinates: 11:47408353; Diff_binding_fc = 1.9; Motif_log2_fc = 3.8.

Experimental evidence of allele-specific enhancer and TF binding activity

BIN1 SNP rs78710909 allele 'C' shows higher enhancer activity. The pGL3-Promoter vector contains an SV40 promoter upstream of the Luciferase gene. DNA fragments containing putative enhancer elements can be inserted either upstream or downstream of the promoter-luc + transcriptional unit. To test the ability of the two alleles (G/C) of the *BIN1* SNP rs78710909 to differentially regulate gene transcription, the DNA fragment (DFrag) flanking the SNP rs7810909 was introduced into the pGL3-Promoter. The DFrag containing the alternate allele 'C' of the SNP rs78710909 demonstrated statistically significant higher enhancer activity compared to the reference allele 'G' (Figure 3).

BIN1 SNP rs78710909 alleles differentially bind p53 and E2F1. The Electrophoretic Mobility Shift Assay (EMSA) shows that the transcription factor p53 binds to DFrag and the binding affinity is higher with the DNA fragment containing the 'C' allele of rs78710909 (Figure 4A). The binding of p53 to the DFrag containing the 'C' allele of SNP rs78710909

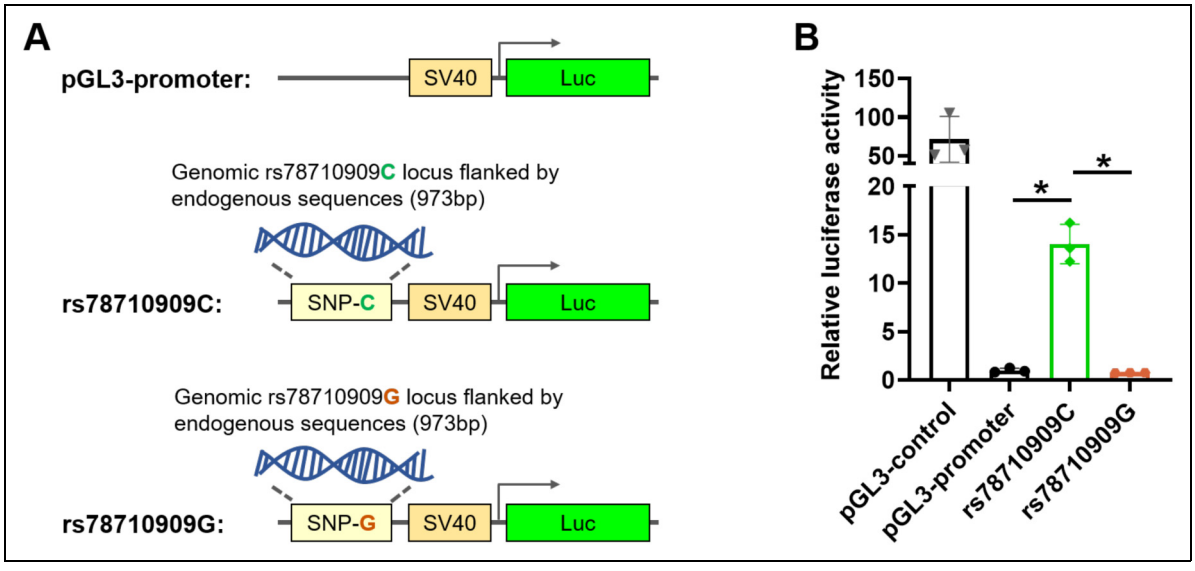

Figure 3. *BIN1* SNP rs78710909C allele, but not rs78710909G allele, showed enhancer activity. (A) Schematic representation of the luciferase reporter construct containing the corresponding SNPs. The indicated genomic sequences upstream of *BIN1* encompassing the corresponding SNPs were inserted into the upstream of pGL3-promoter vector. (B) The genomic fragment containing the *BIN1* SNP rs78710909C allele showed enhancer activity. HEK293T cells were cotransfected with pRL-CMV and a firefly luciferase reporter plasmid containing either the rs78710909C allele or rs78710909G allele. The pRL-CMV vector expressing Renilla luciferase was used as an internal control to calibrate differences in transfection and harvest efficiencies. The firefly luciferase activity of each sample was normalized to the Renilla luciferase activity. The pGL3-control vector was used as positive control.

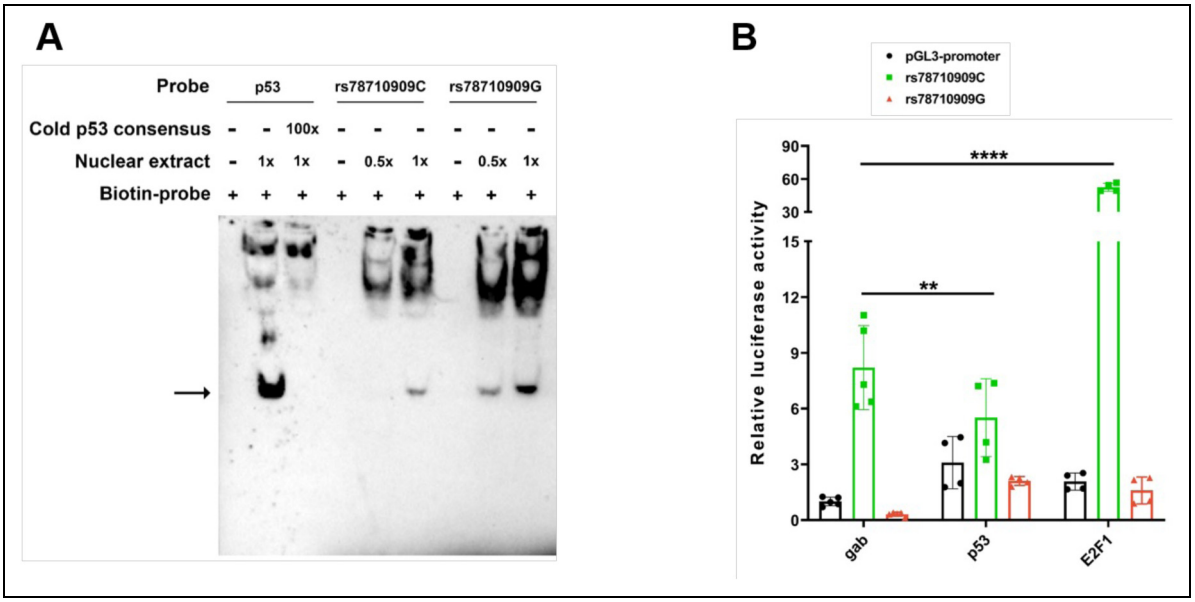

Figure 4. *BIN1* SNP rs78710909 alleles showed differential binding effect with p53 and E2F1. (A) Electrophoretic mobility shift assay for BIN1 SNP rs78710909 alleles. Analysis was performed in the presence (0.5x, 1x) or absence (-) of HEK293T nuclear extract. Each binding reaction contained biotin-labeled p53 (lanes 1–3) or rs78710909C (lanes 4–6) or rs78710909G (lanes 7–9) probes. A 100-fold excess of unlabeled (cold) p53 probes (lane 3) were included in the binding reactions as specific competitors. Labeled probes incubated without the nuclear extracts were included as negative controls (lanes 1, 4, and 7). The arrow indicated the DNA-protein complex. (B) The enhancer activity of BIN1 was differentially affected by p53 and E2F1. HEK293T cells were co-transfected with pRL-CMV, corresponding firefly luciferase reporter, and empty vector (gab), or p53-expressing vector (p53), or E2F1-expressing vector (E2F1). Luciferase activity analysis was performed as in Figure 3B.

seems to dampen its enhancer activity (as measured by relative Luciferase activity) but the binding of E2F1 significantly increases the enhancer activity (Figure 4B).

