Title: Cell-type-specific Alzheimer's disease polygenic risk scores are associated with distinct disease processes in Alzheimer's disease

Authors:

Hyun-Sik Yang[1,2,3,4,*], Ling Teng[1,4], Daniel Kang[2], Vilas Menon[5], Tian Ge[3,4,6], Hilary K. Finucane[3,4,7], Aaron P. Schultz[2,3], Michael Properzi[2], Hans-Ulrich Klein[5], Lori B. Chibnik[2,4,8], Julie A. Schneider[9], David A. Bennett[9], Timothy J. Hohman[10], Richard P. Mayeux[5], Keith A. Johnson[1,2,3,11], Philip L. De Jager[5], Reisa A. Sperling[1,2,3]

[1]Department of Neurology, Brigham and Women's Hospital, Boston, MA, USA

[2]Department of Neurology, Massachusetts General Hospital, Boston, MA, USA

[3]Harvard Medical School, Boston, MA

[4]Broad Institute of MIT and Harvard, Cambridge, MA

[5]Department of Neurology, Columbia University Irving Medical Center, New York, NY, USA

[6]Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA

[7]Department of Medicine, Massachusetts General Hospital, Boston, MA, USA

[8]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[9]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA

[10]Department of Neurology, Vanderbilt University Medical Center, Nashville, TN, USA

[11]Department of Radiology, Massachusetts General Hospital, Boston, MA, USA

*Corresponding Author: Hyun-Sik Yang, MD

E-mail: hyang18@bwh.harvard.edu

Abstract

Alzheimer's disease (AD) heritability is enriched in glial genes, but how and when cell-type-specific genetic risk contributes to AD remains unclear. Here, we derive cell-type-specific AD polygenic risk scores (ADPRS) from two extensively characterized datasets. In an autopsy dataset spanning all stages of AD (n=1,457), astrocytic (Ast) ADPRS was associated with both diffuse and neuritic Aβ plaques, while microglial (Mic) ADPRS was associated with neuritic Aβ plaques, microglial activation, tau, and cognitive decline. Causal modeling analyses further clarified these relationships. In an independent neuroimaging dataset of cognitively unimpaired elderly (n=2,921), Ast-ADPRS were associated with Aβ, and Mic-ADPRS was associated with Aβ and tau, showing a consistent pattern with the autopsy dataset. Oligodendrocytic and excitatory neuronal ADPRSs were associated with tau, but only in the autopsy dataset including symptomatic AD cases. Together, our study provides human genetic evidence implicating multiple glial cell types in AD pathophysiology, starting from the preclinical stage.

Introduction

Alzheimer's disease (AD) is the most common cause of dementia and is among the leading causes of death, but clinically effective disease-modifying intervention has been challenging[1]. A major barrier to AD drug discovery is the complex pathophysiology driven by multiple cell types interacting with amyloid-β (Aβ) and tau proteinopathies[1-4]. Large genome-wide association studies (GWAS) of AD dementia have identified dozens of candidate causal genes, many highly expressed in microglia and astrocytes[5-7]. Further, numerous sub-threshold genetic associations are also enriched in microglial genes[8-13]. Microglia are resident immune cells of the central nervous system (CNS), implicated in Aβ clearance, Aβ-related neuroinflammation, and Aβ- and tau-related neurodegeneration[1,2]. Astrocytes are the hub of lipoprotein and cholesterol metabolism in the CNS, a process closely related to Aβ metabolism, and astrocytes may also exacerbate neurodegeneration[2,14]. Both activated microglia and astrocytes are found in close proximity to neuritic Aβ plaques[1,2,15], a pathologic hallmark of AD. With strong support from human genetics and accumulating experimental evidence, microglia and astrocytes are emerging as promising cellular targets for potential disease-modifying interventions in AD.

However, how and when the AD genetic risk localizing to these cell types contributes to distinct processes in AD remains unclear, making it very difficult to design clinical trials that can precisely target the right cellular program at the right disease stage. AD dementia is the end result of multiple related yet distinct disease processes that gradually progress over more than two decades, with substantial clinical-pathological heterogeneity[3]. There is significant variability in the rate of in vivo tau accumulation in individuals with similar Aβ burden[16], and immune pathways and microglial response might impact tau accumulation above and beyond Aβ[4,17-19]. Further, the rate of cognitive decline is highly variable even in the setting of similar Aβ and tau burden (cognitive resilience)[20-22], partially attributable to the differential cellular responses to neuropathologic insult[17,21,23]. Thus, cell-type-specific heritability analyses based on case-control AD dementia GWAS results alone, that does not account for individual-level AD endophenotype (e.g., Aβ, tau, cognitive decline) variability, do not have the resolution to localize the genetic association to specific AD endophenotype ("how") at a specific phase of disease progression

("when"). While AD endophenotype GWAS is a promising approach to fill this gap, these studies are not yet powered for robust cell-type-specific heritability analyses despite recent growth in sample size[23-26].

Therefore, we need an approach to combine well-powered AD dementia GWAS results with deeply characterized individual-level data in a cell-type-specific manner. Previous studies have identified direct associations of several AD dementia GWAS variants with key AD endophenotypes such as Aβ plaques[27-29], tau tangles[30], and cognitive decline[29,30]. Yet, most AD dementia GWAS loci have small effect sizes that cannot be robustly examined with moderate sample sizes of well-characterized datasets with the AD endophenotype data. Aggregating genetic effects with AD polygenic risk scores (ADPRS) can enable robust detection of overall genetic effects on AD endophenotypes, especially when including both genome-wide significant and sub-threshold genetic associations[31-33], but conventional ADPRS lacks cellular specificity. In this context, a promising alternative is a gene-set-based PRS approach[13,34-38] to capture cell-type-specific AD genetic risk profiles from each individual. Previous studies have applied gene-set-based ADPRS to predict AD dementia in a pathway- or cell-type-specific manner[13,37,38] and reported results largely consistent with the re-analysis of AD dementia GWAS summary statistics[5-13]. Nonetheless, the relationship between cell-type-specific AD genetic risk and AD endophenotypes in specific disease stages remains largely unknown.

Here, we derive single nucleus RNA sequencing (snRNA-seq)-guided cell-type-specific ADPRS from two extensively characterized datasets and clarify how and when AD genetic risk related to specific cell types contributes to distinct disease processes in AD. We first leverage extensive post-mortem neuropathology data, including quantitative AD pathology and histologically observed microglial activation, from two community-based cohorts that span the full pathologic and clinical disease severity spectrum. We observe specific associations of astrocytic (Ast) and microglial (Mic) ADPRS with distinct AD endophenotypes and perform causal modeling analyses. Then, we focus on the preclinical (asymptomatic) stage of AD and examine the association between cell-type-specific ADPRS and neuroimaging AD biomarkers in a clinical trial screening dataset, replicating our key findings and establishing the early role of astrocytic and microglial genetic risk in AD pathogenesis.

Results

Study Participants

Our study participants are from two independent datasets. First, we examined the impact of cell-type-specific ADPRS on the longitudinal cognitive and post-mortem neuropathology data from the Religious Orders Study and the Rush Memory and Aging Project (ROSMAP) (n=1,457, mean age 89.7±6.5, 67% female, 69 % with elevated Aβ, 45% with dementia; **Table 1**). ROS and MAP are community-based cohorts with annual cognitive exams and comprehensive postmortem neuropathologic evaluation, and the full spectrum of pathologic and clinical stages of AD are represented[39]. Second, we analyzed the genetic and phenotypic data from the pre-randomization (screening) phase of the Anti-Amyloid Treatment in Asymptomatic Alzheimer's (A4) study[40], a secondary AD prevention trial. The A4 screening dataset consists of CU older adults with or without elevated Aβ, assessed with florbetapir positron emission tomography (PET) (n=2,921, mean age 71.4±4.8, 60% female, 30% with elevated Aβ; **Table 1**), enabling us to test whether cell-type-specific ADPRS also impacts the in vivo AD endophenotypes in the earliest stages of AD.

Derivation of cell-type-specific ADPRS

We adapted and combined previously described approaches[10,11,13,34-38,41,42] to derive cell-type-specific ADPRS (**Fig. 1a**). We first addressed the linkage disequilibrium (LD) structure using a Bayesian regression approach (PRS-CS)[41] and performed effect size shrinkage on the base AD GWAS summary statistics (Bellenguez et al., stage I)[5] to assign posterior effect sizes for each genetic variant. This LD shrinkage step only uses the GWAS summary statistics for optimization, avoiding overfitting to the target datasets (see Methods). Then, we used a previously published neocortical snRNA-seq dataset from 24 controls[43] (ROSMAP participants with no or very little pathology; no participant overlap with our study) and derived cell-type-specific gene lists from six major brain cell types: excitatory neurons (Ex), inhibitory neurons (In), astrocytes (Ast), microglia (Mic), oligodendrocytes (Oli), and oligodendrocyte

precursor cells (Opc). After removing the *APOE* region (*APOE* ± 1 MB), we defined a cell-type-specific gene list by selecting genes within the top 10% of expression specificity for each cell type (i.e., top 1343 genes), as described in previous studies[10,11]. Only a minor proportion of cell-type-specific genes were specific to two or more brain cell types (**Fig. 1b**), and we allowed this overlap as some genes may have important roles in multiple (but not all) cell types. Each cell-type-specific ADPRS was computed from ROSMAP and A4, using 30 kb margins upstream and downstream of cell-type-specific genes. The cell-type-specific ADPRSs included 7.6% – 10.4% of all examined variants (**Supplementary Table 1**), and they were orthogonal to each other (R^2<0.1), except for the Mic- and Oli-ADPRS pair with a strong positive correlation (R^2=0.31 in ROSMAP, R^2=0.26 in A4) (**Fig. 1c-d**). Much of the shared variance between Mic- and Oli-ADPRSs remained even when the PRSs were derived without genomic margin (R^2=0.21 in ROSMAP, R^2=0.18 in A4), indicating that the correlation between the two is likely due to overlapping genes rather than overlapping genomic margins. Mic-specific and Oli-specific gene sets shared 136 genes (10.1% of each set), which include known AD risk genes[5], such as *ADAM10*, *BIN1*, *CR1*, and *PICALM*.

