Brain transcriptomes of zebrafish and mouse Alzheimer's disease knock-in models imply early disrupted energy metabolism

Karissa Barthelson[1*], Morgan Newman[1], Michael Lardelli[1]

[1]Alzheimer's Disease Genetics Laboratory, School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, SA 5005, Australia

***** Corresponding author

Email: karissa.barthelson@adelaide.edu.au

ORCID: https://orcid.org/0000-0002-4693-8833

Key words: Alzheimer's disease, zebrafish, mouse, RNA-seq, oxidative phosphorylation, brain

Abstract

Energy production is the most fundamentally important cellular activity supporting all other functions, particularly in highly active organs such as brains, Here we summarise transcriptome analyses of young adult (pre-disease) brains from a collection of eleven early-onset familial Alzheimer's disease (EOfAD)-like and non-EOfAD-like mutations in three zebrafish genes. The one cellular activity consistently predicted as affected by only the EOfAD-like mutations is oxidative phosphorylation that produces most of the brain's energy. All the mutations were predicted to affect protein synthesis. We extended our analysis to knock-in mouse models of *APOE* alleles and found the same effect for the late onset Alzheimer's disease risk allele ε4. Our results support a common molecular basis for initiation of the pathological processes leading to both early and late onset forms of Alzheimer's

disease and illustrate the utility of zebrafish and of knock-in, single EOfAD mutation models for understanding the causes of this disease.

Summary statement

Young adult zebrafish mutants and a mouse model of a genetic variant promoting early- and late-onset Alzheimer's disease respectively share changes in brain gene expression indicating disturbance of oxidative phosphorylation.

Introduction

Alzheimer's disease (AD) is a complex and highly heterogenous neurodegenerative disease, defined by the presence of intracellular neurofibrillary tangles (NFTs, primarily consisting of hyperphosphorylated tau proteins), and extracellular plaques mostly consisting of a small peptide, amyloid β (Aβ) (Jack et al., 2018). The pathological basis of AD has been the subject of research for over 100 years (Alzheimer, 1906). Nevertheless, most treatments tested in clinical trials have shown limited therapeutic benefit.

AD has a strong genetic basis (reviewed in (Sims et al., 2020)). In some rare cases, early-onset familial forms of AD (EOfAD, occurring before 65 years of age) arise due to dominant mutations in one of four genes: *PRESENILIN 1* (*PSEN1*), *PRESENILIN 2* (*PSEN2*), *AMYLOID β PRECURSOR PROTEIN* (*APP*), and *SORTILIN-RELATED RECEPTOR 1* (*SORL1*) (reviewed in (Barthelson et al., 2020a; Bertram and Tanzi, 2012; Temitope et al., 2021)). However, most AD cases are sporadic, showing symptom onset after the arbitrarily defined threshold of 65 years (late-onset sporadic AD, LOAD). Genetic variants at many loci have been associated with increased risk of LOAD (Jansen et al., 2019; Kunkle et al., 2019; Lambert et al., 2013). The most potent variant is the ε4 allele of

APOLIPOPROTEIN E (*APOE*) (Farrer et al., 1997) that has been described as “semi-dominant” (Genin et al., 2011).

An understanding of the early cellular stresses on the brain that eventually lead to AD is necessary to advance the development of preventative treatments. This is difficult to achieve through studying living humans, as EOfAD mutations are rare, and access to young, pre-symptomatic brains is limited. Nevertheless, imaging studies have implicated structural and functional changes to the brain long before diagnosis of AD (Iturria-Medina et al., 2016; Quiroz et al., 2015). Brain imaging cannot provide detailed molecular information on these changes. Transcriptome analysis is, currently, the strategy that can provide the highest resolution molecular description of cells and tissues. However, transcriptome analyses of ante-mortem brains carrying EOfAD mutations can only be performed using brain tissue from animal models.

Our group has exploited the zebrafish to generate a collection of knock-in models of EOfAD-like mutations in order to analyse their young brain transcriptomes (Barthelson et al., 2021a; Barthelson et al., 2021b; Barthelson et al., 2020b; Barthelson et al., 2021c; Dong et al., 2021; Hin et al., 2020a; Hin et al., 2020b; Jiang et al., 2020; Newman et al., 2019). Our experimental philosophy has been to replicate, as closely as possible, the single heterozygous mutation state of EOfAD in humans, thereby avoiding possibly misleading assumptions regarding the molecular mechanism(s) underlying the disease. Our overall goal has been to compare a broad-range of EOfAD-like mutations in a number of EOfAD genes to define their shared pathological effects in young adult brains where the long progression to AD begins. To assist in this definition (by exclusion), we also created non-EOfAD-like mutations in the same genes as negative controls, i.e. frameshift mutations in the presenilin genes that do not cause EOfAD (reviewed in (Jayne et al., 2016), the “reading frame-preservation rule”). The presentation of EOfAD and LOAD as a similar diseases (reviewed in (Blennow et al., 2006; Masters et al., 2015)) implies similarity, to some degree, at the cellular and molecular levels. Therefore, despite differences in the genetic variants that promote these two diseases, understanding the molecular effects of heterozygosity for EOfAD mutations may give insight into molecular changes underpinning LOAD.

Here, we summarise our findings of brain transcriptome analyses of EOfAD-like mutations in the zebrafish orthologues of genes implicated in EOfAD: *psen1, psen2* and *sorl1*. EOfAD mutations also exist in *APP*. However, zebrafish express two *APP* "co-orthologous" genes, *appa* and *appb*, complicating analysis of single, heterozygous mutations. Therefore, we re-analysed the best available publicly accessible brain transcriptomic data from a knock-in model of *APP* mutations: the $App^{NL\text{-}G\text{-}F}$ mouse. Finally, we compared whether the brain transcriptome changes occurring due to single heterozygous EOfAD-like mutations in zebrafish are similar to the changes occurring due to the strongest genetic risk factor for LOAD: the ε4 allele of *APOE*, using publicly available brain transcriptome data from a humanised *APOE* targeted-replacement mouse model (APOE-TR) (Sullivan et al., 1997). We identify changes to energy metabolism as the earliest detectable cellular stress due to AD mutations, and demonstrate that knock-in zebrafish models are valuable tools to study the earliest molecular pathological events in this disease.

Results

Transcriptome analysis of zebrafish models of EOfAD

We first collated our findings from our zebrafish models of EOfAD-like mutations in *psen1* (Barthelson et al., 2021a; Hin et al., 2020a; Hin et al., 2020b; Newman et al., 2019), *psen2* (Barthelson et al., 2021b) and *sorl1* (Barthelson et al., 2020b; Barthelson et al., 2021c). An advantage of using zebrafish for RNA-seq analyses is minimisation of genetic and environmental noise through breeding strategies such as that shown in **Fig. 1A**. Large families of synchronous siblings can consist of heterozygous mutant and wild type genotypes allowing direct comparisons of the effects of each mutation. So far, we have performed six brain transcriptomic analyses based on various breeding strategies (summarised in **Table 1** and **Fig. S1-S6**). The detailed analyses can be found in the publications cited above. However, the outcomes are summarised below and in **Fig. 1**.

In our previously published analyses, we found that heterozygosity for most of our EOfAD-like mutations does not result in many differentially expressed (DE) genes in young adult brains (Barthelson et al., 2021a; Barthelson et al., 2021b; Barthelson et al., 2020b; Barthelson et al., 2021c; Newman et al., 2019) (as would be expected for modelling a disease that becomes overt in middle

age). Therefore, we performed gene set enrichment analyses to predict which cellular processes were affected by each of the mutations in each experiment. We used the *KEGG* (Kanehisa and Goto, 2000) gene sets to determine whether changes to gene expression were observed in any of 186 biological pathways/processes. Additionally, we recently proposed that neuronal iron dyshomeostasis may be an effect-in-common of EOfAD mutations in the context of AD pathogenesis (Lumsden et al., 2018). Therefore, we used our recently defined iron responsive element (*IRE*) gene sets (HIN ET AL., 2020B) to test for evidence of iron dyshomeostasis. Biological processes found to be affected in at least two different zebrafish mutants are shown in **Fig. 1B** (the statistical significance of all *KEGG* and *IRE* gene sets in each mutant can be found in **S2 file**). The one gene set consistently altered by all of the EOfAD-like mutations, but not by the non-EOfAD-like mutations examined, is the *KEGG_OXIDATIVE_PHOSPHORYLATION* gene set (**Fig. 1C**), supporting that changes to mitochondrial function are an early cellular stress in EOfAD. The *KEGG_OXIDATIVE_PHOSPHORYLATION* gene set is also affected by heterozygosity for the K97fs mutation of *psen1*. K97fs is a frameshift mutation and so does not follow the "reading-frame preservation rule" (Jayne et al., 2016) of presenilin EOfAD mutations. However, the truncated protein encoded by K97fs resembles an hypoxia-induced isoform of human *PSEN2*, denoted PS2V, which shows increased expression in LOAD brains, (Sato et al., 1999) (and see (Moussavi Nik et al., 2015) for additional explanation). Therefore, K97fs is still an AD-relevant mutation.

