

Title: Clonal hematopoiesis is associated with protection from Alzheimer's disease

Authors: Hind Bouzid¹, Julia A. Belk², Max Jan³, Yanyan Qi¹, Chloé Sarnowski⁴, Sara Wirth¹, Lisa Ma¹, Matthew Chrostek¹, Herra Ahmad^{1,5}, Daniel Nachun¹, Winnie Yao¹, NHLBI Trans-omics for Precision Medicine (TOPMed) Consortium, Alexa Beiser⁶, Alexander G. Bick⁷, Joshua C. Bis⁸, Myriam Fornage⁹, William T. Longstreth, Jr.¹⁰, Oscar L. Lopez¹¹, Pradeep Natarajan^{12,13}, Bruce M. Psaty⁸, Claudia L. Satizabal¹⁴, Joshua Weinstock¹⁵, Eric B. Larson¹⁶, Paul K. Crane¹⁷, C. Dirk Keene¹⁸, Sudha Seshadri¹⁴, Ansuman T. Satpathy¹, Thomas J. Montine¹, Siddhartha Jaiswal^{1,19*}

Affiliations:

¹Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA

²Department of Computer Science, Stanford University, Stanford, CA 94305, USA

³Department of Pathology, Massachusetts General Hospital, Boston, MA 02114, USA

⁴Department of Epidemiology, Human Genetics and Environmental Sciences, The University of Texas Health Science Center at Houston, School of Public Health, Houston, TX 77030, USA

⁵Department of Cardiology, Charité Universitätsmedizin, Berlin, Germany

⁶Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA

⁷Department of Medicine, Vanderbilt University Medical Center, Nashville, TN 37232, USA

⁸ Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology, and Health Services, University of Washington, Seattle, WA 98101, USA

⁹Brown Foundation Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX 77030, USA

¹⁰Departments of Neurology and Epidemiology, University of Washington, Seattle, WA 98195, USA.

¹¹Department of Neurology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, USA

¹²Medical and Population Genetics and Cardiovascular Disease Initiative, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA

¹³Cardiovascular Research Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA

¹⁴Glenn Biggs Institute for Alzheimer's & Neurodegenerative Diseases, University of Texas Health Science Center, San Antonio, TX78229, USA.

¹⁵Center for Statistical Genetics, Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, MI 48109-2029, USA

¹⁶Kaiser Permanente Washington Health Research Institute, Seattle, WA 98101, USA

¹⁷Department of Medicine, University of Washington, Seattle, WA 98101, USA

¹⁸Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA 98195-7110, USA

¹⁹Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

*Corresponding author. Email: sjaiswal@stanford.edu

Abstract: Clonal hematopoiesis of indeterminate potential (CHIP) is a pre-malignant expansion of mutated blood stem cells that also associates with non-hematological disorders. Here, we tested whether CHIP was associated with Alzheimer’s disease (AD). Surprisingly, we found that CHIP carriers had reduced risk of AD dementia or AD neuropathologic features in multiple cohorts. The same mutations found in blood were also detected in the microglia-enriched fraction of brain in 7 out of 8 CHIP carriers. Single-cell chromatin accessibility profiling of brain-derived nuclei in two CHIP carriers revealed that the mutated cells were indistinguishable from microglia and comprised between 42-77% of the total microglial pool. These results suggest a role for mutant, marrow-derived cells in attenuating risk of AD, possibly by supplementing a failing microglial system during aging.

Main Text:

Hematopoietic stem cells (HSCs) randomly accumulate somatic mutations during aging (1). While most of these mutations have no consequence, rare fitness-increasing mutations may allow an HSC to clonally expand. This age-associated expansion is termed clonal hematopoiesis of indeterminate potential (CHIP), is found in 10-30% of those older than 70, most commonly occurs due to loss-of-function mutations in transcriptional regulators such as *DNMT3A*, *TET2*, and *ASXL1*, and can be detected by sequencing of DNA from peripheral blood or bone marrow cells (2). As these mutations are also founding mutations for hematological neoplasms such as acute myeloid leukemia, it is unsurprising that CHIP associates with increased risk of developing these cancers (3–5). However, CHIP also associates with increased risk of atherosclerotic cardiovascular disease and death (6–8). This link is believed to be causal, as mice that are deficient for *Tet2* or *Dnmt3a* in hematopoietic cells develop larger atherosclerotic plaques, presumably due to altered gene expression in mutant macrophages which favors more rapid progression of the lesions (6, 9).

Alzheimer’s disease (AD) remains a leading cause of morbidity and mortality in the aged, but therapies that can effectively slow or halt its progression are lacking. Genome-wide association studies have implicated functional alterations of microglia, the macrophage-like hematopoietic cells in the brain, as a major driver of AD risk (10).

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/) .

Given that the common CHIP-associated mutations are known to alter the development and function of myeloid cells (6, 9, 11), we hypothesized that CHIP may also be associated with risk of AD.

To test this hypothesis, we utilized data from the Framingham Heart Study (FHS) and the Cardiovascular Health Study (CHS), which are two cohorts within the Trans-omics for Precision Medicine (TOPMed) project (12). CHIP variants (Table S1) were identified from blood-derived whole genome sequencing data as previously described (7). Participants in CHS were substantially older on average and a higher proportion were female compared to participants in FHS (Table S2). The FHS subset in TOPMed included some related participants selected for family studies but was otherwise a random selection of the total cohort. The CHS subset was heavily oversampled for coronary heart disease and stroke, conditions indicative of systemic atherosclerosis (see Materials and Methods). Vascular dementia, which can mimic AD clinically, is thought to result from reduced blood flow in the brain in part due to atherosclerosis. Given the established association of CHIP with atherosclerotic cardiovascular disease, we excluded participants in both cohorts who had coronary heart disease or stroke. A diagnosis for AD dementia was made based on criteria from the National Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) for definite, probable, or possible Alzheimer's disease. We excluded anyone with a baseline diagnosis of AD dementia or missing data on incident AD dementia. After these exclusions, there were 2,437 participants in FHS, of whom 92 (3.8%) developed incident AD dementia, and 743 participants in CHS, of whom 166 (22.3%) developed incident AD dementia (Table S2). Since CHIP is associated with increased risk of mortality, confounding due to survivorship bias may occur when performing association studies for CHIP. To mitigate this possibility, we tested for an association between CHIP and AD dementia using competing risks regression, with death as the competing risk, while also adjusting for age, sex, *APOE* genotype, study site (for CHS), and self-reported ancestry (for CHS). Contrary to our expectations, the presence of CHIP was associated with a lower subdistribution hazard ratio (SHR) for incident AD dementia in both cohorts (SHR 0.69, $p=0.13$ in CHS; SHR 0.51, $p=0.060$ in FHS; SHR 0.62, $p=0.021$ in a fixed-effects meta-analysis of the two cohorts) (Fig.

It is made available under a [CC-BY-NC-ND 4.0 International license](#) .

1A, Fig. S1). We obtained similar results using Cox proportional hazards regression models or when including family as a clustered variable for FHS (**Table S4**).

We next sought confirmation of this surprising inverse association in an independent cohort, the Alzheimer’s Disease Sequencing Project (**ADSP**), a case-control study for AD with whole exome sequencing (**WES**) data from brain or blood-derived DNA (*13*). *APOE* genotype is the strongest genetic risk factor for AD (*14*), with *APOE* $\epsilon 2$ alleles conferring protection from disease and *APOE* $\epsilon 4$ alleles conferring increased risk, as compared to those with *APOE* $\epsilon 3\epsilon 3$ (**Fig. S1**). The sample selection strategy for ADSP resulted in cases and controls that were not well-matched for age at WES blood draw for those carrying *APOE* $\epsilon 2$ or $\epsilon 4$ alleles (see **Materials and Methods**). Given CHIP’s strong association to age, this selection bias presented a major source of confounding that precluded the analysis of carriers of these alleles. However, *APOE* $\epsilon 3\epsilon 3$ AD dementia cases and controls were well matched for age (**Table S5**), permitting us to test for an association to CHIP in this set. After excluding those with missing information on age at blood draw, a total of 1,446 controls and 1,104 AD dementia cases with blood-derived whole exome sequencing were available for this analysis (**Table S5**). In this set, there were no overlapping participants between ADSP and TOPMed. CHIP variants were identified in ADSP using an approach previously described (*6*) and the prevalence of CHIP was appropriate for the age of the cohort (**Table S5**). The sequencing depth in ADSP was higher than for TOPMed, which resulted in greater sensitivity to detect smaller clones (**Fig. S2**). Clone size, which is approximated by the variant allele fraction (**VAF**), has previously been shown to be an important predictor of risk for blood cancer (*4, 15*) and cardiovascular outcomes (*6, 7*). In order to directly compare outcomes in ADSP to TOPMed, we limited the definition of CHIP carriers to those with $\text{VAF} > 0.08$ in this analysis—a cutoff that was chosen because it resulted in a VAF distribution that was nearly identical to TOPMed (**Fig. S2**). We found that CHIP with $\text{VAF} > 0.08$ was associated with reduced risk of AD dementia in ADSP (odds ratio [**OR**] 0.66, $p = 5.5 \times 10^{-4}$) (**Fig. 1B, Table S6**). In contrast, having $\text{VAF} \leq 0.08$ had no association to AD dementia (OR 1.25, $p = 0.23$) (**Table S7**), suggestive of a dose-response relationship between the size of the mutant clone and protection from AD dementia. Indeed, higher VAF was also significantly

It is made available under a [CC-BY-NC-ND 4.0 International license](#) .

associated with protection from AD dementia when modeled as a continuous variable (**Table S7**). A meta-analysis of ADSP, CHS, and FHS showed that CHIP-carriers had a significant reduction in risk of AD (OR 0.64, $p=3.0 \times 10^{-5}$) (**Fig. 1C**). In sum, our human genetic association analyses demonstrate that CHIP is associated with protection from AD dementia in multiple cohorts, and that the degree of protection is proportional to the size of the mutant clone.

Mendelian randomization is a form of causal inference in which genetic variants known to influence the risk of a particular exposure or trait (in this case CHIP) are assessed for an association to a clinical outcome (in this case AD). We selected three independent common variants that reached genome-wide significance in a CHIP genetic association study as the instrumental variables for CHIP exposure (rs2853677, rs7726159, rs58322641)(7). Summary statistics from a large AD genome-wide association study (GWAS) and GWAS-by-proxy meta-analysis(16) were then used to perform Mendelian randomization using the inverse-variance weighted method(17). Here, an increase in the genetic risk of CHIP was associated with reduced odds of AD (OR 0.92 per 1 log-odds increase in risk of CHIP, $p=5.6 \times 10^{-3}$, **Fig. S3**).

