Received: 20 November 2023 | Revised: 11 January 2024 | Accepted: 15 January 2024

DOI: 10.1002/alz.13730

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

PERSPECTIVE

Gene replacement-Alzheimer's disease (GR-AD): Modeling the genetics of human dementias in mice

Kellie Benzow[1] | **Kul Karanjeet**[1] | **Adrian L. Oblak**[2] | **Gregory W. Carter**[3] | **Michael Sasner**[3] | **Michael D. Koob**[1]

[1]Laboratory Medicine and Pathology, and Institute for Translational Neuroscience University of Minnesota, Minneapolis, Minnesota, USA

[2]Indiana University School of Medicine, Indianapolis, Indiana, USA

[3]The Jackson Laboratory, Bar Harbor, Maine, USA

Correspondence
Michael D. Koob, Laboratory Medicine and Pathology, and Institute for Translational Neuroscience University of Minnesota, 2101 6th St SE, Minneapolis, MN 55414, USA.
Email: koobx001@umn.edu

Funding information
National Institute on Aging; National Institutes of Health, Grant/Award Numbers: R61/R33NS115089, R56AG067573, RF1AG079125

Abstract

INTRODUCTION: Genetic studies conducted over the past four decades have provided us with a detailed catalog of genes that play critical roles in the etiology of Alzheimer's disease (AD) and related dementias (ADRDs). Despite this progress, as a field we have had only limited success in incorporating this rich complexity of human AD/ADRD genetics findings into our animal models of these diseases. Our primary goal for the gene replacement (GR)-AD project is to develop mouse lines that model the genetics of AD/ADRD as closely as possible.

METHODS: To do this, we are generating mouse lines in which the genes of interest are precisely and completely replaced in the mouse genome by their full human orthologs.

RESULTS: Each model set consists of a control line with a wild-type human allele and variant lines that precisely match the human genomic sequence in the control line except for a high-impact pathogenic mutation or risk variant.

KEYWORDS
APOE-GR, APP-GR, C9ORF72-GR, functional sequence variant, gene-replacement mouse model, MAPT-GR, protective haplotype, PSEN1-GR, PSEN2-GR, risk haplotype, SNCA-GR, TARDBP-GR, TMEM106B-GR, TREM2-GR

1 | INTRODUCTION

The landmark discovery that familial Alzheimer's disease (FAD) can be caused by pathogenic mutations in the *amyloid precursor protein* (*APP*) gene was published in 1991 in the hope that these findings would soon provide the key to identifying linchpin molecular disease mechanisms and effective therapies for AD.[1] This hope has not yet been realized, despite over 30 years of concerted efforts. A critical barrier to progress on these fronts is, in our opinion, the inability of available AD animal models to capture the full complexity of the molecular impacts of these pathogenic mutations. This problem has been highlighted by the repeated failure of therapies developed in available AD mouse models to translate into effective treatments for patients.[2]

2 | TRANSGENIC MOUSE LINES: MODELS OF AD PATHOLOGY BUT NOT AD GENETICS

The standard tools of molecular genetics are still best adapted for working with relatively small DNA fragments. As a result, most of the AD animal models generated by researchers in light of early foundational genetic findings and currently available to the biomedical research community incorporate short, synthetic, and highly simplified cDNA versions of genes expressed from short, exogenous promoter

Kellie Benzow and Kul Karanjeet contributed equally to this work.

fragments. These types of constructs can be powerful tools for testing specific hypotheses and for focusing on defined aspects of a gene's function, but they are not capable of capturing more than a fraction of the subtle complexity of a full human gene. Furthermore, pronuclear injection technologies, which insert constructs in random configurations and locations in the mouse genome, have been the most widely available and, thus, the most widely used means of generating the currently available models. Genomic position effects on transgene expression patterns as well as gene disruption and dysregulation caused by these random insertion events invariably result in dramatic phenotypic variation that overwhelms the true molecular endophenotypes linked to pathogenic mutations in the transgene sequence.

Given that these models differ so dramatically from one another and from AD patients on the genetic level, what criteria have we as a field used to evaluate the clinical relevance of these transgenic lines? An expert advisory panel of academic, industry, and government scientists convened to address this issue recommended in its "best practices" report "choosing models for preclinical studies that exhibit significant and well-characterized pathology relevant to the disease process of interest (that is, amyloid plaques, tau pathology, neuronal loss, oxidative stress/inflammatory changes, and so on)."[2] Perhaps due in part to this recommendation, the rTg4510 mouse model of tauopathy,[3] which exhibits profound premature neurodegeneration and neurofibrillary tangles, became widely used for both basic and preclinical studies. However, our analyses of rTg4510, which was made by random genomic integration of cDNA constructs via pronuclear injection, revealed that the tauopathy phenocopy in rTg4510 requires confounding contributions from (1) an approximately 70-copy TAU_{P301L}-transgene insertion in a 244-kb deletion in *Fgf14*, (2) an approximately seven-copy tTA-transgene insertion in a 508-kb deletion that disrupts another five genes, and (3) high transgene overexpression.[4]