Genes encoding the components of ribosomal subunits, as defined by the gene set *KEGG_RIBOSOME*, were affected by all the EOfAD-like mutations but also by non-EOfAD-like mutations in *psen1* and *psen2* (**Fig. 1D**). Evidence for iron dyshomeostasis was also observed for the relatively severe EOfAD-like mutation *psen1*$^{Q96\_K97del}$/+ (under both normoxia and acute hypoxia conditions), and in transheterozygous *sorl1* mutants (i.e. with complete loss of wild type *sorl1*), as shown by significant changes to the expression of genes possessing IRE(s) in the 3' untranslated regions of their encoded mRNAs (*ire_hq* and *ire_all*).

Transcriptome analysis of the App^{NL-G-F} mouse model

EOfAD is also caused by mutations of the gene *APP.* Modelling of *APP* mutations in zebrafish is complicated by duplication of the *APP*-orthologous gene in this organism. However, brain transcriptome data is available for a knock-in mouse model of EOfAD mutations in *APP*: the *App^{NL-G-F}* mouse model (Castillo et al., 2017). In this model, the murine *App* sequence is modified to carry humanised DNA sequences in the Aβ region, as well as the Swedish, Beyreuther/Iberian, and Arctic EOfAD mutations (Saito et al., 2014). While these mice do not closely reflect the genetic state of heterozygous human carriers of EOfAD mutations of *APP* (as the mice possess a total of six mutations within their modified *App* allele and are usually analysed as homozygotes), they should, at least, not generate artefactual patterns of gene expression change due to overexpression of transgenes (Saito et al., 2016). Castillo et al. performed brain transcriptomic profiling via microarray on the brain cortices of male homozygous *App^{NL-G-F}* mice relative to wild type mice at 12 months of age, as well as a transgenic mouse model of AD, 3xTg-AD mice (Oddo et al., 2003) relative to non-Tg mice (Castillo et al., 2017). All mice used in the study were maintained as inbred lines. However, there is no information on whether any of the mice analysed were littermates. It is highly unlikely that the mice used in each comparison between mutant individuals and their wild type counterparts all arose from the same litter, because obtaining 3 homozygous and 3 wild type male mice in a single litter arising from an in-crossing of heterozygous mutant mice, (expected to produce a wild type : heterozygote : homozygote Mendelian genotype ratio of 1:2:1), would be a rare event as litters of mice generally consist of 5 to 10 pups. Therefore, additional variation was introduced into the analysis through use of mice from different litters and this is likely confounding with genotype. This is important to note, as the results presented here were generated under the assumption that any effects of litter-of-origin are negligible.

Our re-analysis of the microarray dataset of Castillo et al. aimed to address the following questions:

1) Are the *KEGG* gene sets affected in the male homozygous *App^{NL-G-F}* mice similar to those affected in EOfAD-like zebrafish?
2) Is there evidence for iron dyshomeostasis in the brains of these mice?

Initially, we attempted to replicate the results of Castillo et al. using the Affymetrix Transcriptome Analysis Console software. However, we were unable to find sufficient information to achieve this. Therefore, we analysed the microarray dataset in a reproducible manner following a recommended microarray analysis workflow (Carvalho and Irizarry, 2010).

After pre-processing of the raw intensities (**Fig. S7**), we performed principal component analysis (PCA) to explore the overall similarity between samples (**Fig. 2A**). Samples separated across PC1 by genotype, suggesting that the homozygous genotypes in this study result in distinct transcriptome states. Notably, the *App*$^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ samples and their corresponding *App*$^{+/+}$ control samples appear to separate to a greater extent across PC1 than the 3xTg samples and their corresponding non-Tg wild type control samples. Additionally, a differential gene expression analysis revealed 158 and 126 genes to be differentially expressed (DE) in *App*$^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ and 3xTg mice respectively (relative to their corresponding controls, **Fig. S8**). This suggests that the disturbance to the cortex transcriptome in *App*$^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ mice is greater than that in 3xTg mice.

We did not observe alteration of any similar gene sets between the *App*$^{NL\text{-}G\text{-}F}$ mice and our EOfAD-like zebrafish (**Fig. S9-S12**). However, any similarities may well have been masked by the overwhelming effects of greater age, variable environment (mouse litter-of-origin), and the effects of their six *App* mutations on brain cortex cell type proportions and inflammatory processes. The most statistically significantly affected cellular process in 12-month-old *App*$^{NL\text{-}G\text{-}F}$ mice is lysosomal function as represented by the *KEGG_LYSOSOME* gene set (discussed later). Additionally, a plethora of inflammatory gene sets are also affected, with changes in the relative proportions of glial cells, particularly microglia, contributing to the appearance of increased levels of these gene transcripts in the bulk cortex RNA analysed (**Fig. 2B,2C, Fig. S13**). Note that changes to cell-type proportions are not observed in our zebrafish models of EOfAD (Barthelson et al., 2021b; Barthelson et al., 2020b; Barthelson et al., 2021c) (see **Fig. S14** for two examples).

Do changes to gene expression in the oxidative phosphorylation pathway also occur in LOAD?

A puzzling observation from the genome-wide association studies (GWAS) of LOAD, is that none of the risk variants identified fall within the EOfAD genes, *PSEN1*, *PSEN2*, or *APP* (Jansen et al., 2019; Kunkle et al., 2019; Lambert et al., 2013). This has led to speculation that EOfAD and LOAD may be distinct diseases despite their histopathological and cognitive similarities (reviewed in (DeTure and Dickson, 2019; Tellechea et al., 2018)). Only one gene identified by GWAS of LOAD is suspected to harbour mutations causative of EOfAD, *SORL1*. Mutations in *SORL1* cause AD with ages of onset typically later than many mutations in *PSEN1* or *APP* (or may be incompletely penetrant) (Pottier et al., 2012; Thonberg et al., 2017). Nevertheless, as shown in **Fig. 1B**, we identified changes in the *KEGG* gene set for oxidative phosphorylation in young adult zebrafish heterozygous for EOfAD-like mutations in *sorl1,* as well as in zebrafish modelling overexpression of the PS2V isoform that is upregulated in LOAD (K97fs).

The strongest and most common genetic risk factor for LOAD is the ε4 allele of the gene *APOE* (Corder et al., 1993; Genin et al., 2011; Jansen et al., 2019; Kunkle et al., 2019; Lambert et al., 2013; Saunders et al., 1993). Like *APP,* the *APOE* orthologous gene in zebrafish is refractory to analysis due to duplication. Therefore, to compare our zebrafish mutant data to early brain transcriptome changes caused by the ε4 allele of *APOE,* we analysed data from a set of human gene targeted replacement mouse models, APOE-TR (Sullivan et al., 1997). These mouse models transcribe human *APOE* alleles from the endogenous murine *Apoe* promotor: the predominant human allele ε3, the rare AD-protective ε2 allele, and the AD-risk allele, ε4. Zhao et al. performed a comprehensive brain transcriptome profiling experiment across aging in both male and female mice to assess the effect of homozygosity for the ε2 or ε4 alleles relative to the risk-neutral ε3 allele (Zhao et al., 2020). In that analysis, pairwise comparisons between the ε2, or ε4 alleles relative to the ε3 allele were not conducted at each age and sex. Only genes/pathways which were influenced overall by *APOE* genotype, age, sex, and interactions between these factors were reported. Since our aim is to identify the early changes occurring due to AD-related mutations, we re-analysed only the 3-month brain samples from the Zhao et al. dataset (i.e. omitting the samples from 12- and 24-month-old mice) to ask which processes are affected by homozygosity for the ε2, or ε4 alleles relative to the ε3 allele. Hereafter, we refer to these homozygous mice as “APOE2”, “APOE3” and “APOE4”.