The hallmark neuropathological features of AD, regional accumulation of beta-amyloid plaques and tau neurofibrillary tangles, can also be found in some people without a clinical diagnosis of dementia. A neuritic plaque density score developed by the Consortium to Establish a Registry for AD (CERAD) (18) and Braak stage for neurofibrillary tangle distribution (19) are commonly used to assess for these changes at brain autopsy, with an increasing score indicative of more extensive accumulation of pathologic features. A subset of participants in ADSP had brain autopsy performed after death, which allowed us to test whether CHIP was associated with AD-related neuropathologic change (ADNC) in those without dementia (**Table S8**). Here, the presence of CHIP was associated with having a lower CERAD neuritic plaque score (OR 0.50, $p=3.2 \times 10^{-3}$) and Braak stage (OR 0.56, $p=0.015$) using ordinal logistic regression after adjusting for age, sex and *APOE* genotype (**Fig. 1D-E, Table S8**). This suggests that CHIP associates with protection from the burden of neuritic plaque and neurofibrillary tangle formation even in the absence of clinical dementia.

We next asked whether the protection from AD seen in CHIP carriers was influenced by *APOE* genotype. In those age 70 or older, *APOE* genotype was strongly associated with AD dementia risk in those without CHIP ($p=5.8 \times 10^{-5}$ by log-rank test). This effect was not seen in CHIP carriers of the same age ($p=0.70$ by log-rank test), though the smaller sample size in this group may have limited our power to find an association (**Fig. 2A**). In competing risks regression models stratified by *APOE* genotype, there was a similar magnitude of AD dementia risk reduction in CHIP carriers who were *APOE* $\epsilon_3\epsilon_3$ or who carried an *APOE* ϵ_4 allele, but not in those who had the protective *APOE* $\epsilon_2\epsilon_2$ or *APOE* $\epsilon_2\epsilon_3$ genotypes (**Fig. 2B, Table S9**). These findings suggest that the mechanism by which CHIP associates with reduced risk of AD dementia might be redundant with the protection conferred through *APOE* ϵ_2 .

We also assessed whether the risk of AD dementia varied based on the specific mutated gene. Of the most commonly mutated genes in CHIP, all were associated with protection from AD dementia to a similar degree (**Fig. 2C, Table S10**).

We wondered whether cells bearing CHIP-associated mutations could be found in brain, a finding that would strengthen the likelihood of a causal association between CHIP and AD risk. We obtained brain DNA-derived WES data from 1,776 persons in ADSP, of whom 1,462 had AD dementia (82.3%), and assessed for the presence of CHIP-associated variants in these samples. Similar to a prior study (20), we found mutations consistent with CHIP in 17 brain samples, including 15 with AD dementia (**Fig. 3A, Table S11**). Paired blood DNA was not available for the brain exome samples, so we could not determine whether the mutations we identified were indicative of blood-derived cells in brain.

Hematopoietic cells, nearly all of which are microglia, comprise ~1-10% of the total cells in the brain across brain regions (21) and the limit of detection for clonal hematopoiesis by WES at 80X sequencing depth is ~4% of cells

It is made available under a [CC-BY-NC-ND 4.0 International license](#) .

harboring a mutation in a sample (4). Therefore, it would be difficult to detect clonal hematopoiesis mutations from unfractionated brain for the vast majority of CHIP carriers using WES. We hypothesized that most CHIP carriers would have the mutations detectable in the brain if examined using more sensitive methods. To test this hypothesis, we obtained tissue samples from the occipital lobe, and in some cases putamen or cerebellum, of 8 donors from the Adult Changes in Thought (ACT) cohort who were known to have CHIP from blood exome sequencing, as well as 1 person without CHIP (Table S12). All persons were in their 80s and 7 out of 9 were without dementia and had no/low ADNC at the time of death. The 8 CHIP carriers had mutations in *DNMT3A*, *TET2*, *ASXL1*, *SF3B1*, and *GNB1* with the highest frequency in *DNMT3A* (4 out of 8) and *TET2* (3 out of 8) (Table S12), which is representative of the relative proportion of these mutations in the general population (7). In addition, 2 out of the 8 harbored two different CHIP mutations. To determine whether bone marrow derived cells carrying CHIP mutations were present in the brains of these individuals, we digested the frozen brain tissue and isolated intact nuclei, from which we extracted DNA for amplicon sequencing. We detected the same mutations that were present in blood in 6 out of 8 unfractionated brains with VAF ranging from 0.004 to 0.02 (Fig. S4, Table S12). Of note, the two donors where CHIP variants were not identified in unfractionated brain (ACT1, ACT8) both had AD dementia during life. In contrast, all 6 donors where CHIP variants were found in unfractionated brain were without dementia and free of ADNC.

The CHIP variants detected in whole brain DNA might have originated from residual circulating hematopoietic cells in the vasculature, such as granulocytes or lymphocytes. Alternatively, myeloid cells such as microglia in the brain parenchyma could have been the source of the mutations, although microglia are believed to have little contribution from HSC-derived cells in adulthood (22). To distinguish between these possibilities, we conceived a strategy to enrich for these cells from frozen brain tissue. Since the tissue was not viably cryopreserved, isolation of cells based on expression of membrane antigens was not possible. Instead, we used antibodies to nuclear transcription factors to enrich for mononuclear phagocytes, such as macrophages and microglia (Fig. 3B). We stained nuclei for the neuronal-specific transcription factor NeuN (RBFOX3) and c-Maf, a transcription factor expressed in phagocytes as well as some neurons and non-hematopoietic glial cells. We then sorted 4 populations

It is made available under a [CC-BY-NC-ND 4.0 International license](#).

based on the presence or absence of these markers (**Fig. 3B**). The CHIP somatic variants were found in the NeuN⁻ c-Maf⁺ population in 7 out of 8 brains, with a VAF that ranged from 0.02 to 0.28 (representing 4% to 56% of NeuN⁻ c-Maf⁺ nuclei). In contrast, CHIP somatic variants were not detected in the NeuN⁺ c-Maf⁻ neuronal population and were absent or at low levels in the other two populations (**Fig. 3C, Fig. S4**). The VAF was lower in the NeuN⁻ c-Maf⁺ nuclei from cerebellum compared to occipital cortex in the three samples where tissue from both regions was available. In sample ACT8, CHIP variants were robustly detected in the NeuN⁻ c-Maf⁺ population from occipital cortex but not putamen. These results indicate that there is a substantial contribution to the brain mononuclear phagocytic pool from circulating mutated cells, and that there is also regional heterogeneity in the frequency of infiltrating CHIP⁺ cells in brain.

Our flow cytometric analysis indicated that a prominent myeloid population bearing CHIP mutations was present in the brains of most CHIP carriers. However, it was unknown if the mutated cells were similar to endogenous brain microglia or were instead a distinct myeloid population not normally found in brain, such as monocyte-derived macrophages or dendritic cells. To better understand the phenotype of these mutated cells, we performed single-cell assay for transposase-accessible chromatin sequencing (**scATAC-seq**) on brain samples from three ACT participants: a brain donor without CHIP (ACT9), one with *TET2*-mutant CHIP (ACT6), and one with *DNMT3A*-mutant CHIP (ACT2) (**Table S12**). For the ACT6 donor, we analyzed tissue from cerebellum and putamen, whereas occipital cortex was assessed for the other two donors. scATAC-seq was performed on unsorted nuclei, as well as sorted NeuN⁻ c-Maf⁺ nuclei for each sample. After aligning and filtering the scATAC-seq reads, our samples had a median of 12,287 fragments per cell and a median enrichment of fragments in transcription start sites of 9.31, indicating that we recovered high quality scATAC-seq libraries from these archived samples (**Fig. S5**). In total, we recovered high quality scATAC-seq profiles for 38,206 cells. We then aggregated our data with scATAC-seq data from 10 samples (an additional 72,984 cells) from a comprehensive scATAC-seq characterization of the adult human brain (Corces 2020) (21). After clustering and dimensionality reduction, we identified 18 clusters encompassing the major brain cell types (**Fig. 4A-B**), including one cluster that contained previously described microglia as well as myeloid cells from each of our samples (cluster 9) (**Fig. 4B, Fig. S6**,

It is made available under a [CC-BY-NC-ND 4.0 International license](#) .

Table S13). No other hematopoietic cell types were observed in the human brain samples. For cells within cluster 9 grouped by sample, inspecting pseudo-bulk ATAC-seq tracks revealed accessible chromatin at the microglia marker genes *TMEM119*, *P2RY12*, *SALL1*, *CSF1R*, and *SIGLEC8* in each of our samples, which was visually similar to the reference microglia. Cells in this cluster also had high accessibility at the AD-associated genes *APOE*, *TREM2*, *TYROBP*, *AXL*, and *MERTK* (**Fig. 4C**, **Fig. S6**). As an additional control, we examined reference scATAC-seq profiles of blood monocytes and classical dendritic cells (cDCs) (23) and found that these cells had little to no accessibility at the aforementioned microglia genes. Furthermore, the Cluster 9 tracks in all brain samples had low accessibility at *ANPEP* (CD13), *CXCL2*, and *ITGAE* (CD103), in contrast to monocytes or cDCs (**Fig. 4C**, **Fig. S6**). Finally, we quantified the genome-wide similarity of microglia in each pair of samples by considering the number of differential peaks between each pair for cells within cluster 9. The putamen-derived sample showed modest differences from the occipital cortex-derived samples but otherwise essentially no differences were observed, and all comparisons were within the range of variation observed when comparing pairs of the Corces 2020 samples (**Fig. S6**). These results indicate that the cells in cluster 9 are indistinguishable from microglia and unlikely to contain contaminating monocytes or dendritic cells.

Having established that the only hematopoietic cell type present in these brains was microglia, we used the scATAC-seq data to evaluate the effectiveness of our flow cytometric method for enrichment of these cells. We compared the proportion of microglia in the sorted samples to those present in the unsorted samples and observed that microglia were enriched 11 to 33-fold in the sorted samples from occipital cortex and cerebellum, but only modestly in the putamen sample (**Fig. 4D**). The percentage of microglia ranged from 7.2% to 25% in the sorted samples, indicating that there were still large numbers of contaminating non-microglial cells in the NeuN- c-Maf⁺ gate. This suggested that the true fraction of mutated microglia was substantially underestimated by our previous approach using amplicon sequencing of NeuN- c-Maf⁺ nuclei. To estimate the percentage of mutated microglia more accurately, we first assessed the VAF for the CHIP variant in each unsorted brain sample. Since these are heterozygous mutations, multiplying the VAF by 2 gives an estimate of the percentage of mutated cells in the sample. Since microglia were the only hematopoietic cell type present in these brains, we reasoned that we could

It is made available under a [CC-BY-NC-ND 4.0 International license](#) .

divide the percentage of total mutated cells by the percentage of microglia in each unsorted sample to estimate the percentage of mutant microglia. Using this approach, we calculated that 43% of the microglia in ACT2 harbored the *DNMT3A* mutation, as compared to 28% of circulating cells in the blood. For the ACT6 donor, 77% and 42% of putamen and cerebellar microglia harbored the *TET2* mutation, respectively, compared to 28% of circulating blood cells (**Fig. 4E, Table S14**). These results indicate that replacement of endogenous microglia by mutant, marrow-derived cells is widespread in the aging brain. The observation that the proportion of mutated cells is substantially greater in the microglial pool than in the blood also suggests that there is positive selection for the mutant cells in the brain microenvironment.

