Received: 28 April 2024 | Revised: 25 July 2024 | Accepted: 29 July 2024

DOI: 10.1002/alz.14214

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

Transethnic analysis identifies *SORL1* variants and haplotypes protective against Alzheimer's disease

Xiaopu Zhou[1,2,3] | Han Cao[1] | Yuanbing Jiang[1,2] | Yuewen Chen[1,3,4,5] | Huan Zhong[1,2] | Wing Yu Fu[1] | Ronnie Ming Nok Lo[1,2] | Bonnie Wing Yan Wong[1,2] | Elaine Yee Ling Cheng[1,2] | Kin Ying Mok[1,3,6] | Timothy C. Y. Kwok[7] | Vincent C. T. Mok[8] | Fanny C. F. Ip[1,2,3] | Alzheimer's Disease Neuroimaging Initiative | Akinori Miyashita[9] | Norikazu Hara[9] | Takeshi Ikeuchi[9] | John Hardy[2,6,10] | Yu Chen[1,3,4,5] | Amy K. Y. Fu[1,2,3,4] | Nancy Y. Ip[1,2,3,4]

[1]Division of Life Science, State Key Laboratory of Molecular Neuroscience, Molecular Neuroscience Center, The Hong Kong University of Science and Technology, Hong Kong, China

[2]Hong Kong Center for Neurodegenerative Diseases, Hong Kong, China

[3]Guangdong Provincial Key Laboratory of Brain Science, Disease and Drug Development, HKUST Shenzhen Research Institute, Shenzhen, Guangdong, China

[4]SIAT–HKUST Joint Laboratory for Brain Science, Shenzhen, Guangdong, China

[5]The Brain Cognition and Brain Disease Institute, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen–Hong Kong Institute of Brain Science – Shenzhen Fundamental Research Institutions, Shenzhen, Guangdong, China

[6]Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK

[7]Therese Pei Fong Chow Research Centre for Prevention of Dementia, Division of Geriatrics, Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China

[8]Gerald Choa Neuroscience Centre, Lui Che Woo Institute of Innovative Medicine, Therese Pei Fong Chow Research Centre for Prevention of Dementia, Division of Neurology, Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China

[9]Department of Molecular Genetics, Brain Research Institute, Niigata University, Niigata, Japan

[10]Institute for Advanced Study, The Hong Kong University of Science and Technology, Hong Kong, China

Correspondence
Nancy Y. Ip, Division of Life Science, State Key Laboratory of Molecular Neuroscience, Molecular Neuroscience Center, The Hong Kong University of Science and Technology, Hong Kong, China .
Email: boip@ust.hk

The data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of the ADNI and/or provided

Abstract

INTRODUCTION: The *SORL1* locus exhibits protective effects against Alzheimer's disease (AD) across ancestries, yet systematic studies in diverse populations are sparse.

METHODS: Logistic regression identified AD-associated *SORL1* haplotypes in East Asian ($N = 5249$) and European ($N = 8588$) populations. Association analysis between *SORL1* haplotypes and AD-associated traits or plasma biomarkers was conducted. The effects of non-synonymous mutations were assessed in cell-based systems.

RESULTS: Protective *SORL1* variants/haplotypes were identified in the East Asian and European populations. Haplotype Hap_A showed a strong protective effect against AD

Xiaopu Zhou and Han Cao contributed equally to this study.

data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Funding information
National Natural Science Foundation of China (NSFC)/RGC Joint Research Scheme, Grant/Award Numbers: 32061160472, N_HKUST605/20; Fundamental Research Program of Shenzhen Virtual University Park, Grant/Award Number: 2021Szvup137; Research Grants Council of Hong Kong, Grant/Award Number: C6027-19GF; Theme-Based Research Scheme, Grant/Award Number: T13-605/18 W; General Research Fund, Grant/Award Number: HKUST16103122; Areas of Excellence Scheme of the University Grants Committee, Grant/Award Number: AoE/M-604/16; Innovation and Technology Commission, Grant/Award Numbers: InnoHK ITCPD/17-9, ITS/207/18FP, MRP/042/18X, MRP/097/20X; Guangdong Provincial Fund for Basic and Applied Basic Research, Grant/Award Number: 2019B1515130004; Chow Tai Fook Charity Foundation, Grant/Award Number: CTFCF18SC01; Japan Agency for Medical Research and Development, Grant/Award Number: JP23dk0207060; Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, Grant/Award Numbers: 21K07271, 21H03537

in East Asians, linked to less severe AD phenotypes, higher *SORL1* transcript levels, and plasma proteomic changes. A missense variant within Hap_A, rs2282647-C allele, was linked to a lower risk of AD and decreased expression of a truncated SORL1 protein isoform.

DISCUSSION: Our transethnic analysis revealed key *SORL1* haplotypes that exert protective effects against AD, suggesting mechanisms of the protective role of SORL1 in AD.

KEYWORDS
amyloid load, APOE, association, East Asian, European, PET, Pittsburgh compound B, plasma biomarker, protective

Highlights

- We examined the AD-protective mechanisms of *SORL1* in the general population across diverse ancestral backgrounds by jointly analyzing data from three East Asian cohorts (ie, mainland China, Hong Kong, and Japan) and a European cohort.
- Comparative analysis unveiled key ethnic-specific *SORL1* genetic variants and haplotypes. Among these, the *SORL1* minor haplotype, Hap_A, emerged as the primary AD-protective factor in East Asians. Hap_A exerts significant AD-protective effects in both *APOE ε4* carriers and non-carriers.
- *SORL1* haplotype Hap_A is associated with cognitive function, brain volume, and the activity of specific neuronal and immune-related pathways closely connected to AD risk.
- Protective variants within Hap_A are linked to increased SORL1 expression in human tissues.
- We identified an isoform-specific missense variant in Hap_A that modifies the function and levels of a truncated SORL1 protein isoform that is poorly investigated.

1 | BACKGROUND

Alzheimer's disease (AD), characterized by a progressive decline in cognitive and memory functions, primarily affects the elderly population.[1] The anticipated rise in the prevalence of AD due to population aging will impose substantial socioeconomic burdens on both individuals and society.[1,2] However, there are currently few effective strategies to manage or treat AD.[3]

As a heritable disorder, AD is attributed to multiple genetic risk factors.[4–7] Many genes that can affect the likelihood of developing the disease have been discovered.[8–14] This knowledge has provided insights into how AD develops as well as possible intervention methods, including beta amyloid (Aβ)-targeting monoclonal antibodies[15,16] and small molecules that modify apolipoprotein E (*APOE*) gene expression or protein function.[17,18] Besides genetic factors that exert AD risk effects, some genetic factors exhibit protective effects against AD.[19] One such candidate is *SORL1*, which is reported to have a protective effect against AD, highlighted by the identification of common variants linked to reduced AD risk across diverse ancestries.[20–22] However, its underlying genetic mechanism (or mechanisms) remains unclear.

SORL1, a member of the vacuolar protein sorting receptor family, is present in the central nervous system and can bind to apoE protein.[23,24] It is involved in the processing of amyloid precursor protein, clearance of Aβ, and regulation of neuronal functions.[25–27] Association analysis in late-onset AD families, followed by genome-wide association studies (GWASs) joining results from multiple AD cohorts, identified *SORL1* as a pivotal gene associated with AD, highlighted by the discovery of common protective variants against AD.[20,21,28]

Apart from these protective variants, numerous studies have emphasized the enrichment of rare coding variants in AD, including *SORL1* loss-of-function variants.[24,29–32] Specifically, individuals carrying *SORL1* loss-of-function mutations have been found to exhibit more than a 10-fold increased risk of developing AD,[30] suggesting that altered SORL1 function due to haploinsufficiency may contribute to AD pathogenesis.[33] In addition, certain *SORL1* missense variants (including common and rare variants) associated with decreased *SORL1* expression also led to increased AD risk.[20,34,35,36,37] Collectively, these

findings underscore the significant impact of SORL1 protein function and expression levels in AD onset and development, which are likely influenced by *SORL1* variants that contribute to AD risk.

Despite extensive research on SORL1 protein and its associated pathways, there is a limited understanding of AD-protective *SORL1* variants and their biological roles in the human system. Hence, comprehensively understanding the AD-protective variants of *SORL1* along with their associated molecular mechanisms may help uncover molecular targets for AD intervention. Given the variability in the prevalence of *SORL1* genetic variants and haplotypes across people of diverse ancestral backgrounds, distinct variants or haplotypes may confer AD-protective effects in people of different ancestral backgrounds. Therefore, mapping *SORL1* AD-protective variants and haplotypes across diverse ancestral backgrounds may help identify key genetic factors that exert significant protective effects in the general population.

Accordingly, we conducted a comprehensive genetic association analysis using genetic data obtained from three AD cohorts of East Asian individuals from mainland China, Hong Kong, and Japan ($N = 5249$) as well as data from a European cohort ($N = 8588$). Our analysis revealed the presence of variants and haplotypes in these populations that exert significant protective effects against AD. In particular, we identified a minor haplotype located within the *SORL1* locus, Hap_A, which exhibits a major protective effect against AD in the East Asian population ($\beta = -0.379$; meta-$p = 1.32 \times 10^{-10}$). This haplotype is rare among individuals of European ancestry (frequency = 0.001) but is much more common in individuals of East Asian ancestry (frequency = 0.168). Furthermore, we demonstrated that Hap_A also exerts protective effects against AD-associated endophenotypes, including cognitive function, brain volume, and plasma levels of the amyloid, tau, and neurodegeneration (ATN) biomarkers (ie, amyloid, tau, and neurodegeneration). Furthermore, we showed that Hap_A is linked to increased *SORL1* expression and may modulate biological pathways associated with both immune and neuronal functions. Specifically, we identified a coding variant in Hap_A that may regulate the function and levels of a specific truncated SORL1 isoform that has not been extensively investigated.

RESEARCH IN CONTEXT

1. **Systematic review**: The authors reviewed the literature using traditional sources (eg, PubMed and Google Scholar). While some studies report *SORL1* genetic factors related to Alzheimer's disease (AD) in East Asian and European populations separately, no systematic comparative analysis has considered the intrinsic differences in the genomic structure of the *SORL1* locus across different ancestral backgrounds. The study cites relevant sources and incorporates some prior data in the analysis.
2. **Interpretation**: The findings reveal the existence of distinct pools of *SORL1* genetic variants that underlie the protective effects against AD in different ethnic groups; these might be associated with divergent genetic mechanisms. Furthermore, the discovery of *SORL1* Hap_A, which exerts major protective effects against AD while modulating immune and neuronal pathways in human tissues, may be involved in a disease mechanism. Specifically, the observed AD-protective effect of Hap_A, evident in both *APOE* ε4 carriers and non-carriers, makes it a promising intervention target for the disease. In addition, the discovery of the isoform-specific missense *SORL1* variant rs2282647 expands our understanding of the regulation and function SORL1 protein.
3. **Future directions**: The study provides insights into the precise roles of *SORL1* and its dysregulation in AD in the general population. Future research endeavors will be required to understand the roles of the newly identified truncated SORL1 protein and evaluate the viability of targeting SORL1 as a disease intervention strategy.