After pre-processing of the APOE-TR RNA-seq data (**Fig. S15-16**), we performed PCA to visualise the overall similarity between APOE-TR brain transcriptomes. The plot of PC1 against PC2 revealed that samples separated into two distinct clusters of sex across PC2 (**Fig. 3B**). This suggests that the effect of sex on the murine brain transcriptome is substantial and cannot be ignored in the differential gene expression analysis. Among the male samples, APOE4 samples form a cluster distinct from the APOE2 and APOE3 samples, suggesting that the APOE4 genotype has a distinct effect on the transcriptome compared to APOE2 relative to APOE3 in males. This is not observed to the same extent in the female samples. However, the male APOE4 and APOE3 samples appeared to have been taken from distinct litters, as implied from the date-of-birth of each sample (**Fig. 3B**). This confounds the effect of genotype and complicates interpretation of whether any effects observed in a pairwise comparison between male APOE3 and APOE4 mice are due to *APOE* genotype or litter-of-origin (or, most likely, both). Indeed, $\chi 2$ tests for independence revealed that there is a highly significant dependence of *APOE* genotype and litter across the entire 3-month-old dataset ($\chi 2$= 82.7, df = 20, p-value = 1.4e-09), as well as within only male samples ($\chi 2$= 43.1, df = 14, p-value = 8.2e-05) and within only female samples ($\chi 2$= 39.3, df = 14, p-value = 3.0e-04). Some litters did not contain sufficient mice to remove the effect (i.e. some coefficients could not be estimated during the generalised linear model fitting procedure due to the design matrix not having full rank). Therefore, we continued the analysis under the assumption of a negligible effect of litter.

To determine which genes were dysregulated in APOE4 mice and APOE2 mice relative to APOE3 mice, we performed a differential gene expression analysis using *edgeR* (McCarthy et al., 2012; Robinson et al., 2009). Many genes were found to be DE in each comparison, particularly in male APOE4 mice. Additionally, the biases noted by Zhao et al. in the original analysis for increased GC content and longer transcript length among differentially expressed genes was also apparent in our analysis (**Fig. S17**). Therefore, we corrected for these observed biases using conditional quantile normalisation (*cqn*) (Hansen et al., 2012). After *cqn*, many genes were identified as DE in each comparison, and the %GC and gene length biases were decreased (**Fig. S18**).

We next performed enrichment analysis of the *KEGG* (Kanehisa and Goto, 2000) and *IRE* (Hin et al., 2020b) gene sets to determine whether changes are observed in APOE4 mice similar to those in our EOfAD-like model zebrafish (**Fig**. **S19- 23**). We found statistical evidence for significant changes in expression of oxidative phosphorylation and ribosome gene sets in mice homozygous for the humanised ε4 *APOE* allele, consistent with our zebrafish models of EOfAD (although different genes appear to drive the statistical enrichment of the *KEGG_OXIDATIVE_PHOSPHORYALTION* gene set in the two organisms, **Fig. S25**). Interestingly, we observed highly statistically significant changes in the gene set *ire3_all* only in APOE4 male mice, reminiscent of similar signals in some of the young adult EOfAD-related zebrafish (Barthelson et al., 2020b; Hin et al., 2020b) (**Fig. 1B** and **S22C**) and supporting the existence of iron dyshomeostasis. These effects were not observed for the AD-protective ε2 allele (**Fig. 3, Table S2,3**). However, the effects of *APOE* genotype were highly dependent on the litter-of-origin of the samples and changes to cell type proportions were observed in the male APOE4 mice (**Fig. S24**). Therefore, future replication of this analysis with better-controlled transcriptome data is desirable to confirm that the effects observed are due to the *APOE* genotype.

Discussion

Altered gene expression in the oxidative phosphorylation pathway is a transcriptomic signature of genetic variation driving Alzheimer's disease in young adults

Energy production is the most fundamental of cellular activities. Life cannot be sustained without energy, and all other cellular activities depend upon it. The human brain, in particular, has very high energy demands and consumes the majority of the body's glucose when at rest (reviewed in (Zierler, 1999)). Within the brain, the majority of energy use is to maintain the Na^+-K^+ membrane potential of neurons (Attwell and Laughlin, 2001) and neurons are assisted in meeting these energy demands by support from, primarily, astrocytes (e.g. via the astrocyte–neuron lactate shuttle (Pellerin and Magistretti, 1994)). All cells allocate considerable portions of their energy budgets to protein synthesis to maintain their structure and activity (Buttgereit and Brand, 1995). Energy is also required to maintain the low pH and high Ca^{2+} concentration of the lysosome (Christensen et al., 2002), the organelle which mediates uptake and recycling (autophagy) of cellular structural constituents (e.g. the

amino acids for protein synthesis) (reviewed in (Yim and Mizushima, 2020)). Lysosomes are important for uptake and recycling of ferrous iron (Yambire et al., 2019), that is essential for oxidative phosphorylation by mitochondria (Oexle et al., 1999). On the lysosomal membrane, mTOR complexes sense nutrient and energy status to regulate protein synthesis, autophagy, and mitochondrial activity (reviewed in (Lim and Zoncu, 2016)).

The EOfAD genes, *PSEN1*, *PSEN2*, *APP*, and *SORL1*, all encode proteins expressed within the endolysosomal pathway of cells (Andersen et al., 2005; Kawai et al., 1992; Pasternak et al., 2003; Sannerud et al., 2016) and within the mitochondrial associated membranes (MAM) of the endoplasmic reticulum (Area-Gomez et al., 2009; Lim, 2015). MAM are responsible for regulation of ATP production (through Ca^{2+} signalling (Duchen, 1992)), oxidative protein folding (reviewed in (Simmen et al., 2010)), and the initiation of autophagy (Hamasaki et al., 2013). Interestingly, like EOfAD mutant forms of PSEN1 (Lee et al., 2010) and the C99 fragment of APP (Jiang et al., 2019), the ε4 allele of *APOE* has been shown to affect both lysosomal pH (Prasad and Rao, 2018) and the MAM (Tambini et al., 2016). Our analyses of young adult brain transcriptomes in zebrafish have found that five EOfAD-like mutations in a total of three EOfAD gene orthologues (*psen1*, *psen2*, and *sorl1*) all cause statistically significant effects on the expression of genes involved in oxidative phosphorylation, while non AD-related mutations in *psen1* and *psen2* do not. Therefore, effects on oxidative phosphorylation are a common, early "signature" of EOfAD. Intriguingly, we previously observed downregulation of the oxidative phosphorylation genes due to heterozygosity for the *psen1*$^{Q96\_K97del}$ mutation in whole zebrafish larvae at 7 days post fertilisation (dpf) (Dong et al., 2021), suggesting changes to mitochondrial function are a very early cellular stress in EOfAD pathogenesis. Additionally, we observed that the "semi-dominant" ε4 LOAD risk allele (Genin et al., 2011), like EOfAD mutations, also affects the expression of genes involved in oxidative phosphorylation and ribosome function in young adult brains. Thus, changes in oxidative phosphorylation (and so energy production) appear to be a common, early disturbance associated with both early- and late- onset forms of AD.

The majority of the heterozygous EOfAD-like mutations we have studied in zebrafish cause an overall downregulation of the oxidative phosphorylation gene set in young adult brains relative to wild type brains (**Fig. 1C**). Only heterozygosity for the T141_L142delinsMISLISV (reading frame-preserving) mutation of *psen2* has been seen to give overall upregulation of these genes. Another complex, yet probably EOfAD-like mutation in *psen2* we have studied in zebrafish, *psen2*^{S4ter}, (which likely produces Psen2 proteins lacking N-terminal sequences) also showed strong overall upregulation of the oxidative phosphorylation gene set (Jiang et al., 2020), although that dataset contains technical artefacts which complicate interpretation so it was not included in the current analysis. Transheterozygosity for mutations in *sorl1* also results in overall upregulation of oxidative phosphorylation genes. We are uncertain as to why this variability in effects on the oxidative phosphorylation gene set occurs. However, for the *PRESENILIN*s, the single most consistent characteristic of the hundreds of known EOfAD mutations is that they maintain the ability of the genes to produce at least one transcript isoform with the original reading frame (the “reading frame preservation rule” (Jayne et al., 2016)). This strongly supports that all these mutations act via a dominant gain-of-function molecular mechanism (to interfere with a normal cellular function). Alternatively, it may be that the disruption of this gene set that is consistently observed is a product of both genotype and environmental factors, i.e. the mutant fish may be more or less responsive to environmental variation such as changes in water quality, microbiome, handling etc. Also, we note that it can be misleading to infer the direction of change in a particular cell activity, such as oxidative phosphorylation, based on the majority behaviour of a (somewhat arbitrarily) defined set of genes. Obviously, actual measurement of e.g. respiratory rates in the zebrafish mutant brains would be needed to establish, with certainty, how the mutations are affecting oxidative phosphorylation. Note, however, that the subtlety of the gene regulatory effects we have observed in the fish models means that discernment of physiological oxygen consumption differences between mutant fish and their wild type siblings may be challenging. (Simultaneous measurement of differences in the expression levels of the approximately 100 genes in the oxidative phosphorylation gene set gives great statistical sensitivity for detection of subtle differences.)