We show here that, unexpectedly, the presence of CHIP is associated with protection from AD dementia. This effect is seen in multiple cohorts, is not due to survival bias, is seen with several different mutated genes, and is strongest in carriers of *APOE* $\epsilon 3\epsilon 3$ or *APOE* $\epsilon 4$ alleles. The degree of protection from AD dementia seen in CHIP carriers is greater than carrying an *APOE* $\epsilon 2$ allele, which is the most protective common inherited variant for AD (24). CHIP is also associated with lower levels of neuritic plaques and neurofibrillary tangles in those without dementia, indicating a possible modulating effect of CHIP on the underlying pathophysiology of AD. Consistent with this hypothesis, we also detect substantial infiltration of brain by marrow-derived mutant cells which adopt a microglial-like phenotype. We speculate that the mutations associated with CHIP confer circulating precursor cells with an enhanced ability to engraft in the brain, to differentiate into microglia once engrafted, and/or to clonally expand relative to unmutated cells in the brain microenvironment. These non-mutually exclusive possibilities could provide protection from AD by supplementing the phagocytic capacity of the endogenous microglial system during aging. Alternatively, or in addition, the mutations may alter the functionality of the engrafted myeloid cells in a manner that promotes clearance of pathologic beta-amyloid and/or tau. Understanding the interplay between CHIP and the aging brain may yield valuable information about the pathogenesis of AD and provide insights into slowing its progression.

References and Notes

1. H. Lee-Six, N. F. Obro, M. S. Shepherd, S. Grossmann, K. Dawson, M. Belmonte, R. J. Osborne, B. J. P. Huntly, I. Martincorena, E. Anderson, L. O'Neill, M. R. Stratton, E. Laurenti, A. R. Green, D. G. Kent, P. J. Campbell, Population dynamics of normal human blood inferred from somatic mutations. *Nature* (2018), doi:10.1038/s41586-018-0497-0.
2. S. Jaiswal, B. L. Ebert, Clonal hematopoiesis in human aging and disease. *Science*. **366** (2019), doi:10.1126/science.aan4673.
3. G. Genovese, A. K. Kahler, R. E. Handsaker, J. Lindberg, S. A. Rose, S. F. Bakhoun, K. Chambert, E. Mick, B. M. Neale, M. Fromer, S. M. Purcell, O. Svantesson, M. Landen, M. Hoglund, S. Lehmann, S. B. Gabriel, J. L. Moran, E. S. Lander, P. F. Sullivan, P. Sklar, H. Gronberg, C. M. Hultman, S. A. McCarroll, Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. **371**, 2477–87 (2014).
4. S. Jaiswal, P. Fontanillas, J. Flannick, A. Manning, P. V. Grauman, B. G. Mar, R. C. Lindsley, C. H. Mermel, N. Burt, A. Chavez, J. M. Higgins, V. Moltchanov, F. C. Kuo, M. J. Kluk, B. Henderson, L. Kinnunen, H. A. Koistinen, C. Ladenvall, G. Getz, A. Correa, B. F. Banahan, S. Gabriel, S. Kathiresan, H. M. Stringham, M. I. McCarthy, M. Boehnke, J. Tuomilehto, C. Haiman, L. Groop, G. Atzmon, J. G. Wilson, D. Neuberg, D. Altshuler, B. L. Ebert, Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. **371**, 2488–98 (2014).
5. P. Desai, N. Mencia-Trinchant, O. Savenkov, M. S. Simon, G. Cheang, S. Lee, M. Samuel, E. K. Ritchie, M. L. Guzman, K. V. Ballman, G. J. Roboz, D. C. Hassane, Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nat Med*. **24**, 1015–1023 (2018).
6. S. Jaiswal, P. Natarajan, A. J. Silver, C. J. Gibson, A. G. Bick, E. Shvartz, M. McConkey, N. Gupta, S. Gabriel, D. Ardissino, U. Baber, R. Mehran, V. Fuster, J. Danesh, P. Frossard, D. Saleheen, O. Melander, G. K. Sukhova, D. Neuberg, P. Libby, S. Kathiresan, B. L. Ebert, Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N Engl J Med*. **377**, 111–121 (2017).
7. A. G. Bick, J. S. Weinstock, S. K. Nandakumar, C. P. Fulco, E. L. Bao, S. M. Zekavat, M. D. Szeto, X. Liao, M. J. Leventhal, J. Nasser, K. Chang, C. Laurie, B. B. Burugula, C. J. Gibson, A. E. Lin, M. A. Taub, F. Aguet, K. Ardlie, B. D. Mitchell, K. C. Barnes, A. Moscati, M. Fornage, S. Redline, B. M. Psaty, E. K. Silverman, S. T. Weiss, N. D. Palmer, R. S. Vasani, E. G. Burchard, S. L. R. Kardia, J. He, R. C. Kaplan, N. L. Smith, D. K. Arnett, D. A. Schwartz, A. Correa, M. de Andrade, X. Guo, B. A. Konkle, B. Custer, J. M. Peralta, H. Gui, D. A. Meyers, S. T. McGarvey, I. Y. Chen, M. B. Shoemaker, P. A. Peyser, J. G. Broome, S. M. Gogarten, F. F. Wang, Q. Wong, M. E. Montasser, M. Daya, E. E. Kenny, K. E. North, L. J. Launer, B. E. Cade, J. C. Bis, M. H. Cho, J. Lasky-Su, D. W. Bowden, L. A. Cupples, A. C. Y. Mak, L. C. Becker, J. A. Smith, T. N. Kelly, S. Aslibekyan, S. R. Heckbert, H. K. Tiwari, I. V. Yang, J. A. Heit, S. A. Lubitz, J. M. Johnsen, J. E. Curran, S. E. Wenzel, D. E. Weeks, D. C. Rao, D. Darbar, J. Y. Moon, R. P. Tracy, E. J. Buth, N. Rafaels, R. J. F. Loos, P. Durda, Y. Liu, L. Hou, J. Lee, P. Kachroo, B. I. Freedman, D. Levy, L. F. Bielak, J. E. Hixson, J. S. Floyd, E. A. Whitsel, P. T. Ellinor, M. R. Irvin, T. E. Fingerlin, L. M. Raffield, S. M. Armasu, M. M. Wheeler, E. C. Sabino, J. Blangero, L. K. Williams, B. D. Levy, W. H. Sheu, D. M. Roden, E. Boerwinkle, J. E. Manson, R. A. Mathias, P. Desai, K. D. Taylor, A. D. Johnson, N. T.-O. for P. M. Consortium, P. L. Auer, C. Kooperberg, C. C. Laurie, T. W. Blackwell, A. V. Smith, H. Zhao, E. Lange, L. Lange, S. S. Rich, J. I. Rotter, J. G. Wilson, P. Scheet, J. O. Kitzman, E. S. Lander, J. M. Engreitz, B. L. Ebert, A. P. Reiner, S. Jaiswal, G. Abecasis, V. G. Sankaran, S. Kathiresan, P. Natarajan, Inherited causes of clonal haematopoiesis in 97,691 whole genomes. *Nature* (2020), doi:10.1038/s41586-020-2819-2.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/) .

8. D. Nachun, A. T. Lu, A. G. Bick, P. Natarajan, J. Weinstock, M. D. Szeto, S. Kathiresan, G. Abecasis, K. D. Taylor, X. Guo, R. Tracy, P. Durda, Y. Liu, C. Johnson, S. S. Rich, D. Van Den Berg, C. Laurie, T. Blackwell, G. J. Papanicolaou, A. Correa, L. M. Raffield, A. D. Johnson, J. Murabito, J. E. Manson, P. Desai, C. Kooperberg, T. L. Assimes, D. Levy, J. I. Rotter, A. P. Reiner, E. A. Whitsel, J. G. Wilson, S. Horvath, S. Jaiswal, NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, Clonal hematopoiesis associated with epigenetic aging and clinical outcomes. *Aging Cell*. **20**, e13366 (2021).
9. J. J. Fuster, S. MacLauchlan, M. A. Zuriaga, M. N. Polackal, A. C. Ostriker, R. Chakraborty, C. L. Wu, S. Sano, S. Muralidharan, C. Rius, J. Vuong, S. Jacob, V. Muralidhar, A. A. Robertson, M. A. Cooper, V. Andres, K. K. Hirschi, K. A. Martin, K. Walsh, Clonal hematopoiesis associated with Tet2 deficiency accelerates atherosclerosis development in mice. *Science* (2017), doi:10.1126/science.aag1381.
10. A. Nott, I. R. Holtman, N. G. Coufal, J. C. M. Schlachetzki, M. Yu, R. Hu, C. Z. Han, M. Pena, J. Xiao, Y. Wu, Z. Keulen, M. P. Pasillas, C. O'Connor, C. K. Nickl, S. T. Schafer, Z. Shen, R. A. Rissman, J. B. Brewer, D. Gosselin, D. D. Gonda, M. L. Levy, M. G. Rosenfeld, G. McVicker, F. H. Gage, B. Ren, C. K. Glass, Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science*. **366**, 1134–1139 (2019).
11. Q. Zhang, K. Zhao, Q. Shen, Y. Han, Y. Gu, X. Li, D. Zhao, Y. Liu, C. Wang, X. Zhang, X. Su, J. Liu, W. Ge, R. L. Levine, N. Li, X. Cao, Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. *Nature*. **525**, 389–93 (2015).
12. D. Taliun, D. N. Harris, M. D. Kessler, J. Carlson, Z. A. Szpiech, R. Torres, S. A. G. Taliun, A. Corvelo, S. M. Gogarten, H. M. Kang, A. N. Pitsillides, J. LeFaive, S.-B. Lee, X. Tian, B. L. Browning, S. Das, A.-K. Emde, W. E. Clarke, D. P. Loesch, A. C. Shetty, T. W. Blackwell, A. V. Smith, Q. Wong, X. Liu, M. P. Conomos, D. M. Bobo, F. Aguet, C. Albert, A. Alonso, K. G. Ardlie, D. E. Arking, S. Aslibekyan, P. L. Auer, J. Barnard, R. G. Barr, L. Barwick, L. C. Becker, R. L. Beer, E. J. Benjamin, L. F. Bielak, J. Blangero, M. Boehnke, D. W. Bowden, J. A. Brody, E. G. Burchard, B. E. Cade, J. F. Casella, B. Chalazan, D. I. Chasman, Y.-D. I. Chen, M. H. Cho, S. H. Choi, M. K. Chung, C. B. Clish, A. Correa, J. E. Curran, B. Custer, D. Darbar, M. Daya, M. de Andrade, D. L. DeMeo, S. K. Dutcher, P. T. Ellinor, L. S. Emery, C. Eng, D. Fatkin, T. Fingerlin, L. Forer, M. Fornage, N. Franceschini, C. Fuchsberger, S. M. Fullerton, S. Germer, M. T. Gladwin, D. J. Gottlieb, X. Guo, M. E. Hall, J. He, N. L. Heard-Costa, S. R. Heckbert, M. R. Irvin, J. M. Johnsen, A. D. Johnson, R. Kaplan, S. L. R. Kardia, T. Kelly, S. Kelly, E. E. Kenny, D. P. Kiel, R. Klemmer, B. A. Konkle, C. Kooperberg, A. Köttgen, L. A. Lange, J. Lasky-Su, D. Levy, X. Lin, K.-H. Lin, C. Liu, R. J. F. Loos, L. Garman, R. Gerszten, S. A. Lubitz, K. L. Lunetta, A. C. Y. Mak, A. Manichaikul, A. K. Manning, R. A. Mathias, D. D. McManus, S. T. McGarvey, J. B. Meigs, D. A. Meyers, J. L. Mikulla, M. A. Minear, B. D. Mitchell, S. Mohanty, M. E. Montasser, C. Montgomery, A. C. Morrison, J. M. Murabito, A. Natale, P. Natarajan, S. C. Nelson, K. E. North, J. R. O'Connell, N. D. Palmer, N. Pankratz, G. M. Peloso, P. A. Peyser, J. Pleiness, W. S. Post, B. M. Psaty, D. C. Rao, S. Redline, A. P. Reiner, D. Roden, J. I. Rotter, I. Ruczinski, C. Sarnowski, S. Schoenherr, D. A. Schwartz, J.-S. Seo, S. Seshadri, V. A. Sheehan, W. H. Sheu, M. B. Shoemaker, N. L. Smith, J. A. Smith, N. Sotoodehnia, A. M. Stilp, W. Tang, K. D. Taylor, M. Telen, T. A. Thornton, R. P. Tracy, D. J. Van Den Berg, R. S. Vasan, K. A. Viaud-Martinez, S. Vrieze, D. E. Weeks, B. S. Weir, S. T. Weiss, L.-C. Weng, C. J. Willer, Y. Zhang, X. Zhao, D. K. Arnett, A. E. Ashley-Koch, K. C. Barnes, E. Boerwinkle, S. Gabriel, R. Gibbs, K. M. Rice, S. S. Rich, E. K. Silverman, P. Qasba, W. Gan, NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, G. J. Papanicolaou, D. A. Nickerson, S. R. Browning, M. C. Zody, S. Zöllner, J. G. Wilson, L. A. Cupples, C. C. Laurie, C. E. Jaquish, R. D. Hernandez, T. D. O'Connor, G. R. Abecasis, Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature*. **590**, 290–299 (2021).