These findings demonstrate the critical need, when making the next generation of AD mouse models, to (1) fully define the genetic underpinnings of the phenotypes found and (2) make precisely matched control lines to distinguish the true impacts of pathogenic mutations from artifacts of the models. Unfortunately, our ability to determine the disease relevance of endophenotypes in otherwise carefully constructed models can also be confounded by the highly artificial nature of cDNA transgenes used to make these lines. The advanced sets of AD models that we describe here therefore meet a third critically important criterion for the most advanced AD models, and (3) mimic the human genetics of ADRDs as closely as is possible in a mouse.

3 | MODELING THE GENETICS OF AD WITH FULL HUMAN LOCI: GENE-REPLACEMENT (GR) MODELS

We are generating mouse lines in which the genes of interest are precisely and completely replaced in the mouse genome by their full human orthologs, along with all known overlapping, co-regulated non-coding RNAs (eg, *MAPT-AS1*). We describe an example of this approach in Figure 1. Each model set consists of a control line with a wild-type (WT) human allele and variant lines that precisely match the human genomic sequence in the control line except for a high-impact pathogenic mutation or risk variant that we specifically introduce. To make these lines, we developed gene-replacement (GR) technologies that allow us to routinely replace mouse genes with their full human orthologs up to several hundred kilobytes in size. Our GR approach allows us to use full human genes to work toward our goal of generating genetic models of AD, rather than models of AD pathology. We only analyze GR-AD lines that are homozygous for the human genes at these loci, avoiding confounding contributions from mouse allele gene products. When possible, we are also enabling genetics studies by generating matched sets of WT and variant lines at each GR-AD allele.

Our goal in designing GR alleles is to include enough of the potential *cis* regulatory elements for the human gene to function within the mouse genome in a manner that matches its function within the human genome as closely as possible. In practical terms, decisions regarding the precise lengths of the mouse genome to remove and of the human genome to insert are made on a locus-by-locus basis and are guided largely by a direct sequence comparison between the two genomes. In the case of the *APOE*-GR allele described in Figure 1, for example, we found significant genomic sequence differences between the mouse and human genomes throughout the entire *APO* gene cluster and decided to replace the full mouse *Apo* gene cluster with the syntenic human *APO* gene cluster. The transition from the mouse genomic sequence to the human genomic sequence in the *APOE*-GR allele was made in a region of non-conserved sequences just 3′ of the conserved sequences at the end of the flanking *TOMM40/TOMM40* genes. Similarly, the transition from the human genomic sequence to the mouse genome on the other end of the GR allele occurs in a non-conserved region just 3′ of the conserved sequence match at the end of the *APOC2/Apoc2* genes. The goal in each case is to replace the mouse allele with the human allele without disrupting any of the flanking sequences conserved between the two genomes.

Although our knowledge of most *cis* regulatory elements is still limited, by precisely replacing mouse homologs with their human counterparts we embed these genes within a larger genomic context in which the most important distal regulatory elements and overall chromatin structural features will be conserved between the two genomes. For every GR allele that we have studied to date, we have found that the RNA and protein products of these human genes are all expressed in the mouse with the temporal, spatial, and subcellular distribution patterns expected for these human genes.

These precisely matched sets of animal models allow the research community to evaluate the molecular impact of pathogenic mutations within the context of the human genomic sequence in which they occur in patients. The GR-AD mouse lines also contain a full array of potential human therapeutic targets ranging from the full genomic DNA sequences to the multiple RNA transcription isoforms and protein products that they encode. Furthermore, because the genomic

