Please cite this article in press as: Screven et al., Harnessing diversity to study Alzheimer's disease: A new iPSC resource from the NIH CARD and ADNI, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.01.026

Neuron

CellPress

NeuroView

Harnessing diversity to study Alzheimer's disease: A new iPSC resource from the NIH CARD and ADNI

Laurel A. Screven,[1] Caroline B. Pantazis,[1] Katherine M. Andersh,[1] Samantha Hong,[1] Dan Vitale,[1,2,3] Erika Lara,[1] Ray Yueh Ku,[1] Peter Heutink,[4,9] Jason Meyer,[5] Kelley Faber,[6] Kwangsik Nho,[5,7] Andrew J. Saykin,[5,6,7] Tatiana M. Foroud,[5,6] Mike A. Nalls,[1,2,3] Cornelis Blauwendraat,[1,3] Andrew Singleton,[1,3] and Priyanka S. Narayan[1,8,*]

[1]Center for Alzheimer's and Related Dementias (CARD), National Institute on Aging and National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA
[2]DataTecnica LLC, Washington, DC, USA
[3]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA
[4]German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany
[5]Indiana Alzheimer's Disease Research Center, Indiana University School of Medicine, Indianapolis, IN, USA
[6]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA
[7]Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN, USA
[8]Genetics and Biochemistry Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA
[9]Present address: Alector, Inc., South San Francisco, CA, USA
*Correspondence: priyanka.narayan@nih.gov
https://doi.org/10.1016/j.neuron.2024.01.026

The iDA Project (iPSCs to Study Diversity in Alzheimer's and Alzheimer's Disease-related Dementias) is generating 200 induced pluripotent stem cell lines from Alzheimer's Disease Neuroimaging Initiative participants. These lines are sex balanced, include common *APOE* genotypes, span disease stages, and are ancestrally diverse. Cell lines and characterization data will be shared openly.

Over 47 million individuals worldwide live with Alzheimer's disease (AD) and related dementias (ADRDs); this number is rising rapidly with an aging population (Alzheimer's Disease Facts and Figures 2023, https://www.alz.org/). The increase of ADRD incidence is accompanied by an increasing medical, financial, and emotional toll for families and governments around the world. AD is the most common neurodegenerative disease. Although years of research have advanced our understanding of disease progression, only recently have a few disease-modifying immunotherapies emerged for AD.[1] Although these therapies present some hope for slowing disease progression in eligible patients, there is still much more to be understood about the molecular drivers of risk and progression for AD and other related dementias. A more complete understanding of this biology can accelerate the development of preventative and therapeutic approaches for these devastating diseases.

Recent advances in genomic technologies have allowed more detailed research on molecular mechanisms and drivers. For example, the lowered cost of array genotyping and next-generation sequencing has empowered large-scale genetics studies to identify rare and common variants that govern risk and resilience to ADRDs.[2] Single-cell profiling techniques have allowed for the resolution of molecular mosaics of cell types from complex mixtures.[3] Induced pluripotent stem cell (iPSC) technology combined with genome editing enables the exploration of the biochemistry and cell biology of multiple human brain cell types, including neurons, glia, and endothelial cells.[4]

Despite the advances in research methodology, many of the ADRD studies in the past have overlooked differences in genetic, sex, and ancestral diversity.[5] The largest genome-wide association studies identifying novel genetic risk factors have been performed in cohorts of Northern European descent. Smaller epidemiological studies in non-European populations have found that some risk associations diverge between populations with different ancestral backgrounds. A recent meta-analysis of a number of these non-European studies revealed that key genes for ADRDs have ancestry-specific effects.[6] *APOE*, which codes for apolipoprotein E, is one of most notable genes to emerge from this study. The AD risk variant, *APOE4*, has long been touted as the strongest genetic risk factor for AD and impacts risk for several related dementias.[7] However, the magnitude of *APOE4*'s effect on AD risk is dependent on ancestral background; it functions as a strong risk variant in European ancestry populations, while its effects in other populations vary from moderate in populations with African descent to high in South Asian populations.[7] Both epidemiological and genomic studies have suggested that sex-specific differences exist in ADRD incidence, progression, and risk factors.[8] The molecular mechanisms that contribute to these differences are not well understood.

iPSC tools to study ADRDs

iPSC-derived neural cell types offer a tool with which to understand human brain tissue physiology and biochemistry in neurodegenerative diseases. Circumventing the need to obtain human brain biopsies, iPSC-derived neural cell types

Please cite this article in press as: Screven et al., Harnessing diversity to study Alzheimer's disease: A new iPSC resource from the NIH CARD and ADNI, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.01.026

allow researchers to study human brain cells in isolation and in combination with other cell types in two- and three-dimensional cultures. The combination of iPSC-derived cell culture and large-scale sequencing and profiling techniques like single-cell transcriptomics, assay for transposase-accessible chromatin (ATAC) sequencing, genomic sequencing, proteomics, metabolomics, and functional genomics tools have enabled a new era of investigation in molecular ADRD research. Together, these techniques enable researchers to develop testable hypotheses on the cell biology of ADRD causative genes and risk factors.