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/) .

13. G. W. Beecham, J. C. Bis, E. R. Martin, S. H. Choi, A. L. DeStefano, C. M. van Duijn, M. Fornage, S. B. Gabriel, D. C. Kibbitt, D. E. Larson, A. C. Naj, B. M. Psaty, W. Salerno, W. S. Bush, T. M. Foroud, E. Wijmsman, L. A. Farrer, A. Goate, J. L. Haines, M. A. Pericak-Vance, E. Boerwinkle, R. Mayeux, S. Seshadri, G. Schellenberg, The Alzheimer's Disease Sequencing Project: Study design and sample selection. *Neurol Genet.* **3**, e194 (2017).
14. Y. Yamazaki, N. Zhao, T. R. Caulfield, C. C. Liu, G. Bu, Apolipoprotein E and Alzheimer disease: pathobiology and targeting strategies. *Nat Rev Neurol.* **15**, 501–518 (2019).
15. S. Abelson, G. Collord, S. W. K. Ng, O. Weissbrod, N. Mendelson Cohen, E. Niemeyer, N. Barda, P. C. Zuzarte, L. Heisler, Y. Sundaravadanam, R. Luben, S. Hayat, T. T. Wang, Z. Zhao, I. Cirlan, T. J. Pugh, D. Soave, K. Ng, C. Latimer, C. Hardy, K. Raine, D. Jones, D. Hoult, A. Britten, J. D. McPherson, M. Johansson, F. Mbabaali, J. Eagles, J. K. Miller, D. Pasternack, L. Timms, P. Krzyzanowski, P. Awadalla, R. Costa, E. Segal, S. V. Bratman, P. Beer, S. Behjati, I. Martincorena, J. C. Y. Wang, K. M. Bowles, J. R. Quiros, A. Karakatsani, C. La Vecchia, A. Trichopoulou, E. Salamanca-Fernandez, J. M. Huerta, A. Barricarte, R. C. Travis, R. Tumino, G. Masala, H. Boeing, S. Panico, R. Kaaks, A. Kramer, S. Sieri, E. Riboli, P. Vineis, M. Foll, J. McKay, S. Polidoro, N. Sala, K. T. Khaw, R. Vermeulen, P. J. Campbell, E. Papaemmanuil, M. D. Minden, A. Tanay, R. D. Balicer, N. J. Wareham, M. Gerstung, J. E. Dick, P. Brennan, G. S. Vassiliou, L. I. Shlush, Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature.* **559**, 400–404 (2018).
16. J. Schwartzentruber, S. Cooper, J. Z. Liu, I. Barrio-Hernandez, E. Bello, N. Kumasaka, A. M. H. Young, R. J. M. Franklin, T. Johnson, K. Estrada, D. J. Gaffney, P. Beltrao, A. Bassett, Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer's disease risk genes. *Nat. Genet.* **53**, 392–402 (2021).
17. S. Burgess, A. Butterworth, S. G. Thompson, Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **37**, 658–665 (2013).
18. B. T. Hyman, C. H. Phelps, T. G. Beach, E. H. Bigio, N. J. Cairns, M. C. Carrillo, D. W. Dickson, C. Duyckaerts, M. P. Frosch, E. Masliah, S. S. Mirra, P. T. Nelson, J. A. Schneider, D. R. Thal, B. Thies, J. Q. Trojanowski, H. V. Vinters, T. J. Montine, National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement.* **8**, 1–13 (2012).
19. H. Braak, I. Alafuzoff, T. Arzberger, H. Kretschmar, K. Del Tredici, Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathol.* **112**, 389–404 (2006).
20. M. J. Keogh, W. Wei, J. Aryaman, L. Walker, J. van den Amele, J. Coxhead, I. Wilson, M. Bashton, J. Beck, J. West, R. Chen, C. Haudenschild, G. Bartha, S. Luo, C. M. Morris, N. S. Jones, J. Attems, P. F. Chinnery, High prevalence of focal and multi-focal somatic genetic variants in the human brain. *Nat Commun.* **9**, 4257 (2018).
21. M. R. Corces, A. Shcherbina, S. Kundu, M. J. Gloudemans, L. Frésard, J. M. Granja, B. H. Louie, T. Eulalio, S. Shams, S. T. Bagdatli, M. R. Mumbach, B. Liu, K. S. Montine, W. J. Greenleaf, A. Kundaje, S. B. Montgomery, H. Y. Chang, T. J. Montine, Single-cell epigenomic analyses implicate candidate causal variants at inherited risk loci for Alzheimer's and Parkinson's diseases. *Nat. Genet.* **52**, 1158–1168 (2020).
22. F. Ginhoux, S. Lim, G. Hoeffel, D. Low, T. Huber, Origin and differentiation of microglia. *Front Cell Neurosci.* **7**, 45 (2013).

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/) .

23. A. T. Satpathy, J. M. Granja, K. E. Yost, Y. Qi, F. Meschi, G. P. McDermott, B. N. Olsen, M. R. Mumbach, S. E. Pierce, M. R. Corces, P. Shah, J. C. Bell, D. Jhutti, C. M. Nemeč, J. Wang, L. Wang, Y. Yin, P. G. Giresi, A. L. S. Chang, G. X. Y. Zheng, W. J. Greenleaf, H. Y. Chang, Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat. Biotechnol.* **37**, 925–936 (2019).
24. A. Serrano-Pozo, S. Das, B. T. Hyman, APOE and Alzheimer’s disease: advances in genetics, pathophysiology, and therapeutic approaches. *Lancet Neurol.* **20**, 68–80 (2021).
25. T. R. Dawber, W. B. Kannel, The Framingham Study An Epidemiological Approach to Coronary Heart Disease. *Circulation.* **34**, 553–555 (1966).
26. M. Feinleib, W. B. Kannel, R. J. Garrison, P. M. McNamara, W. P. Castelli, The Framingham Offspring Study. Design and preliminary data. *Prev Med.* **4**, 518–25 (1975).
27. G. L. Splansky, D. Corey, Q. Yang, L. D. Atwood, L. A. Cupples, E. J. Benjamin, R. B. D’Agostino Sr., C. S. Fox, M. G. Larson, J. M. Murabito, C. J. O’Donnell, R. S. Vasan, P. A. Wolf, D. Levy, The Third Generation Cohort of the National Heart, Lung, and Blood Institute’s Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol.* **165**, 1328–35 (2007).
28. S. Seshadri, A. Beiser, J. Selhub, P. F. Jacques, I. H. Rosenberg, R. B. D’Agostino, P. W. F. Wilson, P. A. Wolf, Plasma homocysteine as a risk factor for dementia and Alzheimer’s disease. *N. Engl. J. Med.* **346**, 476–483 (2002).
29. O. L. Lopez, L. H. Kuller, A. Fitzpatrick, D. Ives, J. T. Becker, N. Beauchamp, Evaluation of dementia in the cardiovascular health cognition study. *Neuroepidemiology.* **22**, 1–12 (2003).
30. M. T. Heneka, M. J. Carson, J. El Khoury, G. E. Landreth, F. Brosseron, D. L. Feinstein, A. H. Jacobs, T. Wyss-Coray, J. Vitorica, R. M. Ransohoff, K. Herrup, S. A. Frautschy, B. Finsen, G. C. Brown, A. Verkhratsky, K. Yamanaka, J. Koistinaho, E. Latz, A. Halle, G. C. Petzold, T. Town, D. Morgan, M. L. Shinohara, V. H. Perry, C. Holmes, N. G. Bazan, D. J. Brooks, S. Hunot, B. Joseph, N. Deigendesch, O. Garaschuk, E. Boddeke, C. A. Dinarello, J. C. Breitner, G. M. Cole, D. T. Golenbock, M. P. Kummer, Neuroinflammation in Alzheimer’s disease. *Lancet Neurol.* **14**, 388–405 (2015).
31. B. W. Kunkle, B. Grenier-Boley, R. Sims, J. C. Bis, V. Damotte, A. C. Naj, A. Boland, M. Vronskaya, S. J. van der Lee, A. Amlie-Wolf, C. Bellenguez, A. Frizatti, V. Chouraki, E. R. Martin, K. Sleegers, N. Badarinarayan, J. Jakobsdottir, K. L. Hamilton-Nelson, S. Moreno-Grau, R. Olaso, R. Raybould, Y. Chen, A. B. Kuzma, M. Hiltunen, T. Morgan, S. Ahmad, B. N. Vardarajan, J. Epelbaum, P. Hoffmann, M. Boada, G. W. Beecham, J.-G. Garnier, D. Harold, A. L. Fitzpatrick, O. Valladares, M.-L. Moutet, A. Gerrish, A. V. Smith, L. Qu, D. Bacq, N. Denning, X. Jian, Y. Zhao, M. Del Zompo, N. C. Fox, S.-H. Choi, I. Mateo, J. T. Hughes, H. H. Adams, J. Malamon, F. Sanchez-Garcia, Y. Patel, J. A. Brody, B. A. Dombroski, M. C. D. Naranjo, M. Daniilidou, G. Eiriksdottir, S. Mukherjee, D. Wallon, J. Uphill, T. Aspelund, L. B. Cantwell, F. Garzia, D. Galimberti, E. Hofer, M. Butkiewicz, B. Fin, E. Scarpini, C. Sarnowski, W. S. Bush, S. Meslage, J. Kornhuber, C. C. White, Y. Song, R. C. Barber, S. Engelborghs, S. Sordon, D. Voijnovic, P. M. Adams, R. Vandenberghe, M. Mayhaus, L. A. Cupples, M. S. Albert, P. P. De Deyn, W. Gu, J. J. Himali, D. Beekly, A. Squassina, A. M. Hartmann, A. Orellana, D. Blacker, E. Rodriguez-Rodriguez, S. Lovestone, M. E. Garcia, R. S. Doody, C. Munoz-Fernandez, R. Sussams, H. Lin, T. J. Fairchild, Y. A. Benito, C. Holmes, H. Karamujić-Čomić, M. P. Frosch, H. Thonberg, W. Maier, G. Roshchupkin, B. Ghetti, V. Giedraitis, A. Kawalia, S. Li, R. M. Huebinger, L. Kilander, S. Moebus, I. Hernández, M. I. Kamboh, R. Brundin, J. Turton, Q. Yang, M. J. Katz, L. Concari, J. Lord, A. S. Beiser, C. D. Keene, S. Helisalmi, I. Kloszewska, W. A. Kukull, A. M. Koivisto, A. Lynch, L. Tarraga, E. B. Larson, A. Haapasalo, B. Lawlor, T. H. Mosley, R. B. Lipton, V. Solfrizzi, M. Gill, W. T. Longstreth, T. J.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/) .