The mRNA expression of BIN1 neuronal isoforms is lower in AD individuals

Based on the individuals with both WGS and RNA-seq data, the *BIN1* gene is not differentially expressed at a statistically significant level between the AD and Normal states. However, in the RNA-seq data at the transcript level from the ROSMAP study, in AD compared to Normal individuals, the mRNA expression of *BIN1* neuronal isoform transcripts is lower ($p = 0.01$) and the ratio of *BIN1* neuronal to ubiquitous isoform expression is lower ($p = 0.04$, Figure 5).

Excitatory neurons in the superficial neocortical layers expressing lower ratio of the BIN1 neuronal isoforms are more vulnerable

The excitatory neurons in the superficial layers (L2-L4) of the entorhinal cortex and neocortex have been reported to be selectively vulnerable in AD.[11,14–16,45] Moreover, higher expression of *CUX2* has been identified as a marker of excitatory neurons in the L2-L4 superficial layers of the neocortex[19,20,46] and within these *CUX2* + neurons, Cux2+/Rorb + cells are selectively vulnerable in AD.[16] In *CUX2* + neurons, in cells with a 'low' versus 'high' (high: >=0.45) ratio of neuronal isoform to total *BIN1,* 108 genes are upregulated with a minimum 25% difference in percentage of cells expressing the gene (1 – pct.1/pct.2) and adjusted p-value <0.05. Of these 108 genes, 75 genes are associated with pathways related to AD such as synaptic membrane activity, apoptosis, GTPase activity, immune response, vesicle transport, DNA damage response, oxidative stress, and p53 pathway (Supplemental Tables 8 and 9).

The top five upregulated genes based on adjusted p-value, between cells with a 'low' versus 'high' (high: >=0.45) ratio of neuronal isoform to total *BIN1,* are *SSTR2, COL5A2, NEFH, VAMP1, and SMIM10L2B* (Table 2). *SSTR2* is a somatostatin receptor in excitatory neurons and its activation is linked to negative regulation of cell proliferation, immune response, and phosphorylation of tau.[47] *COL5A2* encodes the alpha-chain of collagen-V and its higher expression has been associated with selectively vulnerable neurons early in AD[48] and it is reported to be a marker of poor prognosis in colorectal cancer possibly via the Wnt or mTOR pathway.[49] *NEFH* encodes the heavy neurofilament protein and its higher expression is used as a biomarker for neuronal damage in amyotrophic

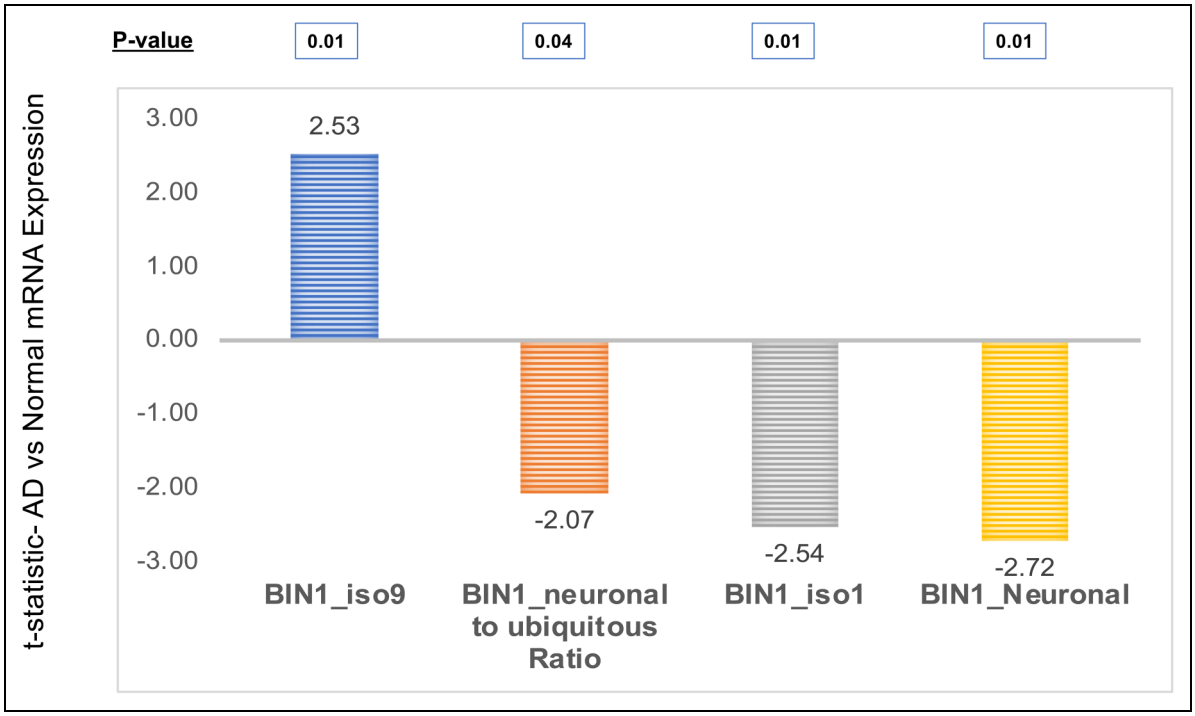

Figure 5. The mRNA expression of *BIN1* neuronal isoforms is lower and of ubiquitous isoforms is higher in patients with AD versus normal. The bulk RNA-Seq data is from the Synapse website for the ROSMAP RNA expression study, 365 Samples selected for analysis based on Age > 60, and AD defined based Braak stage only (AD >=4 and Normal <=1). Target genes chosen based on Supplemental Table 1. See Method for details. *BIN1* neuronal isoforms (12,3) seem to be lower, and the ratio of neuronal to ubiquitous isoforms (5,6,7,9,10,12) seem to be lower in patients with AD.

lateral sclerosis[50] and its expression is higher in AD compared to control patients.[51] *VAMP1*, part of the synaptobrevin family, encodes a vesicle-associated membrane protein which is an important part of the complex involved in docking and/or fusion of synaptic vesicles with the cell membrane and its lower expression is reported to be protective in AD.[52] *SMIM10L2B,* a highly conserved membrane protein, functions as an RNA to regulate enhancers, and *SMIM10L2A/SMIM10L2B* play central roles in the regulation of cell cycle, DNA repair processes and p53 response.[53,54] The Dotplot (Figure 6) and Heatmap (Figure 7, Supplemental Figure 4, Supplemental Table 8) indicate that 34 of the top 40 differentially expressed genes are related to AD pathways and are also upregulated in cells of AD patients with low ratio of the *BIN1* neuronal isoform and suppressed in cells of 'Not-AD' individuals with high expression of the neuronal isoform.