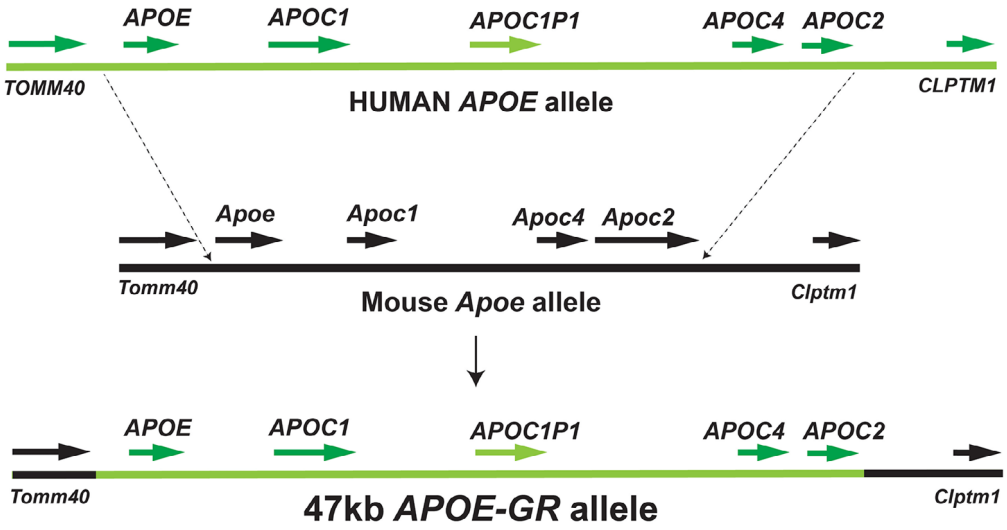

FIGURE 1 Example of the gene-replacement (GR) approach to generating genetic models of human disease in the mouse: *APOE*-GR alleles. For these lines, the mouse genomic region (30 kb) encoding the mouse *Apoe* gene cluster (*Apoe*, *Apoc1*, *Apoc4*, and *Apoc2*) is completely and precisely replaced by the 47-kb syntenic region of the human genome, as shown. For these GR alleles, the transition from the mouse to the human sequence occurs within a non-conserved region just 3′ of the mouse *Tomm40* gene, and the transition from the human to mouse sequence occurs within a non-conserved region just 3′ of the human *APOC2* gene. The 47 kb of the *APOE*-GR allele itself is a fully human *APOE* ε4 haplotype sequence. A matched line with a single nucleotide change introducing the *APOE* ε3 coding change has also been generated in this *APOC*-GR model set.

sequences of these matched sets differ only at sequences specifically changed in each line, any significant molecular differences between matched lines can confidently be attributed to the pathogenic mutation or risk variant in the experimental lines, and any therapeutic agents found to effectively correct these dysfunctions could be expected to have direct therapeutic value to patients.

4 | THE FIRST TEN GR-AD HUMAN GENETIC LOCI IN MICE

Although we will describe GR-AD model sets in detail as we complete their initial characterization studies, here we are providing a brief overview of the first 10 human GR alleles that we have introduced into the mouse genome. In general, these human alleles fall into one of three categories: (1) genes associated with the major AD/ADRD-relevant proteinopathies, (2) genes in which high-impact rare variants cause inherited forms of ADRD, and (3) genes identified as modifiers of the risk of developing dementia. Many but not all these model sets consist of both WT control lines and matched lines with pathogenic/risk variants, for a current total of 27 GR-AD alleles. These mouse lines are also summarized in Table 1.

4.1 | GR alleles associated with AD/ADRD-relevant proteinopathies

We generated GR mouse models that incorporate the full human genes primarily associated with each of the four major AD/ADRD-relevant proteinopathies: (1) **APP/amyloid beta (Aβ)**: *APP* (344 kb of human genomic sequence), *PSEN1* (106 kb), and *PSEN2* (133 kb); (2) **TAU**: *MAPT* (190 kb); (3) **TDP-43/TMEM106B**: *TARDBP* (65 kb) and *TMEM106B* (97 kb); and (4) **α-Synuclein**: *SNCA* (158 kb).

4.1.1 | APP/β-amyloid

Linkage studies of FAD identified pathogenic mutations in only three genes: *APP*,[1] *Presenilin 1* (*PSEN1*),[5] and *Presenilin 2* (*PSEN2*).[5,6] The *APP*-GR allele includes the human genomic region transcribed by the longest APP transcripts (291 kb) as well as the long non-coding *APP*-AS transcription region (47 kb) encompassing the *APP* promoter region. At nearly 350 kb of human genome sequence, the *APP*-GR allele is the largest GR we have made to date. We have not yet generated any matched lines with pathogenic FAD mutations for this WT *APP*-GR allele but are combining this model with other GR models, as described in what follows.

The *PSEN1*-GR alleles consist of a WT control and three variants: G209V,[7] A246E,[5] and V272A.[8] Carriers of the first of these two sequences variants have ages of onset in their fourth or fifth decade, and *post mortem* analysis showed that their neuropathology was typical of an accelerated AD. The V272A variant is associated with a more aggressive disease with an earlier onset, and some carriers exhibited significant Lewy body (α-Synuclein) pathology. We are currently breeding the first generation of a fourth *PSEN1*-GR variant line, E280A,[9] the so-called "Paisa" variant found in an extremely large Colombian kindred.

The *PSEN2*-GR model set consists of a WT allele and a single-variant allele: N141I.[5,6] This variant is found in a large and well-characterized *PSEN2* FAD kindred, with a mean age of onset in the fifth decade

TABLE 1 Summary of GR-AD alleles currently in mice. The gene–replacement (GR) size is that of the human genomic sequence at each allele. The genomic sequence ranges are for the human assembly GRCh38.p14 and the mouse assembly GRCm39.