As testament to the power of iPSCs in ADRD research, many consortia have assembled biobanks of iPSCs focused on a specific disease or disease subset. For example, Answer ALS (answerals.org) has assembled a large collection of iPSCs (over 1,100) derived from amyotrophic lateral sclerosis (ALS) patients and healthy donors. The Tau consortium (tauconsortium.org) has assembled a set of over 30 patient-derived lines from patients with primary tauopathies. National Centralized Repository for Alzheimer's Disease and Related Dementias (NCRAD, ncrad.iu.edu) has collected iPSCs from several labs to assemble a larger centralized bank of over 300 iPSC lines harboring ADRD-causing and risk mutations, as well as healthy control cells. The NINDS Stem Cell Catalog (stemcells.nindsgenetics.org) has banked 421 iPSCs across 14 neurodegenerative diagnoses and their respective controls. Another approach taken is to bank iPSC lines from donors with a wide array of indications. The California Institute of Regenerative Medicine (CIRM, www.cirm.ca.gov/researchers/ipsc-repository/about/) has done just this with 1,556 publicly available iPSC lines in its biobank. These lines span 42 different diagnoses, including 64 lines from patients with AD. Taking a different approach, the NIH Center for Alzheimer's and Related Dementias (CARD) iPSC Neurodegenerative Disease Initiative (iNDI) is engineering multiple iPSC lines, each harboring one of 134 ADRD-associated mutations all in a few well-validated parental-background iPSC lines.[9] iNDI is creating homozygous, heterozygous, and revertant lines to allow comparison of the effects of different ADRD-associated single nucleotide polymorphisms (SNPs) without the confounding effects of different genetic backgrounds and the off-target effects of genetic engineering.

Many of these repositories have tried to ensure access to many of the lines (according to donor wishes) within certain material transfer agreement boundaries. Some of these databases also offer basic characterization data on the lines, including exome sequencing, whole-genome sequencing, genotyping, and RNA-sequencing data. Many repositories contain a wide variety of disease indications depending on donor accessibility. For the repositories curated around a disease indication (Answer ALS and NCRAD), the lines represent multiple genotypes that are causal or confer risk for disease. Due to the paucity of many of these genotypes in the population, the number of iPSC lines available around a single risk gene or causal gene, even in large repositories, is limited. Finally, many of these repositories lack significant representation from non-European ancestry populations. Diverse genetic background representation is especially important in understanding the genetic drivers and risk of neurodegeneration.

The iDA Project

Our goal is to create an ancestrally diverse iPSC repository focused around one of the most important genes in AD risk: *APOE*. The iDA Project cell set is drawn from individuals from the Alzheimer's Disease Neuroimaging Initiative (ADNI) biobank. ADNI participants are classified with one of five increasingly severe diagnoses: control (CN), significant memory concerns (SMCs), early mild cognitive impairment (EMCI), late mild cognitive impairment (LMCI), and AD dementia. Many of these patients have diagnostic imaging, biofluid retrieval, and even brain tissue. We utilized peripheral blood mononuclear cells (PBMCs) from these participants collected by the ADNI Genetics Core and banked at NCRAD. In partnership with the New York Stem Cell Foundation, we are reprogramming these PBMCs into iPSCs. The use of PBMCs to generate iPSCs avoids the additional UV-associated mutational burden that can arise from dermal fibroblast reprogramming. In our cell set, we selected 100 participants with cognitive symptoms of varying degrees (AD dementia, LMCI, EMCI, and SMC) and 100 neurologically healthy control individuals. Within these lines, we used self-declared ancestry status to include all individuals with non-European ancestry whose PMBCs were available. The 200 iDA Project lines were stratified to include 100 male and 100 female lines. Given the incredible influence of the *APOE* genotype on AD risk and progression, we included approximately equal numbers of lines with *APOE3/3*, *APOE3/4*, or *APOE4/4* genotypes. We used ADNI-supplied genotyping array data to determine genetic ancestry using NIH CARD's suite of computational tools[10] (Figure 1).