Ratio of neuronal isoforms to total BIN1 expression is lower with the rs78710909C allele in vulnerable neurons in AD

In bulk RNA-seq data, no statistically significant difference was observed in *BIN1* isoform expression based on the alleles of rs78710909. Based on snRNA-seq data from the ROSMAP study, we measured the celltype specific effect of SNP rs78710909 alleles on *BIN1* isoform expression. Among AD patients, Cux2+/Rorb + excitatory neurons have been reported to be selectively vulnerable[16] and such cells seem to cluster together well in two-dimensional UMAP space (Supplemental Figure 4). We identified 73 genes upregulated (>1 log-fold change) between Cux2 +/Rorb + cells and other L2-L4 excitatory neurons. The top 3 TFs associated with these genes, based on consensus of CHEA and ENCODE from the *Enrichr* webservice,[55] are *SUZ12*, *TP53*, and *NFE2l2 (*Supplemental Table 10). These TFs are associated with neuronal stress response: *SUZ12* is associated with decreased neuronal function,[56] *TP53* regulates cell cycle arrest, apoptosis and DNA damage response,[57] and *NFE2l2* encodes the NRF2 protein which is the master regulator of oxidative stress response.[58]

In excitatory neurons expressing *BIN1's* neuronal isoform, AD related genes which were upregulated in cells with low ratio of *BIN1*'s neuronal isoform, are also upregulated in Cux2+/Rorb + cells but not in other L2-L4 excitatory neurons. Specifically, these AD genes were predominantly expressed in Cux2+/Rorb + cells which had a low ratio of *BIN1*'s neuronal isoform (Figure 8). Within Cux2+/Rorb + cells in AD but not in 'Not-AD' patients, the effect allele 'C' of SNP rs78710909 is associated with lower percentage of cells with high ratio of neuronal

Table 2. In excitatory neurons expressing *CUX2*, AD related genes are the top upregulated genes in cells with 'low' versus 'high' ratio of neuronal isoform to total *BIN1*. Cells with neuronal isoform ratio >=0.45 were labeled 'high' (pct.1) and others 'low' (pct.2). The differential expression analysis was performed using FindMarkers function between these categories of cells with at least 25% of cells expressing the gene and minimum difference in percentage of cells (1 −pct.1/pct.2) between the two conditions of 25%, and $p < 0.05$. Of the 108 differentially expressed genes, 75 are associated with AD related terms (Supplemental Table 8).

	Gene Name	Avg Log2fc	pct.1	pct.2	Adj p	Comments
SSTR2	somatostatin receptor 2	−0.75	0.32	0.58	1.9×10^{-8}	Negative regulation of cell proliferation
COL5A2	collagen type V alpha 2 chain	−0.56	0.51	0.78	4.5×10^{-8}	Marker of vulnerable neurons
SMIM10L2B	small integral membrane protein 10 like 2B	−0.69	0.25	0.46	1.3×10^{-5}	DNA damage response, cell cycle regulation, p53 pathway
NEFH	neurofilament heavy chain	−1.10	0.23	0.45	1.3×10^{-7}	Marker of neurodegeneration
VAMP1	Vesicle associated membrane protein 1	−0.66	0.37	0.59	1.2×10^{-5}	Vesicle transport, synaptic membrane, downregulation is protective

isoform expression ($p < 0.007$) and the average neuronal isoform ratio is lower ($p < 0.01$, Figure 9). No statistically significant difference was observed in inhibitory neurons and oligodendrocytes (Supplemental Figure 5) and the neuronal isoform was not meaningfully expressed in astrocytes and microglia.

Protein expression of BIN1 neuronal isoforms is lower with rs78710909C

Based on the TMT Proteomics dataset from the ROSMAP study, in *APOE*-ϵ4 non-carriers, the regression analysis of BIN1 protein isoforms' expression with the SNP

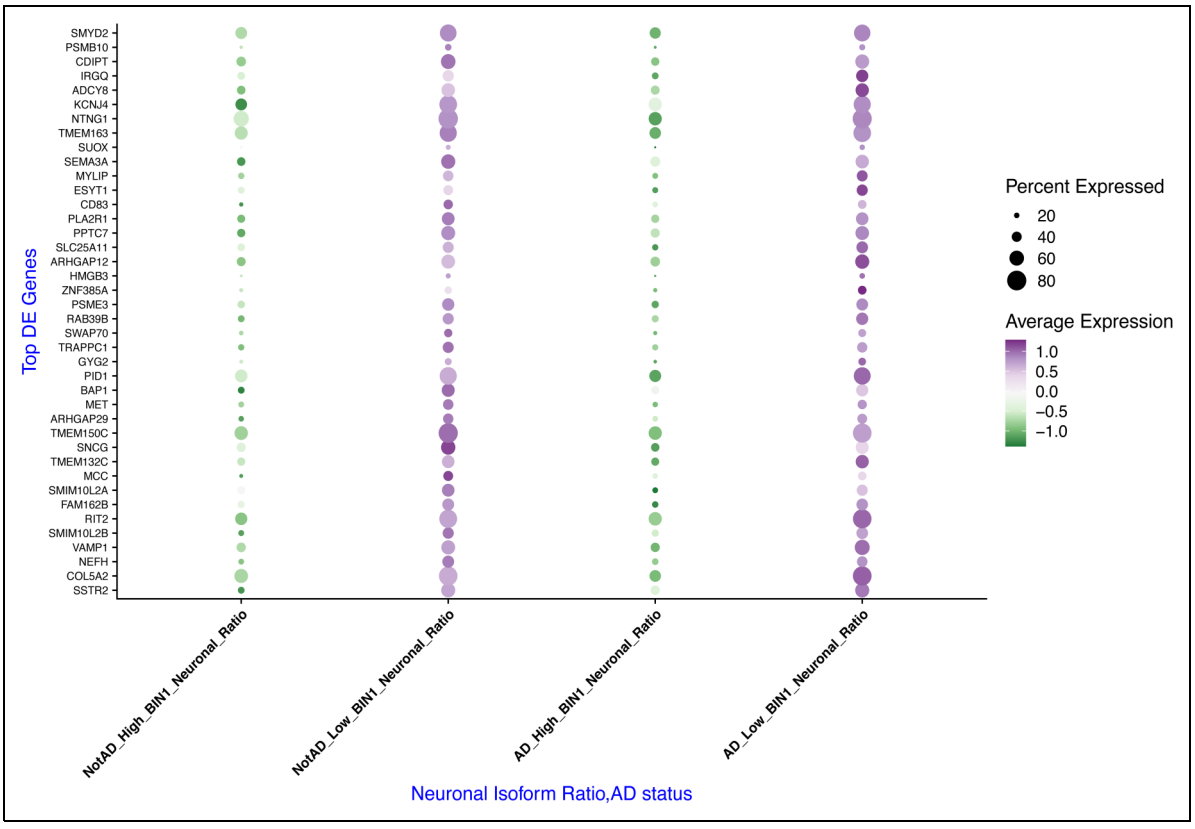

Figure 6. In excitatory neurons expressing *CUX2*, AD related genes are upregulated in cells with 'low' versus 'high' ratio of neuronal isoform to total *BIN1*, thus suggesting they are selectively vulnerable. Cells with neuronal isoform ratio >=0.45 were labeled 'high' and others 'low'. The differential expression analysis was performed using FindMarkers function between these categories of cells with at least 25% of cells expressing the gene and minimum difference in percentage of cells (1 −pct.1/pct.2) between the two conditions of 25%, and $p < 0.05$. The Dotplot shows that the top 40 differentially expressed genes are elevated in cells of AD patients with low ratio of the *BIN1* neuronal isoform and suppressed in cells of not AD patients with high expression of the neuronal isoform. Of the 108 differentially expressed genes, 75 are associated with AD related terms (Supplemental Table 8).