2 | METHODS

2.1 | Study cohorts

This study included three cohorts of East Asian individuals from mainland China, Hong Kong, and Japan, comprising a total of 5249 participants, including 2570 patients with AD and 2679 age- and sex-matched normal controls (NCs). The data for the mainland Chinese and Hong Kong cohorts can be found in our previous publications.[11,38,39] In brief, the mainland Chinese cohort included 2042 participants (1128 patients with AD and 914 NCs), the Hong Kong cohort included 1239 participants (453 patients with AD and 786 NCs), and the Japanese cohort included 1968 participants (989 patients with AD and 979 NCs). We also analyzed data from three cohorts of individuals of European ancestry: the National Institute on Aging Alzheimer's Disease Centers cohort (henceforth ADC cohort, phs000372.v1.p1; 2603 patients with AD and 1470 NCs), the Late-Onset Alzheimer's Disease Family Study cohort (LOAD cohort, phs000168.v2.p2; 1537 patients with AD and 1711 NCs), and the Alzheimer's Disease Neuroimaging Initiative cohort (ADNI cohort, http://adni.loni.usc.edu/; 641 patients with AD and 626 NCs). From those three cohorts we retained only unrelated samples of European ancestry for analysis. In addition, within the ADC cohort, we specifically included patients with a confirmed diagnosis (ie, autopsy-confirmed cases) in the analysis.

The demographic data of these cohorts are presented in Table S1. The details of the quality control and imputation processes are presented in the *Supplementary Methods*. This study was approved by the Clinical Research & Ethics Committees of the Joint Chinese University of Hong Kong – New Territories East Cluster for Prince of Wales Hospital (CREC Reference No. 2015.461), the Kowloon Central Cluster/Kowloon East Cluster for Queen Elizabeth Hospital (KC/KE-15-0024/FR-3), and the Human Participants Research Panel of the Hong Kong University of Science and Technology (CRP No. 180 and

CRP No. 225). The analysis of the Japanese subjects was approved by the Institutional Review Board of Niigata University (G2017-0012 and G2018-0034). All participants provided written informed consent for both study participation and sample collection.

We obtained some of the data used in this study from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership led by Principal Investigator Michael W. Weiner, MD. The primary goal of the ADNI is to determine whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early AD. For up-to-date information, see www.adni-info.org.

2.2 | Identification of *SORL1* variants associated with AD

To identify *SORL1* variants (chr11:121252314-121833763; GRCh38) associated with AD in the three East Asian cohorts, we analyzed the phased, imputed genotype information. We applied logistic regression analysis for variants with a minor allele frequency >1% using PLINK software (version 1.9), with age, sex, and population structure represented by the top five principal components included as covariates; the top five principal components were estimated by PLINK using the "– pca" function.[40] We then conducted meta-analysis using METAL software.[41] We applied multiple test correction with the Bonferroni correction method using the "p.adjust()" function in R. We considered variants with a corrected p value less than .05 to be significantly associated with AD in the East Asian population. We retained variants that passed the significance threshold and had a call rate greater than 0.70 across all analyzed individuals for downstream analysis.

To examine the AD associations of the identified variants in the European population, we obtained the results of a large, recently published GWAS by Kunkle et al.[42] We considered variants with p values below the suggestive threshold of 1×10^{-4} as being associated with AD in the European population and used them for downstream analysis.

2.3 | Prevalence, linkage disequilibrium, and haplotype analysis of variants associated with AD

To determine whether different sets of variants contributed to AD in people of European and East Asian ancestry, we obtained the prevalence of the AD-associated variants identified in the East Asian and European populations from the gnomAD database (version 2.1.1).[43] We calculated the linkage disequilibrium (LD) between the identified variants within the respective populations (ie, 585 East Asian and 632 European participants from the 1000 Genomes Project Phase 3 data) using the "LD()" function in the gaston package in R. For analysis, we included the variants that remained significant after Bonferroni correction in the meta-analysis of the East Asian population as well as the variants surpassing the suggestive threshold in the results of the AD GWAS in the European population.[44]

As the variants included for haplotype and LD analysis were derived from association analysis results using different inclusion criteria for different ancestral backgrounds, they may exhibit different frequencies, haplotype structures, and complex LD patterns. Such complex LD patterns are difficult to resolve through traditional approaches with manual inspection. Accordingly, to identify variants and haplotypes that are key contributors to AD in each ancestral background, we conducted unsupervised k-means clustering on pairwise LD information from the analyzed variants using the "kmeans()" function in the R stats package; thus, we grouped variants that share strong pairwise LD in each ancestral background. We determined the numbers of variant clusters using the "fviz_nbclust()" function in the R factoextra package. To visualize and identify variant clusters, we projected the variants onto a two-dimensional plane by applying uniform manifold approximation and projection (UMAP) to the pairwise LD matrix of the studied variants using the "umap()" function in the R umap package.

To identify AD-associated haplotype structures present in both European and East Asian ancestries or those uniquely present in one ancestral background, we performed haplotype detection by leveraging the phasing information present in the phased Variant Call Format (VCF) file, as previously described.[45] In brief, we selected specific variants within haplotype blocks or variant clusters and extracted their phased information from the VCF files. We then determined the allele composition of each haplotype by analyzing the phased allele information for each variant included in haplotype analysis.

2.4 | Identification of *SORL1* haplotypes associated with AD

We performed multivariate regression analysis to estimate the effects of haplotypes on phenotypes and gene expression levels. First, we generated an $N \times (M + 1)$ matrix, where N is the number of individuals and M is the number of detected haplotypes with frequencies greater than 1% that are detected in the corresponding ancestral background in the 1000 Genomes Phase 3 data. Cells with matrix-stored values of 0, 1, or 2 represent zero, one, or two copies of specific haplotypes, respectively. We included an additional column that sums haplotypes with frequencies less than 1% (labeled "others") to ensure each row sums to 2 to match individuals' diploid genomes. We excluded common haplotypes (usually those containing major alleles) from the regression model during association testing. Thus, the estimated effect sizes (ie, β values) from the model are relative to the common haplotype.

2.5 | Cognitive function, biomarker, and endophenotype data

We measured the cognitive performance of participants from the mainland Chinese and Hong Kong cohorts using the Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA),

respectively. The scores from these tests served as indicators of their cognitive performance. We obtained T1-weighted MRI data from 216 individuals (including 110 individuals with AD and 106 NCs) from Prince of Wales Hospital. To quantify brain region volumes, we subjected the raw imaging files to AccuBrain IV1.2 (BrainNow Medical Technology). We retrieved plasma proteomic data from a previously reported Hong Kong cohort, including $A\beta_{42}$, $A\beta_{40}$, tau, p-tau181, and neurofilament light chain (NfL) levels measured by SIMOA assay for 377 individuals (including 184 individuals with AD and 193 NCs), as well as 1160 proteins quantified by Olink Proteomics for 264 individuals (including 154 individuals with AD and 110 NCs).[46]

We assessed brain Aβ load in 90 participants by PET with Pittsburgh compound B (PiB). Among them, we included 48 individuals with *APOE* ε3 homozygosity for analysis, including 32 individuals carrying *SORL1* haplotype Hap_A and 16 non-carriers. We quantified individual Aβ load as the global (ie, whole brain) retention of PiB 35 min after administration. To investigate the association between brain Aβ load and cognitive function measured by MoCA score, we performed robust regression analysis with age, sex, education level, and *SORL1* haplotype as covariates in both the Hap_A carrier and non-carrier groups.

2.6 | Genotype–endophenotype and genotype–expression association analyses of identified haplotypes

We performed robust multivariate regression analysis to examine the associations of *SORL1* haplotypes with endophenotypes and gene expression levels using the "lmrob()" function in the robustbase package in R. Before regression analysis, we normalized continuous variables that did not follow a normal distribution by rank-based, inverse normal transformation using the "rankNorm()" function in the RNOmni package in R. All association analyses were adjusted for age, sex, and genomic structure represented by the first three principal components. We also adjusted for additional confounding factors for specific endophenotypes, including education level for cognitive performance, intracranial volume for brain volume measured by MRI, and cardiovascular disease history for plasma proteins. For the Genotype-Tissue Expression (GTEx) data, we included the first five principal components, sex, sequencing platform, and sequencing protocol as well as the inferred covariates. We performed Gene Ontology analysis by subjecting a list of identified proteins to the "clusterProfiler()" function in the clusterProfiler R package.[47] Gene Ontology terms with a *q* value (ie, false discovery rate) less than 0.05 are presented.

2.7 | Functional annotation of AD-associated variants

We annotated genetic variants associated with the identified AD-associated variants using candidate cis-regulatory elements (cCREs) sourced from the SCREEN (Search Candidate cis-Regulatory Elements by ENCODE) database.[48] These cCREs were derived from an analysis of the epigenetic profile across multiple tissues and thereby indicated a broad regulatory influence on gene expression. To explore their regulatory effects in the brain, we extended our analysis to include the annotations of genetic variants within the cCREs using a publicly available database of brain single-cell ATAC-seq (assay for transposase accessibility by sequencing).[49]

2.8 | In silico analysis of putative function of Trp15Cys coding variant

We obtained the amino acid sequence for the SORL1 isoform, ENST00000527934.1, from the UniProt database (accession no.: E9PKB0). We visualized the genomic location of the rs2282647 variant in ENST00000527934.1 using the UCSC Genome Browser (https://genome.ucsc.edu/).[50] We conducted in silico prediction of variant function by submitting the amino acid sequence of the wild-type (WT) and mutated proteins to the corresponding database. We predicted disulfate bonds using DIpro software (version 2.0) from Scratch Protein Predictor (https://scratch.proteomics.ics.uci.edu/),[51] disorder score using flDPnn (https://biomine.cs.vcu.edu/servers/flDPnn/; December 2021 version),[52] and signal peptide regions and cleavage sites using SignalP 6.0 (https://services.healthtech.dtu.dk/services/SignalP-6.0/).[53]

2.9 | Confirmation of presence of truncated SORL1 isoform by long-read RNA sequencing

We retrieved the mapped Binary Alignment Map (BAM) file for Nanopore long-read RNA sequencing data for the NA12878 lymphoblastoid cell line from GitHub (https://github.com/nanopore-wgs-consortium).[54] We employed the Integrative Genomics Viewer (version 2.17.4) to visualize the regional BAM file containing mapped reads. We used the UCSC Human BLAT Search tool (https://genome.ucsc.edu/cgi-bin/hgBlat) to align a read representing the candidate truncated SORL1 isoform (ie, SORL1-206 or ENST00000527934.1) to the GRCh38 human reference genome.[55]

2.10 | Overexpression of WT and truncated SORL1 protein in HEK293T cells

We cultured HEK293T cells (ATCC) by seeding 300,000 cells onto 35-mm plates 1 day prior in DMEM (Life Technologies, 12800-017) supplemented with 10% heat-inactivated fetal bovine serum (Corning, 35-079-CV). We transfected HEK293T cells using Lipofectamine 3000 (Life Technologies, L3000001) with 1, 1.5, or 2 µg of either WT and mutant SORL1 plasmids. We generated the SORL1 plasmids by synthesizing the SORL1 sequence and cloned it into the pCDNA3.1-HA vector by Genescript.