Cell-type-specific ADPRSs were associated with distinct AD endophenotypes in ROSMAP

Then, we tested the association of cell-type-specific ADPRSs with seven AD endophenotypes in ROSMAP: autopsy-confirmed AD dementia, immunohistochemistry (IHC)-assessed post-mortem Aβ and paired helical filament tau (PHFtau), Bielschowsky silver stain-assessed diffuse plaque (DP), neuritic plaque (NP), and neurofibrillary tangle (NFT), and longitudinally assessed cognitive decline (**Supplementary Table 2**). IHC enables a molecularly specific and quantitative assessment of Aβ and tau pathology, while the silver stain allows a separate assessment of specific morphological subtypes of Aβ plaques (DP vs. NP). The ADPRS derived from the entire autosomal genome excluding the *APOE* region (All-ADPRS) was associated with all endophenotypes except for DP (**Fig. 2**, **Supplementary Tables 3 - 9**).

Consistent with the previous studies focusing on cell-type-specific heritability of AD dementia[5-13], Mic-ADPRS was the only cell-type-specific ADPRS significantly associated with increased odds of autopsy-confirmed AD dementia (**Fig. 2a** and **Supplementary Table 3**). For comparison, we derived Mic-ADPRS using PRSet[13], a previously published gene-set-based PRS approach that detected AD dementia heritability enrichment in microglial genes in a target dataset with a sample size of >350k[13]. The association between the PRSet-derived Mic-ADPRS and AD dementia was not statistically significant in our moderately-sized dataset (OR=1.16, 95% CI 0.97 to 1.39, p=0.12), indicating that our cell-type-specific ADPRS has an improved statistical power compared to the previously published method.

By contrast, distinct association patterns were observed between the cell-type-specific ADPRSs and other AD endophenotypes. Ast-ADPRS was the only cell-type-specific ADPRS significantly associated with all three tested measures of post-mortem fibrillar Aβ burden: IHC-assessed Aβ and silver-stain assessed DP and NP (**Fig. 2b-d** and **Supplementary Tables 4-6**). Oli-, and Mic-ADPRSs were only associated with the NP burden (**Fig. 2d** and **Supplementary Table 6**). DP is an amorphous aggregation of Aβ with minimal cellular reaction representing early-stage Aβ plaque, while NP contains a dense core with surrounding neuroglial reaction including dystrophic neurites, activated microglia, and reactive astrocytes[44,45]. Thus, our finding suggests that astrocytic genetic programs contribute to Aβ accumulation starting from the early stages of fibrillar Aβ formation, perhaps by shifting the balance between Aβ production and clearance, and has the greatest impact on the overall fibrillar Aβ level. On the other hand, microglial and oligodendrocytic genetic programs may primarily contribute to celluar reaction to Aβ leading to NP formation[46-48].

Multiple cell-type-specific ADPRSs (Mic-, Oli-, Ast-, and Ex-ADPRSs) were associated with IHC-assessed PHFtau burden, and the strongest association was observed with Mic-ADPRS (**Fig. 2e** and **Supplementary Table 7**). Oli-, Ast-, and Ex-ADPRS remained nominally associated ($p<0.05$) with PHFtau even after adjusting for Mic-ADPRS (**Supplementary Table 10**), and Oli-, Ast-, and Ex-ADPRS calculated after excluding genes overlapping with microglia were all associated with PHFtau (**Supplementary Table 11**). Mic-, Oli-, and Ast-ADPRS were also associated with silver stain-assessed

NFT burden, while Ex-ADPRS – NFT association did not reach statistical significance (**Fig. 2f** and **Supplementary Table 8**).

Mic- and Oli-ADPRS were significantly associated with cognitive decline (**Fig. 2g** and **Supplementary Table 9**), but oli-ADPRS was no longer associated with cognitive decline (p=0.41) after adjusting for Mic-ADPRS. We note that cognitive decline in older adults is a complex phenomenon: less than 50% of the variability can be explained even if multiple pathologies and other known contributors are considered together[21,49]. This complexity might have undermined statistical power to detect the weaker associations, leaving Mic-ADPRS as the only cell-type-specific ADPRS independently associated with cognitive decline and dementia after accounting for multiple testing corrections.

In a subset of ROSMAP participants who had morphological assessments of microglial activation (n=201 MAP participants, demographics summarized in **Supplementary Table 12**), we explored whether Mic-ADPRS is associated with the proportion of activated microglia (PAM[50]). Histologically characterized microglial activation from the neocortex has a strong association with AD pathology, but its associations with known AD risk variants, including *APOE* ε4 (p=0.85 in our study), were not significant at the single variant level[50]. Mic-ADPRS was associated with an increased PAM (**Fig. 2h**; beta=9.2×10^{-3}, 95% CI 8.4×10^{-4} to 0.017, p=0.031). Thus, our cell-type-specific ADPRS showed that a higher microglial AD genetic risk may lead to dysfunctional microglial activation, a relationship that was not apparent at the single AD GWAS variant level[50].

All significant associations between cell-type-specific ADPRSs and AD endophenotypes remained similar even when we used different genomic margins for PRS (results from genes±10 kb or ±100kb; **Supplementary Table 13**). None of the observed trait – cell-type-specific ADPRS associations were moderated by age, sex, or *APOE* ε4 dosage.

To summarize, multiple cell-type-specific ADPRS were associated with AD endophenotypes in ROSMAP, consistent with the view that multiple cell types might contribute to AD pathophysiology[1-4]. To further clarify the relationship between cell-type-specific ADPRSs and AD endophenotypes, we performed causal modeling analyses as detailed in the next section.

Causal modeling analyses mapped Mic- and Ast-ADPRS to distinct events in the AD pathophysiologic cascade

We performed causal modeling to map the contribution of each cell-type-specific ADPRS to the sequence of events in AD pathophysiology, focusing on Mic- and Ast-ADPRS that showed significant associations with multiple AD endophenotypes. Although Oli-ADPRS was also strongly associated with NP and tau, we had to exclude it from this modeling approach, given the difficulty in statistically separating its effect from the colinear Mic-ADPRS. Here, we leveraged that genetic variants are assigned randomly at conception and are not subject to reverse causation[51]. We also used the postulated sequence of AD progression as the prior for our model: Aβ accumulation starts as DP, which evolves into NP with surrounding gliosis, triggering tau NFT formation and cognitive decline[4,44].

We first performed causal mediation analyses to distinguish direct and mediated effects among cell-type-specific ADPRSs and AD endophenotypes (**Fig. 3a-d** and **Supplementary Table 14**). Ast-ADPRS had a direct effect on NP not fully mediated by DP (**Fig. 3a**), while the direct Ast-ADPRS – NFT association was no longer significant after considering the NP-mediated effect (**Fig. 3b**). On the other hand, there were significant direct effects of Mic-ADPRS on NFT and cognitive decline even after accounting for the upstream processes (**Fig. 3c-d**).

Synthesizing these results, we constructed a structural equation model (SEM) from n=1,392 ROSMAP participants with no missing data (**Fig. 3e**). This SEM has an excellent model fit and highlights the distinct contribution of each cell-type-specific ADPRS: Ast-ADPRS affects AD pathophysiology mainly through its effect on Aβ (diffuse and neuritic plaques), while Mic-ADPRS has a broader impact on multiple core pathological and clinical endophenotypes of AD (NP, NFT, and cognitive decline). Moreover, Mic-ADPRS influenced cognitive decline above and beyond AD pathology, suggesting the role of microglia in cognitive resilience. We acknowledge that this model derived using post-mortem cross-sectional data cannot prove a causal relationship. Still, our apporach provides a plausible model based on human genetics that can inform future mechanistic studies.

Cell-type-specific ADPRSs were associated with in vivo AD biomarkers in CU older adults

Then, we used in vivo neuroimaging biomarker data from CU older adults in the A4 screening data and assessed the role of cell-type-specific AD genetic risk in preclinical AD. We tested four AD endophenotypes in A4 (Aβ PET, tau PET, hippocampal volume [HV; a marker of neurodegeneration], and Preclinical Alzheimer Cognitive Composite [PACC][52]; **Supplementary Table 15**). All-ADPRS was associated with all tested AD endophenotypes, while cell-type-specific ADPRSs showed distinct association patterns (**Fig. 4**).

Ex-, Ast-, Mic-, and Oli-ADPRS were significantly associated with in vivo Aβ (FDR<0.025; **Fig. 3a** and **Supplementary Table 16**). Ex- and Ast-ADPRS remained nominally associated with Aβ after adjusting for Mic-ADPRS or excluding genes overlapping with ADPRS, while Oli-ADPRS – Aβ association was no longer present (**Supplementary Table 17, 18**). This cell-type-specific ADPRS – Aβ PET association resembled the cell-type-specific ADPRS – NP association from ROSMAP (**Fig. 2d**), and we think this is because the PET radiotracers for Aβ have a greater affinity to NP than DP[53,54]. A larger sample size in A4 likely enabled us to detect the additional Ex-ADPRS –Aβ association.

In a smaller subset with tau PET data (n=302; demographics summarized in **Supplementary Table 19**), only Mic-ADPRS was significantly associated (FDR<0.025) with tau (temporal lobe composite) (**Fig. 4b** and **Supplementary Table 15**). This differs from multiple cell-type-specific ADPRS associations with PHFtau and NFT in ROSMAP. Interestingly, in a subset of ROSMAP participants who were CU (n=454), only Mic-ADPRS was associated with PHFtau (beta=0.16, $p=1.0\times10^{-4}$) while Oli-, Ast-, and Ex-ADPRS were not (p>0.05; **Supplementary Table 21**). Thus, despite important differences in phenotype measurement, results from ROSMAP and A4 suggest a coherent biology: microglia may exacerbate tau pathology starting from the preclinical stage of AD, while other cell types may contribute to tau pathology later in the symptomatic disease stages.