Montine, V. Frisardi, M. Diez-Fairen, F. Rivadeneira, R. C. Petersen, V. Deramecourt, I. Alvarez, F. Salani, A. Ciaramella, E. Boerwinkle, E. M. Reiman, N. Fievet, J. I. Rotter, J. S. Reisch, O. Hanon, C. Cupidi, A. G. Andre Uitterlinden, D. R. Royall, C. Dufouil, R. G. Maletta, I. de Rojas, M. Sano, A. Brice, R. Cecchetti, P. S. George-Hyslop, K. Ritchie, M. Tsolaki, D. W. Tsuang, B. Dubois, D. Craig, C.-K. Wu, H. Soininen, D. Avramidou, R. L. Albin, L. Fratiglioni, A. Germanou, L. G. Apostolova, L. Keller, M. Koutroumani, S. E. Arnold, F. Panza, O. Gkatzima, S. Asthana, D. Hannequin, P. Whitehead, C. S. Atwood, P. Caffarra, H. Hampel, I. Quintela, Á. Carracedo, L. Lannfelt, D. C. Rubinsztein, L. L. Barnes, F. Pasquier, L. Frölich, S. Barral, B. McGuinness, T. G. Beach, J. A. Johnston, J. T. Becker, P. Passmore, E. H. Bigio, J. M. Schott, T. D. Bird, J. D. Warren, B. F. Boeve, M. K. Lupton, J. D. Bowen, P. Proitsi, A. Boxer, J. F. Powell, J. R. Burke, J. S. K. Kauwe, J. M. Burns, M. Mancuso, J. D. Buxbaum, U. Bonuccelli, N. J. Cairns, A. McQuillin, C. Cao, G. Livingston, C. S. Carlson, N. J. Bass, C. M. Carlsson, J. Hardy, R. M. Carney, J. Bras, M. M. Carrasquillo, R. Guerreiro, M. Allen, H. C. Chui, E. Fisher, C. Masullo, E. A. Crocco, C. DeCarli, G. Bisceglia, M. Dick, L. Ma, R. Duara, N. R. Graff-Radford, D. A. Evans, A. Hodges, K. M. Faber, M. Scherer, K. B. Fallon, M. Riemenschneider, D. W. Fardo, R. Heun, M. R. Farlow, H. Kölsch, S. Ferris, M. Leber, T. M. Foroud, I. Heuser, D. R. Galasko, I. Giegling, M. Gearing, M. Hüll, D. H. Geschwind, J. R. Gilbert, J. Morris, R. C. Green, K. Mayo, J. H. Growdon, T. Feulner, R. L. Hamilton, L. E. Harrell, D. Driche, L. S. Honig, T. D. Cushion, M. J. Huentelman, P. Hollingworth, C. M. Hulette, B. T. Hyman, R. Marshall, G. P. Jarvik, A. Meggy, E. Abner, G. E. Menzies, L.-W. Jin, G. Leonenko, L. M. Real, G. R. Jun, C. T. Baldwin, D. Grozeva, A. Karydas, G. Russo, J. A. Kaye, R. Kim, F. Jessen, N. W. Kowall, B. Vellas, J. H. Kramer, E. Vardy, F. M. LaFerla, K.-H. Jöckel, J. J. Lah, M. Dichgans, J. B. Leverenz, D. Mann, A. I. Levey, S. Pickering-Brown, A. P. Lieberman, N. Klopp, K. L. Lunetta, H.-E. Wichmann, C. G. Lyketsos, K. Morgan, D. C. Marson, K. Brown, F. Martiniuk, C. Medway, D. C. Mash, M. M. Nöthen, E. Masliah, N. M. Hooper, W. C. McCormick, A. Daniele, S. M. McCurry, A. Bayer, A. N. McDavid, J. Gallacher, A. C. McKee, H. van den Bussche, M. Mesulam, C. Brayne, B. L. Miller, S. Riedel-Heller, C. A. Miller, J. W. Miller, A. Al-Chalabi, J. C. Morris, C. E. Shaw, A. J. Myers, J. Wiltfang, S. O'Bryant, J. M. Olichney, V. Alvarez, J. E. Parisi, A. B. Singleton, H. L. Paulson, J. Collinge, W. R. Perry, S. Mead, E. Peskind, D. H. Cribbs, M. Rossor, A. Pierce, N. S. Ryan, W. W. Poon, B. Nacmias, H. Potter, S. Sorbi, J. F. Quinn, E. Sacchinelli, A. Raj, G. Spalletta, M. Raskind, C. Caltagirone, P. Bossù, M. D. Orfei, B. Reisberg, R. Clarke, C. Reitz, A. D. Smith, J. M. Ringman, D. Warden, E. D. Roberson, G. Wilcock, E. Rogaeva, A. C. Bruni, H. J. Rosen, M. Gallo, R. N. Rosenberg, Y. Ben-Shlomo, M. A. Sager, P. Mecocci, A. J. Saykin, P. Pastor, M. L. Cuccaro, J. M. Vance, J. A. Schneider, L. S. Schneider, S. Slifer, W. W. Seeley, A. G. Smith, J. A. Sonnen, S. Spina, R. A. Stern, R. H. Swerdlow, M. Tang, R. E. Tanzi, J. Q. Trojanowski, J. C. Troncoso, V. M. Van Deerlin, L. J. Van Eldik, H. V. Vinters, J. P. Vonsattel, S. Weintraub, K. A. Welsh-Bohmer, K. C. Wilhelmsen, J. Williamson, T. S. Wingo, R. L. Woltjer, C. B. Wright, C.-E. Yu, L. Yu, Y. Saba, A. Pilotto, M. J. Bullido, O. Peters, P. K. Crane, D. Bennett, P. Bosco, E. Coto, V. Boccardi, P. L. De Jager, A. Lleo, N. Warner, O. L. Lopez, M. Ingelsson, P. Deloukas, C. Cruchaga, C. Graff, R. Gwilliam, M. Fornage, A. M. Goate, P. Sanchez-Juan, P. G. Kehoe, N. Amin, N. Ertekin-Taner, C. Berr, S. DeBette, S. Love, L. J. Launer, S. G. Younkin, J.-F. Dartigues, C. Corcoran, M. A. Ikram, D. W. Dickson, G. Nicolas, D. Champion, J. Tschanz, H. Schmidt, H. Hakonarson, J. Clarimon, R. Munger, R. Schmidt, L. A. Farrer, C. Van Broeckhoven, M. C. O'Donovan, A. L. DeStefano, L. Jones, J. L. Haines, J.-F. Deleuze, M. J. Owen, V. Gudnason, R. Mayeux, V. Escott-Price, B. M. Psaty, A. Ramirez, L.-S. Wang, A. Ruiz, C. M. van Duijn, P. A. Holmans, S. Seshadri, J. Williams, P. Amouyel, G. D. Schellenberg, J.-C. Lambert, M. A. Pericak-Vance, Alzheimer Disease Genetics Consortium (ADGC), European Alzheimer's Disease Initiative (EADI), Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium (CHARGE), Genetic and Environmental Risk in AD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium (GERAD/PERADES), Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019).

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

32. W. A. Kukull, R. Higdon, J. D. Bowen, W. C. McCormick, L. Teri, G. D. Schellenberg, G. van Belle, L. Jolley, E. B. Larson, Dementia and Alzheimer disease incidence: a prospective cohort study. *Arch. Neurol.* **59**, 1737–1746 (2002).
33. A. Matevosian, S. Akbarian, Neuronal nuclei isolation from human postmortem brain tissue. *J. Vis. Exp. JoVE* (2008), doi:10.3791/914.
34. J. M. Granja, M. R. Corces, S. E. Pierce, S. T. Bagdatli, H. Choudhry, H. Y. Chang, W. J. Greenleaf, ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.* **53**, 403–411 (2021).