Following transfection, we lysed HEK293T cells (>3×10^5 cells per well, approximately 0.3 mg protein per well) in 1× RIPA buffer containing protease inhibitors for 30 min using a Ferris wheel

apparatus. We determined the protein concentration in the lysate using the Bradford dye method (Bio-Rad, 5000006). We mixed the protein lysate with sample buffer and 1× RIPA and then denatured the sample by boiling at 100°C for 5 min. We subsequently loaded 20 µg denatured protein per sample onto SDS-acrylamide gels and performed electrophoreses at 50 A (25 A per gel) for 2 h. We then transferred the resolved proteins onto nitrocellulose membranes. We carried out immunodetection using anti-HA tag antibody (Cell Signaling Technology, 2367S, 1:500) and anti-α-tubulin antibody (Sigma, T9026, 1:5000). We visualized the signals using enhanced chemiluminescence substrate and blotting using a ChemiDoc MP Imaging System (Bio-Rad). We quantified protein abundance using Adobe Photoshop (CS6) to analyze the mean intensity as summarized in the histogram panels. We conducted background subtraction by considering the intensity from surrounding regions in the gel blots. The cDNA sequences for both the WT and mutated proteins as well as a list of reagents and their respective sources are presented in the *Supplementary Data* and the *Supplementary Methods* sections in the *Supplementary Information*.

2.11 | Data visualization

We generated regional plots to visualize the results of the *SORL1* association analysis using Locuszoom (https://my.locuszoom.org/). We generated volcano plots and scatter plots for variant locations using the R ggplot2 package. We generated LD plots using the "LD.plot()" function in the R gaston package. Finally, we generated dot plots, heatmaps, and bar charts using GraphPad Prism (version 8). We obtained the multitissue expression quantitative trait loci (eQTL) plot from the GTEx website (https://www.gtexportal.org/).

3 | RESULTS

3.1 | Identification of *SORL1* variants protective against AD in East Asian population

SORL1 genetic variants exerting AD-protective effects are present across diverse populations. However, most studies involve populations of European ancestry. Therefore, we first systematically screened for variants in the *SORL1* locus that are associated with AD in the East Asian population. Meta-analysis of three East Asian AD cohorts (N = 5249, including 2570 patients with AD and 2679 NCs; Table S1) revealed robust AD-protective signals in the *SORL1* locus (ie, chr11:121252314-121833763; GRCh38), tagged by the sentinel variant, rs11604897 (β = −0.354, meta-p = 3.80×10^{-11}; Figure 1A). An additional 75 variants passing Bonferroni correction were also identified (corrected meta-p < .05, Table S2).

While AD-protective variants in the *SORL1* locus are found in individuals of European ancestry,[12,42] the sentinel AD-associated variant of *SORL1* is rs11218343. However, in the East Asian population, we identified rs11604897, which is located in the intronic region of the gene, as the primary AD-associated variant of *SORL1* (Figure 1A). This suggests that East Asians harbor a set of AD-protective variants in *SORL1* that differs from that in individuals of European ancestry. Therefore, to investigate whether East Asian and European populations had distinct sets of AD-protective variants and how they exerted their effects in AD, we obtained AD GWAS data from a recent study conducted in a population of European ancestry ("European AD GWAS" population hereafter).[42] Intriguingly, the sentinel AD-associated *SORL1* variant identified in the East Asian population, rs11604897, did not pass the suggestive threshold ($p < 1 \times 10^{-4}$) in the European AD GWAS population but was still significantly associated with AD ($p = 4.81 \times 10^{-3}$; Figure 1B, Table S3). Moreover, of the 75 variants associated with AD in the East Asian population, only five variants – rs11218343, rs9665907, rs3781832, rs1784920, and rs3824969 – reached the suggestive threshold ($p < 1 \times 10^{-4}$) in the European AD GWAS population (Figure 1B, Table S3). These findings suggest that distinct sets of AD-protective genetic variants are present in populations of East Asian and European ancestry.

3.2 | Characterization of *SORL1* variants protective against AD in East Asian and European populations

The presence of distinct pools of AD-protective *SORL1* variants in the East Asian and European populations suggests the existence of different *SORL1* haplotype structures that are associated with AD in both ancestral backgrounds. To investigate how these distinct *SORL1* haplotypes conferred AD-protective effects with respect to ancestral background, we selected variants associated with AD in both groups for subsequent analysis. Specifically, we retained variants that were significantly associated with AD from the meta-analysis of the East Asian population or the European AD GWAS data for subsequent comparative analysis (Table S3). We stratified those variants according to the significance of the association with AD in each group. This yielded five, 70, and 20 variants significantly associated with AD in both groups, only in the East Asian population, and only in the European population, respectively (Table S3).

To compare the possible variations of those AD-associated variants between the East Asian and European populations, we queried the gnomAD database for the allele frequencies of the 95 identified AD-associated *SORL1* variants and compared their prevalence between ancestral backgrounds.[38] The results revealed notable differences in the allele frequencies of these variants. Remarkably, some of the newly discovered AD-associated variants identified in the East Asian population were substantially more common in the East Asian population (frequency > 0.10) than in the European population (frequency < 0.01) (Figure S1, Table S4).

We then conducted our LD analysis of the 95 identified AD-associated variants within each population using genomic data from the 1000 Genomes Project.[56] We excluded five variants from the analysis owing to their low frequency (minor allele frequency < 0.01) in the East Asian population (Table S4). Accordingly, we recovered distinct LD blocks among these 90 variants in both populations and observed complex and different LD patterns between groups (Figures 2A and

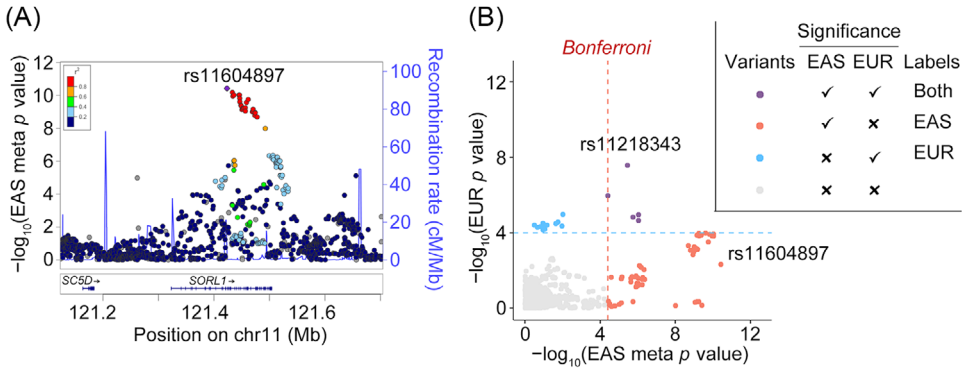

FIGURE 1 Comparative analysis of *SORL1* genetic variants associated with Alzheimer's disease (AD) in the East Asian and European populations. (A) Regional association plot of meta-analysis results for variants in SORL1 locus in East Asian population ($N = 5249$). The sentinel variant, rs11604897, is highlighted. (B) Comparison of effects of variants on modifying AD risks between East Asian (EAS) and European (EUR) populations. The *p* values for AD associations in European population are from Genome-Wide Association Study summary statistics from Kunkle et al.[42] Blue and red dashed lines represent Bonferroni-corrected *p* value cutoffs of .05 in EUR population and meta-*p* value cutoff of 1×10^{-4}, respectively.

S2). Of note, specific AD-associated variants exhibited stronger LD in the East Asian population than in the European population (Figure 2A). This suggests that the AD-associated *SORL1* variants may form unique variant clusters as well as different haplotypes in each population, which may account for the differing AD-protective effects of those *SORL1* variants between populations.

Therefore, to examine the clustering of *SORL1* AD-associated variants in each population, we conducted unsupervised *k*-means clustering analysis to stratify those 90 AD-associated variants into variant clusters according to their LD patterns (ie, pairwise correlations measured by R^2). Based on the *k*-means clustering results, we classified the 90 variants into four variant clusters in the East Asian population (designated "EA_1" to "EA_4") and six variant clusters in the European population (designated "EU_1" to "EU_6") (Figure 2A, Table S5). Comparative analysis of these *SORL1* variant clusters in the East Asian and European populations showed that the variants within a given cluster in the East Asian population were present in different variant clusters in the European population (eg, variants in cluster EA_3 were present in clusters EU_1, EU_4, or EU_5) (Figure S3, Table S5). These results suggest that the haplotype structures in the *SORL1* locus exhibit complex patterns between populations. Taken together, our results suggest that *SORL1* is an AD-protective genetic factor in both the East Asian and European populations, while different sets of variants contribute to the AD-protective mechanism in these populations.

3.3 | Analysis of variants and haplotypes associated with protective effect of *SORL1* against AD in East Asian and European populations

The results discussed above indicate that AD-associated variants in the *SORL1* locus form distinct variant clusters in the East Asian and European populations (Figure S2). Therefore, to investigate which variant clusters (or variants) play vital roles in AD-protective effects, we conducted LD and haplotype analysis to independently identify common haplotypes (ie, minor allele frequency > 1%) present in the variant clusters in the two populations using the 1000 Genomes Project data (Table S6).

Accordingly, in the East Asian population, LD analysis identified five haplotype blocks within the variant clusters EA_1 to EA_4. We excluded two variants owing to their low LD with other variants in the East Asian population (Table S5). Notably, we subdivided variant cluster EA_4 into EA_4a and EA_4b based on the strong LD patterns observed within each cluster (Figures 2A and S4). Subsequent haplotype and association analysis revealed that variant clusters EA_1, EA_2, EA_3, and EA_4a exhibited AD-protective effects as indicated by significant associations of the minor haplotypes within each variant cluster with AD (effect size: −0.365 to −0.238, Bonferroni-corrected meta-$p < .05$; Figure 2B, Tables S7 and S8). Notably, among the four variant clusters, the minor haplotype in variant cluster EA_3 exerted the strongest AD-protective effect ($\beta = -0.365$, Bonferroni-corrected meta-$p = 3.76 \times 10^{-9}$; Table 1, Figure 2B, Tables S7 and S8). This finding is corroborated by the results of multivariate lasso regression, wherein the minor haplotype in EA_3 consistently exhibited dominant AD-protective effects over other minor haplotypes from other variant clusters (Figure S5, Table S9). Thus, we determined that the variants in cluster EA_3 exerted major protective effects against AD in the East Asian population.