On the other hand, there was a limited cell-type-specific ADPRS association with neurodegeneration (HV) or cognition (PACC) in A4. In a subset with structural MRI data (n=1,266;

demographics summarized in **Supplementary Table 22**), HV was not associated with any cell-type-specific ADPRSs (**Fig. 4c** and **Supplementary Table 23**). PACC, a sensitive cognitive composite optimized to detect early Aβ-related cognitive decline[52], was only associated with Ast-ADPRS (**Fig. 4d** and **Supplementary Table 24**), and this association remained similar even after adjusting for PET-measured Aβ (beta=-0.10, p=0.015). The A4 study is likely underpowered to detect the impact of cell-type-specific AD genetic risk on neurodegeneration or cognitive impairment because all participants in the A4 screening dataset were CU without extensive AD-related neurodegeneration or cognitive decline. Nonetheless, the Aβ-independent association between Ast-ADPRS and PACC hints at a possible effect of Ast-ADPRS on early cognitive decline above and beyond AD pathology.

All significant associations between cell-type-specific ADPRS and AD endophenotypes in A4 were robust to the size of genomic margins (genes±10 kb or ±100kb; **Supplementary Table 25**). None of the observed trait – cell-type-specific ADPRS associations were significantly moderated by age, sex, or *APOE* ε4 dosage.

Discussion

We derived cell-type-specific ADPRSs in two independent and well-characterized datasets to show that AD genetic risk localizing to different neuroglial cell types makes distinct contributions to AD pathophysiology. Our findings provide human genetics evidence to support the disease model where astrocytes play an important early role in Aβ clearance before plaque maturation, while microglia are primarily involved in later phases of Aβ plaque maturation (i.e., NP formation) and abnormal tau accumulation. Lipoprotein and cholesterol metabolism, primarily driven by astrocytes in the CNS, are thought to be linked to Aβ metabolism[2]. Further, Aβ clearance occurs through the blood-brain barrier (BBB) and perivascular circulation, and astrocytes are among the main constituents of the BBB[2]. Thus, AD genetic risk localizing to astrocyte-specific genes may collectively undermine the Aβ metabolism and perivascular Aβ clearance leading to initial parenchymal fibrillar Aβ accumulation (DP). Then, dysfunctional microglial activation—which the aggregate microglial AD genetic risk may drive—and

ineffective fibrillar Aβ removal may lead to NP formation, with additional contributions from astrocytes. On the other hand, tau pathology accumulation is likely to be initially driven by microglia in the preclinical stage, with later contributions of oligodendrocytes and cell-autonomous actions of excitatory neurons. Interestingly, a recent spatial transcriptomics study from mice[15] showed that Aβ plaques are surrounded immediately by microglia and more distantly by astrocytes, while hyperphosphorylated tau affects excitatory neurons in an environment enriched with oligodendrocytes, providing a landscape coherent with what our human genetics study suggests.

It is important to note that the association of each cell-type-specific ADPRS with AD endophenotypes was much weaker than that of *APOE* loci, and the current cell-type-specific ADPRS is unlikely to be a clinically useful tool for disease risk stratification. Nonetheless, our study demonstrates that cell-type-specific PRS can be used to gain deeper pathophysiologic insights from well-characterized cohorts and guide future mechanistic and clinical-translational studies. For example, cell-type-specific PRS could be leveraged for a genetically guided sampling of induced pluripotent stem cell (iPSC) lines for specific cell type differentiation, or it can be used for cell-type-specific pharmacogenomic studies of anti-Aβ immunotherapies. Further, our study provides genetic support to use in vivo Aβ and tau PET—both associated with Mic-ADPRS in preclinical AD—as intermediate biomarker read-outs in future AD prevention trials modulating microglia.

Our approach significantly extends and improves previously published methods to determine cell type specificity in AD dementia heritability[5-11] in the following aspects. First, we derive individual-level PRS to not only assess cell-type-specific genetic contributions to the final outcome of AD dementia but also examine the impact of each cell-type-specific AD genetic risk on key pathophysiologic events in AD. This is a significant advancement toward the genetic dissection of distinct AD endophenotypes in humans. Second, by combining summary statistics-driven optimization of Bayesian LD shrinkage (PRS-CS) with cell-type-specific partitioning of PRS, our approach improved statistical power while avoiding overfitting during the target dataset-based p-value thresholding. Third, our approach allows straightforward adjustment of other cell-type-specific ADPRS or AD endophenotypes in the analysis by

simply including them as a covariate, which is another significant advantage over purely summary statistics-based computational methods. This strength enabled us to perform complex multivariate causal modeling that included multiple cell-type-ADPRSs and AD endophenotypes in the same model, thereby clarifying the role of each cell-type-specific ADPRS in AD progression.

Several limitations should be considered in interpreting our results. First, our results establish associations between the cell-type-informed genetic risk and AD endophenotypes but do not directly identify the mechanisms of these associations. Second, our results may have been affected by the parameters and base datasets we chose, such as the number of genes to be considered cell-type-specific (e.g., top 10%), genomic margin near the target genes (e.g., 30 kb), the snRNA dataset, and the base GWAS. To maximize comparability, we followed the convention adopted by previous studies on cell-type-specific heritability of AD dementia[9-11] (e.g., using top 10% of the cell-type-specific genes). The rationales for choosing the particular snRNA dataset[43] and the base GWAS[5] are detailed in Methods, and we showed that the association between cell-type-specific ADPRS and AD endophenotypes were robust to the genomic margins selected. Third, our cell-type-informed ADPRS excludes the larger *APOE* region and cannot assess the cell-type-specific impact of genes within this region. Also, our study focuses on common variants and does not consider rare variants that may have strong effects on AD endophenotypes. Fourth, our study focuses on comparing across major cell types, but it is well known that diverse cell states (subtypes) exist within each cell type, including disease-associated cell states[43,55]. Also, we could not examine AD genetic risk localizing to endothelial cells and pericytes in this study, given the small number of vascular niche cell types profiled in a typical snRNA-seq study. With increasing sample size and newer methods[56], we hope to assess cell-state-specific AD risk targeting rarer cell states—including endothelial and pericytes subtypes—in the near future. Finally, we limited our study to participants of European ancestry, as well-powered AD GWAS summary statistics were only available from individuals of European ancestry. Our current results may not generalize to other ancestries, and well-powered AD GWAS from non-European ancestries are urgently required to address the racial and ethnic disparities in AD genomics research.

Despite these limitations, our study leveraged two well-characterized datasets to reveal robust and coherent direct associations of AD genetic risk localizing to different glial cell type with distinct disease processes in AD, including in the preclinical stage. Further, our cell-type-specific PRS can be extended to any other phenotypes beyond AD, as far as there are available GWAS summary statistics and well-characterized target datasets. Future studies combining cell-type-specific polygenic approaches with large-scale multimodal data from deeply phenotyped cohorts and model systems could enable further causal dissection of the cellular contributions to AD pathogenesis.

Methods

ROSMAP: participants and phenotypic characterization

The Religious Orders Study (ROS) started in 1994 and is enrolling Catholic priests, brothers, and nuns across religious communities in the United States[39]. The Rush Memory and Aging Project (MAP) started in 1997 and is enrolling diverse participants from northern Illinois[39]. ROS and MAP were approved by an Institutional Review Board (IRB) of Rush University Medical Center. Each participant signed an informed consent, Anatomic Gift Act, and Repository Consent allowing their data to be repurposed. Both studies enrolled older participants who did not have known dementia at enrollment and agreed to organ donation after death (overall autopsy rate > 85%). ROS and MAP (ROSMAP) were designed for combined analyses, and the same team of investigators at Rush Alzheimer's Disease Center (RADC) perform coordinated clinical and neuropathological assessments. By the time of their death, ROSMAP participants exhibit a broad and continuous spectrum of cognitive and functional impairment (ranging from cognitively unimpaired to dementia) and neuropathology burden (ranging from no pathology to severe neurodegenerative/cerebrovascular pathology), representing the general aging population[39]. Deceased participants of European ancestry who had quality-controlled genome-wide genetic data and immunohistochemistry evaluation of either Aβ or tau burden were included in our study (n=1,457).

Each participant got a comprehensive annual cognitive evaluation including the following 19 tests spanning multiple cognitive domains[39]: Word List Memory/Recall/Recognition, East Boston Immediate/Delayed Recall, Logical memory immediate/delayed, Boston Naming Test, Category Fluency, reading test (10 items), Digit Span forward/backward/ordering, Judgment of Line Orientation, Standard Progressive Matrices, Symbol Digit Modalities Test, Number Comparison, Stroop Color Naming, and Stroop Word Reading. Each participant's annual global cognitive function was defined as the average z-scores from these tests (standardized to baseline measures). Longitudinal cognitive decline was captured by a random slope of global cognitive function from a linear mixed-effect model adjusting for baseline age, sex, and years of education and their time interaction terms[39,57]. Final clinical diagnoses (cognitively unimpaired, mild cognitive impairment, and dementia) were assigned by a neurologist using all available antemortem clinical data without access to the postmortem neuropathologic evaluation[58].