After reprogramming, iPSCs are subject to a thorough quality-control pipeline, including testing for sterility, mycoplasma, karyotype abnormalities, array genotyping to match identity to the parent PBMCs, pluripotency, and differentiation capacity into the three germ layers. These iPSCs will then be banked and distributed to the research community through NCRAD. NIH CARD is also planning to generate foundational companion data for these lines. Initially, we will differentiate these iPSCs into two neural cell types with significant involvement in AD pathogenesis: astrocytes and neurons. Following these initial two cell types, we plan to expand our data generation for microglia, a cell type which is known to play a unique role in AD depending on its *APOE* genotype. For each of these differentiations, we plan to generate single-cell-level data on the transcriptome and the chromatin state, as well as bulk proteomics and long-read DNA sequencing. These data will be shared openly with the research community. Taken together, these cells and accompanying data will be a transformative resource for those interested in ADRD research and, specifically, in the role of *APOE* variants on disease pathogenesis across genetic ancestries.

In addition to their well-established association with AD risk, *APOE* variants also govern risk and resilience for other neurological diseases like dementia with Lewy bodies, macular degeneration, and cerebral amyloid angiopathy and non-neurological conditions like COVID-19 and cardiovascular disease.[7] Therefore, the iDA Project lines and data could serve as a resource for researchers across multiple disease indications.

Please cite this article in press as: Screven et al., Harnessing diversity to study Alzheimer's disease: A new iPSC resource from the NIH CARD and ADNI, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.01.026