Acknowledgements

We thank Marilynn Miller, Kelley Faber, Amanda Kuzma, William Lee and KatieRose Richmire for providing age at blood draw for ADSP participants, Aimee Schantz for compiling the neuropathology data from ACT, Walter Kukull and Jessica Culhane for compiling neuropathology data from NACC/ADRCs, and Ben Ebert and Ryan Corces for helpful discussions. Most importantly, we thank the studies and participants who provided biological samples and data.

The following are study specific acknowledgements:

TOPMed

WGS for the TOPMed program was supported by the National Heart, Lung and Blood Institute (NHLBI). Centralized read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1; contract HHSN268201800002I). Phenotype harmonization, data management, sample-identity quality control and general study coordination were provided by the TOPMed Data Coordinating Center (R01HL-120393; U01HL-120393; contract HHSN268201800001I). Genome sequencing for “NHLBI TOPMed: Whole Genome Sequencing and Related Phenotypes in the Framingham Heart Study” (phs000974.v1.p1) was performed at the Broad Institute Genomics Platform (3R01HL092577-06S1, 3U54HG003067-12S2). Genome sequencing for “NHLBI TOPMed: Whole Genome Sequencing Project: Cardiovascular Health Study” (phs001369.v2.p2) was performed at the Broad Institute Genomics Platform (3R01HL092577-06S1, 3U54HG003067-12S2) and Baylor University (HHSN268201600033I, 3U54HG003273-12S2, HHSN268201500015C). We gratefully acknowledge the studies and participants who provided biological samples and data for TOPMed.

Framingham Heart Study: The Framingham Heart Study (FHS) acknowledges the support of contracts NO1-HC-25195, HHSN268201500001I and 75N92019D00031 from the National Heart, Lung and Blood Institute and grant supplement R01 HL092577-06S1 for this research. We also acknowledge the dedication of the FHS study participants without whom this research would not be possible. Dr. Vasani is supported in part by the Evans Medical Foundation and the Jay and Louis Coffman Endowment from the Department of Medicine, Boston University School of Medicine.

The FHS was also supported by grants from the National Institutes of Aging (AG054076 and AG058589).

Cardiovascular Health Study: This research was supported by contracts HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, 75N92021D00006, and grants U01HL080295 and U01HL130114 from the National Heart, Lung, and Blood Institute (NHLBI), with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ADSP

The Alzheimer's Disease Sequencing Project (ADSP) is comprised of two Alzheimer's Disease (AD) genetics consortia and three National Human Genome Research Institute (NHGRI) funded Large Scale Sequencing and Analysis Centers (LSAC). The two AD genetics consortia are the Alzheimer's Disease Genetics Consortium (ADGC) funded by NIA (U01 AG032984), and the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) funded by NIA (R01 AG033193), the National Heart, Lung, and Blood Institute (NHLBI), other National Institute of Health (NIH) institutes and other foreign governmental and non-governmental organizations. The Discovery Phase analysis of sequence data is supported through U01AG047133 (to Drs. Schellenberg, Farrer, Pericak-Vance, Mayeux, and Haines); U01AG049505 to Dr. Seshadri; U01AG049506 to Dr. Boerwinkle; U01AG049507 to Dr. Wijsman; and U01AG049508 to Dr. Goate and the Discovery Extension Phase analysis is supported through U01AG052411 to Dr. Goate, U01AG052410 to Dr. Pericak-Vance and U01 AG052409 to Drs. Seshadri and Fornage. Data generation and harmonization in the Follow-up Phases is supported by U54AG052427 (to Drs. Schellenberg and Wang).

The ADGC cohorts include: Adult Changes in Thought (ACT) funded by NIA (U01AG006781 and U19AG066567), the Alzheimer's Disease Research Centers (ADRCs) including the University of Washington ADRC funded by NIA (P50AG005136 and P30AG066509), the Chicago Health and Aging Project (CHAP), the Memory and Aging Project (MAP), Mayo Clinic (MAYO), Mayo Parkinson's Disease controls, University of Miami, the Multi-Institutional Research in Alzheimer's Genetic Epidemiology Study (MIRAGE), the National Cell Repository for Alzheimer's Disease (NCRAD), the National Institute on Aging Late Onset Alzheimer's Disease Family Study (NIA-LOAD), the Religious Orders Study (ROS), the Texas Alzheimer's Research and Care Consortium (TARC), Vanderbilt University/Case Western Reserve University (VAN/CWRU), the Washington Heights-Inwood Columbia Aging Project (WHICAP) and the Washington University Sequencing Project (WUSP), the Columbia University Hispanic- Estudio Familiar de Influenza Genetica de Alzheimer (EFIGA), the University of Toronto (UT), and Genetic Differences (GD).

The CHARGE cohorts are supported in part by National Heart, Lung, and Blood Institute (NHLBI) infrastructure grant HL105756 (Psaty), RC2HL102419 (Boerwinkle) and the neurology working group is supported by the National Institute on Aging (NIA) R01 grant AG033193. The CHARGE cohorts participating in the ADSP include the following: Austrian Stroke Prevention Study (ASPS), ASPS-Family study, and the Prospective Dementia Registry-Austria (ASPS/PRODEM-Aus), the Atherosclerosis Risk in Communities (ARIC) Study, the Cardiovascular Health Study (CHS), the Erasmus Rucphen Family Study (ERF), the Framingham Heart Study (FHS), and the Rotterdam Study (RS). ASPS is funded by the Austrian Science Fond (FWF) grant number P20545-P05 and P13180 and the Medical University of Graz. The ASPS-Fam is funded by the Austrian Science Fund (FWF) project I904, the EU Joint Programme - Neurodegenerative Disease Research (JPND) in frame of the BRIDGET project (Austria, Ministry of Science) and the Medical University of Graz and the Steiermärkische Krankenanstalten Gesellschaft. PRODEM-Austria is supported by the Austrian Research Promotion agency (FFG) (Project No. 827462) and by the Austrian National Bank (Anniversary Fund, project 15435. ARIC research is carried out as a collaborative study supported by NHLBI contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C). Neurocognitive data in ARIC is collected by U01 2U01HL096812, 2U01HL096814, 2U01HL096899, 2U01HL096902, 2U01HL096917 from the NIH (NHLBI, NINDS, NIA and NIDCD), and with previous brain MRI examinations funded by R01-HL70825 from the NHLBI. CHS research was supported by contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and grants U01HL080295 and U01HL130114 from the NHLBI with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by R01AG023629, R01AG15928, and R01AG20098 from the NIA. FHS research is supported by NHLBI contracts N01-HC-25195 and HHSN268201500001I. This study was also supported by additional grants from the NIA (R01s AG054076, AG049607 and AG033040 and NINDS (R01 NS017950). The ERF study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002-01254). High-throughput analysis of the ERF data was supported by a joint grant from the Netherlands Organization for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, the Netherlands Organization for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the municipality of Rotterdam. Genetic data sets are also supported by the Netherlands Organization of Scientific Research NWO Investments (175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), and the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project 050-060-810. All studies are grateful to their participants, faculty and staff. The content of these manuscripts is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the U.S. Department of Health and Human Services.

The four LSACs are: the Human Genome Sequencing Center at the Baylor College of Medicine (U54 HG003273), the Broad Institute Genome Center (U54HG003067), The American Genome Center at the Uniformed Services University of the Health Sciences (U01AG057659), and the Washington University Genome Institute (U54HG003079).

Biological samples and associated phenotypic data used in primary data analyses were stored at Study Investigators institutions, and at the National Cell Repository for Alzheimer's Disease (NCRAD, U24AG021886) at Indiana University funded by NIA. Associated Phenotypic Data used in primary and secondary data analyses were provided by Study Investigators, the NIA funded Alzheimer's Disease Centers (ADCs), and the National Alzheimer's Coordinating Center (NACC, U01AG016976) and the National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS, U24AG041689) at the University of Pennsylvania, funded by NIA, and at the Database for Genotypes and Phenotypes (dbGaP) funded by NIH. This research was supported in part by the Intramural Research Program of the National Institutes of Health, National Library of Medicine. Contributors to the Genetic Analysis Data included Study Investigators on projects that were individually funded by NIA, and other NIH institutes, and by private U.S. organizations, or foreign governmental or nongovernmental organizations.

NACC

The NACC database is funded by NIA/NIH Grant U24 AG072122. NACC data are contributed by the NIA-funded ADCs: P30 AG019610 (PI Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P50 AG005134 (PI Bradley Hyman, MD, PhD), P50 AG016574 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Thomas Wisniewski, MD), P30 AG013854 (PI Robert Vassar, PhD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P50 AG005131 (PI James Brewer, MD, PhD), P50 AG023501 (PI Bruce Miller, MD), P30 AG035982 (PI Russell Swerdlow, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG049638 (PI Suzanne Craft, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P50 AG033514 (PI Sanjay Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD).

Funding: This work was supported by grants to S.J. from the Burroughs Wellcome Fund Career Award for Medical Scientists, Fondation Leducq (TNE-18CVD04), the Ludwig Center for Cancer Stem Cell Research at

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/) .

Stanford University, and the National Institutes of Health (DP2-HL157540), to S.J. and H.B. from the National Institutes of Health (R01-HL148565), to T.J.M. from the National Institutes of Health (AG053959), to C.D.K. from the Nancy and Buster Alvord Endowment, to A.T.S. from the National Institutes of Health K08CA230188 and the Burroughs Wellcome Fund Career Award for Medical Scientists, to P.N. from the National Institutes of Health (R01HL142711, R01HL148050, R01HL151283, R01HL148565, R01HL135242, R01HL151152, R01DK125782), Fondation Leducq (TNE-18CVD04), and MGH (Fireman Endowed Chair in Vascular Medicine).

Author contributions: H.B. conceptualized experimental design and performed all research investigations on brain tissue. J.A.B. performed single-cell ATAC-seq experiments and analysis with assistance from Y.Q. and W.Y. M.J. performed analyses and data curation of ADSP data. S.W., L.M., M.C., H.A., and D.N. provided genotyping and variant calling for CHIP from human samples. A.B., J.C.B., M.F., W.T.L., O.L.L., B.M.P., C.L.S., and S.S. contributed to neurocognitive phenotype curation in CHS and FHS and provided guidance on epidemiological study design. A.G.B., P.N., and J.W. performed CHIP calling in TOPMed. E.B.L., P.K.C., and C.D.K. provided brain tissues and neuropathology data from ACT. A.T.S. provided analysis, resources, and supervision for single-cell ATAC-seq. T.J.M. provided supervision, conceptualization, and access to resources for studies involving brain autopsy tissue. S.J. conceived the research plan and supervision for all analyses. H.B., J.A.B., and S.J. wrote the original draft with review from all authors

Competing interests: A.T.S. is a founder of Immunai and Cartography Biosciences and receives research funding from 10x Genomics, Arsenal Biosciences, and Allogene Therapeutics. S.J. is a scientific advisor to Novartis, Roche Genentech, AVRO Bio, and Foresite Labs. P.N. reports investigator-initiated grants from Amgen, Apple, AstraZeneca, Novartis, and Boston Scientific, personal fees from Blackstone Life Sciences, Apple, AstraZeneca, Novartis, Genentech, and Foresite Labs, and spousal employment at Vertex, all unrelated to the present work.