We conducted a similar analysis in the European population using data from the ADNI, LOAD, and ADC cohorts ($N = 8588$ individuals, including 4781 patients with AD and 3807 NCs; Table S1). As variants from variant cluster EU_1 did not exhibit LD in individuals of European ancestry (Figure S2), we only assessed the AD associations of the five key variants – rs11218343, rs9665907, rs3781832, rs1784920, and rs3824969 – which are significantly associated with AD in both the East Asian and European populations (Figures S2 and S4). Meta-analysis indicated that variant rs11218343 and one minor haplotype in variant cluster EU_5 exerted significant AD-protective effects in individuals of European ancestry (meta-$p < .05$; Table 1, Figure 2C, Table S10). Thus, we successfully replicated the AD-protective effects of

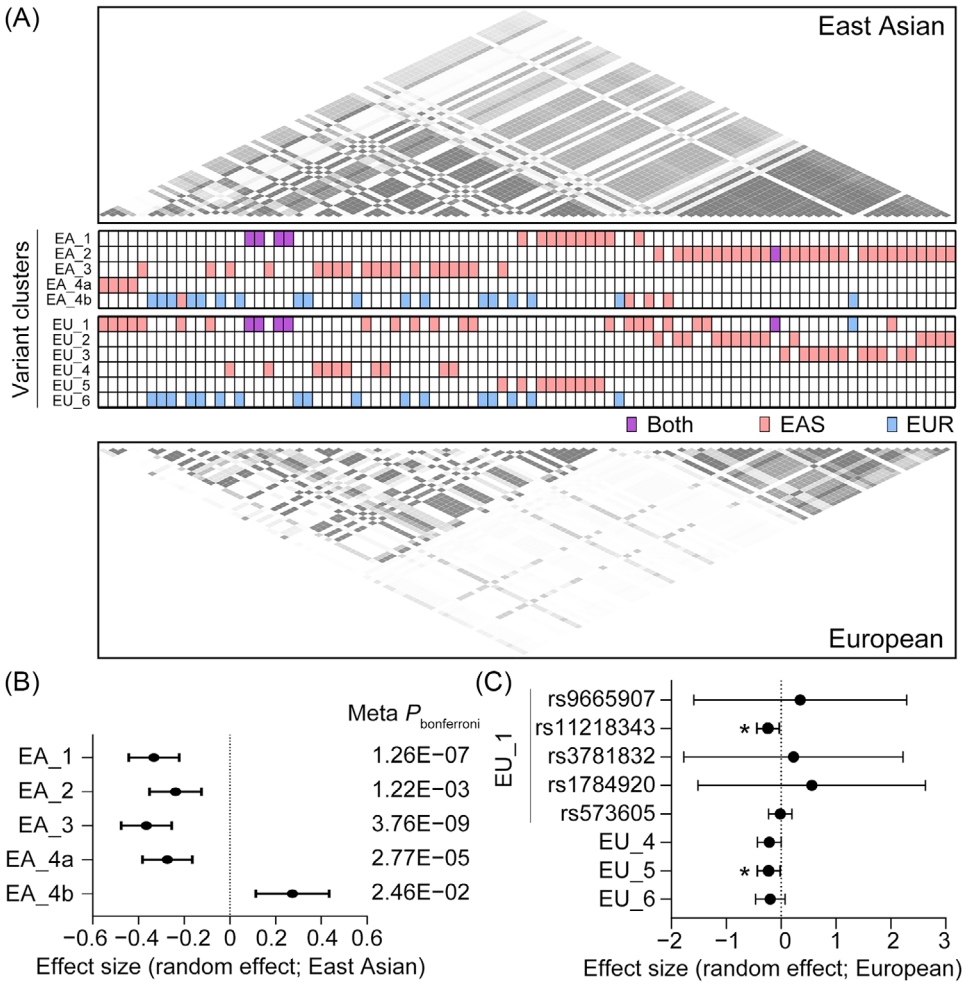

FIGURE 2 Identification of distinct variant clusters in *SORL1* locus associated with AD in East Asian and European populations. (A) Linkage disequilibrium plots of candidate variants in East Asian population (upper panel) and European population (lower panel). Data are from high-coverage, whole-genome sequencing data from the 1000 Genomes Project Phase 3. Colors denote the pairwise R^2 values among the selected variants. Middle panel: Attributes of variants from clusters EA_1 to EA_4 in East Asian population and clusters EU_1 to EU_6 in European population, specifically their associations with AD. Cluster EA_4 is subdivided into EA_4a and EA_4b owing to linkage disequilibrium patterns in corresponding variants. (B) Meta-analysis results of effects of minor haplotypes with the lowest *p* values from each variant cluster on AD ($N = 5249$). Meta-*p* values after Bonferroni correction calculated from Han and Eskin's random effects model are shown. (C) Meta-analysis results of effects of minor haplotypes, with lowest *p* values from each variant cluster and five variants on AD association (*$p < .05$, Han and Eskin's random effects model). (B, C) Data are mean ± 95% confidence intervals. Adj, adjusted; Both, variants significantly associated with Alzheimer's disease in both East Asian and European populations; Chr, chromosome; EAS, East Asian population; EUR, European population; kb, kilobase.

variant rs11218343 in the European population. Moreover, we identified a distinct variant cluster, EU_5, containing a minor haplotype that confers protection against AD in individuals of European ancestry (Table 1, Figure 2C, Tables S8 and S10).

3.4 | Identification of *SORL1* AD-protective haplotypes

We revealed the specific variant clusters within the *SORL1* locus that are associated with major protective effects against AD in the East Asian population (ie, EA_3) and European population (ie, EU_5 and rs11218343) (Figure 2). To comprehensively evaluate the genetic factors of *SORL1* that protect against AD in the general population, we combined the findings from both populations, that is, we combined the variants from variant clusters that exert dominant AD-protective effects in both populations. In addition, we kept variants rs9665907, rs3781832, and rs1784920 from EU_1, which are in LD with rs11218343 in the East Asian population and are significantly associated with AD in both the East Asian and European populations. Thus, we selected 31 variants for downstream analysis, including rs2282647, which is present in both EA_3 and EU_5 (Figure 3A, Table S11).

LD analysis among the 31 selected variants revealed that most variants exhibited strong LD in the East Asian population, suggesting the presence of haplotypes associated with AD (Figure 3A). Again, these

TABLE 1 Associations between Alzheimer's disease and key variants or haplotypes in each *SORL1* variant cluster.

			Random effects model					Heterogeneity	
Cluster	Haplotype/SNP	Study #	β	SE	OR (95% CI)	*p* (RE)	*p* (RE2)	I^2 (%)	*p* (Q)
East Asian									
EA_1	ACTAAGGTCCGTGA	3	−0.332	0.056	0.717 (0.643–0.801)	3.79E−09	5.46E−09	0	0.896
EA_2	ACGATGCGTTTCTATA CTACTTCATCTTG	3	−0.238	0.058	0.788 (0.704–0.883)	3.95E−05	5.30E−05	0	0.456
EA_3	TTAACGGTACAGATT AGAC	3	−0.365	0.057	0.694 (0.621–0.776)	1.05E−10	1.64E−10	0	0.983
EA_4a	AGCG	3	−0.274	0.056	0.760 (0.681–0.849)	8.82E−07	1.21E−06	0	0.505
EA_4b	GGGGTCTTCGGGTT CCCGGGGA	3	0.274	0.082	1.315 (1.120–1.545)	8.12E−04	1.07E−03	0	0.949
European									
EU_1	rs9665907	3	0.349	0.989	1.418 (0.204–9.850)	7.24E−01	7.68E−01	0	0.999
EU_1	rs11218343	3	−0.238	0.105	0.788 (0.642–0.968)	2.26E−02	2.86E−02	0	0.732
EU_1	rs3781832	3	0.226	1.019	1.254 (0.170–9.237)	8.25E−01	8.69E−01	0	0.999
EU_1	rs1784920	3	0.559	1.058	1.749 (0.220–13.911)	5.97E−01	6.48E−01	0	0.998
EU_1	rs573605	3	−0.015	0.108	0.985 (0.797–1.217)	8.92E−01	9.96E−01	26.1	0.258
EU_4	AACGGTCATA	3	−0.214	0.11	0.807 (0.651–1.002)	5.29E−02	6.54E−02	0	0.760
EU_5	CAGGTCCGT	3	−0.226	0.107	0.798 (0.647–0.984)	3.50E−02	4.37E−02	0	0.726
EU_6	AAAAGCCTAATGCTTTA	3	−0.196	0.137	0.822 (0.628–1.075)	1.52E−01	1.82E−01	0	0.541

Abbreviations: AD, Alzheimer's disease; CI, confidence interval; OR, odds ratio; *Q*, Cochran's *Q* statistic; RE, random effects meta-analysis model; RE2, Han and Eskin's random effects meta-analysis model; SE, standard error; SNP, single-nucleotide polymorphism.

variants exhibited different LD patterns in the East Asian and European populations, suggesting that the types and frequencies of haplotypes defined by those 31 variants vary between these groups (Figure 3A).

Accordingly, we conducted haplotype analysis using the phased genotypes of the 31 variants from the studied AD cohorts and identified eight common haplotypes (ie, frequency > 1%) that are present in the East Asian and/or European population (Figure 3B, Table S12). Notably, haplotype Hap_A comprises all of the AD-protective alleles of the 31 variants (Figure 3B). Considering the possible aggregate effects of these protective variants, Hap_A may exhibit the strongest AD-protective effects in the general population and may therefore be a primary contributor to the AD-protective effect of *SORL1* (Figure 3B).

Therefore, to investigate the AD-protective effects of the identified haplotypes, we again conducted multivariate logistic regression analysis to examine their associations with AD in individual AD cohorts (Table S12). In the East Asian population, Hap_A exhibited consistent and significant AD-protective effects in all three AD cohorts as expected (nominal *p* values = 3.08×10^{-5}, 2.80×10^{-3}, and 6.19×10^{-5} in the mainland Chinese, Hong Kong, and Japanese cohorts, respectively; Figure 3C, Table S12). Meta-analysis of the three East Asian cohorts confirmed the significant AD-protective effects of Hap_A (β = −0.379, meta-*p* = 1.32×10^{-10}; Tables 2 and S13). Thus, the results indicate that Hap_A is the primary genetic factor that exerts AD-protective effects in the East Asian population (Figure 3C, Table S13). In contrast, Hap_A is significantly less common in the European population (frequency = 0.001) than in the East Asian population (frequency = 0.168). In addition, in the European population, haplotypes Hap_C and Hap_H showed trending relationships with AD (meta-*p* = 0.086 and 0.096, respectively; Figure 3C, Table S13). Meta-analysis summarizing the results from all six cohorts again showed that only Hap_A was significantly associated with AD (β = −0.379, meta-*p* = 1.15×10^{-10}, Table S13). Notably, the AD-protective effect of Hap_A remained significant in individuals after stratifying by *APOE* ε4 genotype in the East Asian population (β = −0.508 and −0.319, meta-*p* = 5.30×10^{-5} and 1.64×10^{-10} for *APOE*-ε4 carriers and non-carriers, respectively; Tables 2 and S14). Hence, our findings suggest that the AD-protective effects of *SORL1* haplotype Hap_A persist irrespective of *APOE* ε4 genotype.

3.5 | *SORL1* haplotype Hap_A is associated with AD-associated endophenotypes

The identification of *SORL1* haplotype Hap_A provides valuable insights into the genetic basis of the AD-protective effects of *SORL1* (Figure 3). Nevertheless, it is unclear how Hap_A exerts this protective effect and which biological processes are affected. Therefore, to investigate the biological effects of Hap_A on the human system, we examined the associations between the allele dosage of Hap_A and multiple AD-associated endophenotypes, including cognitive function, brain region volumes, and the plasma ATN biomarkers (ie, Aβ, p-tau181, and NfL) (Figures 4A and S6). Accordingly, the presence of Hap_A was associated with better cognitive performance as measured by the MMSE score in the mainland China cohort (β = 0.088, $p = 2.28 \times 10^{-2}$).