A comprehensive post-mortem neuropathological evaluation was performed to quantify AD pathology (amyloid-β [Aβ] plaques and tau neurofibrillary tangles)[59,60]. Immunohistochemistry was used to assess the overall Aβ and paired helical filament tau (PHFtau) across eight brain regions (hippocampus, entorhinal cortex, mid-frontal cortex, inferior temporal cortex, angular gyrus, calcarine cortex, anterior cingulate cortex, and superior frontal cortex). Quantitative Aβ and PHFtau burdens were defined as the percentage area occupied by each pathology, averaged across the eight brain regions. In addition, we also analyzed distinct subtypes of Aβ plaques (diffuse plaques and neuritic plaques) — that reflect varying decree of local neuroglial reaction to Aβ accumulation[44,45] — and neurofibrillary tangles (NFT; tau pathology) assessed with silver-stained slides from five brain regions (entorhinal cortex, hippocampus [CA1], mid temporal cortex, inferior parietal cortex, and midfrontal cortex). Then, the count from each region was scaled with the corresponding standard deviation and averaged to derive quantitiave summary measures of diffuse plaque and neuritic plaque burdens. All quantitative AD pathology variables were square-rooted for further analyses given their positively skewed distributions. "Elevated Aβ" was defined as the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) neuritic plaque score of "definite" or "probable." A pathologic AD diagnosis was made using the modified

National Institute on Aging-Reagan Institute criteria[61]. A diagnosis of AD with dementia (autopsy-confirmed AD dementia) was made when a participant with pathologic AD also had the final clinical diagnosis of dementia[3]. The control group was defined as individuals without pathologic AD, who were also deemed to be cognitively unimpaired (CU) per the final clinical diagnosis.

Microglial density was assessed using microscopic examination in a subset of MAP participants, with the following morphologic criteria[50]: stage 1, not activated (thin, ramified processes); stage 2, activated without macrophage-like appearance (rounded cell body >14 μm with thickened processes); stage 3, activated with macrophage-like appearance (cell body >14 μm). The proportion of activated microglia (PAM) was defined as the square root of the proportion of stage 3 microglia[50]. We used average PAM from two neocortical regions (inferior temporal and mid-frontal) that showed significant association with AD pathology in a previous study[50].

A4 screening data: participants and phenotypic characterization

The A4 study is a secondary prevention trial that enrolled CU older adults (between age 65 and 85) with evidence of cortical Aβ accumulation on PET imaging from 67 sites in the United States, Australia, Canada, and Japan[40,62]. The study protocol was approved by IRBs at each participating site, and all participants signed informed consent before the study procedures. Inclusion criteria to select CU older adults included Clinical Dementia Rating global score of 0, Mini-Mental State Examination (MMSE) score of 25 to 30, and Logical Memory Delayed Recall-Iia (LMDR-Iia) score of 6 to 18. 6,763 participants underwent cognitive screening, and 4,486 participants who met the cognitive inclusion criteria (i.e., CU) had a screening positron emission tomography (PET) to quantify fibrillar Aβ burden. Although only those determined to have elevated Aβ were eligible to be randomized in the A4 clinical trial, we used all available data from the screening visits for this study. Participants of European ancestry with florbetapir PET and genetic data were included in our study (n=2,921).

The Preclinical Alzheimer Cognitive Composite (PACC)[52], a cognitive composite optimized to detect early Aβ-related cognitive changes, was calculated by averaging z-scores of the following four

cognitive tests at screening: MMSE, the Free and Cued Selective Reminding Test (total recall), LMDR-Iia, and the Digit Symbol Substitution Test.

We used ^{18}F-florbetapir PET to quantify Aβ. The florbetapir PET was acquired between 50 and 70 minutes after injecting florbetapir, and a mean cortical standardized uptake value ratio (SUVR) was calculated using a whole cerebellar reference. "Elevated Aβ" was defined as with florbetapir PET cortical SUVR ≥1.15, or 1.15 > SUVR > 1.10 and a positive visual read[62]. We used ^{18}F-flortaucipir (FTP) PET to quantify tau. The FTP PET was acquired between 80 and 110 minutes after injecting FTP, and tau was quantified using FTP SUVR from the bilateral temporal composite region of interest (ROI), which includes the entorhinal cortex, parahippocampal gyrus, fusiform gyrus, and inferior temporal cortex. These are among the earliest regions to accumulate cerebral tau pathology in AD, and the tau burden in these regions is associated with cognitive decline[16,63,64]. We did not apply partial volume correction (PVC). 3D T1-weighted brain MRI was done, and the image was processed with NeuroQuant (http://www.cortechslabs.com/neuroquant) for automated segmentation and segmental volume calculation. We used bilateral hippocampal volume (HV) as a marker of neurodegeneration, adjusting for intracranial volume (ICV).

Genetic data acquisition, processing, and study participant selection

In ROSMAP, DNA was extracted from blood or postmortem brain tissue. Codons 112 and 158 from *APOE* exon 4 were sequenced to determine *APOE* haplotypes (ε2, ε3, or ε4). Genome-wide genotyping was performed on the Affymetrix GeneChip 6.0 platform (n=1,878), the Illumina OmniQuad Express platform (n=566), or the (Illumina) Infinium Global Screening Array (n=494)[65]. Data from each genotyping platform were processed using the same quality control (QC) pipeline, using PLINK[66]: genotype call rate (SNP) > 95%, minor allele frequency (MAF) > 0.01, non-random missingness (SNP) $p<10^{-9}$, Hardy-Weinberg equilibrium $p<1.0\times10^{-6}$, genotype success rate (individual) < 95%, concordant sex, and no excess heterozygosity (individual). Closely related individuals (identity-by-descent pi-hat>0.1) were excluded. Principal components (PCs) of the genotype covariance matrix were derived

using EIGENSTRAT[67], and population outliers (including all participants of non-European descent) were removed (using the default setting) to avoid confounding by population structure. This resulted in a total of n=2,496 participants with quality-controlled genetic data. The variants were phased using Eagle v2.4[68], and palindromic variants were removed before imputation. Imputation was performed separately for each genotyping platform using Michigan Imputation Server[69] and the Haplotype Reference Consortium (HRC) reference panel (r1.1 2016, lifted-over to GRCh38 coordinates)[70], and variants with MAF>0.01 and imputation quality R^2>0.8 from all three platforms were retained. We merged imputed data from all three genotyping platforms, and after removing EIGENSTRAT outliers from the merged dataset, we had n=2,385 participants with 6,577,494 genetic variants. We included n=1,457 deceased participants with *APOE* genotypes and immunohistochemistry quantification of AD pathology (Aβ or tau), who were not part of the single nucleus RNA-sequencing (snRNA-seq) dataset[43] used to define cell-type-specific gene sets.

In the A4 screening data, DNA was extracted from blood. Targeted genotyping was used to derive *APOE* ε2, ε3, and ε4 haplotypes. Genome-wide genotyping was performed for 3,465 consenting A4 screen participants using Illumina Global Screening Array, resulting in 700,078 genotyped variants[26]. We used the same genetic data QC pipeline as ROSMAP. We limited our analyses to non-Hispanic Whites and removed ancestry outliers identified with the genotype PCs. We used the same imputation procedures as ROSMAP, resulting in 7,269,997 variants (GRCh38) in 3,025 participants. Among these participants, 2,921 participants who also had florbetapir PET and *APOE* genotypes were included in our analysis.

Imputed genotype dosages from both datasets were rounded to integers (0, 1, or 2) before use in further analyses.

Derivation of cell-type-specific ADPRS

Derivation of cell-type-specific ADPRS requires (1) base GWAS summary statistics, (2) LD reference panel, (3) cell-type-specific gene sets, and (4) a target dataset with individual-level genotype data.

(1) Base GWAS summary statistics and (2) LD reference panel: We used the summary statistics from a large genome-wide association study of AD dementia and AD dementia by proxy (parental history of AD dementia)[5]. Among multiple AD GWASs that were recently published, we chose the study by Bellenguez et al.[5] because it used the European Alzheimer & Dementia Biobank (EADB) and the UK Biobank (UKBB) datasets for stage I, and thus ROSMAP or A4 was not a part of its stage I summary statistics (i.e., no sample overlap). Further, the Bellenguez et al. study[5] included the largest number of cases and identified the most genome-wide significant loci among the GWASs published before the time of our analysis. We used PRS-CS[41], a Bayesian regression approach using continuous shrinkage prior, to perform effect estimate shrinkage and derive the posterior effect size of each single nucleotide polymorphism (SNP) included in the AD GWAS summary statistics (stage I of Bellenguez et al.[5]) while minimizing information loss. We used the 1000 Genome Project Phase 3 European subset (1000G EUR)[71] as the LD reference panel and used the "PRS-CS-auto" option to estimate the global shrinkage parameter (φ)—that reflects the sparsity of the genetic architecture—for each chromosome. From the base AD GWAS'[5] stage I summary statistics (limited to the variants assessed in >50% of the stage I participants), φ was $1.2\times10^{-4} \pm 1.7\times10^{-5}$ (calculated per each chromosome). This parameter optimization step only uses the base GWAS summary statistics and the LD reference panel, independent of the target dataset characteristics, thereby avoiding an overfitting problem.

We only used the HapMap3[72] SNPs (~1 million SNPs) included in the 1000G EUR reference panel for a computationally tractable estimation of φ from the AD GWAS summary statistics, as described in the original implementation of PRS-CS[41]. While HapMap3 is a relatively small reference panel, it effectively captures heritability attributable to common haplotypes when used in Bayesian genetic effect size shrinkage methods: in a previous study, PRS-CS has outperformed other PRS

approaches that used the full 1000G EUR panel (>4 million SNPs), and the PRS-CS model performance only slightly improved even when a denser reference panel was used[41].

(3) Cell-type-specific gene sets: We used published snRNA-seq data of 34,987 cells (nuclei) from the prefrontal cortex of n=24 control participants from ROSMAP (no to very little pathology at autopsy)[43] to derive cell-type-specific gene sets. At the time of the analyses, this dataset was one of the largest snRNA datasets published from older adults with no to little pathology. These participants were excluded from our study to prevent reserve causation. We used the cell type annotation from the snRNA-seq study reported this data[43], and included six major brain cell types in our analyses: excitatory neurons (Ex), inhibitory neurons (In), astrocytes (Ast), microglia (Mic), oligodendrocytes (Oli), and oligodendrocyte precursor cells (Opc). Similar to the previous study[43], we excluded endothelial cells and pericytes from analyses as only a small number of cells were profiled from these cell types. We adapted a previously published method[10,11] to derive the top decile of the genes specifically expressed in each cell type. First, we identified 13,438 autosomal genes expressed in >1% of cells from one or more cell types. We excluded genes within the larger *APOE* region (*APOE* ± 1 Mb: GRCh38 19:43,905,781 – 45,909,393), as *APOE* ε4 has a disproportionately large effect size that would dwarf the effects of other common variants. Second, a gene expression specificity metric ("Sg") was calculated for gene *i* in cell type *j*, by dividing the average expression of gene *i* in cell type *j* (E_{ij}) by the summation of E_{ij} across all 6 cell types (i.e., $Sg_{ij} = E_{ij}/(\sum_j E_{ij})$). Third, we rank-ordered Sg of genes expressed in each cell type and defined genes within the top decile (n=1,343) as cell-type-specific genes.