Importantly, male mice homozygous for LOAD risk allele *APOE* ε4, showed altered expression of the oxidative phosphorylation gene set, while female APOE4 mice showed a similar trend that did not reach the threshold for statistical significance. In fact, male APOE4 mice appeared to have more disrupted brain transcriptomes than female APOE4 mice (**Fig. 3B**) including alteration in the abundance of transcripts with IREs in their 3' UTRs, supporting the possibility of early brain iron dyshomeostasis (as was previously observed in *psen*$^{Q96\_K97del}$/+ fish and those lacking wild type *sorl1* function). This was unexpected, as human females are more susceptible to the effects of *APOE* ε4 than males (Farrer et al., 1997; Wang et al., 2020). However, other sources of variation (i.e. litter-of-origin, changes to cell type proportions, and possibly transcript length) may be masking the true effects of *APOE* genotype in these mice. Reassuringly, a recent single-cell RNA-seq analysis (where discrepancies due to differences in cell-type proportions are overcome) of female APOE4 mice showed that expression of oxidative phosphorylation genes is decreased at 12 months-of-age, particularly in astrocytes. Further dissection of the metabolic phenotype of female APOE4 mice revealed a shift away from oxidative phosphorylation and towards glycolysis for ATP production (Farmer et al., 2021). Remarkably, this result was consistent with observations in human ε4-carrying females, who show lower energy expenditure, decreased oxygen consumption, and alterations to their plasma metabolomes indicative of increased glycolysis (Farmer et al., 2021).

The changes to gene expression in the oxidative phosphorylation pathway in both EOfAD-like zebrafish and APOE4 mice demonstrates the similarity, at the molecular level, between the cellular effects of genetic variants causing EOfAD and the most significant variant promoting LOAD. The differences in disease onset age between EOfAD and LOAD may be due to the severity of the effects on energy metabolism of the different genetic variants that promote each disease in concert with environmental variables. This is consistent with the fact that classification of AD into these two subtypes appears arbitrary since there is no discernible early age-dependent peak in the population prevalence of dementia (of which AD contributes the majority of cases). The observed molecular pathway similarity between knock-in models of EOfAD and LOAD genetic variation supports the validity and utility of analysing early molecular events in AD pathogenesis using knock-in models in both zebrafish and mice, and that analysis of endogenous EOfAD mutations can contribute information valuable for understanding LOAD pathogenesis. The mystery of why GWAS has failed to detect variation in *PSEN1*, *PSEN2*, or *APP* in LOAD remains, although some mutations in *PSEN2*

and *APP* genes do cause later onset familial forms of the disease (Cruchaga et al., 2012) and/or show incomplete penetrance (Finckh et al., 2000; Rossor et al., 1996; Sherrington et al., 1996; Thordardottir et al., 2018) or recessive inheritance (Di Fede et al., 2009; Tomiyama et al., 2008).

Brain hypometabolism is a diagnostic criterion for AD and can be visualised by 2-deoxy-2-(^{18}F)fluoro-D-glucose positron emission tomography (^{18}FDG-PET) (reviewed in (Marcus et al., 2014)). This technique has been used previously to investigate adults at risk of developing AD (e.g. (Ou et al., 2019)) while blood oxygenation level dependent (BOLD) functional magnetic resonance imaging (fMRI) has identified brain regional changes in activity (and hence energy consumption) in child carriers of the *PSEN1* E280A ("Paisa") EOfAD mutation (Quiroz et al., 2015). This suggests that ^{18}FDG-PET, or techniques such as dynamic glucose-enhanced (DGE) MRI (Tolomeo et al., 2018), might enable screening of individuals to determine risk of later development of AD and be useful tools in investigating early energy changes in animal models of AD. Indeed, such techniques have been exploited to monitor energy consumption in the brains of transgenic mouse models (Huang et al., 2020; Luo et al., 2012; Poisnel et al., 2012; Tolomeo et al., 2018). However, they have not yet been applied to mice with knock-in EOfAD mutations. The observations of change in *in vivo* brain energy metabolism discussed above are consistent with the observed changes in the expression of genes of the oxidative phosphorylation pathway in the post-mortem brains of early, and late AD subjects relative to age-matched controls (Manczak et al., 2004). Neuronal cells derived from human induced pluripotent stem cells (hiPSCs) of LOAD patients also show increased expression of oxidative phosphorylation proteins and oxidative stress (Birnbaum et al., 2018) and neurons derived from a patient carrying the *PSEN1*^{S170F} EOfAD mutation showed mitochondrial abnormalities (Li et al., 2020).

Our findings regarding EOfAD mutations in zebrafish were not consistent with findings from our analysis of transcriptome data from homozygous *App*$^{NL-G-F}$ mice. However, this transcriptome data was generated from "middle-aged" (12 month old) mice rather than young adults, and the endogenous *App* gene of the mouse was altered with a total of six mutations (three that humanise the sequence of the Aβ region and three EOfAD mutations), motivated by the idea that the more aggregation-prone human Aβ sequence plays a critical role in the pathogenic mechanism of AD. Therefore, it is not directly comparable with our zebrafish EOfAD models, which contain single, EOfAD-like mutations within single alleles of endogenous genes. To our knowledge, a transcriptome

analysis has not been performed on *App*$^{NL\text{-}G\text{-}F}$ mice at a younger age. However, the expression of genes involved in lysosomal function (*KEGG_LYSOSOME*) was observed to be highly significantly upregulated in homozygous *App*$^{NL\text{-}G\text{-}F}$ brains. This is not unexpected, as acidification of the endo-lysosomal system is impaired by increased levels of the β-CTF fragment of APP (also known as C99 and generated by β-secretase cleavage of APP (Jiang et al., 2019)). Increased β-CTF has been observed in the brains of *App*$^{NL\text{-}G\text{-}F}$ mice (Saito et al., 2014). In a mouse model of a lysosomal storage disorder (Glycogen storage disease type 2, which most seriously affects muscle (Yambire et al., 2019)), lysosomes failed to become sufficiently acidic and this resulted in an intracellular ferrous iron deficiency and a pseudo-hypoxic response, mitochondrial dysfunction, and inflammation (Yambire et al., 2019). (Degradation of HIF1-α, the master transcriptional regulator of the cellular response to hypoxia, is dependent on both oxygen and ferrous iron (Ivan et al., 2001).) Additionally, the *App*$^{NL\text{-}G\text{-}F}$ mouse model shows increased levels of Aβ from a young age (Saito et al., 2014), and the deposition of Aβ into plaques was shown to be associated with the increased expression of genes in the complement system in a comprehensive spatial transcriptomics study of aging in the *App*$^{NL\text{-}G\text{-}F}$ mouse model (Chen et al., 2020), providing another avenue for these mutations to trigger inflammation. Mitochondrial dysfunction has not been observed directly in *App*$^{NL\text{-}G\text{-}F}$ mice. However, increased levels of oxidative stress have been observed at 12 months of age (Izumi et al., 2020), suggestive of increased reactive oxygen species (ROS) that can be generated by dysfunction of mitochondrial respiration. Therefore, we suspect that similar processes are being affected in *App*$^{NL\text{-}G\text{-}F}$ mice to those in our zebrafish models and the APOE4 knock-in mice. However, subtle signs of mitochondrial dysfunction in the transcriptome may be obscured by noise from the strong inflammatory signals in the bulk brain transcriptomic data (as well as confounding influences on the transcriptome analysis such as litter-of-origin effects).

mTOR signalling can regulate ribosomal gene set expression

All of the mutations studied have also resulted in changes in the levels of transcripts required for ribosome formation. Protein translation is one of the most energy-costly processes within a cell (Buttgereit and Brand, 1995), and so expression of ribosomal proteins is modulated by the mammalian target of rapamycin (mTOR) system that surveys cellular nutrient status to adjust cellular

metabolism (reviewed in (Mayer and Grummt, 2006; Zhou et al., 2015)). mTOR signalling, which appears to be increased in late-stage AD brains compared to controls (Griffin et al., 2005; Li et al., 2005; Sun et al., 2014), is regulated by growth factors, nutrients, energy levels and stress. In addition to ribosome biogenesis, mTOR signalling plays a role in various other cellular processes also implicated in AD pathogenesis, such as autophagy and lipid metabolism (reviewed in (Saxton and Sabatini, 2017)).