	Model name	GR size	Human genome region in mouse	Mouse genome region replaced
1	***MAPT*-GR** wt(H2.1)	190 kb	Chr17: 45838648–46029144	Chr11:104066949–104223491
2	***MAPT*-GR** wt(H1)	190 kb	Chr17: 45838648–46029144	Chr11:104066949–104223491
3	***MAPT*-GR***P301L	190 kb	Chr17: 45838648–46029144	Chr11:104066949–104223491
4	***MAPT*-GR***N279K	190 kb	Chr17: 45838648–46029144	Chr11:104066949–104223491
5	***MAPT*-GR***10IVS+16T	190 kb	Chr17: 45838648–46029144	Chr11:104066949–104223491
6	***MAPT*-GR***L266V	190 kb	Chr17: 45838648–46029144	Chr11:104066949–104223491
7	***MAPT*-GR***G272V	190 kb	Chr17: 45838648–46029144	Chr11:104066949–104223491
8	***APP*-GR** wt	344 kb	Chr21:25880264–26224036	Chr16:84750107–85019011
9	***PSEN1*-GR** wt	106 kb	Chr14:73121400–73227703	Chr12:83730110–83793335
10	***PSEN1*-GR** *V272A	106 kb	Chr14:73121400–73227703	Chr12:83730110–83793335
11	***PSEN1*-GR** *G209V	106 kb	Chr14:73121400–73227703	Chr12:83730110–83793335
12	***PSEN1*-GR** *A246E	106 kb	Chr14:73121400–73227703	Chr12:83730110–83793335
13	***PSEN1*-GR** *E280A	106 kb	Chr14:73121400–73227703	Chr12:83730110–83793335
14	***PSEN2*-GR** wt	113 kb	Chr1:226823270–226935953	Chr1:180028072–180100006
15	***PSEN2*-GR** *N141I	113 kb	Chr1:226823270–226935953	Chr1:180028072–180100006
16	***APOE*-GR***ε4	47 kb	Chr19:44904076–44951132	Chr7: 19404448–19434653
17	***APOE*-GR***ε3	47 kb	Chr19:44904076–44951132	Chr7: 19404448–19434653
18	***SNCA*-GR** wt	158 kb	Chr4:89707185–89865027	Chr6:60683985–60848702
19	***SNCA*-GR** *A53T	158 kb	Chr4:89707185–89865027	Chr6:60683985–60848702
20	***SNCA*-GR** *E46K	158 kb	Chr4:89707185–89865027	Chr6:60683985–60848702
21	***TARDBP*-GR** wt	65 kb	Chr1: 10989074–11054422	Chr4:148679229–148723697
22	***TMEM106B*-GR** wt(r)	97 kb	Chr7:1218525312282003	Chr6: 13033006–13107770
23	***TMEM106B*-GR** wt(p)	97 kb	Chr7:1218525312282003	Chr6: 13033006–13107770
24	***TREM2*-GR** wt	280 kb	Chr6:41099491–41379205	Chr17:48404707–48693648
25	***TREM2*-GR***R47H	280 kb	Chr6:41099491–41379205	Chr17:48404707–48693648
26	***TREM2*-GR** rs9357347-C	280 kb	Chr6:41099491–41379205	Chr17:48404707–48693648
27	***C9ORF72*-GR** wt	103 kb	Chr9:27515238–27618526	Chr4:35145138–35274047

and neuropathologies typical of AD.[10] We have not yet bred mice to be homozygous for *PSEN2*-GR alleles, and subsequent quality control (QC)-level validations of these alleles are pending.

4.1.2 | TAU

Our *microtubule associated protein tau-gene replacement* (*MAPT*-GR) matched set of lines currently consists of seven alleles total: two WT control lines (H2 or H1 haplotype)[11] and lines with one of five pathogenic variants in the H1 haplotype. Over 60% of frontotemporal dementia (FTD) tauopathy cases are caused by three of these mutations: the P301L and N279K mutations in the second microtubule-binding domain encoded in exon 10 of *MAPT* and the C to T mutation at the 16th base pair of the intron after Exon 10 (denoted "10IVS+16 C > T"). The other two mutations in the *MAPT*-GR series, L266V[12] and G272V,[13] are more closely linked to Pick's disease, an ADRD characterized by the pathological presence of Pick's bodies.

4.1.3 | TDP-43/TMEM106B

AD and other dementias frequently involve TAR DNA binding protein of 43 kDa (TDP-43) pathology.[14] Our *TAR DNA binding protein*-gene replacment (*TARDBP*-GR) allele currently consists of a WT control allele without matching variant allele. The *TMEM106B* locus was initially identified as an allele that plays a significant role in modifying the risk of developing frontotemporal lobar degeneration (FTLD) with TDP-43 inclusions (FTLD-TDP).[15] Recently, the amyloid fibrils in this disease were found to be formed primarily of TMEM106B protein fragments and not TDP-43.[16–18] One of the two *TMEM106B*-GR alleles incorporates a risk haplotype and the other a protective haplotype,[15] as described in more detail in Rodney et al. 2023.[19] We have not

yet bred mice to be homozygous for *TMEM106B*-GR alleles, and subsequent QC-level validations of these alleles are pending.