(4) PRS calculation in target datasets: We used PLINK 1.90[66,73] to calculate PRS. All PRSs excluded the larger *APOE* region. Cell-type-specific ADPRSs were computed using the subset of the variants located within each cell-type-specific gene set ± 30 kb (per GRCh38 reference coordinates). We chose a 30 kb margin upstream and downstream to include most cis-regulatory variants of a given gene and to allow for mapping errors due to the LD structure[74,75]. Then, to ensure that our results are robust to the choice of the genomic margin, we compared our results with the results derived from PRSs using a 10 kb margin or a

100 kb margin. We also calculated conventional ("All") ADPRS for the ROSMAP and A4 participants to compare with cell-type-specific ADPRSs. Each PRS was standardized before further analyses.

Statistical Analysis

All statistical analyses were done with R version 4.2 (https://cran. r-project.org/). Participants with missing values were excluded from each analysis, and we indicated the number of participants included in each analysis. We used the UpSetR package[76] to visualize the membership of cell-type-specific genes. Correlations between cell-type-specific PRSs were examined with Pearson's correlation and were summarized with a heatmap colored by R^2.

In ROSMAP, we tested the association of cell-type-specific ADPRS with seven phenotypes: AD with dementia (AD dementia, binary), Aβ (continuous), diffuse plaque (DP), neuritic plaque (NP), PHFtau (continous), NFT (continuous), and cognitive decline (CogDec, continuous). Models with AD dementia as an outcome used logistic regression controlling for *APOE* ε4 dosage (0, 1, or 2), *APOE* ε2 dosage (0, 1, or 2), age at death, sex, years of education, genotyping platform, and the first three genotype principal components (PC1-3). Models with neuropathology as an outcome used linear regression controlling for *APOE* ε4 dosage, *APOE* ε2 dosage, age at death, sex, genotyping platform, and PC1-3. Models with CogDec as an outcome used linear regression controlling for *APOE* ε4 dosage, *APOE* ε2 dosage, genotyping platform, and PC1-3; Age, sex, and years of education were already accounted for when deriving the CogDec variable from the longitudinal cognitive data. We also assess the association of Mic-ADPRS with PAM, adjusting for *APOE* ε4 dosage, *APOE* ε2 dosage, age at death, sex, genotyping platform, and PC1-3.

In A4, we tested the association of cell-type-specific ADPRS with four phenotypes: screening neocortical Aβ (continuous), baseline temporal tau (continuous), baseline HV (continuous), and screening PACC (continuous). Models with Aβ or tau as an outcome used linear regression controlling for *APOE* ε4 dosage, *APOE* ε2 dosage, age, sex, and PC1-3. Models with HV as an outcome used linear regression controlling for *APOE* ε4 dosage, *APOE* ε2 dosage, age, sex, intracranial volume (ICV), and PC1-3.

Models with PACC as an outcome used linear regression controlling for *APOE* ε4 dosage, *APOE* ε2 dosage, age, sex, years of education, and PC1-3.

We calculated the false discovery rate (FDR) for each dataset (separately for ROSMAP [49 analyses] and A4 [28 analyses]) and used FDR<0.025 (=0.05/2) as the statistical significance threshold for the main discovery analyses (Fig. 2 and 4) given two independent datasets being used for testing. Reported effect sizes of *APOE* ε4 and ε2 were assessed in models including All-ADPRS. Mic-independent cell-type-specific associations using the following two approaches: (1) adjusting for Mic-ADPRS, and (2) deriving cell-type-specific ADPRS excluding genes overlapping with Mic-ADPRS. We also performed sensitivity analyses using different genomic margins for PRS derivation (genes±10 kb or ±100 kb) to ensure that the significant results were not driven by our choice of the genomic margin (genes±30 kb). For the association of Mic-ADPRS with PAM, we used a nominal p-value threshold of p<0.05, and this was a targeted, post hoc analysis. We examined the moderating effect of age, sex, and *APOE* ε4 by examining the association between the interaction terms ([age, sex, or *APOE* ε4]×[cell-type-specific ADPRS]) with AD endophenotypes and used p<0.017 (=0.05/3 potentially moderating variables) as our significance threshold.

For causal mediation analysis, we used the widely accepted sequence of AD pathophysiology progression[4,44] as our prior: DP→NP→NFT→CogDec. We used R package 'mediation[77]' for causal mediation analysis using a non-parametric bootstrap option with 10,000 simulations. Then, we used R package 'lavaan' for structural equation modeling (SEM). The model was fitted for individuals with non-missing data (n=1,392) using default lavaan settings except that we used bootstrapping (10,000 iterations), using residualized variables (i.e., the residual after regressing out *APOE* ε4 and ε2, age at death, sex, genotyping platform, and PC1-3 from a linear model). Model fit was assessed with multiple indices, including Comparative Fit Index (CFI), Tucker Lewis Index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR).

Reference

1 Long, J. M. & Holtzman, D. M. Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. *Cell* **179**, 312-339, doi:10.1016/j.cell.2019.09.001 (2019).

2 De Strooper, B. & Karran, E. The Cellular Phase of Alzheimer's Disease. *Cell* **164**, 603-615, doi:10.1016/j.cell.2015.12.056 (2016).

3 Jack, C. R., Jr. *et al.* NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* **14**, 535-562, doi:10.1016/j.jalz.2018.02.018 (2018).

4 Selkoe, D. J. & Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med* **8**, 595-608, doi:10.15252/emmm.201606210 (2016).

5 Bellenguez, C. *et al.* New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet* **54**, 412-436, doi:10.1038/s41588-022-01024-z (2022).

6 Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet* **51**, 404-413, doi:10.1038/s41588-018-0311-9 (2019).

7 Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet* **51**, 414-430, doi:10.1038/s41588-019-0358-2 (2019).

8 Gagliano, S. A. *et al.* Genomics implicates adaptive and innate immunity in Alzheimer's and Parkinson's diseases. *Ann Clin Transl Neurol* **3**, 924-933, doi:10.1002/acn3.369 (2016).

9 Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet* **50**, 621-629, doi:10.1038/s41588-018-0081-4 (2018).

10 Bryois, J. *et al.* Genetic identification of cell types underlying brain complex traits yields insights into the etiology of Parkinson's disease. *Nat Genet* **52**, 482-493, doi:10.1038/s41588-020-0610-9 (2020).

11 Skene, N. G. *et al.* Genetic identification of brain cell types underlying schizophrenia. *Nat Genet* **50**, 825-833, doi:10.1038/s41588-018-0129-5 (2018).

12 Jagadeesh, K. A. *et al.* Identifying disease-critical cell types and cellular processes across the human body by integration of single-cell profiles and human genetics. *bioRxiv*, doi:10.1101/2021.03.19.436212 (2021).

13 Choi, S. W. *et al.* PRSet: Pathway-based polygenic risk score analyses and software. *PLoS Genet* **19**, doi:10.1371/journal.pgen.1010624 (2023).

14 Liddelow, S. A. *et al.* Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* **541**, 481-487, doi:10.1038/nature21029 (2017).

15 Zeng, H. *et al.* Integrative in situ mapping of single-cell transcriptional states and tissue histopathology in a mouse model of Alzheimer's disease. *Nat Neurosci*, doi:10.1038/s41593-022-01251-x (2023).

16 Hanseeuw, B. J. *et al.* Association of Amyloid and Tau With Cognition in Preclinical Alzheimer Disease: A Longitudinal Study. *JAMA Neurol* **76**, 915-924, doi:10.1001/jamaneurol.2019.1424 (2019).

17 Green, G. S. *et al.* Cellular dynamics across aged human brains uncover a multicellular cascade leading to Alzheimer's disease. *bioRxiv*, doi:10.1101/2023.03.07.531493 (2023).

18 Pascoal, T. A. *et al.* Microglial activation and tau propagate jointly across Braak stages. *Nat Med* **27**, 1592-1599, doi:10.1038/s41591-021-01456-w (2021).

19 Yang, H. S. *et al.* Plasma IL-12/IFN-gamma axis predicts cognitive trajectories in cognitively unimpaired older adults. *Alzheimers Dement* **18**, 645-653, doi:10.1002/alz.12399 (2022).

20 Montine, T. J. *et al.* Concepts for brain aging: resistance, resilience, reserve, and compensation. *Alzheimers Res Ther* **11**, 22, doi:10.1186/s13195-019-0479-y (2019).

21 White, C. C. *et al.* Identification of genes associated with dissociation of cognitive performance and neuropathological burden: Multistep analysis of genetic, epigenetic, and transcriptional data. *PLoS Med* **14**, e1002287, doi:10.1371/journal.pmed.1002287 (2017).

22 Yu, L. *et al.* Residual decline in cognition after adjustment for common neuropathologic conditions. *Neuropsychology* **29**, 335-343, doi:10.1037/neu0000159 (2015).

23 Dumitrescu, L. *et al.* Genetic variants and functional pathways associated with resilience to Alzheimer's disease. *Brain* **143**, 2561-2575, doi:10.1093/brain/awaa209 (2020).

24 Beecham, G. W. *et al.* Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. *PLoS Genet* **10**, e1004606, doi:10.1371/journal.pgen.1004606 (2014).

25 Damotte, V. *et al.* Plasma amyloid beta levels are driven by genetic variants near APOE, BACE1, APP, PSEN2: A genome-wide association study in over 12,000 non-demented participants. *Alzheimers Dement* **17**, 1663-1674, doi:10.1002/alz.12333 (2021).