Figure 7. In excitatory neurons expressing *CUX2*, AD related genes are the top upregulated genes in cells with 'low' versus 'high' ratio of neuronal isoform to total *BIN1*. Cells with neuronal isoform ratio >=0.45 were labeled 'high' (pct.1) and others 'low' (pct.2). The differential expression analysis was performed using FindMarkers function between these categories of cells with at least 25% of cells expressing the gene and minimum difference in percentage of cells (1 –pct.1/pct.2) between the two conditions of 25%, and $p < 0.05$. Of the 108 differentially expressed genes, 75 are associated with AD related terms (Supplemental Table 8).

rs78710909 shows that the expression of non-neuronal isoforms is higher ($p < 0.03$) and the ratio of the expression of neuronal to non-neuronal isoforms is lower with the effect allele 'C' of the SNP rs78710909 with a p-value of 0.007 (Supplemental Table 11).

Discussion

In this study, we systematically explored the association of AD-GWAS risk variants with DNA repair hotspots and the underlying mechanisms by which prioritized SNPs and genes may affect AD pathology in selectively vulnerable neuronal populations. We filtered the *BIN1* SNP rs78710909, which is located within a significant AD-GWAS locus, a DNA repair hotspot, and the *BIN1* promoter/enhancer region. Our analysis revealed that the 'C' allele of rs78710909 confers a 1.52-fold higher risk of AD in *APOE* ϵ4 non-carriers and exhibits a 5.4-fold difference in transcription factor binding affinity compared to the reference 'G' allele. *In vitro* experiments further demonstrated that the 'C' allele is associated with increased enhancer activity, weaker binding of the transcription factor p53, but stronger binding of E2F1. These findings suggest that the effect allele may disrupt the balance between p53-mediated repression and E2F1-driven activation of *BIN1* expression. Importantly, our study provides novel insights into the isoform-specific effects of the rs78710909 SNP on *BIN1* expression. We found that the 'C' allele is associated with a lower ratio of the neuronal isoforms of *BIN1* in selectively vulnerable Cux2+/Rorb + excitatory neurons within the superficial layers of the dorsolateral prefrontal cortex. This reduction in the neuronal *BIN1* isoform ratio was specific to AD patients, indicating a disease-relevant mechanism. Further analysis revealed that genes upregulated in neurons with

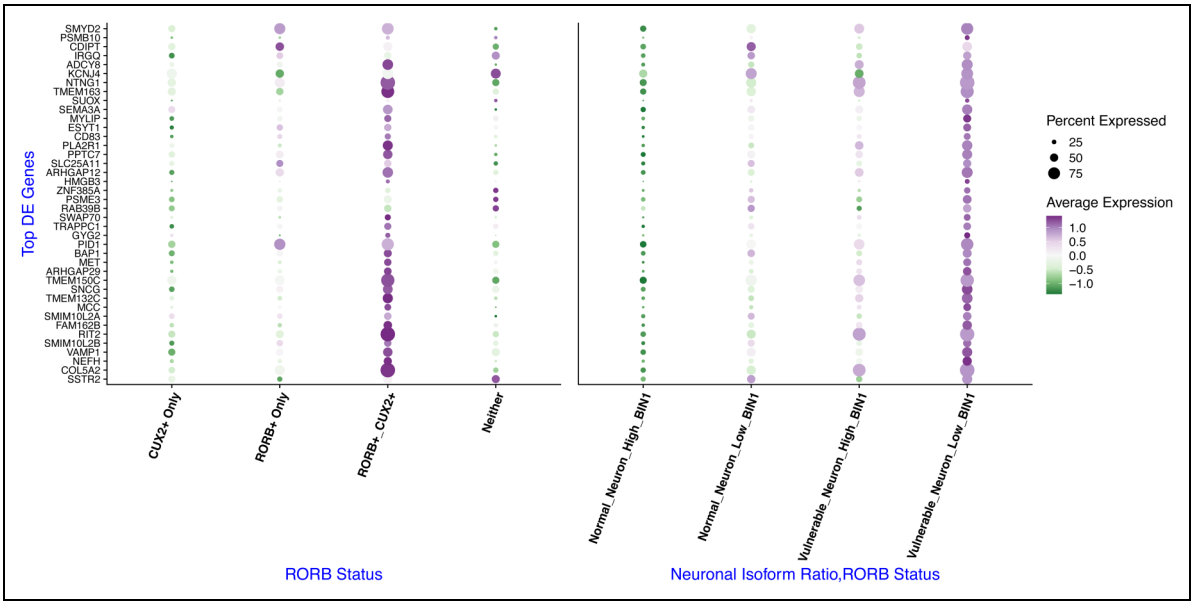

Figure 8. The expression of AD related genes upregulated in cells with lower ratio of the *BIN1* neuronal isoform are upregulated only in putatively vulnerable cells expressing both *CUX2* and *RORB*. The snRNA-seq data is from the ROSMAP single nucleus RNA-seq (snRNA) of 424 individuals available from the Synapse website. AD was defined based on Braak-Stage (AD:>=4), and "high" neuronal ratio was defined as >=0.45. The L2-L4 excitatory neurons expressing both *CUX2* and *RORB* have been reported to be the most vulnerable in AD. Such cells have been labeled "Vulnerable" and rest "Normal".