4.1.4 | α-Synuclein

Accumulations of α-Synuclein (encoded in the *SNCA* gene) are hallmarks of dementia with Lewy bodies (DLB), Parkinson's disease (PD), Parkinson's disease dementia (PDD), and multiple system atrophy (MSA) but are also often found in AD patients.[20] The *SNCA*-GR model set currently consists of a WT *SNCA*-GR control allele and two matched alleles with pathogenic variants (A53T[21] and E46K[22]).

4.2 | GR alleles associated with other inherited forms of ADRD

4.2.1 | *C9ORF72*-GR (109 kb)

This is currently the only GR allele associated with an inherited ADRD that is also not primarily responsible for a corresponding proteinopathy.[23,24] In patients, hexanucleotide expansions of between 700 and 1600 repeats were found in affected family members with FTD/amyotrophic lateral sclerosis.[24] Generating this size of repeat expansion in the germline of mouse is technically challenging. A matched variant line with an expansion of the hexanucleotide repeat in this gene is in progress but not yet complete.

4.3 | GR alleles that modify the risk of developing dementia

4.3.1 | *APOE*-GR (47 kb five-gene cluster)

The primary AD risk gene encoded in this gene cluster replacement allele encodes the full *apolipoprotein E* (*APOE*) regulatory and transcribed region encoding all known *APOE* transcript isoforms. This GR allele gene cluster also encodes the *APOC1*, *APOC2*, and *APOC4* genes, as well as the *APOC1P1* pseudogene 1. **Two lines**: (*APOE* ε4, ε3 variant on the same ε4 haplotype background).

4.3.2 | *TREM2*-GR (280-kb seven-gene cluster)

The primary AD risk gene encoded in this gene-cluster replacement allele encodes the *triggering receptor expressed on myeloid cells 2* (*TREM2*) gene. This GR allele gene cluster also encodes the human genes *TREM1*, *TREML1*, *TREML2*, *TREML4*, and *NCR2*, as well as several pseudogenes and non-coding RNAs. This region has diverged significantly from the syntenic mouse region that it replaces (288 kb of mouse genome replaced). **Three lines**: (WT, R47H risk variant in *TREM2*, and rs9357347-C non-coding protective variant[25,26]). We have not yet bred mice to be homozygous for *TREM2*-GR alleles, and subsequent QC-level validations of these alleles are pending.

5 | MOUSE LINES WITH MULTIPLE HUMAN GR-AD ALLELES

Many of the human protein products of these GR-AD alleles are known to interact with one another (eg, APP protein is processed by PSEN1 or PSEN2). A mouse that carries these full sets of multiple human genes is likely to model this human biology more accurately than those that carry a single human gene (eg, human *APP* in a line with mouse *Psen1*). We have therefore generated a set of 3xGR-AD lines that are homozygous for three human genes: *PSEN1*-GR, *APP*-GR, and *MAPT*(H2.1)-GR. We have dubbed this 3xGR-AD line "$P_1A\tau$" (*PSEN1*, *APP*, and tau). Note that in this nomenclature, the "3×" refers to the number of human GR alleles in the line and does not indicate the presence or absence of mutations or variants in these genes. The 3×GR-AD set of "$P_1A\tau$" lines currently consists of a control line with three WT alleles ("$P_1A\tau$") and three variant lines that carry the G209V, V272A, or A246E variant in the *PSEN1* allele ($P_1{}^{G209V}$ Aτ, $P_1{}^{V272A}$ Aτ, or $P_1{}^{A246E}$ Aτ), along with WT *APP* and *MAPT*. These types of combined GR lines allow us to evaluate the impacts of the *PSEN1* pathogenic variants, not only on the human PSEN1 protein but also on the dynamics of the WT human APP and MAPT proteins. A 5xGR-AD that also incorporates *PSEN2*-GR and *APOE4*-GR is in progress. Similarly, we are also working toward breeding mouse lines with the genetic potential to develop all four of the major AD/ADRD-relevant proteinopathies. These lines will be homozygous for at least seven of the human GR alleles described here (*PSEN1*, *PSEN2*, *APP*, *MAPT*, *SNCA*, *TARDBP*, and *TMEM106B*).