26 Raghavan, N. S. *et al.* Association Between Common Variants in RBFOX1, an RNA-Binding Protein, and Brain Amyloidosis in Early and Preclinical Alzheimer Disease. *JAMA Neurol* **77**, 1288-1298, doi:10.1001/jamaneurol.2020.1760 (2020).

27 Apostolova, L. G. *et al.* Associations of the Top 20 Alzheimer Disease Risk Variants With Brain Amyloidosis. *JAMA Neurol* **75**, 328-341, doi:10.1001/jamaneurol.2017.4198 (2018).

28 Bradshaw, E. M. *et al.* CD33 Alzheimer's disease locus: altered monocyte function and amyloid biology. *Nat Neurosci* **16**, 848-850, doi:10.1038/nn.3435 (2013).

29 Chibnik, L. B. *et al.* CR1 is associated with amyloid plaque burden and age-related cognitive decline. *Ann Neurol* **69**, 560-569, doi:10.1002/ana.22277 (2011).

30 Franzmeier, N., Rubinski, A., Neitzel, J., Ewers, M. & Alzheimer's Disease Neuroimaging, I. The BIN1 rs744373 SNP is associated with increased tau-PET levels and impaired memory. *Nat Commun* **10**, 1766, doi:10.1038/s41467-019-09564-5 (2019).

31 Tan, C. H. *et al.* Polygenic hazard score, amyloid deposition and Alzheimer's neurodegeneration. *Brain* **142**, 460-470, doi:10.1093/brain/awy327 (2019).

32 Mormino, E. C. *et al.* Polygenic risk of Alzheimer disease is associated with early- and late-life processes. *Neurology* **87**, 481-488, doi:10.1212/WNL.0000000000002922 (2016).

33 Ge, T. *et al.* Dissociable influences of APOE epsilon4 and polygenic risk of AD dementia on amyloid and cognition. *Neurology* **90**, e1605-e1612, doi:10.1212/WNL.0000000000005415 (2018).

34 Rammos, A. *et al.* The role of polygenic risk score gene-set analysis in the context of the omnigenic model of schizophrenia. *Neuropsychopharmacology* **44**, 1562-1569, doi:10.1038/s41386-019-0410-z (2019).

35 Yao, Y. *et al.* Cell type-specific and cross-population polygenic risk score analyses of MIR137 gene pathway in schizophrenia. *iScience* **24**, 102785, doi:10.1016/j.isci.2021.102785 (2021).

36 Zeng, Y. *et al.* A Combined Pathway and Regional Heritability Analysis Indicates NETRIN1 Pathway Is Associated With Major Depressive Disorder. *Biol Psychiatry* **81**, 336-346, doi:10.1016/j.biopsych.2016.04.017 (2017).

37 Tesi, N. *et al.* Immune response and endocytosis pathways are associated with the resilience against Alzheimer's disease. *Transl Psychiatry* **10**, 332, doi:10.1038/s41398-020-01018-7 (2020).

38 Bellou, E. *et al.* Age-dependent effect of APOE and polygenic component on Alzheimer's disease. *Neurobiol Aging* **93**, 69-77, doi:10.1016/j.neurobiolaging.2020.04.024 (2020).

39 Bennett, D. A. *et al.* Religious Orders Study and Rush Memory and Aging Project. *J Alzheimers Dis* **64**, S161-S189, doi:10.3233/JAD-179939 (2018).

40 Sperling, R. A. *et al.* The A4 study: stopping AD before symptoms begin? *Sci Transl Med* **6**, 228fs213, doi:10.1126/scitranslmed.3007941 (2014).

41 Ge, T., Chen, C. Y., Ni, Y., Feng, Y. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun* **10**, 1776, doi:10.1038/s41467-019-09718-5 (2019).

42 Watanabe, K., Umicevic Mirkov, M., de Leeuw, C. A., van den Heuvel, M. P. & Posthuma, D. Genetic mapping of cell type specificity for complex traits. *Nat Commun* **10**, 3222, doi:10.1038/s41467-019-11181-1 (2019).

43 Mathys, H. *et al.* Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332-337, doi:10.1038/s41586-019-1195-2 (2019).

44 Mackenzie, I. R., Hao, C. & Munoz, D. G. Role of microglia in senile plaque formation. *Neurobiol Aging* **16**, 797-804, doi:10.1016/0197-4580(95)00092-s (1995).

45 Serrano-Pozo, A., Frosch, M. P., Masliah, E. & Hyman, B. T. Neuropathological alterations in Alzheimer disease. *Cold Spring Harb Perspect Med* **1**, a006189, doi:10.1101/cshperspect.a006189 (2011).

46 Hyman, B. T. *et al.* National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement* **8**, 1-13, doi:10.1016/j.jalz.2011.10.007 (2012).

47 Mrak, R. E. Microglia in Alzheimer brain: a neuropathological perspective. *Int J Alzheimers Dis* **2012**, 165021, doi:10.1155/2012/165021 (2012).

48 Duyckaerts, C., Delatour, B. & Potier, M. C. Classification and basic pathology of Alzheimer disease. *Acta Neuropathol* **118**, 5-36, doi:10.1007/s00401-009-0532-1 (2009).

49 Boyle, P. A. *et al.* To what degree is late life cognitive decline driven by age-related neuropathologies? *Brain* **144**, 2166-2175, doi:10.1093/brain/awab092 (2021).

50 Felsky, D. *et al.* Neuropathological correlates and genetic architecture of microglial activation in elderly human brain. *Nat Commun* **10**, 409, doi:10.1038/s41467-018-08279-3 (2019).

51 Lawlor, D. A., Harbord, R. M., Sterne, J. A., Timpson, N. & Davey Smith, G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* **27**, 1133-1163, doi:10.1002/sim.3034 (2008).

52 Donohue, M. C. *et al.* The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. *JAMA Neurol* **71**, 961-970, doi:10.1001/jamaneurol.2014.803 (2014).

53 Chapleau, M., Iaccarino, L., Soleimani-Meigooni, D. & Rabinovici, G. D. The Role of Amyloid PET in Imaging Neurodegenerative Disorders: A Review. *J Nucl Med* **63**, 13S-19S, doi:10.2967/jnumed.121.263195 (2022).

54 Ikonomovic, M. D. *et al.* Post-mortem analyses of PiB and flutemetamol in diffuse and cored amyloid-beta plaques in Alzheimer's disease. *Acta Neuropathol* **140**, 463-476, doi:10.1007/s00401-020-02175-1 (2020).

55 Cain, A. *et al.* Multi-cellular communities are perturbed in the aging human brain and with Alzheimer's disease. *bioRxiv*, 2020.2012.2022.424084, doi:10.1101/2020.12.22.424084 (2020).

56 Yang, A. C. *et al.* A human brain vascular atlas reveals diverse mediators of Alzheimer's risk. *Nature* **603**, 885-892, doi:10.1038/s41586-021-04369-3 (2022).

57 De Jager, P. L. *et al.* A genome-wide scan for common variants affecting the rate of age-related cognitive decline. *Neurobiol Aging* **33**, 1017 e1011-1015, doi:10.1016/j.neurobiolaging.2011.09.033 (2012).

58 Schneider, J. A., Arvanitakis, Z., Bang, W. & Bennett, D. A. Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology* **69**, 2197-2204, doi:10.1212/01.wnl.0000271090.28148.24 (2007).

59 Bennett, D. A. *et al.* Apolipoprotein E epsilon4 allele, AD pathology, and the clinical expression of Alzheimer's disease. *Neurology* **60**, 246-252, doi:10.1212/01.wnl.0000042478.08543.f7 (2003).

60 Bennett, D. A., Schneider, J. A., Wilson, R. S., Bienias, J. L. & Arnold, S. E. Neurofibrillary tangles mediate the association of amyloid load with clinical Alzheimer disease and level of cognitive function. *Arch Neurol* **61**, 378-384, doi:10.1001/archneur.61.3.378 (2004).

61 Consensus recommendations for the postmortem diagnosis of Alzheimer's disease. The National Institute on Aging, and Reagan Institute Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease. *Neurobiol Aging* **18**, S1-2 (1997).

62 Sperling, R. A. *et al.* Association of Factors With Elevated Amyloid Burden in Clinically Normal Older Individuals. *JAMA Neurol* **77**, 735-745, doi:10.1001/jamaneurol.2020.0387 (2020).

63 Braak, H., Alafuzoff, I., Arzberger, T., Kretzschmar, H. & Del Tredici, K. Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathol* **112**, 389-404, doi:10.1007/s00401-006-0127-z (2006).

64 Johnson, K. A. *et al.* Tau positron emission tomographic imaging in aging and early Alzheimer disease. *Ann Neurol* **79**, 110-119, doi:10.1002/ana.24546 (2016).

65 De Jager, P. L. *et al.* A multi-omic atlas of the human frontal cortex for aging and Alzheimer's disease research. *Sci Data* **5**, 180142, doi:10.1038/sdata.2018.142 (2018).

66 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795 (2007).

67 Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-909, doi:10.1038/ng1847 (2006).

68 Loh, P. R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet* **48**, 1443-1448, doi:10.1038/ng.3679 (2016).

69 Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat Genet* **48**, 1284-1287, doi:10.1038/ng.3656 (2016).

70 McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* **48**, 1279-1283, doi:10.1038/ng.3643 (2016).

71 The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68-74, doi:10.1038/nature15393 (2015).

72 International HapMap Consortium *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52-58, doi:10.1038/nature09298 (2010).

73 Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7, doi:10.1186/s13742-015-0047-8 (2015).

74 Vosa, U. *et al.* Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet* **53**, 1300-1310, doi:10.1038/s41588-021-00913-z (2021).

75 Wu, Y., Zheng, Z., Visscher, P. M. & Yang, J. Quantifying the mapping precision of genome-wide association studies using whole-genome sequencing data. *Genome Biol* **18**, 86, doi:10.1186/s13059-017-1216-0 (2017).