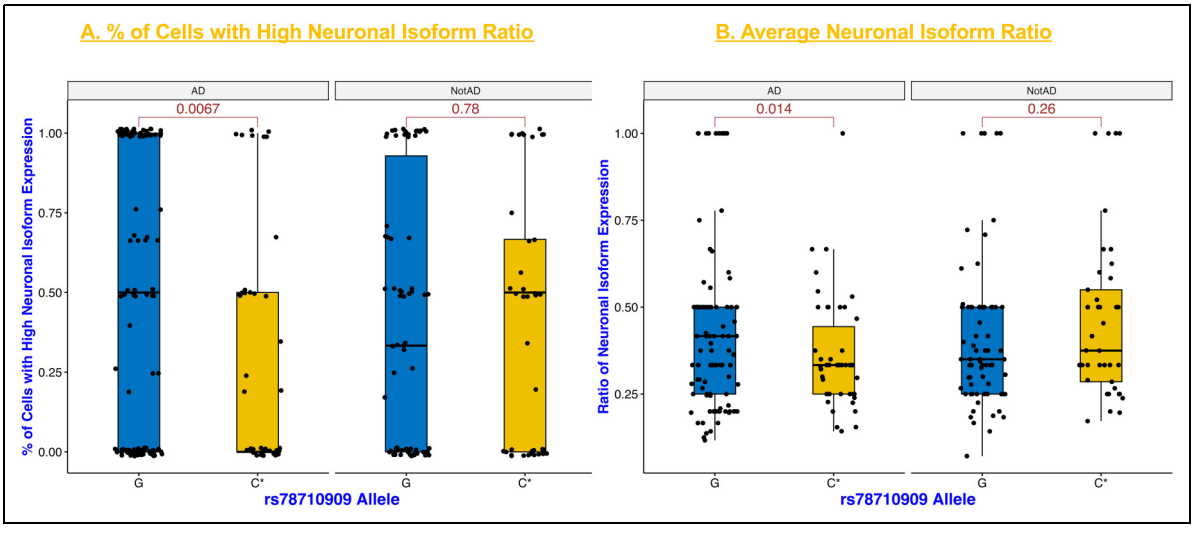

Figure 9. With the effect allele C of the SNP rs78710909, the expression of *BIN1* neuronal isoforms is lower in vulnerable excitatory neurons in AD patients. The snRNA-seq data is from the ROSMAP single nucleus RNA-seq (snRNA) of 424 individuals available from the Synapse website. C* allele represents (GC*+C*C*) and G allele represents GG. t-test was used to compare means. AD was defined based on Braak-Stage (AD:>=4), and "high" neuronal ratio was defined as >=0.45. The L2-L4 *CUX2* and *RORB* expressing excitatory neurons have been reported to be the most vulnerable in AD. With the effect allele C, across all cells in this subgroup with non-zero neuronal isoform ratio expression for each individual, AD patients have lower expression of (A) % of cells with high ratio of neuronal isoform and (B) the average ratio of neuronal isoforms. No significant differential expression was observed in 'Not AD' patient.

a lower *BIN1* neuronal isoform ratio are associated with AD-related pathways, including synaptic function, markers of neurodegeneration, apoptosis, immune response, DNA damage, and oxidative stress. Notably, these stress-related genes were selectively upregulated in the vulnerable Cux2+/Rorb + excitatory neurons, but not in other layer 2–4 excitatory neuronal populations. Taken together, our findings suggest a novel pathogenic

pathway linking the *BIN1* SNP rs78710909, reduced expression of *BIN1*'s neuronal isoforms, and the vulnerability of specific cortical excitatory neuron subtypes to AD-related stress and neurodegeneration.

We selected the *BIN1* SNP rs78710909 for further investigation (Figure 1, Table 1) because the locus of *BIN1* SNP rs78710909 (G/C) satisfied several criteria. This SNP locus is significant at a genome-wide level[4] and has been identified in a finemapping study,[4] it is located within a DNA repair hotspot and *BIN1* promoter/enhancer region GH02J127099 (Supplemental Figure 2),[31] the allele 'C' of the SNP rs78710909 is associated with higher AD risk based on WGS data with AD defined based on Braak-stage only (Table 1), and there is a 5.4× difference in the binding affinity between the reference and effect allele to specific transcription factors (Figure 2). Among the SNPs in high LD (~1) with rs78710909, we decided to focus on the SNP rs78710909 because phylogenetic and epigenetic analysis indicates that this SNP is more evolutionarily conserved and likely to be more functionally significant (Table 1). This SNP is not in LD (r2 or D') with the top SNPs in the *BIN1* locus, rs6733839 and rs744373[5,6] (Supplemental Figure 3) and the functional mechanism maybe distinct from that of rs6733839 and rs744373.

Our *in vitro* experimental data suggests that the 'C' allele of rs78710909 has higher enhancer activity compared to the 'G' allele (Figure 3). The TF p53 binds preferentially to the allele 'G' (Figure 4A). Interestingly, the enhancer activity of the 'C' allele was suppressed by binding to p53 but was significantly stronger upon binding with the transcription factor E2F1 (Figure 4B). These results indicate that p53 binding could be inhibiting an E2F1-dependent increase in BIN1 activity by competitively binding the same enhancer and this inhibitory role of p53 could be negatively affected due to weaker binding with the effect allele 'C' compared to the major allele 'G' of rs78710909.

The understanding of *BIN1*'s function in AD is complicated by the observation that the expression of the 14 *BIN1* isoforms is not only different by celltype but also by AD status.[7,8,59] Research indicates that the expression of the ubiquitous *BIN1* isoforms is elevated in AD, and that of the neuronal isoforms is reduced.[8,60,61] It has been proposed that the BIN1 neuronal isoforms (1–3) are localized to the cytoplasm and may serve neuroprotective functions. These isoforms potentially facilitate endocytosis of key molecules in the Aβ pathway,[62] negatively regulate the amyloidogenic processing of APP by BACE1, mitigate excitotoxicity associated with calcium channels, and sequester tau protein.[63–68] In contrast, the ubiquitous isoforms (6,9,10,12) are predominantly localized in the nucleus and may negatively influence DNA repair and damage response pathways, promote various aspects of the development and propagation of tauopathy,[69,70] and apoptosis in both neuronal and glial cells through E2F1, c-Myc, and PARP-1 dependent mechanisms.[71–77] The role of the ubiquitous BIN1 isoforms in regulating the inflammatory response seems to be nuanced. In the early stages of AD, lower *BIN1* microglial expression may lead to the accumulation of AD stress due to inadequate microglial response. In the later stages of AD, a higher level of microglial *BIN1* may be involved in triggering an excessive inflammatory response and by favoring the exocytosis of phagocytosed tau resulting in phosphorylated tau propagation.[59,70]

The rs78710909-C allele is associated with a higher risk of AD (Table 1). Given the salience of isoform-specific expression of *BIN1*, we verified that the mRNA expression of neuronal isoforms is lower and ubiquitous isoforms is higher in AD compared to Normal individuals (Figure 5). While we did not find allele-specific differential isoform expression in bulk RNA-seq data, our analysis has revealed celltype specific effects in AD. In Cux2+/Rorb + excitatory neurons, the effect allele 'C' of the *BIN1* SNP rs78710909 was associated with the lower average ratio of the neuronal isoform to total *BIN1* expression and a lower percentage of cells expressing a high ratio of the neuronal isoform of *BIN1*. These results were observed in individuals diagnosed with AD but not others, thus suggesting a disease-relevant mechanism (Figure 9). No difference was observed in other excitatory neurons, inhibitory neurons and oligodendrocytes (Supplemental Figure 6). This implies an underlying mechanism specific to AD pathology and in Cux2 +/Rorb + neurons which have previously been identified as selectively vulnerable in AD.[16,17,20,22]

The role of rs78710909-C in alternative splicing of *BIN1* is supported by the lower expression of BIN1 neuronal protein isoforms in *APOE* ϵ4 non-carriers (Supplemental Table 11). The 'C' allele is associated with alternative splicing that favors the exclusion of the exon specific to the neuronal isoforms, particularly exon '7,' leading to diminished expression of these isoforms. Our analyses suggest this may involve differential binding of splicing factors like ILF2, RBM8A, and EWSR1 to the genomic region surrounding the SNP rs78710909 (Supplemental Table 12). The precise mechanisms underlying allele-specific alternative splicing remain to be elucidated. It is plausible that mechanisms leading to the altered isoform ratio of BIN1 are triggered by neuronal stress conditions that are experienced during AD progression. Notably, splicing factors ILF2, RBM8A, and EWSR1 are implicated in pathways commonly disrupted in AD, including DNA damage response, autophagy, and apoptosis.[78–80]

Importantly, we observed that in excitatory neurons in the superficial layers (L2-L4) of the dorsolateral prefrontal cortex, the genes associated with hallmarks of AD pathology were upregulated in neurons expressing a "low" versus "high" ratio of neuronal isoform to total *BIN1* (Figure 6, Figure 7, Table 2). Further, we found that these neuronal stress related genes are upregulated only in Cux2+/Rorb + neurons[16,17,20] but not in other excitatory

neurons (Figure 8, Supplemental Figure 4). This selective expression of neuronal stress related genes suggests that the cells with "low" expression of *BIN1*'s neuronal isoform are a vulnerable population in AD.