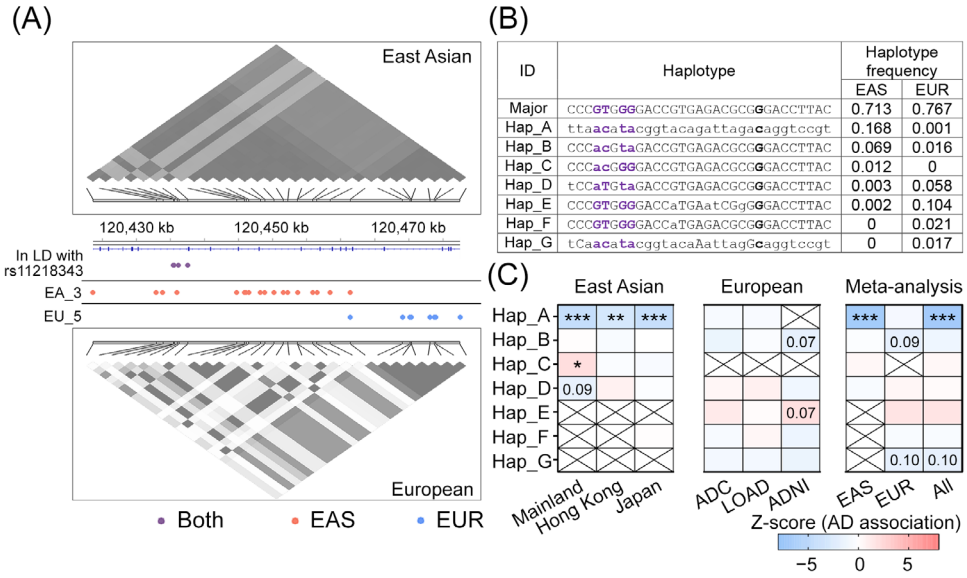

FIGURE 3 Identification of AD-protective *SORL1* haplotype Hap_A. (A) Linkage disequilibrium plots in East Asian population (upper panel) and European population (lower panel) for 31 variants from variant clusters EA_3 and EU_5, four variants from EU_1 (including rs11218343), and an additional three variants (ie, rs9665907, rs3781832, and rs1784920) in linkage disequilibrium with rs11218343. Data are high-coverage, whole-genome sequencing data from the 1000 Genomes Project. Color denotes pairwise R^2 values among the 31 selected variants. (B) Common haplotypes (ie, frequency > 1%) defined by the selected 31 variants identified in East Asian and European populations (ie, "Major" and Hap_A to Hap_G); the "Major" haplotype denotes the most common haplotype present in the general population. Lowercase letters denote the AD-protective alleles of corresponding variants. Purple letters denote variants associated with AD in both the East Asian and European populations. Boldface letters denote the rs2282647 variant, which is present in both variant clusters EA_3 and EU_5. (C) Heatmap of association results of identified haplotypes in individual cohorts as well as meta-analysis results in different ethnic groups and across all cohorts. Colors indicate association *Z*-scores calculated as effect size divided by standard error. For individual cohorts, nominal *p* values obtained from logistic regression are displayed; for meta-analysis results, *p* values from Han and Eskin's random effects model are displayed (***$p < .001$, **$p < .01$, *$p < .05$; *p* values from .05 to .1 are displayed as digits). AD, Alzheimer's disease; Both, variants exhibiting significant associations with AD in both East Asian and European populations; EAS, East Asian population; EUR, European population; kb, kilobase; LD, linkage disequilibrium.

TABLE 2 Associations between Alzheimer's disease and *SORL1* haplotype Hap_A stratified by *APOE* ε4 genotype.

		Random effects model					Heterogeneity	
Group	Study no.	β	SE	OR (95% CI)	*p* (RE)	*p* (RE2)	I^2 (%)	*p* (Q)
All participants	3	−0.379	0.058	0.685 (0.611–0.768)	8.42E−11	1.32E−10	0	.896
APOE ε4 carriers	3	−0.508	0.115	0.602 (0.481–0.753)	3.95E−05	5.30E−05	0	.355
APOE ε4 non-carriers	3	−0.319	0.097	0.727 (0.601–0.878)	1.05E−10	1.64E−10	0	.183

Note: Data are shown for the East Asian population.
Abbreviations: AD, Alzheimer's disease; CI, confidence interval; OR, odds ratio; *Q*, Cochran's *Q* statistic; RE, random effects meta-analysis model; RE2, Han and Eskin's random effects meta-analysis model; SE, standard error.

In addition, we observed a similar trend of association between Hap_A allele dosage and MoCA score in the Hong Kong cohort ($\beta = 0.067$, $p = 1.35 \times 10^{-1}$; Figures 4A and S6, Table S15). Concordantly, in the Hong Kong cohort, Hap_A was also significantly associated with a larger volume of the caudate, which is affected during AD progression ($\beta = 0.313$, $p = 2.18 \times 10^{-2}$), elevated plasma $A\beta_{42/40}$ ratio ($\beta = 0.279$, $p = 1.22 \times 10^{-2}$), and decreased p-tau181 level ($\beta = -0.204$, $p = 3.36 \times 10^{-2}$) as well as a trend of elevated $A\beta_{42}$ ($\beta = 0.175$, $p = 7.69 \times 10^{-2}$) (Figures 4A and S6, Table S15). Overall, these findings corroborate the observed AD-protective effects of Hap_A, which further suggests possible associated molecular mechanisms.

It is noteworthy that in patients with AD in the Hong Kong cohort, the presence of Hap_A was also associated with larger volumes of the caudate ($\beta = 0.482$, $p = 1.84 \times 10^{-2}$) and amygdala ($\beta = 0.648$, $p = 1.32 \times 10^{-3}$) as well as a lower level of plasma tau protein ($\beta = -0.295$, $p = 3.49 \times 10^{-2}$) (Figures 4A and S6, Table S15). These findings suggest that besides modifying AD risk, Hap_A may also modulate AD progression after onset.

To test this hypothesis, we examined the association between Aβ load measured by PET using PiB and cognitive performance measured by MoCA score in an independent AD cohort from Hong Kong comprising individuals homozygous for *APOE* ε3 ($n = 48$). Interestingly,

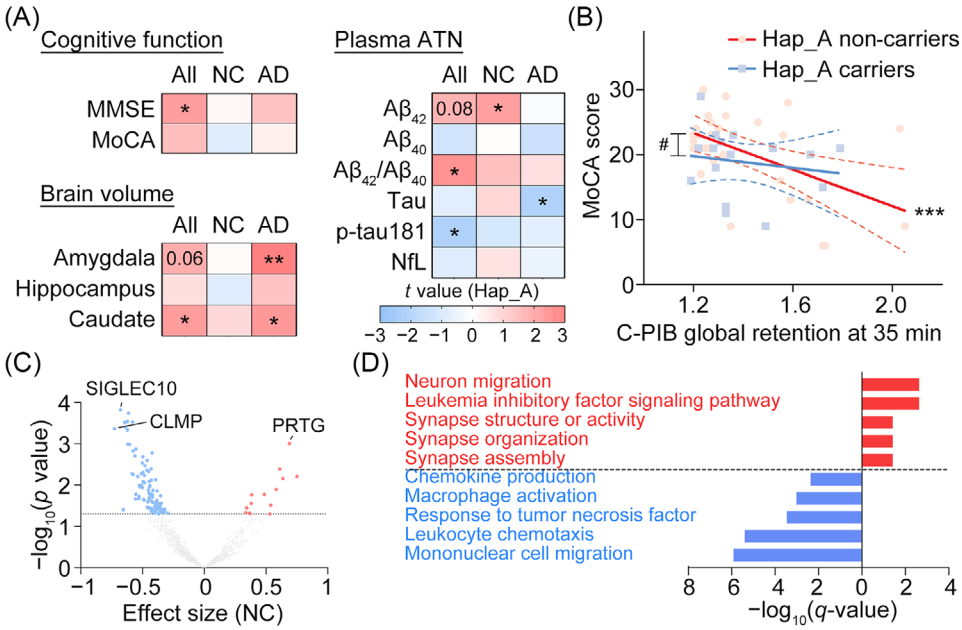

FIGURE 4 Effects of *SORL1* AD-protective haplotype on AD endophenotypes and biological processes. (A) Heatmap summarizing associations between AD-protective haplotype (Hap_A) and AD-associated endophenotypes, including cognitive function (MMSE: $n = 1392$ from mainland Chinese cohort, including 1082 individuals with AD and 310 NCs; MoCA: $n = 1210$ from Hong Kong population, including 429 individuals with AD and 781 NCs), volumes of specific brain regions ($n = 216$, including 110 individuals with AD and 106 NCs), and levels of plasma ATN biomarkers ($n = 377$, including 184 individuals with AD and 193 NCs). The color scale corresponds to *t* values obtained from the association analysis (robust regression, $^{*}p < .05$; *p* values from .05 to .1 are displayed as digits). (B) Association between MoCA score and Aβ brain load detected by PET in *APOE* ε3 homozygous individuals harboring haplotype Hap_A ($n = 32$) or without haplotype Hap_A ($n = 16$) (robust regression, $^{***}p < .001$; *Z*-score test, $^{\#}p < .05$). (C) Volcano plot showing association between plasma proteomes and allele dosage of haplotype Hap_A in NCs ($n = 110$). Proteins with a nominal *p* value less than .05 (robust regression) are marked in red (positive association) or blue (negative association). Key genes that exhibit strong associations or significant level changes are marked in the plot. (D) Bar charts displaying biological processes in plasma proteome that are associated with haplotype Hap_A. Red and blue denote biological processes that are positively or negatively associated with haplotype Hap_A, respectively. Aβ, beta amyloid; AD, Alzheimer's disease; ATN, amyloid, tau, and neurodegeneration; C-PiB, carbon-11 – labeled Pittsburgh compound B; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; NC, normal control; NfL, neurofilament light chain; p-tau181, phosphorylated tau protein at threonine 181; PET, positron emission tomography.

brain Aβ load was strongly negatively associated with cognitive performance in AD patients without Hap_A ($\beta = -15.489$, $p = 5.85 \times 10^{-4}$). However, the cognitive performance of individuals with Hap_A was not associated with increased brain Aβ load ($p = .56$; Figure 4B, Table S16). Moreover, there was a significant difference in the slope between Hap_A carriers and non-carriers regarding the association between cognitive performance and Aβ load, suggesting that individuals with Hap_A are less affected by brain Aβ load in terms of their cognitive function ($p = 4.04 \times 10^{-2}$; Figure 4B, Table S16). These results collectively indicate that *SORL1* haplotype Hap_A is associated with lower AD severity and accompanying cognitive decline.