The mTOR proteins are localised at lysosomes within the mTORC1 and mTORC2 protein complexes (Sancak et al., 2010). Intriguingly, the v-ATPase complex that acidifies the endolysosomal pathway is required for mTORC1 activation (Zoncu et al., 2011). Proper assembly of the v-ATPase at the lysosome requires the PSEN1 protein (and this process is impaired in EOfAD patient fibroblasts) (Lee et al., 2010). Stimulation of mTOR signalling has been observed in response to accumulation of Aβ (Caccamo et al., 2010), while hyperactivation of mTOR is observed in Down's syndrome (where the dosage of the *APP* gene is increased because it resides on Chromosome 21 and early-onset AD is common) (Bordi et al., 2019; Iyer et al., 2014). Intriguingly, Bordi et al. observed that inhibition of mTOR signalling (specifically, mTORC1) rescues auto- and mito-phagy defects in the fibroblasts of Down's syndrome individuals (Bordi et al., 2019). Among our transcriptome analyses of AD models, we only observed statistically significant changes to the expression of genes in the *KEGG* gene set for mTOR signalling in transheterozygous *sorl1* mutants, in *psen1*$^{Q96\_K97del}$/+ mutant zebrafish after acute hypoxia exposure, and in both male and female APOE4 mice. However, the majority of regulation of mTOR signalling occurs at the protein level, so that it is perhaps unsurprising that, in the EOfAD mutants, we could only detect significant changes in the transcriptional response to altered mTOR signalling rather than in the mTOR gene set itself.

Advantages and disadvantages of zebrafish for analysis of genetic variants driving AD

In a highly sensitive analysis method such as RNA-seq, minimising external sources of variation increases resolving power. Our analysis has revealed that zebrafish can be highly advantageous for transcriptome profiling in the context of RNA-seq, as large numbers of progeny can be produced from a single pair mating, and these can subsequently be raised together in a single aquarium system,

thus reducing both genetic and environmental variation. This has allowed us to observe subtle effects due to the EOfAD-like mutations we have analysed. In contrast, a female mouse can only birth relatively small litters of 5-10 pups, making it particularly difficult to obtain sufficient numbers of synchronous sibling samples (particularly when genotypes of interest are produced by crossing of heterozygotes). Our re-analysis of the APOE-TR mouse brain transcriptomes was unable to distinguish with great certainty whether the effects we observed were due to *Apoe* genotype or litter-of-origin. Also, information on whether $App^{NL\text{-}G\text{-}F}$ mice were littermates was not available, and this required us to assume that the effects of litter were negligible in order to perform the analysis. Another contrast between brain transcriptome analysis in zebrafish compared to mice is the influence of sex. Mouse brain transcriptomes show very significant differences due to sex, while sex has a negligible effect on bulk brain transcriptomes from zebrafish (Barthelson et al., 2021a; Barthelson et al., 2021b; Barthelson et al., 2020b; Barthelson et al., 2021c; Drew et al., 2012), and can generally be ignored in a differential expression analysis. We also found evidence for changes to cell type proportions in both APOE-TR and $App^{NL\text{-}G\text{-}F}$ mice, a phenomenon that can create the artefactual appearance of gene expression change. We have not observed cell-type proportion differences in 6 month or 24 month old zebrafish (Barthelson et al., 2021a; Barthelson et al., 2021b; Barthelson et al., 2020b; Barthelson et al., 2021c; Hin et al., 2020a; Hin et al., 2020b) (**Fig. S14**), possibly associated with the resistance to damage of the highly regenerative zebrafish brain (Kroehne et al., 2011). While this regenerative ability may hinder use of zebrafish for studying overt neurodegeneration, it can facilitate analysis of young, bulk brain transcriptomes before overt pathological processes would be expected.

The advantages of zebrafish for analysing the early effects of EOfAD mutations are countered, occasionally, by disadvantages. The teleost lineage in which zebrafish arose underwent an early whole-genome duplication event (reviewed in (Meyer and Van de Peer, 2005)), such that many human genes are represented by duplicate “co-orthologues” in zebrafish (e.g. the co-orthologues of *APP* and *APOE* in zebrafish are *appa* / *appb* and *apoea* / *apoeb* respectively). This complicates interpretation of the effects of mutations in these genes. Additionally, zebrafish have never been shown, definitively, to be capable of producing Aβ, a pathological hallmark of AD. The β-secretase (BACE) site of human APP does not appear conserved in zebrafish Appa and Appb (Moore et al., 2014). Whether Aβ accumulation is a cause or consequence of AD pathological processes continues

as a matter of debate within the AD research community (reviewed in (Morris et al., 2018)). If zebrafish cannot produce Aβ, then the changes we have observed in the brains of our zebrafish models may illuminate Aβ-independent effects of EOfAD mutations.

Knock-in models of Alzheimer's disease mutations may model more accurately early AD-associated pathological changes

Knock-in mouse models of single EOfAD mutations were generated 15-20 years ago (Guo et al., 1999; Kawasumi et al., 2004) but their brain transcriptomes have never been analysed in detail. This is likely because these mice showed only very mild cognitive phenotypes and lacked the AD histopathology currently used to define the disease (Aβ deposition and neurofibrillary tangles of tau protein (Jack et al., 2018)). By expressing multiple mutant forms of EOfAD genes in transgenic mice, Aβ plaques can be detected and cognitive changes observed (reviewed in (Esquerda-Canals et al., 2017; Myers and McGonigle, 2019)). However, experience with use of many such "mouse models of Alzheimer's disease" has shown a lack of correlation of cognitive changes with Aβ levels (Foley et al., 2015) (Aβ levels do not closely correlate with cognitive changes in humans either (Giannakopoulos et al., 2003)) and transcriptome analysis of their brains has shown little to no concordance with transcriptomes from post-mortem AD brains, or between the models themselves (Hargis and Blalock, 2017). In two papers in 2014 and 2016, Saito and colleagues described phenotypic disparities between transgenic and *APP* EOfAD mutation knock-in mouse models (Saito et al., 2014; Saito et al., 2016). In the 2016 paper they went so far as to declare that,

> "*We recently estimated using single App knock-in mice that accumulate amyloidβ peptide without transgene overexpression that 60% of the phenotypes observed in Alzheimer's model mice overexpressing mutant amyloid precursor protein (APP) or APP and presenilin are artifacts (Saito et al., 2014). The current study further supports this estimate by invalidating key results from papers that were published in Cell. These findings suggest that more than 3000 publications based on APP and APP/PS overexpression must be reevaluated*."

Nevertheless, since 2016 thousands more papers have been published using transgenic mouse models of Alzheimer's disease. In this light, we were surprised to find that *App*$^{NL-G-F}$ homozygous mice display a young adult brain transcriptome that is more severely disturbed than in the multiply transgenic 3xTg-AD model - although that apparent disturbance is likely somewhat artefactual and due to changes in the relative proportions of different cell types in the model. As mentioned above, changes to cell type proportions are not observed in our zebrafish models (Barthelson et al., 2021b; Barthelson et al., 2020b; Barthelson et al., 2021c) (**Fig.S14**),

Frustration with the difficulties of exploiting both transgenic and knock-in models of EOfAD mutations in mice has contributed to the drive for examining knock-in mouse models of LOAD risk variants, such as now conducted by the MODEL-AD Consortium (Oblak et al., 2020). The brain transcriptome similarities seen between our single mutation, heterozygous EOfAD mutation-like knock-in zebrafish models and the knock-in *APOE* ε4 mice strongly support the informative value of these models and imply that heterozygous EOfAD mutation knock-in mouse models offer a path forward, particularly in understanding the earliest molecular events that lead to Alzheimer's disease.