Taken together, our results suggest that the *BIN1* SNP rs78710909 contributes to AD pathogenesis by shifting the balance of *BIN1* isoform expression, with reduced levels of the neuroprotective neuronal isoforms and increased levels of the ubiquitous isoforms involved in deleterious processes like tauopathy and apoptosis. This effect is particularly observed in selectively vulnerable Cux2+/Rorb + excitatory neuron population in AD individuals and thus suggesting a disease-relevant mechanism. It is possible that the underlying mechanism is triggered during neuronal stress conditions that prevail in AD pathology.

Conclusions and next steps

The *BIN1* locus is the second most significant in AD-GWAS.[6] We propose that the *BIN1* SNP rs78710909 may be involved in the underlying mechanism linking *BIN1* and selectively vulnerable neurons to AD pathology. The 'C' allele of the SNP rs78710909 is associated with lower expression of neuronal isoforms of *BIN1* in selectivley vulnerable Cux2+/Rorb + excitatory neurons in AD patients. Further, genes related to hallmarks of AD pathology are upregulated in cells with lower ratio of *BIN1*'s neuronal isoform, and the same genes are upregulated in Cux2+/Rorb + neurons but not other types of excitatory neurons. Our evidence supports the hypothesis that the *BIN1* SNP rs78710909 lowers the ratio of *BIN1*'s neuronal isoform in Cux2+/Rorb + excitatory neurons and increases their vulnerability to neuronal stress in AD.

The key strength of our study is that it utilizes *in vitro* studies and integrates information from large publicly available bulk RNA-seq, proteomics, and snRNA-seq databases to investigate allele and cell-type-specific effects. Future studies can help to clarify the effect of the 'C' allele of the *BIN1* SNP rs78710909 on *BIN1*'s isoform and celltype specific expression and whether this is affected by Aβ accumulation, tau, age, oxidative stress, and DNA damage.

Acknowledgments

The results published here are in whole or in part based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.org). Data generation was supported by the following NIH grants: P30AG10161, P30AG72975, R01AG15819, R01AG17917, R01AG036836, U01AG46152, U01AG61356, U01AG046139, P50 AG016574, R01 AG032990, U01AG046139, R01AG018023, U01AG006576, U01AG006786, R01AG025711, R01AG017216, R01AG003949, R01NS080820, U24NS072026, P30AG19610, U01AG046170, RF1AG057440, and U24AG061340, and the Cure PSP, Mayo and Michael J Fox foundations, Arizona Department of Health Services and the Arizona Biomedical Research Commission. We thank the participants of the Religious Order Study and Memory and Aging projects for the generous donation, the Sun Health Research Institute Brain and Body Donation Program, the Mayo Clinic Brain Bank, and the Mount Sinai/JJ Peters VA Medical Center NIH Brain and Tissue Repository. Data and analysis contributing investigators include Nilüfer Ertekin-Taner, Steven Younkin (Mayo Clinic, Jacksonville, FL), Todd Golde (University of Florida), Nathan Price (Institute for Systems Biology), David Bennett, Christopher Gaiteri (Rush University), Philip De Jager (Columbia University), Bin Zhang, Eric Schadt, Michelle Ehrlich, Vahram Haroutunian, Sam Gandy (Icahn School of Medicine at Mount Sinai), Koichi Iijima (National Center for Geriatrics and Gerontology, Japan), Scott Noggle (New York Stem Cell Foundation), Lara Mangravite (Sage Bionetworks). Study data were provided through the Accelerating Medicine Partnership for AD (U01AG046161 and U01AG061357) based on samples provided by the Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago. Data collection was supported through funding by NIA grants P30AG10161, R01AG15819, R01AG17917, R01AG30146, R01AG36836, U01AG32984, U01AG46152, the Illinois Department of Public Health, and the Translational Genomics Research Institute. Study data were generated from postmortem brain tissue provided by the Religious Orders Study and Rush Memory and Aging Project (ROSMAP) cohort at Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago. This work was funded by NIH grants U01AG061356 (De Jager/Bennett), RF1AG057473 (De Jager/Bennett), and U01AG046152 (De Jager/Bennett) as part of the AMP-AD consortium, as well as NIH grants R01AG066831 (Menon) and U01AG072572 (De Jager/St George-Hyslop). We express our thanks to the kind support provided by Dr Yuanhua Huang and Ms. Ruiyan Hou for their guidance regarding BIN1 isoform and celltype specific analysis.

ORCID iDs

Georgy Sapozhnikov https://orcid.org/0009-0002-6884-0797
Shoutang Wang https://orcid.org/0000-0003-1049-9127
You-Qiang Song https://orcid.org/0000-0001-9407-2256

Statements and declarations

Author contributions

Rajesh Ranganathan (Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Validation; Visualization; Writing – original draft; Writing – review & editing); Siwen Li (Formal analysis; Investigation; Methodology; Validation; Visualization; Writing – original draft; Writing – review & editing); Georgy Sapozhnikov (Data curation; Formal analysis; Investigation; Methodology; Writing – original draft; Writing – review & editing); Shoutang Wang (Conceptualization; Project administration; Supervision; Writing – review & editing); You-Qiang Song (Conceptualization; Data curation; Formal analysis; Funding

acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Writing – review & editing).

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported in part by grants from the Innovation and Technology Commission – Hong Kong (MRP/056/21, PiH/234/22, InP/053/23 and PiH/086/22 received by Dr Youqiang Song).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability

The data supporting the findings of this study were derived from the resources available in the AD Knowledge Portal (https://adknowledgeportal.org). The data is available for download after obtaining applicable approval directly from the portal. The specific data files used in this study have been described in the Methods section of the manuscript.

Supplemental material

Supplemental material for this article is available online.