76 Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938-2940, doi:10.1093/bioinformatics/btx364 (2017).

77 Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. mediation: R package for causal mediation analysis. *Journal of Statistical Software* **59** (2014).

Acknowledgments: We thank the participants and study staff of the Religious Orders Study (ROS), the Rush Memory and Aging Project (MAP), and the A4 Study. This work was funded by the United States National Institutes of Health (K23AG062750 [H.-S.Y.]). ROSMAP is supported by NIH grants P30AG010161 [D.A.B.], P30AG072975 [D.A.B.], R01AG015819 [D.A.B.], R01AG017917 [D.A.B], U01AG046152 [P.L.D, D.A.B], and U01AG061356 [P.L.D, D.A.B]. The A4 study (NCT02008357) was funded by the National Institute on Aging (grants U19AG010483 [R.A.S. and others] and R01AG063689 [R.A.S. and others]), Eli Lilly and Co, and several philanthropic organizations. Parts of the figures were created with BioRender.com.

Author Contributions:

H.-S.Y.: Study design and conception; analysis and interpretation of data; drafting and substantially revising the work.

L.T.: Analysis and interpretation of data; substantially revising the work.

D.K.: Analysis and interpretation of data; substantially revising the work.

V.M.: Analysis and interpretation of data; substantially revising the work.

T.G.: Study design and conception; interpretation of data; substantially revising the work.

H.K.F.: Study design and conception; interpretation of data; substantially revising the work.

A.P.S.: Analysis and interpretation of data; substantially revising the work.

M.P.: Analysis and interpretation of data; substantially revising the work.

H.-U.K.: Interpretation of data; substantially revising the work.

L.B.C.: Interpretation of data; substantially revising the work.

J.A.S.: Acquisition and interpretation of data; substantially revising the work.

D.A.B.: Acquisition and interpretation of data; substantially revising the work.

T.J.H.: Interpretation of data; substantially revising the work.

R.P.M.: Acquisition and interpretation of data; and substantially revising the work.

K.A.J.: Acquisition and interpretation of data; and substantially revising the work.

P.L.D.: Study design and conception; acquisition and interpretation of data; substantially revising the work.

R.A.S.: Study design and conception; acquisition and interpretation of data; substantially revising the work.

Competing Interests:

The authors declare no competing interests in this work.

Materials & Correspondence: Correspondence and requests for materials should be addressed to H.-S.Y. (e-mail: hyang18@bwh.harvard.edu)

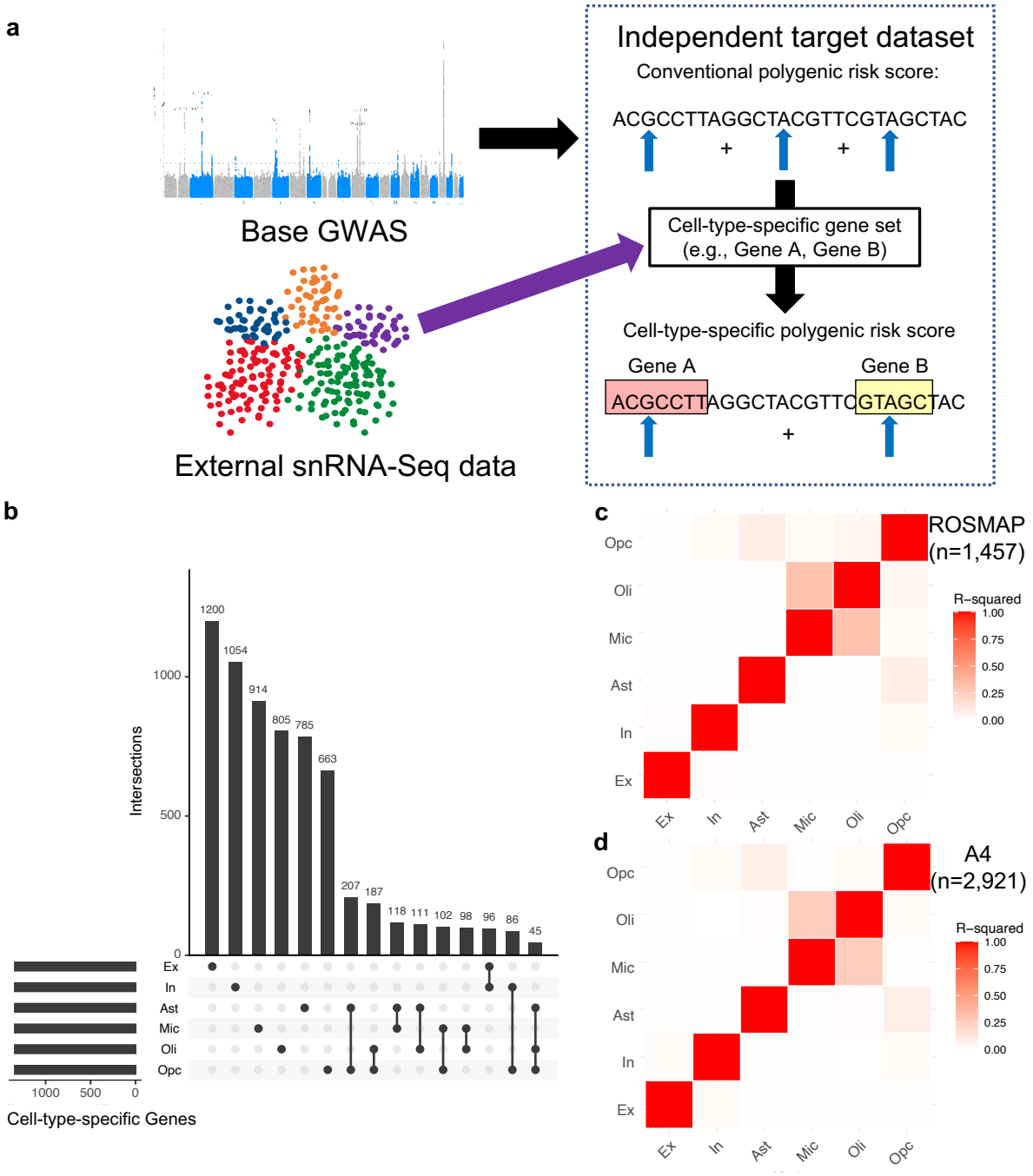
Tables.

Table 1. Study Participant Characteristics.

	ROSMAP (n=1457)	A4 (n=2921)
Mean Age, years (SD)	89.7 (6.5)	71.4 (4.8)
Female (%)	973 (67)	1740 (60)
Mean Education, years (SD)	16.3 (3.6)	16.7 (2.7)[a]
APOE ε4 carrier (%)	376 (26)	1038 (36)
Mean Florbetapir, cortical SUVR (SD)	NA	1.10 (0.19)
Elevated Aβ (%)	1008 (69)	890 (30)[b]
Pathological diagnosis of AD	954 (65)	NA
Median MMSE (IQR)	24 (13)	29 (2)
All-cause dementia (%)	661 (45)	0 (0)
AD dementia (%)	539 (37)	0 (0)

Abbreviations: AD dementia, AD with dementia; APOE, apolipoprotein E; IQR, interquartile range; MMSE, Mini-Mental State Examination; SD, standard deviation; SUVR, standardized uptake value ratio (whole cerebellar reference). [a]n=2919 with data. [b]n=2920 with data.

1 **Figures**



2

3 **Fig. 1. Cell-type-specific Alzheimer's disease polygenic risk scores (ADPRS). a,** A schematic of cell-

4 type-specific PRS derivation. **b,** An UpSetR plot of cell-type-specific gene sets used to define cell-type-

5 specific ADPRS. Each cell-type-specific gene set includes genes within the top 10% of cell-type

6 specificity (n=1,343). Each row of the matrix represents each cell-type-specific gene set, and each column

of the matrix represents an intersection of one or more sets. Gene sets in each intersection were indicated by filled black circles connected by a black vertical line. The vertical bar graph on the top shows the number of genes in each intersection. The 15 most frequent intersections were visualized. **c-d,** Correlation matrix among cell-type-specific ADPRS (**c,** ROSMAP; **d,** A4). Pearson's correlation coefficient was positive for all pairs, and the square of Pearson's correlation coefficient (R^2) between pairs of cell types was visualized.

Abbreviations: Ast, astrocytes; Ex, excitatory neurons; In, inhibitory neurons; Mic, microglia; Oli, oligodendrocytes; Opc, oligodendroglial progenitor cells; snRNA-Seq, single nucleus RNA-sequencing.