Methods

Analysis of knock-in zebrafish models of EOfAD

For analysis of 6 month and 24 month old genome-edited Tübingen zebrafish (sample sizes and sexes are indicated in **Table 1 and Fig S1-S6**), we obtained the differential gene expression analysis outputs and harmonic mean p-values (statistical significance of gene sets) from each individual analysis (see (Barthelson et al., 2021a; Barthelson et al., 2021b; Barthelson et al., 2020b; Barthelson et al., 2021c)). For these analyses, differential gene expression analysis was performed using *edgeR* (Robinson et al., 2009), and enrichment analysis was performed by calculation of the harmonic mean p-value (Wilson, 2019) of the raw p-values of three methods of ranked-list based enrichment analyses: *fry* (Wu et al., 2010), *camera* (Wu and Smyth, 2012), and *GSEA* (Subramanian et al., 2005) (as implemented in the *fgsea* R package (Sergushichev, 2016)). We used the harmonic mean p-value to determine the overall significance of changes to gene expression within gene sets because this method does not assume that component p-values are independent (Wilson, 2019). We have

previously validated the use of the harmonic mean p-value on simulated RNA-seq datasets (Barthelson et al., 2020b). We considered a gene set to be significantly altered if the FDR-adjusted harmonic mean p-value remained below 0.05 after FDR adjustment. The gene sets used for enrichment analysis were the *KEGG* (Kanehisa and Goto, 2000) gene sets, to determine whether changes to gene expression were observed in any of 186 biological pathways/processes. Additionally, we used our recently defined iron responsive element (*IRE*) gene sets (HIN ET AL., 2020B) to test for evidence of iron dyshomeostasis. For the K97fs and Q96_K97del analyses, enrichment analysis was not performed on the *KEGG* gene sets in the original analyses. Therefore, we performed the enrichment analysis as described above for these datasets. For the K97fs analysis, we obtained the gene-level counts and the results of the differential gene expression analysis described in (Hin et al., 2020a) from https://github.com/UofABioinformaticsHub/k97fsZebrafishAnalysis. For the Q96_K97del analysis, we obtained the gene-level counts and the results of the differential gene expression analysis described in (Hin et al., 2020b) from the first author of the cited paper. Note that, for each zebrafish analysis, the sample size was usually n = 6 zebrafish per genotype, based on our previous calculation that this sample size should give approximately 70% power to detect the majority of expressed transcripts in a zebrafish brain transcriptome at a fold change > 2 and at a false discovery rate of 0.05 (Barthelson et al., 2020b).

App^{NL-G-F} microarray re-analysis

The raw .CEL files were obtained from GEO and analysed with *R* (R Core Team, 2019). Pre-processing was performed using the *rma* (Irizarry et al., 2003) method as implemented in the *oligo* package (Carvalho and Irizarry, 2010). We omitted any probesets which contained a median log2 intensity value of < 3.5 (lowly expressed) and also any probesets assigned to multiple genes. Differential gene expression analysis was performed using *limma* (Ritchie et al., 2015), specifying pairwise contrasts between the *App^{NL-G-F}* homozygous mice, or the 3xTg homozygous mice with their respective controls by using a contrasts matrix. We considered a probeset to be differentially expressed in each contrast if the FDR adjusted p-value was < 0.05. For over-representation of the *KEGG* and *IRE* gene sets within the DE genes, we used *kegga* (Young et al., 2010). We also performed ranked-list based enrichment analysis using the harmonic mean p-value as described for the zebrafish analyses.

APOE-TR RNA-seq re-analysis

We obtained the raw fastq files for the entire APOE-TR RNA-seq experiment from AD Knowledge Portal (accession number syn20808171, https://adknowledgeportal.synapse.org/). The raw reads were first processed using *AdapterRemoval* (version 2.2.1) (Schubert et al., 2016), setting the following options: *--trimns*, *--trimqualities, --minquality 30, --minlength 35.* Then, the trimmed reads were aligned to the *Mus musculus* genome (Ensembl GRCm38, release 98) using *STAR* (version 2.7.0) (Dobin et al., 2013) using default parameters to generate .bam files. These bam files were then sorted and indexed using *samtools* (version 1.10) (Li et al., 2009). The gene expression counts matrix was generated from the .bam files using *featureCounts* (version 1.5.2) (Liao et al., 2014). We only counted the number of reads which uniquely aligned to, strictly, exons with a mapping quality of at least 10 to predict expression levels of genes in each sample.

We then imported the output from *featureCounts* (Liao et al., 2014) for analysis with *R* (R Core Team, 2019). We first omitted genes which are lowly expressed (and are uninformative for differential expression analysis). We considered a gene to be lowly expressed if it contained, at most, 2 counts per million (CPM) in 8 or more of the 24 samples we analysed. We also assessed whether the sex of each sample was correctly classified by examining the expression of genes which are located on the Y-chromosome. Three samples appeared to be classified incorrectly and this was subsequently corrected (**Fig. S16**).

To determine which genes were dysregulated in APOE4 and APOE2 mice relative to APOE3, we performed a differential gene expression analysis using a generalised linear model and likelihood ratio tests using *edgeR* (McCarthy et al., 2012; Robinson et al., 2009). We chose a design matrix which specifies the *APOE* genotype and sex of each sample. The contrasts matrix was specified to compare the effect of APOE2 or APOE4 relative to APOE3 in males and in females. In this analysis, we considered a gene to be differentially expressed (DE) if the FDR-adjusted p-value was < 0.05. A bias for longer transcript length and higher %GC content was observed in this dataset. Therefore, we corrected for this bias using conditional quantile normalisation (*cqn*) (Hansen et al., 2012). We calculated the average transcript length per gene, and a weighted (by transcript length) average %GC content per gene as input to *cqn* to produce the offset to correct for the bias. This offset was then included in an additional generalised linear model and likelihood ratio tests in *edgeR* with the same

design and contrast matrices. For over-representation of the *KEGG* and *IRE* gene sets within the DE genes, we used *goseq* (Young et al., 2010), specifying average transcript length to generate the probability weighting function, which corrects for the probability that a gene is classified as DE based on its transcript length (average transcript length per gene) alone. We also performed ranked-list based enrichment analysis as described for the zebrafish analysis.

Visualisation of gene expression data throughout this analysis was performed with *ggplot2* (Wickham, 2016), *pheatmap* (Kolde, 2019) and *pathview* (Luo et al., 2017). The code used to perform the analysis in this study can be found at https://github.com/karissa-b/AD-signature.

Acknowledgements

The authors would like to thank Dr. Nhi Hin for providing the Q96_K97del and K97fs zebrafish datasets and original analyses. We also wish to thank Dr, Stephen Pederson for his advice and guidance throughout the project. This work was supported with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide. The results for the young APOE-TR mice were based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.synapse.org/). We thank Drs. Patrick Sullivan and Nobuyo Maeda for generating human APOE-TR mice and providing access through Taconic.

Competing interests

No competing interests declared.

Funding

K.B. was supported by an Australian Government Research Training Program Scholarship, and by funds from the Carthew Family Trust. M.L. and M.N. were both supported by grants from the National Health and Medical Research Council of Australia (NHMRC, GNT1061006 and GNT1126422). M.L. is an academic employee of the University of Adelaide.

Data availability

All zebrafish datasets are available from the Gene Expression Omnibus (GEO,), or in the European Nucleotide Archive (ENA) repositories. For accession numbers for each of these six datasets, see Table 1. The mouse microarray data is available from GEO under the accession number GSE36980. The APOE-TR RNA-seq data is available from the AD Knowledge Portal (accession number syn20808171, https://adknowledgeportal.synapse.org/).

References

Alzheimer, A. (1906). Über einen eigenartigen schweren erkrankungsprozeβ der hirnrincle. *Neurol Cent* **25**.

Andersen, O. M., Reiche, J., Schmidt, V., Gotthardt, M., Spoelgen, R., Behlke, J., von Arnim, C. A., Breiderhoff, T., Jansen, P., Wu, X. et al. (2005). Neuronal sorting protein-related receptor sorLA/LR11 regulates processing of the amyloid precursor protein. *Proc Natl Acad Sci U S A* **102**, 13461-6.

Area-Gomez, E., de Groof, A. J., Boldogh, I., Bird, T. D., Gibson, G. E., Koehler, C. M., Yu, W. H., Duff, K. E., Yaffe, M. P., Pon, L. A. et al. (2009). Presenilins are enriched in endoplasmic reticulum membranes associated with mitochondria. *American Journal of Pathology* **175**, 1810-6.

Attwell, D. and Laughlin, S. B. (2001). An Energy Budget for Signaling in the Grey Matter of the Brain. *Journal of Cerebral Blood Flow & Metabolism* **21**, 1133-1145.

Barthelson, K., Dong, Y., Newman, M. and Lardelli, M. (2021a). PRESENILIN 1 Mutations Causing Early-Onset Familial Alzheimer's Disease or Familial Acne Inversa Differ in Their Effects on Genes Facilitating Energy Metabolism and Signal Transduction. *Journal of Alzheimer's Disease* **Preprint**, 1-21.

Barthelson, K., Newman, M. and Lardelli, M. (2020a). Sorting Out the Role of the Sortilin-Related Receptor 1 in Alzheimer's Disease. *J Alzheimers Dis Rep* **4**, 123-140.