Data and materials availability:

All genetic data from TOPMed and ADSP are available on dbGaP for investigators with approval protocols. Single-cell ATAC-seq data from human brain samples will be deposited in Gene Expression Omnibus.

Figure 1

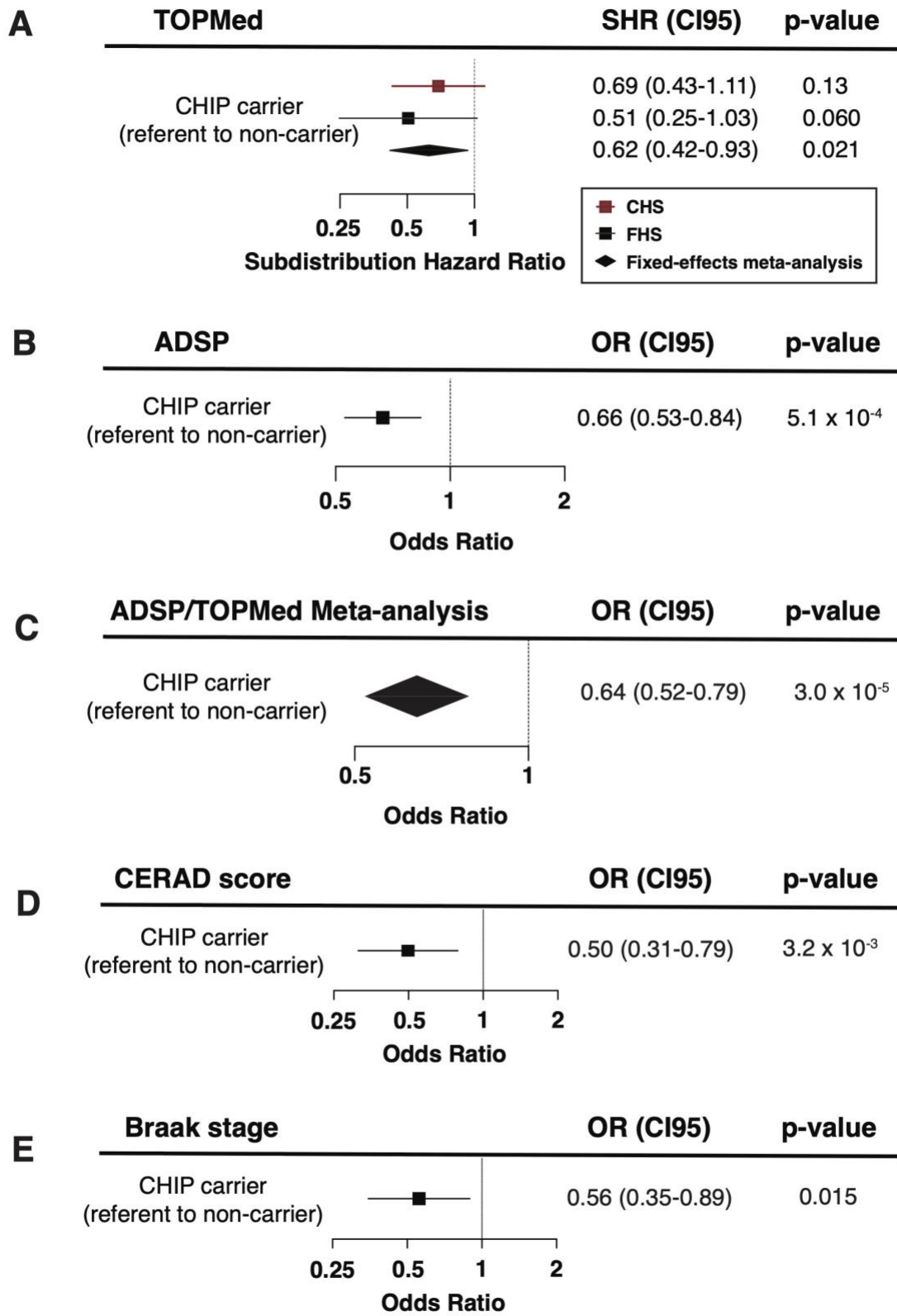


Fig. 1. CHIP is associated with protection from AD dementia and ADNC in those without dementia

A) Forest plot for risk of incident AD in CHIP carriers from Cardiovascular Health Study (CHS) and Framingham Heart Study (FHS) relative to non-carriers. Subdistribution hazard ratios (SHR), 95% confidence intervals (CI95) and Wald p-values were calculated for each covariate from competing risks regression models which included age at blood draw, sex, and *APOE* genotype as covariates. Results from CHS and FHS were then meta-analyzed using a fixed-effects model for the two cohorts (see Figure S1 for full regression results).

B) Forest plot for risk of AD in CHIP carriers relative to non-carriers from the AD Sequencing Project (ADSP) with *APOE* $\epsilon 3\epsilon 3$ genotype. Odds ratio (OR), 95% confidence interval (CI95) and Wald p-value were calculated from a logistic regression model that also included age at time of blood draw and sex as covariates (see Table S6 for full regression results).

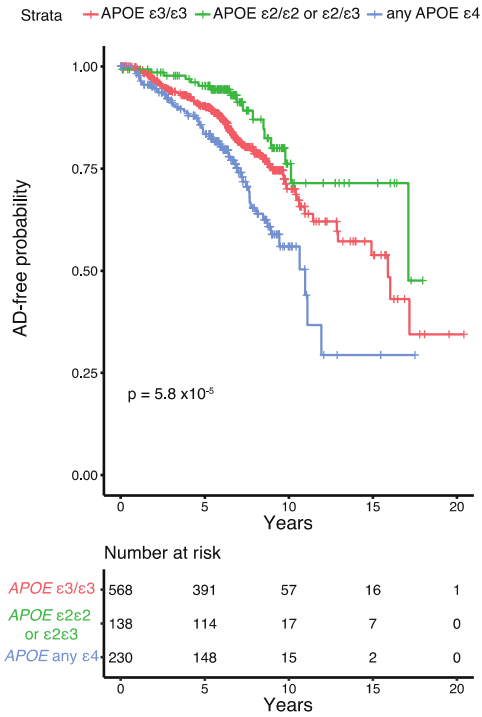
C) Fixed-effects meta-analysis for risk of AD in CHIP carriers using logistic regression in ADSP, FHS, and CHS.

D) Plot of odds ratio (OR) and 95% confidence interval (CI95) for increased CERAD neuritic plaque score in ADSP participants without a dementia diagnosis from an ordinal logistic regression model. The covariates included in the model are age at autopsy, *APOE* genotype, sex, and CHIP status. The p-values were calculated by comparing the t-statistic for each covariate against a standard normal distribution. Full regression results are in Table S8.

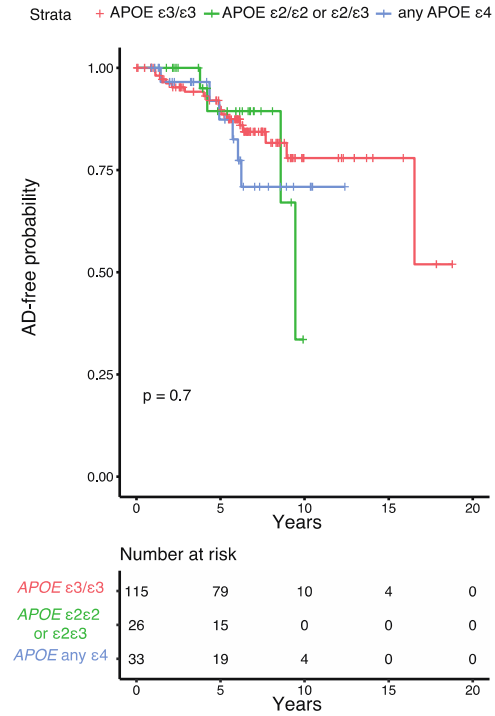
E) Plot of odds ratio (OR) and 95% confidence interval (CI95) for increased Braak stage in ADSP participants without a dementia diagnosis from an ordinal logistic regression model. The covariates included in the model are age at autopsy, *APOE* genotype, sex, and CHIP status. The p-values were calculated by comparing the t-statistic for each covariate against a standard normal distribution. Full regression results are in Table S8.

Figure 2

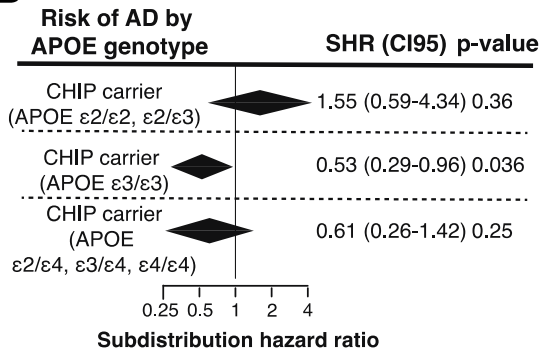
A No CHIP (Age >70)



CHIP (Age >70)



B



C

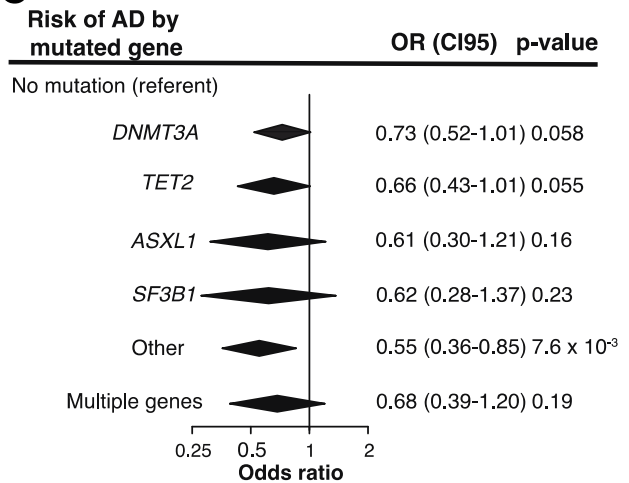


Fig. 2. Associations of CHIP to AD by *APOE* genotype and mutated driver gene

A) Kaplan-Meier curve showing AD-free probability in CHIP non-carriers (left) and carriers (right), stratified by *APOE* genotype. Analysis was restricted to those older than 70 at time of blood draw.