6 | VALIDATION OF GR MOUSE MODELS BY MODEL-AD PROGRAM

The goals of the MODEL-AD Consortium (model-ad.org) are to develop novel, translatable animal models of late-onset AD (LOAD) and to use clinically relevant assays to align models to corresponding patient populations and stages of disease, establish a rigorous preclinical testing paradigm using these models, and ensure widespread availability of models and data. All models are based on human genetics, as we use knock-in or GR approaches to introduce risk variants associated with human LOAD into the mouse genome. The majority of this genetic risk likely lies in non-coding regions associated with genes expressed in microglia.[27] The GR approach described here will enable the MODEL-AD program to include variants expressed in non-conserved regulatory regions (eg, *APOE*, *TMEM106B*, and *INPP5D* regulatory variants) and in entire loci that are not conserved between human and mouse (eg, *CR1*, *MS4A*). To increase the utility of the GR-AD models to the research community, each set of models will be characterized using the existing MODEL-AD phenotyping pipeline, which has been designed to focus on key measures that can be used for preclinical testing.[28]

This pipeline consists of a staged process with go/no go gates to prioritize models for further aging studies within MODEL-AD. Critically, extensive aging is included in the primary screen step, such that age-dependent phenotypes can be detected. The initial phase includes frailty analysis, fluid biomarkers (neurofilament light chain, glial fibrillary acidic protein), neuropathology (amyloid and tau pathology, analysis of microglia and astrocytes, as well as quantification of neuronal number), and multiomics including proteomics and transcriptomics at 4 and 18 months of age. Those models deemed most clinically relevant advance to a second phase that includes additional ages to define potential therapeutic windows and more clinically relevant phenotyping modalities such as in vivo imaging (eg,magnetic resonance imaging for volumetric analysis, positron emission tomography/computed tomography). These GR models will also serve as platform models for relatively rapid CRISPR-based modeling of additional risk variants. For example, we plan to engineer additional TAU risk variants into the *MAPT*(H1.0)-GR model described here and to use the humanized *APOE*-GR model to compare the effects of non-coding *APOE* risk variants found in different human populations.

7 | INCORPORATION OF GR MODELS INTO MODEL-AD

The GR models shown to be most promising (as determined by AD-relevant differences between the risk and control models) will be incorporated into complex models expressing complementary variants on the background of a "base model" consisting of alleles thought to be necessary but not sufficient to drive LOAD-like phenotypes. Currently, the favored base model expresses the gene replacement of the WT *MAPT* (TAU) of the H1 haplotype and knock-ins to humanize APOE ε4 and the three amino acids that vary between human and mouse in the Aβ1-42 region in the mouse *App* locus ("LOAD3," JAX#37778). One long-term goal of the MODEL-AD program is to develop and define models that represent the heterogeneity of the disease. For example, only about a quarter of AD patients in the ROSMAP study had pure Aβ and TAU pathology, with the rest also having at least one of TDP-43 pathology, Lewy bodies, hippocampal sclerosis, or cerebrovascular compromise. Presumably, adding *SNCA-GR*, *TARDBP-GR*, and/or *TMEM106B-GR* loci with disease risk variants would drive some of these other pathologies and thus create novel models of LOAD subtypes.

8 | ACCELERATED RELEASE OF GR-AD LINES WITHOUT RESTRICTIONS

GR-AD models are being made available through the Jackson Laboratory mouse repository as soon as is practical for academic and commercial use without restrictions. In general, molecular characterization (validation of genomic insert and appropriate expression patterns) will be completed, but extensive aging or characterization of phenotypes will not be required before making new models widely available. Currently five of the *MAPT*-GR are available: *MAPT*(H2.1)-GR (JAX #33668), *MAPT*(H1.0)-GR (JAX #35398), *MAPT*(H1.0*N279K)-GR (JAX #35794), *MAPT*(H1.0*10IVS+16 C > T)-GR (JAX #36664), and *MAPT*(H1.0*P301L)-GR (JAX #37420). An additional nine GR-AD lines are currently in the distribution pipeline: *APOE*-GR ε4 (JAX #38919), *APOE*-GR ε3 (JAX #38918), *APP*-GR (JAX #38920), *SNCA*-GR (JAX #39166), *SNCA**A53T-GR (JAX #39167), *SNCA**E46K-GR (JAX #39168), *TARDBP*-GR (JAX #39101), *PSEN1*-GR (JAX #39177), and 3xGR-AD($P_1A\tau$) (JAX #39175)

9 | DISCUSSION

The foundational line for each GR-AD model set carries a WT human gene or gene cluster that precisely replaces the orthologous gene(s) in the mouse genome. Although these mouse orthologs are similar to their human counterparts, human and mouse genes are typically significantly different in their genomic DNA sequence. These DNA sequence differences can cause differences in gene expression patterns, in the ratios and content of the transcription splicing isoforms generated, and in intrinsic protein properties, any or all of which might prove to play significant roles in the etiology and progression of ADRD-relevant disease processes.

To enable genetic studies, most of the GR-AD model sets introduced here also consist of mouse lines that carry a GR-AD allele with a functional sequence variant that is known either (1) to be pathogenic or (2) to modify the risk of developing dementia. In most cases, these sequence variants are single nucleotide changes in the WT GR-AD sequence, but in two instances to date we have introduced full alternative haplotypes as variant GR-AD alleles. The two major families of haplotypes of the human *MAPT* gene are designated H1 and H2[11] and diverged from each other as the result of an ancient genomic inversion event that inhibited subsequent interchromosomal recombinations between these inverted genomic segments.[29] The genomic sequence of the *MAPT* (H2.1)-GR allele is primarily that of an H2 *MAPT* allele, but this GR allele does not involve the much larger genomic inversion found in the human genome. In the case of the *TMEM106B*-GR alleles, we generated alternative GR alleles consisting of either the protective or risk haplotype because the functional sequence differences between these haplotypes have not yet been fully defined.