Barthelson, K., Pederson, S. M., Newman, M., Jiang, H. and Lardelli, M. (2021b). In-Frame and Frameshift Mutations in Zebrafish Presenilin 2 Affect Different Cellular Functions in Young Adult Brains. *Journal of Alzheimer's Disease Reports* **5**, 395-404.

Barthelson, K., Pederson, S. M., Newman, M. and Lardelli, M. (2020b). Brain transcriptome analysis reveals subtle effects on mitochondrial function and iron homeostasis of mutations in the SORL1 gene implicated in early onset familial Alzheimer's disease. *Molecular Brain* **13**, 142.

Barthelson, K., Pederson, S. M., Newman, M. and Lardelli, M. (2021c). Brain Transcriptome Analysis of a Protein-Truncating Mutation in Sortilin-Related Receptor 1 Associated With Early-Onset Familial Alzheimer's Disease Indicates Early Effects on Mitochondrial and Ribosome Function. *Journal of Alzheimer's Disease* **79**, 1105-1119.

Bertram, L. and Tanzi, R. E. (2012). The genetics of Alzheimer's disease. *Progress in molecular biology and translational science* **107**, 79.

Birnbaum, J. H., Wanner, D., Gietl, A. F., Saake, A., Kündig, T. M., Hock, C., Nitsch, R. M. and Tackenberg, C. (2018). Oxidative stress and altered mitochondrial protein expression in the absence of amyloid-β and tau pathology in iPSC-derived neurons from sporadic Alzheimer's disease patients. *Stem Cell Res* **27**, 121-130.

Blennow, K., de Leon, M. J. and Zetterberg, H. (2006). Alzheimer's disease. *The Lancet* **368**, 387-403.

Bordi, M., Darji, S., Sato, Y., Mellén, M., Berg, M. J., Kumar, A., Jiang, Y. and Nixon, R. A. (2019). mTOR hyperactivation in Down Syndrome underlies deficits in autophagy induction, autophagosome formation, and mitophagy. *Cell Death & Disease* **10**, 563.

Buttgereit, F. and Brand, M. D. (1995). A hierarchy of ATP-consuming processes in mammalian cells. *The Biochemical journal* **312 (Pt 1)**, 163-167.

Caccamo, A., Majumder, S., Richardson, A., Strong, R. and Oddo, S. (2010). Molecular interplay between mammalian target of rapamycin (mTOR), amyloid-beta, and Tau: effects on cognitive impairments. *J Biol Chem* **285**, 13107-13120.

Carvalho, B. S. and Irizarry, R. A. (2010). A framework for oligonucleotide microarray preprocessing. *Bioinformatics (Oxford, England)* **26**, 2363-2367.

Castillo, E., Leon, J., Mazzei, G., Abolhassani, N., Haruyama, N., Saito, T., Saido, T., Hokama, M., Iwaki, T., Ohara, T. et al. (2017). Comparative profiling of cortical gene expression in Alzheimer's disease patients and mouse models demonstrates a link between amyloidosis and neuroinflammation. *Scientific Reports* **7**, 17762.

Chen, W.-T., Lu, A., Craessaerts, K., Pavie, B., Sala Frigerio, C., Corthout, N., Qian, X., Laláková, J., Kühnemund, M., Voytyuk, I. et al. (2020). Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. *Cell* **182**, 976-991.e19.

Christensen, K. A., Myers, J. T. and Swanson, J. A. (2002). pH-dependent regulation of lysosomal calcium in macrophages. *Journal of Cell Science* **115**, 599.

Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., Roses, A. D., Haines, J. L. and Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* **261**, 921-3.

Cruchaga, C., Haller, G., Chakraverty, S., Mayo, K., Vallania, F. L. M., Mitra, R. D., Faber, K., Williamson, J., Bird, T., Diaz-Arrastia, R. et al. (2012). Rare variants in APP, PSEN1 and PSEN2 increase risk for AD in late-onset Alzheimer's disease families. *PLoS ONE* **7**, e31039-e31039.

DeTure, M. A. and Dickson, D. W. (2019). The neuropathological diagnosis of Alzheimer's disease. *Molecular Neurodegeneration* **14**, 32.

Di Fede, G., Catania, M., Morbin, M., Rossi, G., Suardi, S., Mazzoleni, G., Merlin, M., Giovagnoli, A. R., Prioni, S. and Erbetta, A. (2009). A recessive mutation in the APP gene with dominant-negative effect on amyloidogenesis. *Science* **323**, 1473-1477.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* **29**, 15-21.

Dong, Y., Newman, M., Pederson, S. M., Barthelson, K., Hin, N. and Lardelli, M. (2021). Transcriptome analyses of 7-day-old zebrafish larvae possessing a familial Alzheimer's disease-like mutation in psen1 indicate effects on oxidative phosphorylation, ECM and MCM functions, and iron homeostasis. *BMC Genomics* **22**, 211.

Drew, R. E., Settles, M. L., Churchill, E. J., Williams, S. M., Balli, S. and Robison, B. D. (2012). Brain transcriptome variation among behaviorally distinct strains of zebrafish (Danio rerio). *BMC Genomics* **13**, 323.

Duchen, M. R. (1992). Ca(2+)-dependent changes in the mitochondrial energetics in single dissociated mouse sensory neurons. *The Biochemical journal* **283 (Pt 1)**, 41-50.

Esquerda-Canals, G., Montoliu-Gaya, L., Güell-Bosch, J. and Villegas, S. (2017). Mouse Models of Alzheimer's Disease. *Journal of Alzheimer's Disease* **57**, 1171-1183.

Farmer, B. C., Williams, H. C., Devanney, N. A., Piron, M. A., Nation, G. K., Carter, D. J., Walsh, A. E., Khanal, R., Young, L. E. A., Kluemper, J. C. et al. (2021). APOE4 lowers energy expenditure in females and impairs glucose oxidation by increasing flux through aerobic glycolysis. *Molecular Neurodegeneration* **16**, 62.

Farrer, L. A., Cupples, L. A., Haines, J. L., Hyman, B., Kukull, W. A., Mayeux, R., Myers, R. H., Pericak-Vance, M. A., Risch, N. and van Duijn, C. M. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA* **278**, 1349-56.

Finckh, U., Alberici, A., Antoniazzi, M., Benussi, L., Fedi, V., Giannini, C., Gal, A., Nitsch, R. M. and Binetti, G. (2000). Variable expression of familial Alzheimer disease associated with presenilin 2 mutation M239I. *Neurology* **54**, 2006-8.

Foley, A. M., Ammar, Z. M., Lee, R. H. and Mitchell, C. S. (2015). Systematic review of the relationship between amyloid-β levels and measures of transgenic mouse cognitive deficit in Alzheimer's disease. *J Alzheimers Dis* **44**, 787-95.

Genin, E., Hannequin, D., Wallon, D., Sleegers, K., Hiltunen, M., Combarros, O., Bullido, M. J., Engelborghs, S., De Deyn, P., Berr, C. et al. (2011). APOE and Alzheimer disease: a major gene with semi-dominant inheritance. *Mol Psychiatry* **16**, 903-7.

Giannakopoulos, P., Herrmann, F. R., Bussière, T., Bouras, C., Kövari, E., Perl, D. P., Morrison, J. H., Gold, G. and Hof, P. R. (2003). Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. *Neurology* **60**, 1495-1500.

Griffin, R. J., Moloney, A., Kelliher, M., Johnston, J. A., Ravid, R., Dockery, P., O'Connor, R. and O'Neill, C. (2005). Activation of Akt/PKB, increased phosphorylation of Akt substrates and loss and altered distribution of Akt and PTEN are features of Alzheimer's disease pathology. *Journal of Neurochemistry* **93**, 105-117.

Guo, Q., Fu, W., Sopher, B. L., Miller, M. W., Ware, C. B., Martin, G. M. and Mattson, M. P. (1999). Increased vulnerability of hippocampal neurons to excitotoxic necrosis in presenilin-1 mutant knock-in mice. *Nat Med* **5**, 101-6.

Hamasaki, M., Furuta, N., Matsuda, A., Nezu, A., Yamamoto, A., Fujita, N., Oomori, H., Noda, T., Haraguchi, T., Hiraoka, Y. et al. (2013). Autophagosomes form at ER-mitochondria contact sites. *Nature* **495**, 389-393.

Hansen, K. D., Irizarry, R. A. and Wu, Z. (2012). Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics* **13**, 204-216.