References

1. Jack CR, Bennett DA, Blennow K, et al. NIA-AA Research framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement* 2018; 14: 535–562.
2. 2021 Alzheimer's disease facts and figures. *Alzheimers Dement* 2021; 17: 327–406.
3. Breijyeh Z and Karaman R. Comprehensive review on Alzheimer's disease: causes and treatment. *Molecules* 2020; 25: 5789.
4. Schwartzentruber J, Cooper S, Liu JZ, et al. Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer's disease risk genes. *Nat Genet* 2021; 53: 392–402.
5. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* 2013; 45: 1452–1458.
6. Bellenguez C, Grenier-Boley B, Holmans PA, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet* 2022; 54: 412–414.
7. Nott A, Holtman IR, Coufal NG, et al. Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science* 2019; 366: 1134–1139.
8. Taga M, Petyuk VA, White C, et al. BIN1 Protein isoforms are differentially expressed in astrocytes, neurons, and microglia: neuronal and astrocyte BIN1 are implicated in tau pathology. *Mol Neurodegener* 2020; 15: 44–44.
9. De Rossi P, Buggia-Prevot V, Clayton BLL, et al. Predominant expression of Alzheimer's disease-associated BIN1 in mature oligodendrocytes and localization to white matter tracts. *Mol Neurodegener* 2016; 11: 59.
10. Keren-Shaul H, Spinrad A, Weiner A, et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* 2017; 169: 1276–1290.e1217.
11. Stranahan AM and Mattson MP. Selective vulnerability of neurons in layer ii of the entorhinal cortex during aging and Alzheimer's disease. *Neural Plasticity* 2010; 2010: 108190–108198.
12. Muratore CR, Zhou C, Liao M, et al. Cell-type dependent Alzheimer's disease phenotypes: probing the biology of selective neuronal vulnerability. *Stem Cell Reports* 2017; 9: 1868–1884.
13. Fu H, Possenti A, Freer R, et al. A tau homeostasis signature is linked with the cellular and regional vulnerability of excitatory neurons to tau pathology. *Nat Neurosci* 2019; 22: 47–56.
14. Gomez-Isla T, Price JL, McKeel Jr DW, et al. Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. *J Neurosci* 1996; 16: 4491–4500.
15. Braak H and Del Tredici K. Spreading of tau pathology in sporadic Alzheimer's disease along cortico-cortical top-down connections. *Cereb Cortex* 2018; 28: 3372–3384.
16. Leng K, Li E, Eser R, et al. Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nat Neurosci* 2021; 24: 276–287.
17. Otero-Garcia M, Mahajani SU, Wakhloo D, et al. Molecular signatures underlying neurofibrillary tangle susceptibility in Alzheimer's disease. *Neuron* 2022; 110: 2929–2948.e8.
18. Koffie RM, Meyer-Luehmann M, Hashimoto T, et al. Oligomeric amyloid β associates with postsynaptic densities and correlates with excitatory synapse loss near senile plaques. *Proc Natl Acad Sci U S A* 2009; 106: 4012–4017.
19. Lake BB, Ai R, Kaeser GE, et al. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* 2016; 352: 1586–1590.
20. Fujita M, Gao Z, Zeng L, et al. Cell subtype-specific effects of genetic variation in the Alzheimer's disease brain. *Nat Genet* 2024; 56: 605–614.
21. Hodge RD, Bakken TE, Miller JA, et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* 2019; 573: 61–68.
22. Roussarie J-P, Yao V, Rodriguez-Rodriguez P, et al. Selective neuronal vulnerability in Alzheimer's disease: a network-based analysis. *Neuron* 2020; 107: 821–835.e812.
23. Lodato MA, Rodin RE, Bohrson CL, et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* 2018; 359: 555–559.
24. Ranganathan R, Sapozhnikov G, Ni W, et al. Recent developments in the role of DNA damage response and understanding its implications for new therapeutic approaches in Alzheimer's disease. *Transl Med Aging* 2023; 7: 52–65.
25. Yankner BA, Lu T, Pan Y, et al. Gene regulation and DNA damage in the ageing human brain. *Nature* 2004; 429: 883–891.

26. Reid DA, Reed PJ, Schlachetzki JCM, et al. Incorporation of a nucleoside analog maps genome repair sites in postmitotic human neurons. *Science* 2021; 372: 91–94.
27. Stott RT, Kritsky O and Tsai L-H. Profiling DNA break sites and transcriptional changes in response to contextual fear learning. *PLoS One* 2021; 16: e0249691.
28. Nouspikel T and Hanawalt PC. Terminally differentiated human neurons repair transcribed genes but display attenuated global DNA repair and modulation of repair gene expression. *Mol Cell Biol* 2000; 20: 1562–1570.
29. Watanabe K, Taskesen E, van Bochoven A, et al. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017; 8: 1826–1811.
30. Oscanoa J, Sivapalan L, Gadaleta E, et al. SNPnexus: a web server for functional annotation of human genome sequence variation (2020 update). *Nucleic Acids Res* 2020; 48: W185–W192.
31. Machiela MJ and Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 2015; 31: 3555–3557.
32. Fishilevich S, Nudel R, Rappaport N, et al. Genehancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)* 2017; 2017: bax028.
33. Boytsov A, Abramov S, Aiusheeva AZ, et al. ANANASTRA: annotation and enrichment analysis of allele-specific transcription factor binding at SNPs. *Nucleic Acids Res* 2022; 50: W51–W56.
34. Paz I, Kosti I, Ares M, et al. RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucl Acids Res* 2014; 42: W361–W367.
35. Greenwood AK, Montgomery KS, Kauer N, et al. The AD knowledge portal: a repository for multi-omic data on Alzheimer's disease and aging. *Curr Protoc Hum Genet* 2020; 108: e105.
36. De Jager PL, Ma Y, McCabe C, et al. A multi-omic atlas of the human frontal cortex for aging and Alzheimer's disease research. *Sci Data* 2018; 5: 180142.
37. Wang M, Beckmann ND, Roussos P, et al. The Mount Sinai cohort of large-scale genomic, transcriptomic and proteomic data in Alzheimer's disease. *Sci Data* 2018; 5: 180185.
38. Allen M, Carrasquillo MM, Funk C, et al. Human whole genome genotype and transcriptome data for Alzheimer's and other neurodegenerative diseases. *Sci Data* 2016; 3: 160089.
39. Higginbotham L, Ping L, Dammer EB, et al. Integrated proteomics reveals brain-based cerebrospinal fluid biomarkers in asymptomatic and symptomatic Alzheimer's disease. *Sci Adv* 2020; 6: eaaz9360.
40. Zhou J and Troyanskaya OG. Predicting effects of noncoding variants with deep learning-based sequence model. *Nat Methods* 2015; 12: 931–934.
41. Kennedy RE, Cutter GR and Schneider LS. Effect of APOE genotype status on targeted clinical trials outcomes and efficiency in dementia and mild cognitive impairment resulting from Alzheimer's disease. *Alzheimers Dement* 2014; 10: 349–359.
42. McLure KG and Lee PWK. How p53 binds DNA as a tetramer. *EMBO J* 1998; 17: 3342–3350.
43. Ono R, Kaisho T and Tanaka T. PDLIM1 Inhibits NF-kappa B-mediated inflammatory signaling by sequestering the p65 subunit of NF-kappa B in the cytoplasm. *Sci Rep* 2015; 5: 18327.
44. Melsted P, Booeshaghi AS, Liu L, et al. Modular, efficient and constant-memory single-cell RNA-Seq preprocessing. *Nat Biotechnol* 2021; 39: 813–818.
45. Zhao R, Grunke SD, Wood CA, et al. Activity disruption causes degeneration of entorhinal neurons in a mouse model of Alzheimer's circuit dysfunction. *eLife* 2022; 11: e83813.
46. Tasic B, Yao Z, Graybuck LT, et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature* 2018; 563: 72–78.
47. Song Y-H, Yoon J and Lee S-H. The role of neuropeptide somatostatin in the brain and its application in treating neurological disorders. *Exp Mol Med* 2021; 53: 328–338.
48. Gazestani V, Kamath T, Nadaf NM, et al. Early Alzheimer's disease pathology in human cortex involves transient cell states. *Cell* 2023; 186: 4438–4453.e4423.
49. Wang J, Jiang Y-H, Yang P-Y, et al. Increased collagen type V α2 (COL5A2) in colorectal cancer is associated with poor prognosis and tumor progression. *Onco Targets Ther* 2021; 14: 2991–3002.
50. Campos-Melo D, Hawley ZCE and Strong MJ. Dysregulation of human NEFM and NEFH mRNA stability by ALS-linked miRNAs. *Mol Brain* 2018; 11: 43.
51. Arzouni N, Matloff W, Zhao L, et al. Identification of dysregulated genes for late-onset Alzheimer's disease using gene expression data in brain. *J Alzheimers Dis Parkinsonism* 2020; 10: 498.
52. Sevlever D, Zou F, Ma L, et al. Genetically-controlled vesicle-associated membrane protein 1 expression may contribute to Alzheimer's pathophysiology and susceptibility. *Mol Neurodegener* 2015; 10: 18.
53. Chaudhary R and Lal A. Long noncoding RNAs in the p53 network. *Wiley Interdiscip Rev RNA* 2017; 8: 10.1002/wrna.1410.
54. da Silveira WA, Renaud L, Hazard ES, et al. miRNA and lncRNA expression networks modulate cell cycle and DNA repair inhibition in senescent prostate cells. *Genes* 2022; 13: 208.
55. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucl Acids Res* 2016; 44: W90–W97.
56. Patel AO, Caldwell AB, Ramachandran S, et al. Endotype characterization reveals mechanistic differences across brain regions in sporadic Alzheimer's disease. *J Alzheimers Dis Rep* 2023; 7: 957–972.
57. Fischer M. Census and evaluation of p53 target genes. *Oncogene* 2017; 36: 3943–3956.
58. Qu Z, Sun J, Zhang W, et al. Transcription factor NRF2 as a promising therapeutic target for Alzheimer's disease. *Free Radic Biol Med* 2020; 159: 87–102.