3.6 | Biological processes modulated by *SORL1* haplotype Hap_A

The foregoing results suggest that Hap_A has modulatory effects on various AD-related biological events, especially protective effects on cognitive performance in the presence of increased Aβ burden. To understand the mechanisms whereby *SORL1* haplotype Hap_A exerts such effects, we subsequently identified which biological processes it influenced.

Accordingly, we collected high-throughput plasma protein data from 110 cognitively NCs, including data on the abundance of 1160 plasma proteins assayed by proximity extension assay in the Hong Kong AD cohort. Genotype–expression association analysis identified 112 proteins that were significantly associated with Hap_A, including 14 and 98 exhibiting positive and negative associations, respectively ($p < .05$; Figure 4C, Table S17). The key proteins whose plasma protein levels are modulated by Hap_A (ie, low *p* values or large effect sizes) include SIGLEC10 ($\beta = -0.678$, $p = 1.52 \times 10^{-4}$), which is involved in immune cell signaling, CLMP ($\beta = -0.727$, $p = 4.39 \times 10^{-4}$), which is involved in osteoblast differentiation and angiogenesis, and PRTG ($\beta = 0.693$, $p = 9.85 \times 10^{-4}$), which is involved in the regulation of cell division and growth (Figure 4C, Table S17).

To further characterize the biological processes that are modulated by Hap_A, we performed Gene Ontology analysis of the 112 dysregulated proteins. The results indicate that the corresponding genes of the 98 downregulated plasma proteins in Hap_A carriers are involved in multiple immune-related pathways, including "mononuclear cell

migration" ($q = 1.20 \times 10^{-6}$), "leukocyte chemotaxis" ($q = 3.94 \times 10^{-6}$), and "response to tumor necrosis factor" ($q = 3.48 \times 10^{-4}$) (Figure 4D, Table S18). Meanwhile, the corresponding genes of the 14 plasma proteins upregulated in Hap_A carriers are involved in neuronal-related pathways, including "neuronal migration" ($q = 2.32 \times 10^{-3}$), "synaptic structure and activity" ($q = 3.85 \times 10^{-2}$), and "synapse organization" ($q = 3.85 \times 10^{-2}$) (Figure 4D, Table S18). These findings and the results indicating the association between Hap_A and volumetric changes in specific brain regions collectively suggest that Hap_A modulates neuronal pathways. Moreover, the results imply the involvement of *SORL1* in the modulation of immune-related processes.

3.7 | *SORL1* variants protective against AD are linked to increased *SORL1* expression

Recent studies suggest a correlation between increased *SORL1* expression and decreased AD risk.[20–22] Given that most variants within Hap_A are located in non-coding regions, they may be associated with altered gene expression of *SORL1*. Therefore, we examined whether any variants of Hap_A were located within the regulatory regions using the SCREEN database. Interestingly, some of these AD-protective variants, such as rs75279208, reside in the pre-annotated cCREs, which also occupy open chromatin regions within brain cells (Figure 5A, Table S19).

Next, to investigate whether the AD-protective alleles were linked to altered *SORL1* transcript levels, we queried the GTEx database for seven variants from Hap_A that reside in cCREs (Table S19). Notably, the protective alleles of variants rs3781831 and rs75279208, which are both located in Hap_A, were associated with increased *SORL1* transcript levels in human neural tissues (Figure 5B,C, Table S19). Extending the analysis to *SORL1* AD risk variants identified from the meta-analysis results of the East Asian population uncovered four *SORL1* AD risk variants showing suggestive associations with AD (nominal meta *p* values from 1.68×10^{-4} to 4.48×10^{-4}, Table S19); all of the risk alleles of these variants were associated with decreased *SORL1* transcript levels in human tissues (Figure 5D,E, Table S19). Therefore, our analysis suggests a potential mechanism underlying the AD-protective effects of Hap_A, specifically upregulation of the basal expression of *SORL1* in human tissues.

3.8 | Isoform-specific Trp15Cys coding variant in Hap_A

The foregoing results indicate a protective role of Hap_A in mitigating AD risk. To investigate the underlying biological mechanisms, we scrutinized the functional aspects of variants within Hap_A. Of note, we identified variant rs2282647 as an isoform-specific coding variant situated in the exonic region of a specific SORL1 isoform, ENST00000527934.1 (also termed "SORL1-206"), which encodes a truncated SORL1 protein. This variant results in an amino acid substitution from tryptophan to cysteine at the 15th position (Trp15Cys, Figure S7A).

Subsequent in silico analysis of the WT and mutated proteins indicated that the rs2282647 variant (ie, Trp15Cys) had the potential to introduce new disulfide bonds between the 11th and 15th residues in the mutated protein (Figure S7B). Meanwhile, the variant appears to contribute to the stabilization of the secondary structure of the SORL1 protein isoform, as suggested by a lower disorder score in the mutated protein compared to the WT protein in the variant proximity regions (Figure S7C). Furthermore, given the proximity of the variant to the N-terminal of the protein isoform, we conducted additional in silico analyses for the signal peptide and cleavage site prediction in both the WT and mutated proteins. Intriguingly, we identified a signal peptide and corresponding SPase I cleavage site in the mutated protein that was not present in the WT protein (Figure S7D). Thus, the Trp15Cys mutation may modify the localization or stability of the truncated SORL1 protein isoform.

We extended our investigation to the Nanopore long-read RNA sequencing data derived from the NA12878 lymphoblastoid cell line and confirmed the existence of this truncated isoform in the human system (Figure S8). In addition, to assess the influence of the Trp15Cys mutation in a biological context, we created vectors expressing distinct HA-tagged versions of both WT and truncated mutated SORL1 proteins each harboring the Trp15Cys mutation. We then transfected these constructs into HEK293T cells. Interestingly, the level of truncated mutated SORL1 protein exhibited a noteworthy decrease, being approximately 17% lower than that of the WT protein ($p = .0094$; Figure 6A,B). Collectively, these results suggest the putative roles of rs2282647 (Trp15Cys) in conferring a protective effect against AD by specifically altering protein function and the level of one SORL1 isoform.

4 | DISCUSSION

Genetic analysis of AD can help predict disease risk, assist diagnosis, and develop interventions. Many recent studies focus on genetic factors associated with an increased risk of AD, such as *APOE* and *TREM2*.[57–59] Of note, recent research on AD-protective variants, including *SPI1* haplotypes and *APOE* coding mutations (eg, the Christchurch and Jacksonville mutations), have revealed biological pathways that may be associated with a reduced risk of AD.[60–63] These findings provide valuable insights into the development of intervention strategies for AD. Of note, *SORL1* is a leading AD GWAS candidate whose common variants exert AD-protective effects in multiple ethnic groups.[20–22] Accordingly, in the present study, we conducted transethnic, fine-mapping analysis of *SORL1* AD-protective genetic factors in both East Asian and European populations.[38] Our analysis identified variants and haplotypes that exert strong AD-protective effects but have not been covered by large-scale GWASs.[64] Specifically, we identified *SORL1* haplotype Hap_A, which confers dominant AD-protective effects in the general population. Meanwhile, the results of our meta-analysis of the East Asian population corroborate

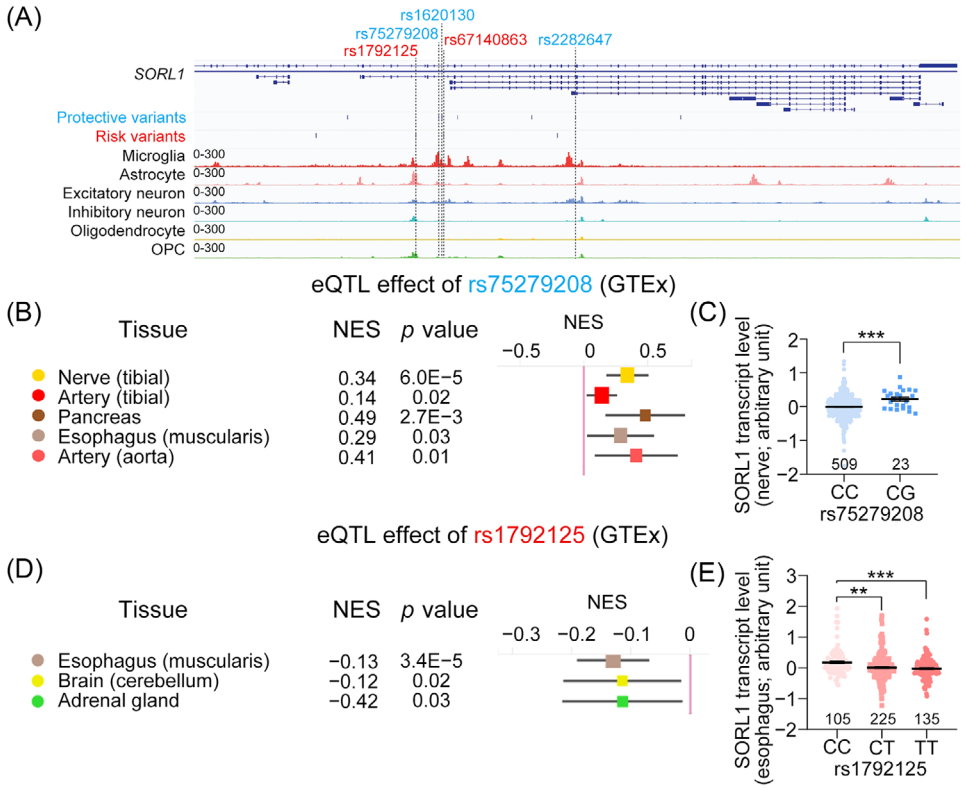

FIGURE 5 Effects of *SORL1* AD-protective and risk alleles on modulation of *SORL1* transcript expression. (A) Visualization of *SORL1* AD risk and AD-protective variants that sit in candidate cis-regulatory regions and brain cell active chromatin regions. The AD risk and protective variants are denoted in red and blue, respectively. The track visualizes the SORL1 isoforms, location of AD-protective and risk variants, and signals of brain single-cell ATAC-seq data retrieved from Corces et al.[49] (B) Association of AD-protective variant, rs75279208, with elevated *SORL1* transcript levels across different tissues in GTEx dataset. (C) *SORL1* transcript levels in nerve tissue among rs75279208 carriers ($n = 23$) and non-carriers ($n = 509$) (robust linear regression, $^{***}p < .001$). (D) Association of AD risk variant, rs1792125, with reduced *SORL1* transcript levels across different tissues in GTEx dataset. (E) *SORL1* transcript levels in esophageal tissue stratified by rs1792125 genotype ($n = 105$, 225, and 135 for homozygous non-carriers, heterozygous carriers, and homozygous carriers, respectively; robust linear regression, $^{**}p < .01$, $^{***}p < .001$). AD, Alzheimer's disease; ATAC-seq, assay for transposase accessibility by sequencing; eQTL, expression quantitative trait loci; GTEx, Genotype-Tissue Expression Project; NES, normalized effect size.