An underlying tenet of molecular genetics is the idea that functional sequence variants in a gene result in significant changes to the properties of the gene and its products that in turn result in quantifiable observed phenotypes. The long-term functional consequences of the sequence variants introduced into the GR-AD alleles have been firmly established in dementia patients. However, the precise molecular events directly altered by these sequence changes and their role in the initiation and progression of AD/ADRD-relevant dysfunction and degeneration have not yet been identified. The GR-AD model sets introduced here are intended to provide an experimental system in which the molecular impacts of these sequence changes within their human genomic sequence context can be fully explored by careful

comparisons between animals from the WT control and matched experimental variant lines.

We do not have a priori knowledge of the full range of the molecular impacts of these variants and expect the GR-AD systems to provide insights into disease mechanisms that have not been evident in classic transgenic models. The initial characterization studies outlined here are those currently conducted for new lines evaluated by the MODEL-AD consortium and are not meant to be a comprehensive list of the types of evaluations that should performed on the forthcoming GR-AD lines. By making these lines freely available to all researchers on an expedited basis and without restriction, our intention is that they will enable experiments encompassing the full range of expertise and interests of the ADRD research community.

As with any mammalian genetic system, GR-AD mice can be either homozygous for the variant human GR allele or in a heterozygous mix with the WT human allele. The functional variants introduced in GR alleles are typically linked to dominant traits, with dementia patients most often heterozygous for these often rare variants. Mice with heterozygous GR-AD alleles are therefore typically a better genetic match to dementia patients, whereas homozygous GR animals can offer advantages for breeding schemes and in some instances may accelerate or accentuate the impacts of the variant. The impacts of homozygous mutations may have significant differences with those of heterozygous mutations, however, and so the ploidy of GR-AD alleles should be carefully considered in experimental designs.

The impact of many AD/ADRD-relevant sequence variants can often be better assessed in the context of multiple human genes, and the GR-AD system offers the flexibility to combine several GR-AD alleles in the same animal. For example, we are currently evaluating *PSEN1*-GR variants in animals that also carry WT *APP*-GR and *MAPT*-GR human alleles (ie, 3xGR-AD "PAτ" mice). Brain pathologies that had been thought to primarily involve TDP-43 (the protein product of the *TARDBP* gene) were recently found to be composed of TMEM106B fragments.[16–18] We will therefore evaluate *TMEM106B*-GR variants in lines that also carry the human *TARDBP*-GR allele. Similarly, the impacts of variants that modify the risk of developing dementia (eg, *APOE4*-GR) will typically be evaluated in animals that also carry GR-AD alleles associated with one or more ADRD-relevant proteinopathies.

Selection of these first 10 GR-AD alleles briefly described here was based on strong genetic links between these genes and AD/ADRDs and, in many cases, a direct involvement in proteinopathies characteristic of these diseases. These initial GR-AD alleles are not yet a comprehensive or balanced representation of the underlying genetic causes of human dementia but present only the first of many steps in this process. The bias in this first set of GR-AD alleles toward AD/ADRD protein pathologies is currently being balanced by ongoing projects generating additional GR-AD human alleles linked to other key aspects of AD/ADRD pathophysiology. We actively encourage input in these choices going forward as we develop this experimental resource for the AD/ADRD research community.

ACKNOWLEDGMENTS

The work reported here was supported by the National Institute on Aging and the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under awards R61/R33NS115089, R56AG067573, and RF1AG079125.