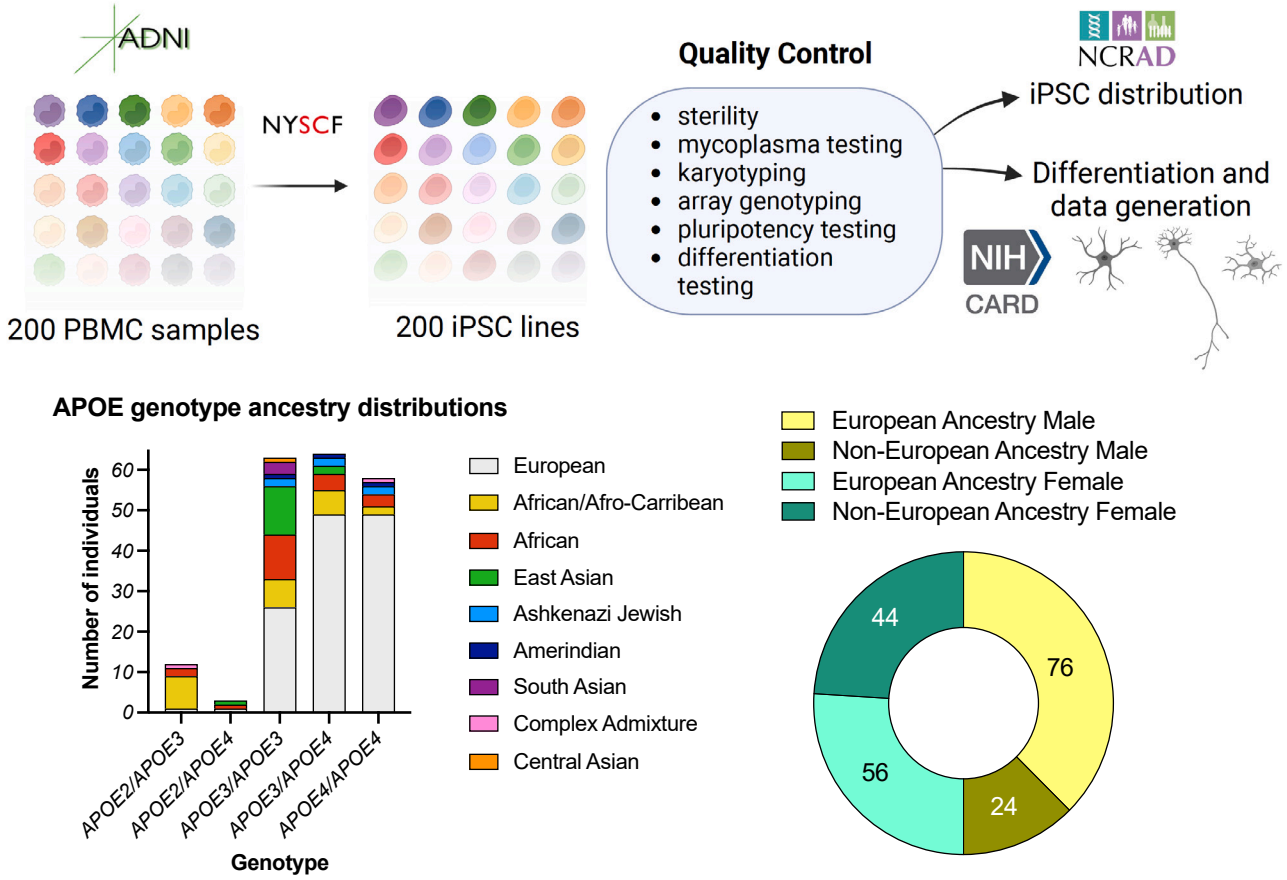

Figure 1. Overview of the iDA Project

The iDA Project will generate 200 iPSC lines from PBMCs from ADNI participants. Once reprogrammed, these lines are subject to rigorous quality-control testing and then shared for public distribution through NCRAD. At the NIH, these lines will be differentiated into central nervous system cell types like astrocytes, neurons, and microglia. Genomic, transcriptomic, and proteomic data will be generated for each of these differentiated lines and shared openly with the research community. These lines are primarily of the *APOE3* homozygous, *APOE3/APOE4* heterozygous, and *APOE4* homozygous genotypes. The lines contain all the non-European ancestry participants whose samples were available through ADNI and contain equal numbers of male and female samples. EUR, European; AAC, African or Afro-Caribbean; AFR, African; EAS, East Asian; AJ, Ashkenazi Jewish; AMR, Amerindian; SAS, South Asian; CAH, complex admixture history; CAS, Central Asian. These data, along with the iPSC lines, will serve as an important resource to examine the interactions among cognitive impairment, sex, *APOE* genotype, and ancestral background. Created with BioRender.com.

Discussion

Despite the incredible potential for iPSC research in uncovering the etiology of ADRDs, existing iPSCs that have previously been used for research do not typically reflect the ancestral and sex diversity of the patient populations within the US and worldwide. The iDA Project repository is a step toward increasing representation of non-European ancestry lines to iPSC neurodegenerative disease research and studying the effects of *APOE* genotypes in the context of ancestral diversity. NIH CARD has initiated another large iPSC initiative, the iNDI project, where a large set of iPSCs harboring ADRD mutations is engineered into a few well-validated background parental lines. The iDA Project takes a complementary approach; although both projects create hundreds of iPSC lines, the iDA Project focuses on genetic diversity across individuals and generates lines with variants of the same gene in different genetic backgrounds.

Our foundational data will not only enable the research of those using lines from the iDA Project but will also allow the study of the interactions between sex, ancestral background, and *APOE* genotype. For this first set of iDA Project lines, we were limited by the participants present within the ADNI cohort (Figure 1). This includes 126 lines of European ancestry, 44 lines of African or Afro-Caribbean ancestry, 19 lines of Asian ancestry (East, South, and Central Asian), 6 lines with Ashkenazi Jewish ancestry, 3 lines with Amerindian ancestry, and 2 lines of a complex admixture of ancestries. We realize that we may not yet have enough lines to perform a fully powered analysis of the interaction between ancestral background and molecular phenotypes. We also recognize that some minoritized groups who are severely impacted by ADRDs may not be represented well in this initial iDA line set. For example, there are no lines from individuals with Latine/Hispanic origin in our dataset and very few lines from individuals with Amerindian ancestry. There is the opportunity to expand into lines that examine risk factors other than *APOE* and to establish further

lines that fully capture the diversity of the US ADRD patient population. The creation of these lines and expansion of the iDA Project depends on availability of research participant material and access. Working with historically marginalized communities to expand these banks must be done carefully. Initially, trust of the biomedical enterprise must be established with these communities, who have often been the victims of past abuses. However, we hope that our initial set of iPSCs will be a springboard for future line creation and curation with diversity in mind.

We are delighted to introduce the iDA Project to the research community, not only as a powerful scientific resource but also as an example of conscious inclusion within ADRD research. We hope that our effort will be the start of model curation and data generation that better represents all the communities impacted by ADRDs.

ACKNOWLEDGMENTS

We would like to thank all the research participants for taking part in the ADNI study.

This work was supported in part by the Intramural Research Program of the National Institute on Aging (NIA) (AG000534), and the Center for Alzheimer's and Related Dementias (CARD), within the Intramural Research Program of the NIA and the National Institute of Neurological Disorders and Stroke (AG000546). This work utilized the computational resources of the National Institutes of Health (NIH) HPC Biowulf cluster (http://hpc.nih.gov). M.A.N. and D.V.'s participation in this project was part of a competitive contract awarded to Data Tecnica International LLC by the NIH to support open science research.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the NIA and the National Institute of Biomedical Imaging and Bioengineering and through generous contributions from the following: AbbVie; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company, Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

The National Centralized Repository for Alzheimer's Disease and Related Dementias (NCRAD) receives government support under a cooperative agreement grant (U24 AG021886) awarded by the NIA.