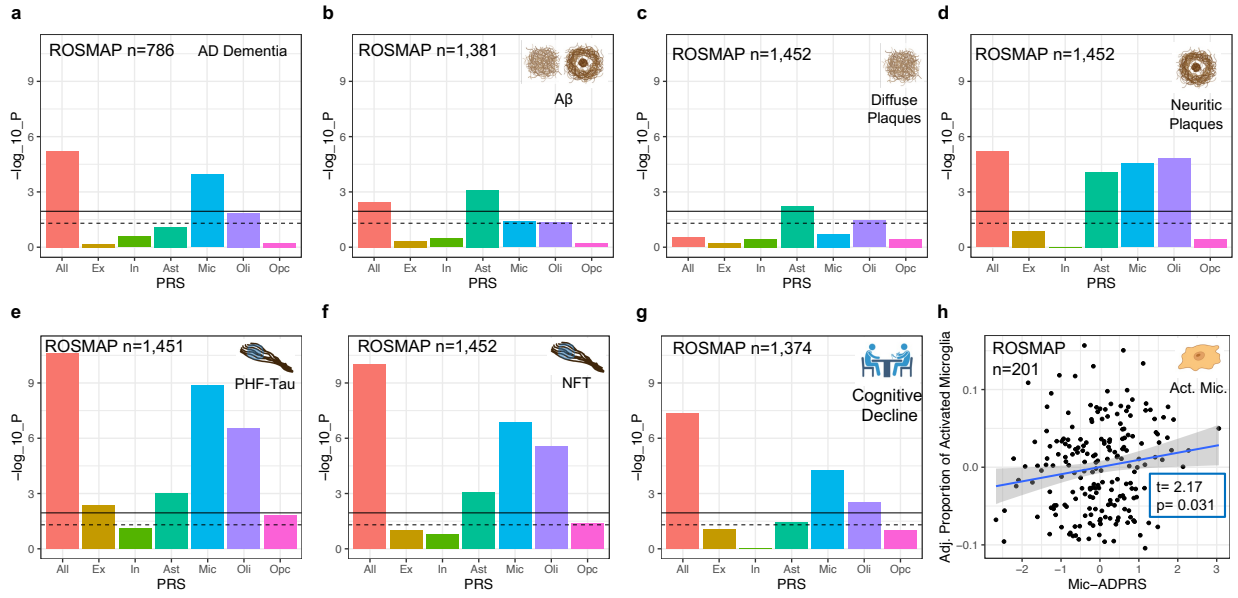


Fig. 2. Association of cell-type-specific AD polygenic risk scores in ROSMAP. a, Association of cell-type-specific ADPRSs with the odds of AD with dementia (case: n=538, control: n=248). **b**, Association of cell-type-specific ADPRSs with overall Aβ burden. **c**, Association of cell-type-specific ADPRSs with overall diffuse plaque (DP) burden. **b**, Association of cell-type-specific ADPRSs with overall neuritic plaque (NP) burden. **e**, Association of cell-type-specific ADPRSs with overall paired-helical-filament tau (PHFtau) burden. **e**, Association of cell-type-specific ADPRSs with overall neurofibrillary tangle (NFT) burden. **f**, Association of cell-type-specific ADPRSs with cognitive decline (the slope of annual change in antemortem measures of global cognitive composite). y-axis indicates -$\log_{10}$(p-value) of each association. For **a-g**, the black solid horizontal line indicates the p-value corresponding to statistical significance (FDR=0.025), and the black dashed horizontal line indicates p=0.05. **h,** Association of Mic-ADPRS (x-axis) and the proportion of activated microglia (PAM, y-axis). On y-axis, residual PAM values adjusting for covariates (*APOE* ε4, *APOE* ε2, age at death, sex, genotyping platform, and the first three genotyping principal components) were shown. Abbreviations: Act. Mic., activated microglia; All, full autosomal genome; Ast, astrocytes; Ex, excitatory neurons; In, inhibitory neurons; Mic, microglia; Oli, oligodendrocytes; Opc, oligodendroglial progenitor cells.

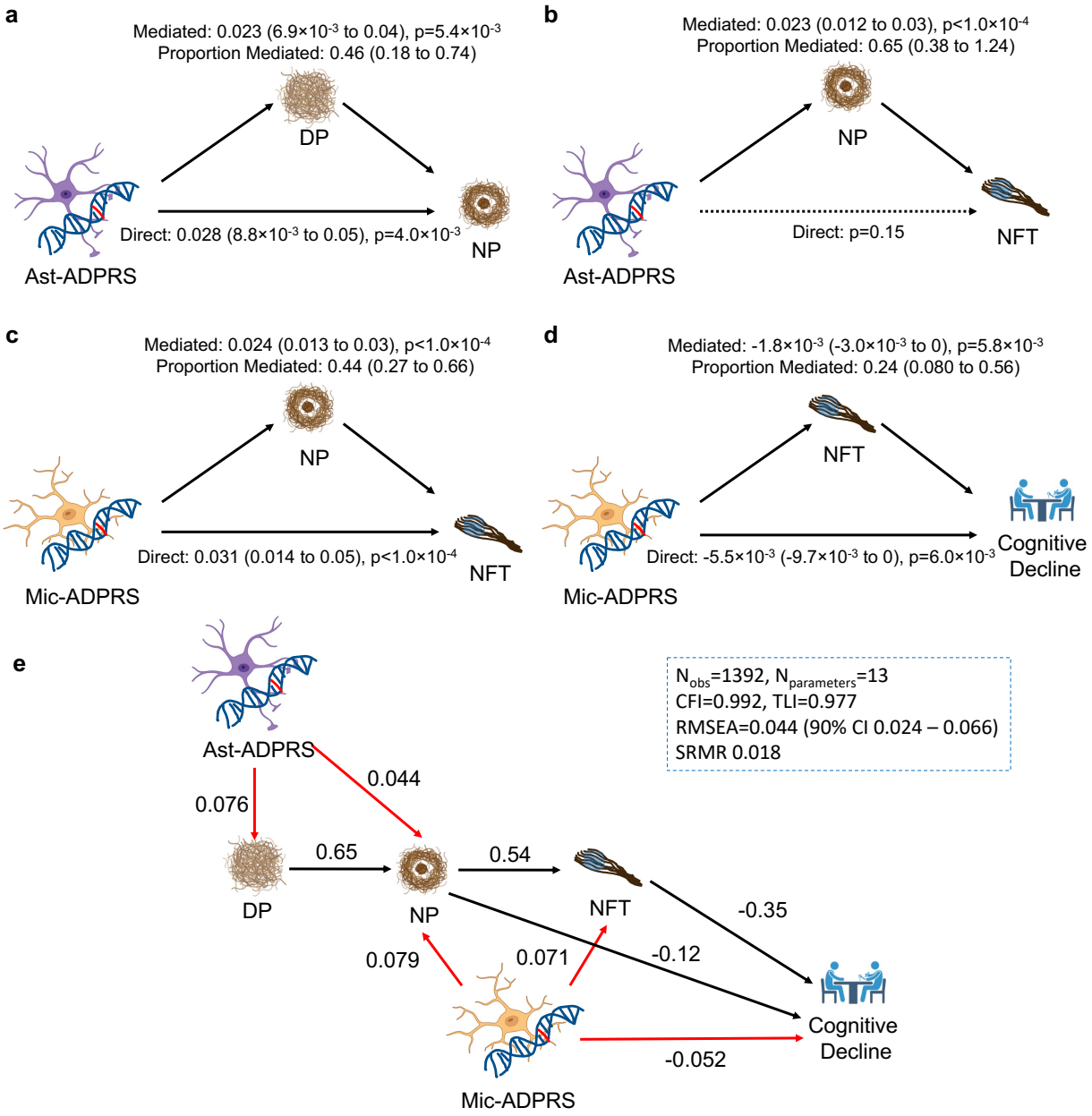


Fig. 3. Causal mediation analyses ans structural equation modeling of cell-type-specific ADPRS and AD endophenotypes. a, DP partially mediates Ast-ADPRS – NP association (n=1,452). **b,** NP mediates most of the Ast-ADPRS – NFT association (n=1,474), and the direct effect of Ast-ADPRS on NFT is not significant. **c,** NP partially mediates Mic-ADPRS – tau association (n=1,474). **d,** NFT partially mediates Mic-ADPRS – cognitive decline association (n=1,392). The model included NP burden as a covariate. **e,** Structural equation modeling (SEM) shows a probable causal relationship between cell-type-specific

ADPRS and AD endophenotypes. Black solid arrows indicate phenotype-phenotype associations, and red solid arrows indicate genotype-phenotype associations. All depicted associations were nominally significant ($p<0.05$). Numbers adjacent to each arrow indicate completely standardized solutions (relative strength of the effect). Model fit metrics indicate an excellent model fit. All models in **a-e** are adjusted for age, sex, education (for cognitive decline slope), *APOE* ε2 count, *APOE* ε4 count, genotype batch, and first three genotype principal components. Also see **Supplementary Table 14**. Abbreviations: CFI, comparative fit index; DP, diffuse plaque; N_{obs}, number of observations (participants); $N_{parameter}$, number of model parameters; NP, neuritic plaque; n.s., not significant; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; TLI, Tucker Lewis Index

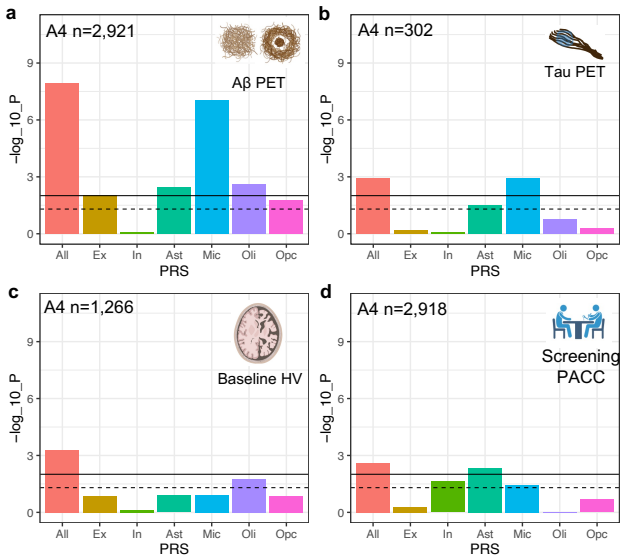


Fig. 4. Association of cell-type-specific AD polygenic risk scores in A4. a, Association of cell-type-specific ADPRSs with cortical Aβ (florbetapir PET cortical composite SUVR). **b**, Association of cell-type-specific ADPRSs with temporal lobe tau (flortaucipir PET temporal lobe composite SUVR). **c**, Association of cell-type-specific ADPRSs with hippocampal volume (HV). **d,** Association of cell-type-specific ADPRSs with screening Preclinical Alzheimer Cognitive Composite (PACC). y-axis indicates -$\log_{10}$(p-value) of each association. The black solid horizontal line indicates the p-value corresponding to statistical significance (FDR=0.025), and the black dashed horizontal line indicates p=0.05. Abbreviations: SUVR, standardized uptake value ratio.

Supplementary Information:

Supplementary Tables 1-25.

Data Availability:

ROSMAP resources can be requested at https://www.radc.rush.edu. The A4/LEARN screening (pre-randomization) data can be requested at https://ida.loni.usc.edu/.