CONFLICT OF INTEREST STATEMENT

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

REFERENCES

1. Goate A, Chartier-Harlin M-C, Mullan M, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimers-disease. *Nature*. 1991;349(6311):704-706.
2. Shineman DW, Basi GS, Bizon JL, et al. Accelerating drug discovery for Alzheimer's disease: best practices for preclinical animal studies. *Alzheimers Res Ther*. 2011;3(5):28.
3. Ramsden M, Kotilinek L, Forster C, et al. Age-dependent neurofibrillary tangle formation, neuron loss, and memory impairment in a mouse model of human tauopathy (P301L). *J Neurosci*. 2005;25(46):10637-10647.
4. Gamache J, Benzow K, Forster C, et al. Factors other than hTau overexpression that contribute to tauopathy-like phenotype in rTg4510 mice. *Nat Commun*. 2019;10(1):2479.
5. Rogaev EI, Sherrington R, Rogaeva EA, et al. Familial Alzheimers-disease in kindreds with missense mutations in a gene on chromosome-1 related to the Alzheimers-disease type-3 gene. *Nature*. 1995;376(6543):775-778.
6. Levy-Lahad E, Wasco W, Poorkaj P, et al. Candidate gene for the chromosome-1 familial Alzheimers-disease locus. *Science*. 1995;269(5226):973-977.
7. Poorkaj P, Sharma V, Anderson L, et al. Missense mutations in the chromosome 14 familial Alzheimer's disease presenilin 1 gene. *Hum Mutat*. 1998;11(3):216-221.
8. Jimenez-Escrig A, Rabano A, Guerrero C, et al. New V272A presenilin 1 mutation with very early onset subcortical dementia and parkinsonism. *Eur J Neurol*. 2004;11(10):663-669.
9. Lopera F, Ardilla A, Martínez A, et al. Clinical features of early-onset Alzheimer disease in a large kindred with an E280A presenilin-1 mutation. *JAMA*. 1997;277(10):793-799.
10. Jayadev S, Leverenz JB, Steinbart E, et al. Alzheimer's disease phenotypes and genotypes associated with mutations in presenilin 2. *Brain*. 2010;133:1143-1154.
11. Baker M, Litvan I, Houlden H, et al. Association of an extended haplotype in the tau gene with progressive supranuclear palsy. *Hum Mol Genet*. 1999;8(4):711-715.
12. Hogg M, Grujic ZM, Baker M, et al. The L266V tau mutation is associated with frontotemporal dementia and Pick-like 3R and 4R tauopathy. *Acta Neuropathol*. 2003;106(4):323-336.
13. Groen JJ, Endtz LJ. Hereditary Pick's disease: second re-examination of the large family and discussion of other hereditary cases, with particular reference to electroencephalography, a computerized tomography. *Brain*. 1982;105(Pt 3):443-459.
14. Wilson RS, Yu L, Trojanowski JQ, et al. TDP-43 pathology, cognitive decline, and dementia in old age. *JAMA Neurol*. 2013;70(11):1418-1424.
15. Van Deerlin VM, Sleiman PMA, Martinez-Lage M, et al. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet*. 2010;42(3):234-239.
16. Chang A, Xiang X, Wang J, et al. Homotypic fibrillization of TMEM106B across diverse neurodegenerative diseases. *Cell*. 2022;185(8):1346-1355. e15.

17. Jiang YX, Cao Q, Sawaya MR, et al. Amyloid fibrils in FTLD-TDP are composed of TMEM106B and not TDP-43. *Nature*. 2022;605(7909):304-309.
18. Schweighauser M, Arseni D, Bacioglu M, et al. Age-dependent formation of TMEM106B amyloid filaments in human brains. *Nature*. 2022;605(7909):310-314.
19. Rodney A, Karanjeet K, Benzow K, Koob MD. A common Alu element insertion in the 3'UTR of TMEM106B is associated with an increased risk of dementia. *BioRXIV preprint*. 2023. doi:10.1101/2023.09.07.555370
20. Kotzbauer PT, Tojanowski JQ, Lee VMV. Lewy body pathology in Alzheimer's disease. *J Mol Neurosci*. 2001;17(2):225-232.
21. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science*. 1997;276(5321):2045-2047.
22. Zarranz JJ, Alegre J, Gómez-Esteban JC, et al. The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia. *Ann Neurol*. 2004;55(2):164-173.
23. Dejesus-Hernandez M, Mackenzie IR, Boeve BF, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*. 2011;72(2):245-256.
24. Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*. 2011;72(2):257-268.
25. Carrasquillo MM, Allen M, Burgess JD, et al. A candidate regulatory variant at the TREM gene cluster associates with decreased Alzheimer's disease risk and increased TREML1 and TREM2 brain gene expression. *Alzheimers Dement*. 2017;13(6):663-673.
26. Tian M-L, Ni X-N, Li J-Q, et al. A candidate regulatory variant at the TREM gene cluster confer Alzheimer's disease risk by modulating both amyloid-beta pathology and neuronal degeneration. *Front Neurosci*. 2019;13:742.
27. Romero-Molina C, Garretti F, Andrews SJ, Marcora E, Goate AM. Microglial efferocytosis: diving into the Alzheimer's disease gene pool. *Neuron*. 2022;110(21):3513-3533.
28. Oblak AL, Lin PB, Kotredes KP, et al. Comprehensive evaluation of the 5XFAD mouse model for preclinical testing applications: a MODEL-AD study. *Front Aging Neurosci*. 2021;13:713726.
29. Stefansson H, Helgason A, Thorleifsson G, et al. A common inversion under selection in Europeans. *Nat Genet*. 2005;37(2):129-137.

SUPPORTING INFORMATION

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

How to cite this article: Benzow K, Karanjeet K, Oblak AL, Carter GW, Sasner M, Koob MD. Gene replacement-Alzheimer's disease (GR-AD): Modeling the genetics of human dementias in mice. *Alzheimer's Dement*. 2024;20:3080–3087. https://doi.org/10.1002/alz.13730