P.N. is supported by the Intramural Research Program of the National Institute for Diabetes and Digestive and Kidney Diseases. A.J.S. receives support from multiple NIH grants (P30 AG010133, P30 AG072976, R01 AG019771, R01 AG057739, U19 AG024904, R01 LM013463, R01 AG068193, T32 AG071444, U01 AG068057, U01 AG072177, and U19 AG074879). K.N. received support from an NIH grant (R01 LM012535).

DECLARATION OF INTERESTS

A.J.S. has received support from Avid Radiopharmaceuticals, a subsidiary of Eli Lilly (in-kind contribution of PET tracer precursor), and has participated in scientific advisory boards (Bayer Oncology, Eisai, Novo Nordisk, and Siemens Medical Solutions USA, Inc) and an observational study monitoring board (MESA, NIH NHLBI). He also serves as editor-in-chief for *Brain Imaging and Behavior*, a Springer-Nature journal, and serves on four other NIA external advisory committees. M.A.N. currently serves on the scientific advisory board for Character Bio Inc. and is a scientific founder at Neuron23 Inc.

REFERENCES

1. Widera, E.W., Brangman, S.A., and Chin, N.A. (2023). Ushering in a New Era of Alzheimer Disease Therapy. JAMA *330*, 503–504. https://doi.org/10.1001/jama.2023.11701.

2. Bellenguez, C., Küçükali, F., Jansen, I.E., Kleineidam, L., Moreno-Grau, S., Amin, N., Naj, A.C., Campos-Martin, R., Grenier-Boley, B., Andrade, V., et al. (2022). New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat. Genet. *54*, 412–436. https://doi.org/10.1038/s41588-022-01024-z.

3. Mathys, H., Peng, Z., Boix, C.A., Victor, M.B., Leary, N., Babu, S., Abdelhady, G., Jiang, X., Ng, A.P., Ghafari, K., et al. (2023). Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology. Cell *186*, 4365–4385.e27. https://doi.org/10.1016/j.cell.2023.08.039.

4. Penney, J., Ralvenius, W.T., and Tsai, L.H. (2020). Modeling Alzheimer's disease with iPSC-derived brain cells. Mol. Psychiatr. *25*, 148–167. https://doi.org/10.1038/s41380-019-0468-3.

5. Jonson, C., Levine, K.S., Lake, J., Hertslet, L., Jones, L., Patel, D., Kim, J., Bandres-Ciga, S., Terry, N., Mata, I.F., et al. (2024). Assessing the lack of diversity in genetics research across neurodegenerative diseases: a systematic review of the GWAS Catalog and literature. Preprint at medRxiv. https://doi.org/10.1101/2024.01.08.24301007.

6. Lake, J., Warly Solsberg, C., Kim, J.J., Acosta-Uribe, J., Makarious, M.B., Li, Z., Levine, K., Heutink, P., Alvarado, C.X., Vitale, D., et al. (2023). Multi-ancestry meta-analysis and fine-mapping in Alzheimer's disease. Mol. Psychiatr. *28*, 3121–3132. https://doi.org/10.1038/s41380-023-02089-w.

7. Yang, L.G., March, Z.M., Stephenson, R.A., and Narayan, P.S. (2023). Apolipoprotein E in lipid metabolism and neurodegenerative disease. Trends Endocrinol. Metabol. *34*, 430–445. https://doi.org/10.1016/j.tem.2023.05.002.

8. Neu, S.C., Pa, J., Kukull, W., Beekly, D., Kuzma, A., Gangadharan, P., Wang, L.-S., Romero, K., Arneric, S.P., Redolfi, A., et al. (2017). Apolipoprotein E Genotype and Sex Risk Factors for Alzheimer Disease: A Meta-analysis. JAMA Neurol. *74*, 1178–1189. https://doi.org/10.1001/jamaneurol.2017.2188.

9. Ramos, D.M., Skarnes, W.C., Singleton, A.B., Cookson, M.R., and Ward, M.E. (2021). Tackling neurodegenerative diseases with genomic engineering: a new stem cell initiative from the NIH. Neuron *109*, 1080–1083.

10. Vitale, D., Koretsky, M., Kuznetsov, N., Hong, S., Song, Y., and Levine, K. (2023). GenoTools. Zenodo Version 1.0.2. https://doi.org/10.5281/zenodo.10443258.