B) Forest plot for effect of CHIP on AD risk in participants from CHS and FHS stratified by *APOE* genotype. Participants were binned into those with neutral (*APOE* $\epsilon 3\epsilon 3$), low risk (*APOE* $\epsilon 2\epsilon 2$ and $\epsilon 2\epsilon 3$), and high-risk (any *APOE* $\epsilon 4$ allele) groups. Subdistribution hazard ratios (SHR), 95% confidence intervals (CI95) and Wald p-values were calculated for each covariate (age at time blood draw for sequencing, sex, CHIP carrier status) from competing risks regression models, and results from FHS and CHS were then meta-analyzed using a fixed-effects model (see Table S9 for full regression results).

C) Plot of odds ratios for effect of mutated CHIP gene on AD in participants from the TOPMed cohorts (CHS and FHS) and ADSP. Odds ratios (OR), 95% confidence intervals (CI95) and Wald p-values were calculated for each covariate (age at time blood draw for sequencing, sex, cohort, and *APOE* genotype) from logistic regression models, and results from the TOPMed cohorts and ADSP were then meta-analyzed using a fixed-effects model (see Table S10 for full regression results).

Figure 3

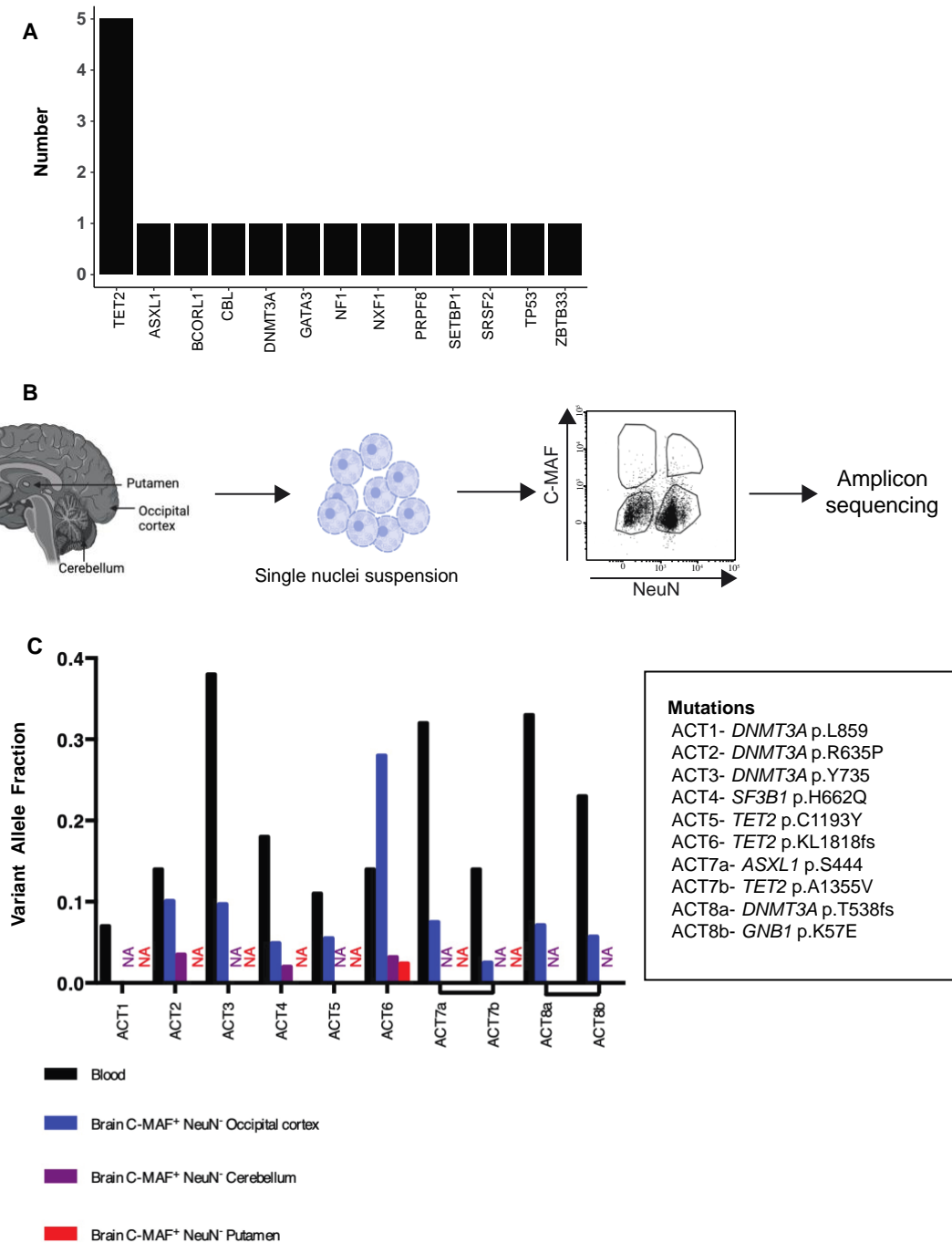


Fig. 3. CHIP variants can be found in the microglia-enriched fraction of brain

A) Barplot of putative CHIP mutations identified from whole exome sequencing of brain DNA from 1,775 persons in ADSP. Full details on the variants identified are in Table S11.

B) Schematic of experimental workflow. Autopsy samples from occipital cortex, cerebellum and putamen were digested to prepare single nuclei suspensions. Nuclei were then stained and sorted using antibodies to C-Maf⁺ (marker of myeloid cells) and NeuN⁺ (Marker of Neuronal cells), followed by amplicon sequencing for CHIP variants.

It is made available under a [CC-BY-NC-ND 4.0 International license](#) .

C) Barplot of the variant allele fraction (VAF) of the CHIP variants from 8 donors (ACT1 to ACT8). For each sample, the VAF in the blood and in the brain C-Maf⁺ NeuN⁻ population are shown. Occipital cortex was available for all 8 donors. A bar for cerebellum or putamen is shown if available, otherwise NA in the corresponding color designates lack of an available sample (purple for cerebellum and red for putamen). The CHIP mutations carried by each participant are reported in the box on the right of the barplot.

Figure 4

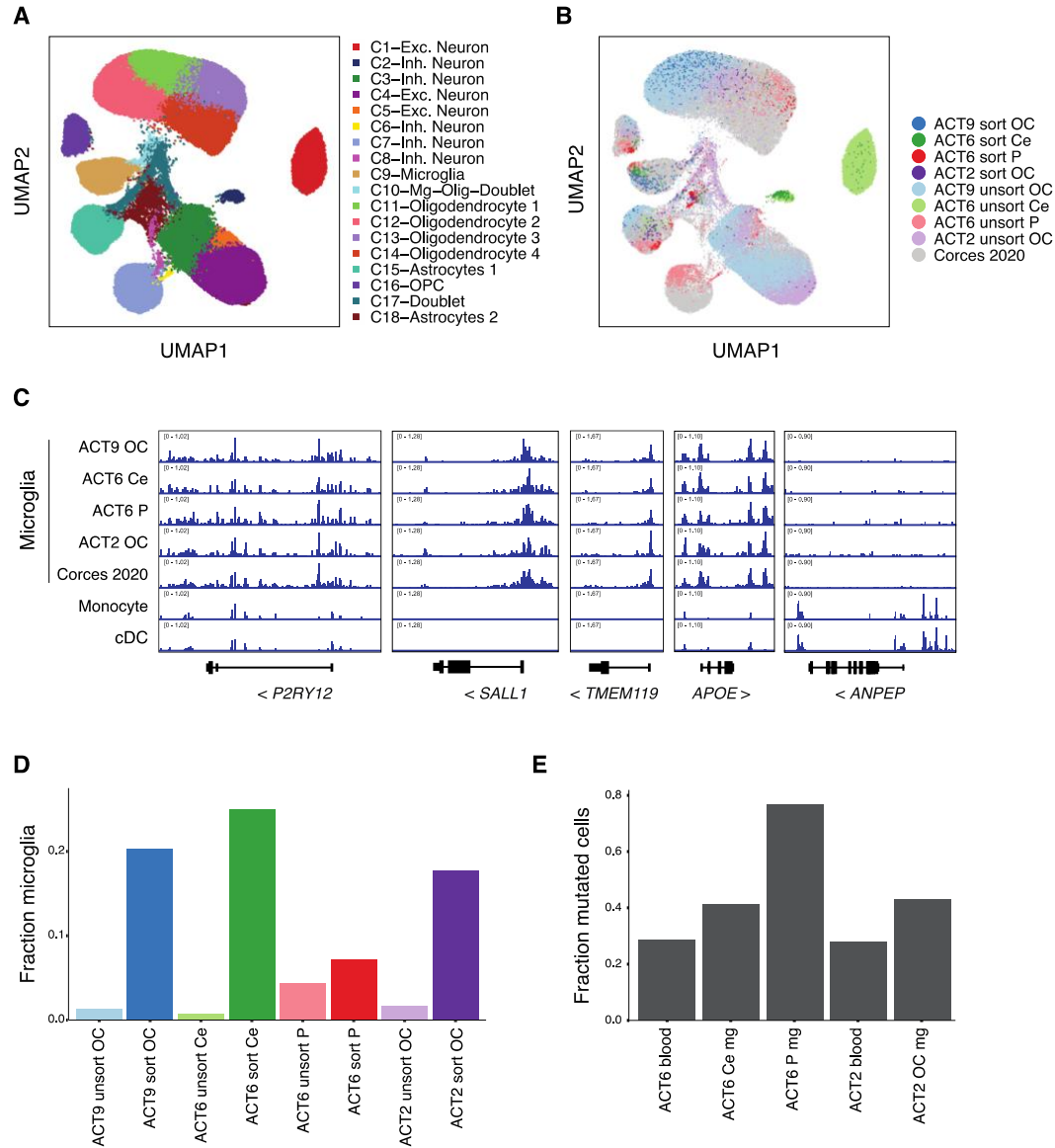


Fig. 4. scATAC-seq of brain samples from CHIP carriers reveals that the mutated cells are similar to microglia and comprise a large proportion of the microglial pool

A) scATAC-seq profiles of 111,190 cells from our dataset and the Corces 2020 adult human brain dataset. Each dot represents the scATAC-seq profile of one cell and is colored by its assigned cluster.

B) scATAC-seq profiles of all cells colored by which sample it originated from. Samples from Corces 2020 are aggregated and shown in grey. Sorted samples were from the c-Maf⁺ NeuN⁻ gate.

C) Pseudo-bulk tracks for selected gene loci. The top 5 tracks show scATAC-seq coverage of cells from the indicated sample (or aggregated Corces 2020 samples) within C9, the microglia cluster. The monocyte and classical dendritic cell (cDC) tracks are from the Satpathy 2019 hematopoiesis dataset.

D) Fraction of cells in cluster C9 (microglia) for sorted versus unsorted brain samples.

E) Proportion of microglia (mg) bearing a CHIP mutation in each sample, calculated by dividing the percentage of cells in cluster 9 in each unsorted sample by 2 times the VAF of the CHIP mutation for that sample.

The brain regions are abbreviated as Ce for Cerebellum, OC for Occipital cortex, and P for putamen.