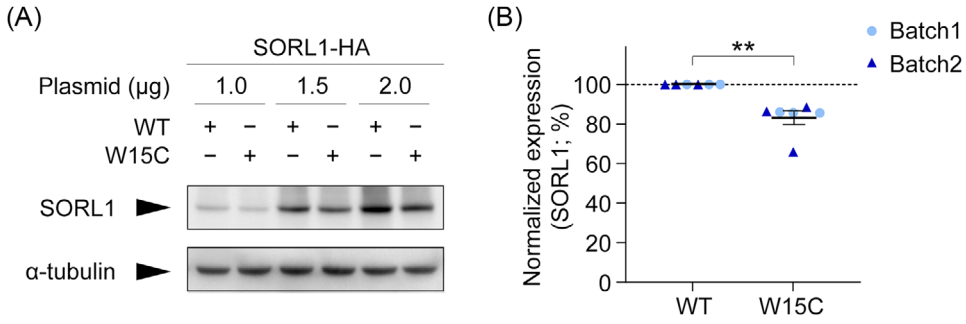

FIGURE 6 Impact of isoform-specific Trp15Cys coding variant on modulation of truncated SORL1 protein levels. (A) Western blots depicting expression levels of WT and mutated (W15C) HA-tagged truncated SORL1 proteins transfected into HEK293T cells with 1-, 1.5-, or 2-μg plasmids. The truncated SORL1 protein was detected using HA antibody, with α-tubulin concurrently blotted as loading control. (B) Dot plot illustrating relative expression levels between WT and W15C HA-tagged truncated SORL1 proteins. The expression of mutated SORL1 protein is normalized to both α-tubulin and WT SORL1 protein. A mixed-effects model (ie, REML) was employed to examine the impacts of both batch effects and mutation on the modulation of the truncated SORL1 protein. $^{**}p < .01$ indicates a significant effect of the Trp15Cys mutation on the modulation of protein level. SORL1-HA, HA-tagged truncated SORL1 protein; W15C, mutated truncated SORL1 protein with Trp15Cys mutation; WT, wild type.

findings reported in the literature, revealing significant associations at many of the reported AD-associated sites (Tables S20 and S21). Taken together, our analysis reveals specific genetic factors in the *SORL1* locus that confer AD-protective effects.

Previous functional studies of SORL1 protein in the context of AD primarily focus on its role in modulating amyloid precursor protein trafficking.[27] SORL1 directs the trafficking of amyloid precursor protein toward recycling pathways, thereby influencing the production of Aβ.[20] Concordantly, a recent study using *SORL1*-depleted human induced pluripotent stem cells suggests that *SORL1* is crucial for endosomal trafficking in human neurons.[65] Notably, our genotype–proteome association analysis suggests that *SORL1* modulates both synaptic- and immune-associated pathways (Figure 4D). These findings corroborate previous studies indicating the presence of novel SORL1 isoforms in neuronal dendrites[66] as well as the roles of SORL1 in the human microglial endolysosomal system.[26] Therefore, further investigation is required to elucidate how *SORL1* impacts synaptic and immune functions as well as its potential relationships with AD pathogenesis and progression.

Notably, our analysis revealed the isoform-specific coding variant rs2282647, resulting in a tryptophan-to-cysteine substitution at the 15th position within a less-explored truncated SORL1 protein (Figure S7). Importantly, this variant stands out as the sole variant present in minor haplotypes associated with AD in both the European population (ie, in variant cluster EU_5) and East Asian population (ie, in variant cluster EA_3; Figure 3), underscoring its potentially pivotal role in conferring AD-protective effects. Furthermore, our in silico analysis indicates a probable role of rs2282647 in modulating the structure and function of this relatively unexplored SORL1 isoform (Figure S7). Subsequent examination of long-read sequencing data supportedd the presence of this isoform in the human system (Figure S7), and our cell-based experiment further elucidated its role in modulating the levels of truncated SORL1 proteins (Figure 6). Hence, studies involving precise functional assessment will be required to delineate the roles of this SORL1 isoform in both normal physiological and disease contexts.

The FDA-approved Aβ monoclonal antibodies, aducanumab and lecanemab, bring side effects specifically in *APOE* ε4 carriers, who constitute a sizable population with an elevated risk of developing AD.[16,67] For instance, in the CLARITY AD trial for lecanemab, the incidence of amyloid-related imaging abnormalities varied among *APOE* ε4 non-carriers (5.4%), *APOE* ε4 heterozygotes (10.9%), and *APOE* ε4 homozygotes (32.6%).[68] This highlights the need for further exploration of alternative intervention strategies tailored specifically to *APOE* ε4 carriers. Our current analysis demonstrates a consistent protective effect of *SORL1* haplotype Hap_A irrespective of *APOE* ε4 genotype (Table S19). Notably, the AD risk associated with *APOE* ε4 appears to be lower among Hap_A carriers than non-carriers, with odds ratios of 3.39 and 3.77, respectively (Table S19). This protective effect may be due to the modulation of *SORL1* expression or function of SORL1 protein. Consequently, targeting *SORL1* expression or related pathways is emerging as a promising intervention strategy for AD, particularly among *APOE* ε4 carriers, who face heightened risks when undergoing monoclonal antibody therapy.

Accumulating evidence underscores the importance of studying genetic disease risks in people who are not of European ancestry.[69–71] Interactions between diverse genomic structures and ethnic backgrounds can produce genetic findings that are not generalizable to other populations. In addition to uncovering previously unknown genetic risks that contribute to diseases, the present study also showcases the utility of transethnic genetic analysis, which refines analyses of known genetic risk loci and reveals key population-specific genetic factors. Consequently, further investigation of AD genetic risks in more diverse populations will advance our understanding of currently known genetic risk factors. In turn, this may facilitate the integration of genetic data into routine clinical practice and therapeutic development.

AUTHOR CONTRIBUTIONS

X.Z., Yu C., A.K.F., and N.Y.I. conceived of the project. Yu C., T.C.K., V.C.M., and F.C.I. organized patient recruitment and sample collection. Yuewen C., R.M.L., B.W.W., and E.Y.C. organized the clinical data and prepared the samples for sequencing analysis. A.M., N.H., and T.I. contributed data for the Japanese cohort. X.Z. and H.C. set up the data-processing pipelines. W.Y.F. designed and coordinated the wet-lab experiment. X.Z., H.C., Y.J., H.Z., K.Y.M., A.M., N.H., T.I., J.H., A.K.F., and N.Y.I. analyzed the data. X.Z., H.C., A.K.F., and N.Y.I. wrote the manuscript with the input from all co-authors.

ACKNOWLEDGMENTS

We thank Pauline Kwan, Hazel Mok, Dr. Phillip Y.C. Chan, Choi Ying Ling, and Estella P.S. Tong for coordinating the collection of clinical data. We also thank Yifan Fu, Samuel San Yuen Chan, Tiffany T.W. Mak, Sarah Ma, Ka Chun Lok, and Cara Wing Si Kwong for their excellent technical assistance as well as other members of the Ip Laboratory for many helpful discussions. Data collection and sharing for the ADNI data were funded by the ADNI (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense Award W81XWH-12-2-0012). The ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and generous contributions from the following: AbbVie, Alzheimer's Association, Alzheimer's Drug Discovery Foundation, Araclon Biotech, BioClinica, Biogen, Bristol-Myers Squibb Company, CereSpir, Cogstate, Eisai, Elan Pharmaceuticals, Eli Lilly and Company, EuroImmun, F. Hoffmann-La Roche and its affiliated company Genentech, Fujirebio, GE Healthcare, IXICO, Janssen Alzheimer Immunotherapy Research & Development, Johnson & Johnson Pharmaceutical Research & Development, Lumosity, Lundbeck, Merck & Co., Meso Scale Diagnostics, NeuroRx Research, Neurotrack Technologies, Novartis Pharmaceuticals Corporation, Pfizer, Piramal Imaging, Servier, Takeda Pharmaceutical Company, and Transition Therapeutics. The Canadian Institutes of Health Research provides funds to support ADNI clinical sites in Canada. Private-sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. Data from the ADNI are disseminated by the

Laboratory for Neuro Imaging at the University of Southern California. This work was supported by the National Natural Science Foundation of China (NSFC)/RGC Joint Research Scheme (32061160472 and N_HKUST605/20), the Fundamental Research Program of Shenzhen Virtual University Park (2021Szvup137), the Research Grants Council of Hong Kong (the Collaborative Research Fund [C6027-19GF], the Theme-Based Research Scheme [T13-605/18 W], the General Research Fund [HKUST16103122]), the Areas of Excellence Scheme of the University Grants Committee (AoE/M-604/16), the Innovation and Technology Commission (InnoHK ITCPD/17-9, ITS/207/18FP, MRP/042/18X, and MRP/097/20X), the Guangdong Provincial Fund for Basic and Applied Basic Research (2019B1515130004), the Chow Tai Fook Charity Foundation (CTFCF18SC01), the Japan Agency for Medical Research and Development (AMED, Grant JP23dk0207060 [T.I.]), and Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan (Grants 21K07271 and 21H03537 [A.M.]).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. Author disclosures are available in the Supporting information.

DATA AVAILABILITY STATEMENT

The data and results of the genetic and AD-associated endophenotypic analyses are provided in the *Supplementary Information*. Regarding the raw data from the Chinese cohorts, the consent form signed by each participant states that the research content will remain private under the supervision of the hospital and research team. Therefore, these data will only be made available and shared in the context of a formal collaboration; applications for data sharing and project collaboration will be processed and reviewed by a review panel hosted at the Hong Kong University of Science and Technology. Researchers may contact sklneurosci@ust.hk for further details on project collaboration and the sharing of data from this study.

ETHICS STATEMENT

Y.J. and F.C.I. are cofounders of Cognitact. J.H. has served as a consultant for Eli Lilly and Eisai.