59. Sudwarts A, Ramesha S, Gao T, et al. BIN1 Is a key regulator of proinflammatory and neurodegeneration-related activation in microglia. *Mol Neurodegener* 2022; 17: 33.
60. Marques-Coelho D, Iohan L, Melo de Farias AR, et al. Differential transcript usage unravels gene expression alterations in Alzheimer's disease human brains. *NPJ Aging Mech Dis* 2021; 7: 2.
61. Holler CJ, Davis PR, Beckett TL, et al. Bridging integrator 1 (BIN1) protein expression increases in the Alzheimer's disease brain and correlates with neurofibrillary tangle pathology. *J Alzheimers Dis* 2014; 42: 1221–1227.
62. Perdigão C, Barata MA, Burrinha T, et al. Alzheimer's disease BIN1 coding variants increase intracellular aβ levels by interfering with BACE1 recycling. *J Biol Chem* 2021; 297: 101056.
63. Sartori M, Mendes T, Desai S, et al. BIN1 Recovers tauopathy-induced long-term memory deficits in mice and interacts with tau through Thr348 phosphorylation. *Acta Neuropathol* 2019; 138: 631–652.
64. Miyagawa T, Ebinuma I, Morohashi Y, et al. BIN1 Regulates BACE1 intracellular trafficking and amyloid-beta production. *Hum Mol Genet* 2016; 25: 2948–2958.
65. Bhattacharyya R, Teves CAF, Long A, et al. The neuronal-specific isoform of BIN1 regulates β-secretase cleavage of APP and aβ generation in a RIN3-dependent manner. *Sci Rep* 2022; 12: 3486.
66. Lambert E, Saha O, Soares Landeira B, et al. The Alzheimer susceptibility gene BIN1 induces isoform-dependent neurotoxicity through early endosome defects. *Acta Neuropathol Commun* 2022; 10: 4.
67. De Rossi P, Nomura T, Andrew RJ, et al. Neuronal BIN1 regulates presynaptic neurotransmitter release and memory consolidation. *Cell Rep* 2020; 30: 3520–3535.e3527.
68. Glennon EB, Lau DHW, Gabriele RMC, et al. Bridging integrator 1 protein loss in Alzheimer's disease promotes synaptic tau accumulation and disrupts tau release. *Brain Commun* 2020; 2: fcaa011.
69. Thomas S, Hoxha K, Tran A, et al. Bin1 antibody lowers the expression of phosphorylated Tau in Alzheimer's disease. *J Cell Biochem* 2019; 120: 18320–18331.
70. Crotti A, Sait HR, McAvoy KM, et al. BIN1 Favors the spreading of Tau via extracellular vesicles. *Sci Rep* 2019; 9: 9477–9420.
71. Pyndiah S, Tanida S, Ahmed KM, et al. c-MYC suppresses BIN1 to release poly(ADP-ribose) polymerase 1: a mechanism by which cancer cells acquire cisplatin resistance. *Sci Signal* 2011; 4: ra19.
72. Elliott K, Ge K, Du W, et al. The c-Myc-interacting adaptor protein Bin1 activates a caspase-independent cell death program. *Oncogene* 2000; 19: 4669–4684.
73. Cassimere EK, Pyndiah S and Sakamuro D. The c-MYC-interacting proapoptotic tumor suppressor BIN1 is a transcriptional target for E2F1 in response to DNA damage. *Cell Death Differ* 2009; 16: 1641–1653.
74. Kinney EL, Tanida S, Rodrigue AA, et al. Adenovirus E1A oncoprotein liberates c-Myc activity to promote cell proliferation through abating Bin1 expression via an Rb/E2F1-dependent mechanism. *J Cell Physiol* 2008; 216: 621–631.
75. Folk WP, Kumari A, Iwasaki T, et al. Loss of the tumor suppressor BIN1 enables ATM ser/thr kinase activation by the nuclear protein E2F1 and renders cancer cells resistant to cisplatin. *J Biol Chem* 2019; 294: 5700–5719.
76. Ramalingam A, Farmer GE, Stamato T, et al. Bin1 interacts with and restrains the DNA end-binding protein complex ku. *Cell Cycle* 2007; 6: 1914–1918.
77. Kumari A, Iwasaki T, Pyndiah S, et al. Regulation of E2F1-induced apoptosis by poly(ADP-ribosyl)ation. *Cell Death Differ* 2015; 22: 311–322.
78. Lee J, Nguyen PT, Shim HS, et al. EWSR1, A multifunctional protein, regulates cellular function and aging via genetic and epigenetic pathways. *Biochim Biophys Acta Mol Basis Dis* 2019; 1865: 1938–1945.
79. Marchesini M, Ogoti Y, Fiorini E, et al. ILF2 Is a regulator of RNA splicing and DNA damage response in 1q21-amplified multiple myeloma. *Cancer Cell* 2017; 32: 88–100.e106.
80. Zou D, Li R, Huang X, et al. Identification of molecular correlations of RBM8A with autophagy in Alzheimer's disease. *Aging (Albany, NY)* 2019; 11: 11673–11685.