ORCID

Nancy Y. Ip https://orcid.org/0000-0002-2763-8907

REFERENCES

1. 2022 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2022;18:700-789.
2. Li X, Feng X, Sun X, Hou N, Han F, Liu Y. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2019. *Front Aging Neurosci*. 2022;14:937486.
3. Long JM, Holtzman DM. Alzheimer disease: an update on pathobiology and treatment strategies. *Cell*. 2019;179:312-339.
4. Bellenguez C, Grenier-Boley B, Lambert J-C. Genetics of Alzheimer's disease: where we are, and where we are going. *Curr Opin Neurobiol*. 2020;61:40-48.
5. Bertram L, Tanzi RE. Alzheimer disease risk genes: 29 and counting. *Nat Rev Neurol*. 2019;15:191-192.
6. Sims R, Hill M, Williams J. The multiplex model of the genetics of Alzheimer's disease. *Nat Neurosci*. 2020;23:311-322.
7. Andrews SJ, Fulton-Howard B, Goate A. Interpretation of risk loci from genome-wide association studies of Alzheimer's disease. *Lancet Neurol*. 2020;19:326-335.
8. Lambert J-C, 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(12):1452-1458. doi:10.1038/ng.2802
9. Guerreiro R, Wojtas A, Bras J, et al. TREM2 variants in Alzheimer's disease. *N Engl J Med*. 2013;368(2):117-127. doi:10.1056/NEJMoa1211851
10. Jonsson T, Stefansson H, Steinberg S, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. *N Engl J Med*. 2013;368(2):107-116. doi:10.1056/NEJMoa1211103
11. Zhou X, Chen Yu, Mok KY, et al. Identification of genetic risk factors in the Chinese population implicates a role of immune system in Alzheimer's disease pathogenesis. *Proc Natl Acad Sci*. 2018;115:1697-1706.
12. Jansen IE, Savage JE, Watanabe K, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet*. 2019;51:404-413.
13. Wightman DP, Jansen IE, Savage JE, et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat Genet*. 2021;53:1276-1282.
14. Bellenguez C, Küçükali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet*. 2022;54:412-436.
15. Sevigny J, Chiao P, Bussière T, et al. The antibody aducanumab reduces Aβ plaques in Alzheimer's disease. *Nature*. 2016;537:50-56.
16. Van Dyck CH, Swanson CJ, Aisen P, et al. Lecanemab in early Alzheimer's disease. *N Engl J Med*. 2023;388:9-21.
17. Cramer PE, Cirrito JR, Wesson DW, et al. ApoE-directed therapeutics rapidly clear β-amyloid and reverse deficits in AD mouse models. *Science (1979)*. 2012;335:1503-1506.
18. Wang C, Najm R, Xu Q, et al. Gain of toxic apolipoprotein E4 effects in human iPSC-derived neurons is ameliorated by a small-molecule structure corrector. *Nat Med*. 2018;24:647-657.
19. Seto M, Weiner RL, Dumitrescu L, Hohman TJ. Protective genes and pathways in Alzheimer's disease: moving towards precision interventions. *Mol Neurodegener*. 2021;16:29.
20. Rogaeva E, Meng Y, Lee JH, et al. The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nat Genet*. 2007;39:168-177.
21. Miyashita A, Koike A, Jun G, et al. SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians. *PLoS One*. 2013;8:e58618.
22. Zhang C-C, Wang H-Fu, Tan M-S, et al. SORL1 is associated with the risk of late-onset Alzheimer's disease: a replication study and meta-analyses. *Mol Neurobiol*. 2017;54:1725-1732.
23. Lee JH, Barral S, Reitz C. The neuronal sortilin-related receptor gene SORL1 and late-onset Alzheimer's disease. *Curr Neurol Neurosci Rep*. 2008;8:384-391.
24. Campion D, Charbonnier C, Nicolas G. SORL1 genetic variants and Alzheimer disease risk: a literature review and meta-analysis of sequencing data. *Acta Neuropathol*. 2019;138:173-186.
25. Knupp A, Mishra S, Martinez R, et al. Depletion of the AD risk gene SORL1 selectively impairs neuronal endosomal traffic independent of amyloidogenic APP processing. *Cell Rep*. 2020;31:107719.
26. Mishra S, Knupp A, Young JE, Jayadev S. Depletion of the AD risk gene SORL1 causes endo-lysosomal dysfunction in human microglia. *Alzheimer's & Dementia*. 2022;18:e068943.
27. Yin R-H, Yu J-T, Tan L. The role of SORL1 in Alzheimer's disease. *Mol Neurobiol*. 2015;51:909-918.

28. Naj AC, Jun G, Beecham GW, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet*. 2011;43:436-441.
29. Holstege H, de Waal MWJ, Tesi N, et al. Effect of prioritized SORL1 missense variants supports clinical consideration for familial Alzheimer's Disease. *medRxiv* 2023. doi:10.1101/2023.07.13.23292622
30. Holstege H, Hulsman M, Charbonnier C, et al. Exome sequencing identifies rare damaging variants in ATP8B4 and ABCA1 as risk factors for Alzheimer's disease. *Nat Genet*. 2022;54:1786-1794.
31. Raghavan NS, Brickman AM, Andrews H, et al. Whole-exome sequencing in 20,197 persons for rare variants in Alzheimer's disease. *Ann Clin Transl Neurol*. 2018;5:832-842.
32. Bellenguez C, Charbonnier C, Grenier-Boley B, et al. Contribution to Alzheimer's disease risk of rare variants in TREM2, SORL1, and ABCA7 in 1779 cases and 1273 controls. *Neurobiol Aging*. 2017;59:220.e1-220.e9.
33. Andersen OM, Monti G, Jensen AMG, et al. Relying on the relationship with known disease-causing variants in homologous proteins to predict pathogenicity of *SORL1* variants in Alzheimer's disease. 2023.
34. Liu G, Sun J-Y, Xu M, Yang X-Y, Sun B-L. SORL1 variants show different association with early-onset and late-onset Alzheimer's disease risk. *J Alzheimers Dis*. 2017;58:1121-1128.
35. Hung C, Tuck E, Stubbs V, et al. SORL1 deficiency in human excitatory neurons causes APP-dependent defects in the endolysosome-autophagy network. *Cell Rep*. 2021;35(11):109259.
36. Fazeli E, Fazeli E, Fojtík P, Holstege H, Andersen OM. Functional characterization of SORL1 variants in cell-based assays to investigate variant pathogenicity. *Philos Trans R Soc B Biol Sci*. 2024;379(1899):20220377.
37. Tejada Moreno JA, Villegas Lanau A, Madrigal Zapata L, et al. Mutations in SORL1 and MTHFDL1 possibly contribute to the development of Alzheimer's disease in a multigenerational Colombian Family. *PLoS One*. 2022;17(7):e0269955.
38. Zhou X, Chen Y, Ip FCF, et al. Genetic and polygenic risk score analysis for Alzheimer's disease in the Chinese population. *Alzheimers Dement*. 2020;12:e12074.
39. Zhou X, Chen Yu, Ip FCF, et al. Deep learning-based polygenic risk analysis for Alzheimer's disease prediction. *Commun Med*. 2023;3:49.
40. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559-575.
41. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190.
42. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat Genet*. 2019;51:414-430.
43. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581:434.
44. Kunkle BW, Schmidt M, Klein H-U. Novel Alzheimer disease risk loci and pathways in African American individuals using the African Genome Resources Panel: a meta-analysis. *JAMA Neurol*. 2021;78:102-113.
45. Zhou X, Chen Yu, Mok KY, et al. Non-coding variability at the APOE locus contributes to the Alzheimer's risk. *Nat Commun*. 2019;10:3310.
46. Jiang Y, Zhou X, Ip FC, et al. Large-scale plasma proteomic profiling identifies a high-performance biomarker panel for Alzheimer's disease screening and staging. *Alzheimers Dement*. 2022;18:88-102.
47. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation*. 2021;2:100141.
48. ENCODE Project Consortium *et al.* Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*. 2020;583:699-710.
49. Corces MR, Shcherbina A, Kundu S, et al. Single-cell epigenomic analyses implicate candidate causal variants at inherited risk loci for Alzheimer's and Parkinson's diseases. *Nat Genet*. 2020;52:1158-1168.
50. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12:996-1006.
51. Cheng J, Randall AZ, Sweredoski MJ, Baldi P. SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res*. 2005;33:W72.
52. Hu G, Katuwawala A, Wang K, et al. flDPnn: accurate intrinsic disorder prediction with putative propensities of disorder functions. *Nat Commun*. 2021;12:1-8.
53. Teufel F, Almagro Armenteros JJ, Johansen AR, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol*. 2022;40:1023-1025.
54. Workman RE, Tang AD, Tang PS, et al. Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nat Methods*. 2019;16:1297-1305.
55. Kent WJ. BLAT – the BLAST-like alignment tool. *Genome Res*. 2002;12:656-664.
56. Byrska-Bishop M, Evani US, Zhao X, et al. High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios. *Cell*. 2022;185:3426-3440.e19.
57. Serrano-Pozo A, Das S, Hyman BT. APOE and Alzheimer's disease: advances in genetics, pathophysiology, and therapeutic approaches. *Lancet Neurol*. 2021;20:68-80.
58. Carmona S, Zahs K, Wu E, Dakin K, Bras J, Guerreiro R. The role of TREM2 in Alzheimer's disease and other neurodegenerative disorders. *Lancet Neurol*. 2018;17:721-730.
59. Shi Y, Holtzman DM. Interplay between innate immunity and Alzheimer disease: APOE and TREM2 in the spotlight. *Nat Rev Immunol*. 2018;18:759-772.
60. Huang K-L, Marcora E, Pimenova AA, et al. A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. *Nat Neurosci*. 2017;20:1052-1061.
61. Cao H, Zhou X, Chen Yu, et al. Association of SPI1 haplotypes with altered SPI1 gene expression and Alzheimer's disease risk. *J Alzheimers Dis*. 2022. doi:10.3233/JAD-215311
62. Arboleda-Velasquez JF, Lopera F, O'hare M, et al. Resistance to autosomal dominant Alzheimer's disease in an APOE3 Christchurch homozygote: a case report. *Nat Med*. 2019;25:1680-1683.
63. Liu C-C, Murray ME, Li X, et al. APOE3-Jacksonville (V236E) variant reduces self-aggregation and risk of dementia. *Sci Transl Med*. 2021;13:eabc9375.
64. Rajabli F, Benchek P, Tosto G, et al. Multi-ancestry genome-wide meta-analysis of 56,241 individuals identifies LRRC4C, LHX5-AS1 and nominates ancestry-specific loci PTPRK, GRB14, and KIAA0825 as novel risk loci for Alzheimer's disease: the Alzheimer's Disease Genetics Consortium. *medRxiv* 2023. doi:10.1101/2023.07.06.23292311
65. Knupp A, Mishra S, Martinez R, et al. Depletion of the AD risk gene SORL1 selectively impairs neuronal endosomal traffic independent of amyloidogenic APP processing. *Cell Rep*. 2020;31:107719.
66. Monti G, Kjolby M, Jensen AMG, et al. Expression of an alternatively spliced variant of SORL1 in neuronal dendrites is decreased in patients with Alzheimer's disease. *Acta Neuropathol Commun*. 2021;9:43.
67. Salloway S, Chalkias S, Barkhof F, et al. Amyloid-related imaging abnormalities in 2 phase 3 studies evaluating aducanumab in patients with early Alzheimer disease. *JAMA Neurol*. 2022;79:13-21.
68. Cummings J, Apostolova L, Rabinovici GD, et al. Lecanemab: appropriate use recommendations. *J Prev Alzheimers Dis*. 2023;10:362-377.
69. Sleiman PM, Qu HQ, Connolly JJ, et al. Trans-ethnic genomic informed risk assessment for Alzheimer's disease: an International Hundred K+ Cohorts Consortium study. *Alzheimers Dement*. 2023;13:16.
70. Iwatsubo T, Iwata A, Suzuki K, et al. Japanese and North American Alzheimer's Disease Neuroimaging Initiative studies:

harmonization for international trials. *Alzheimers Dement*. 2018;14:1077-1087.

71. Lake J, Warly Solsberg C, Kim JJ, et al. Multi-ancestry meta-analysis and fine-mapping in Alzheimer's disease. *Mol Psychiatry*. 2023;28(7):3121-3132. doi:10.1038/s41380-023-02089-w

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhou X, Cao H, Jiang Y, et al. Transethnic analysis identifies *SORL1* variants and haplotypes protective against Alzheimer's disease. *Alzheimer's Dement*. 2024;1-17. https://doi.org/10.1002/alz.14214