Hargis, K. E. and Blalock, E. M. (2017). Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us? *Behavioural Brain Research* **322**, 311-328.

Hin, N., Newman, M., Kaslin, J., Douek, A. M., Lumsden, A., Nik, S. H. M., Dong, Y., Zhou, X.-F., Mañucat-Tan, N. B., Ludington, A. et al. (2020a). Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2. *PLoS ONE* **15**, e0227258.

Hin, N., Newman, M., Pederson, S. M. and Lardelli, M. M. (2020b). Iron Responsive Element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease. *bioRxiv*, 2020.05.01.071498.

Huang, J., van Zijl, P. C. M., Han, X., Dong, C. M., Cheng, G. W. Y., Tse, K. H., Knutsson, L., Chen, L., Lai, J. H. C., Wu, E. X. et al. (2020). Altered d-glucose in brain parenchyma and cerebrospinal fluid of early Alzheimer's disease detected by dynamic glucose-enhanced MRI. *Sci Adv* **6**, eaba3884.

Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B. and Speed, T. P. (2003). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research* **31**, e15-e15.

Iturria-Medina, Y., Sotero, R., Toussaint, P., Mateos-Pérez, J., Evans, A. and Initiative, A. s. D. N. (2016). Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nature Communications* **7**.

Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J. M., Lane, W. S. and Kaelin Jr, W. G. (2001). HIFα Targeted for VHL-Mediated Destruction by Proline Hydroxylation: Implications for O2 Sensing. *Science* **292**, 464.

Iyer, A. M., van Scheppingen, J., Milenkovic, I., Anink, J. J., Adle-Biassette, H., Kovacs, G. G. and Aronica, E. (2014). mTOR Hyperactivation in Down Syndrome Hippocampus Appears Early During Development. *Journal of Neuropathology & Experimental Neurology* **73**, 671-683.

Izumi, H., Sato, K., Kojima, K., Saito, T., Saido, T. C. and Fukunaga, K. (2020). Oral glutathione administration inhibits the oxidative stress and the inflammatory responses in AppNL−G-F/NL−G-F knock-in mice. *Neuropharmacology* **168**, 108026.

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J. et al. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* **14**, 535-562.

Jansen, I. E., Savage, J. E., Watanabe, K., Bryois, J., Williams, D. M., Steinberg, S., Sealock, J., Karlsson, I. K., Hägg, S., Athanasiu, L. et al. (2019). Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nature Genetics* **51**, 404-413.

Jayne, T., Newman, M., Verdile, G., Sutherland, G., Munch, G., Musgrave, I., Moussavi Nik, S. H. and Lardelli, M. (2016). Evidence for and Against a Pathogenic Role of Reduced gamma-Secretase Activity in Familial Alzheimer's Disease. *Journal of Alzheimer's Disease*.

Jiang, H., Pederson, S. M., Newman, M., Dong, Y., Barthelson, K. and Lardelli, M. (2020). Transcriptome analysis indicates dominant effects on ribosome and mitochondrial function of a premature termination codon mutation in the zebrafish gene psen2. *PLoS ONE* **15**, e0232559.

Jiang, Y., Sato, Y., Im, E., Berg, M., Bordi, M., Darji, S., Kumar, A., Mohan, P. S., Bandyopadhyay, U., Diaz, A. et al. (2019). Lysosomal Dysfunction in Down Syndrome Is APP-Dependent and Mediated by APP-βCTF (C99). *The Journal of Neuroscience* **39**, 5255.

Kanehisa, M. and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* **28**, 27-30.

Kawai, M., Cras, P., Richey, P., Tabaton, M., Lowery, D. E., Gonzalez-DeWhitt, P. A., Greenberg, B. D., Gambetti, P. and Perry, G. (1992). Subcellular localization of amyloid precursor protein in senile plaques of Alzheimer's disease. *Am J Pathol* **140**, 947-58.

Kawasumi, M., Chiba, T., Yamada, M., Miyamae-Kaneko, M., Matsuoka, M., Nakahara, J., Tomita, T., Iwatsubo, T., Kato, S., Aiso, S. et al. (2004). Targeted introduction of V642I mutation in amyloid precursor protein gene causes functional abnormality resembling early stage of Alzheimer's disease in aged mice. *Eur J Neurosci* **19**, 2826-38.

Kolde, R. (2019). pheatmap: Pretty Heatmaps, vol. .

Kroehne, V., Freudenreich, D., Hans, S., Kaslin, J. and Brand, M. (2011). Regeneration of the adult zebrafish brain from neurogenic radial glia-type progenitors. *Development* **138**, 4831.

Kunkle, B. W. Grenier-Boley, B. Sims, R. Bis, J. C. Damotte, V. Naj, A. C. Boland, A. Vronskaya, M. Van Der Lee, S. J. Amlie-Wolf, A. et al. (2019). Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nature Genetics* **51**, 414-430.

Lambert, J.-C. Ibrahim-Verbaas, C. A. Harold, D. Naj, A. C. Sims, R. Bellenguez, C. Jun, G. DeStefano, A. L. Bis, J. C. Beecham, G. W. et al. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genetics* **45**, 1452.

Lee, J. H., Yu, W. H., Kumar, A., Lee, S., Mohan, P. S., Peterhoff, C. M., Wolfe, D. M., Martinez-Vicente, M., Massey, A. C., Sovak, G. et al. (2010). Lysosomal proteolysis and autophagy require presenilin 1 and are disrupted by Alzheimer-related PS1 mutations. *Cell* **141**, 1146-58.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* **25**, 2078-9.

Li, L., Kim, H. J., Roh, J. H., Kim, M., Koh, W., Kim, Y., Heo, H., Chung, J., Nakanishi, M., Yoon, T. et al. (2020). Pathological manifestation of the induced pluripotent stem cell-derived cortical neurons from an early-onset Alzheimer's disease patient carrying a presenilin-1 mutation (S170F). *Cell Proliferation* **53**, e12798.

Li, X., Alafuzoff, I., Soininen, H., Winblad, B. and Pei, J.-J. (2005). Levels of mTOR and its downstream targets 4E-BP1, eEF2, and eEF2 kinase in relationships with tau in Alzheimer's disease brain. *The FEBS Journal* **272**, 4211-4220.

Liao, Y., Smyth, G. K. and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)* **30**, 923-930.

Lim, A. H. L. (2015). Analysis of the subcellular localization of proteins implicated in Alzheimer's Disease. In *Genetics and Evolution*, vol. Doctor of Philosophy (PhD), pp. 235: University of Adelaide.

Lim, C.-Y. and Zoncu, R. (2016). The lysosome as a command-and-control center for cellular metabolism. *Journal of Cell Biology* **214**, 653-664.

Lumsden, A. L., Rogers, J. T., Majd, S., Newman, M., Sutherland, G. T., Verdile, G. and Lardelli, M. (2018). Dysregulation of Neuronal Iron Homeostasis as an Alternative Unifying Effect of Mutations Causing Familial Alzheimer's Disease. *Frontiers in Neuroscience* **12**, 533-533.

Luo, F., Rustay, N. R., Ebert, U., Hradil, V. P., Cole, T. B., Llano, D. A., Mudd, S. R., Zhang, Y., Fox, G. B. and Day, M. (2012). Characterization of 7- and 19-month-old Tg2576 mice using multimodal in vivo imaging: Limitations as a translatable model of Alzheimer's disease. *Neurobiology of aging* **33**, 933-944.

Luo, W., Pant, G., Bhavnasi, Y. K., Blanchard, S. G., Jr. and Brouwer, C. (2017). Pathview Web: user friendly pathway visualization and data integration. *Nucleic Acids Research* **45**, W501-W508.

Manczak, M., Park, B. S., Jung, Y. and Reddy, P. H. (2004). Differential expression of oxidative phosphorylation genes in patients with Alzheimer's disease. *NeuroMolecular Medicine* **5**, 147-162.

Marcus, C., Mena, E. and Subramaniam, R. M. (2014). Brain PET in the Diagnosis of Alzheimer's Disease. *Clinical Nuclear Medicine* **39**.

Masters, C. L., Bateman, R., Blennow, K., Rowe, C. C., Sperling, R. A. and Cummings, J. L. (2015). Alzheimer's disease. *Nature Reviews Disease Primers* **1**, 15056.

Mayer, C. and Grummt, I. (2006). Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene* **25**, 6384-6391.

McCarthy, D. J., Chen, Y. and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**, 4288-4297.

Meyer, A. and Van de Peer, Y. (2005). From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *Bioessays* **27**, 937-45.

Moore, D. B., Gillentine, M. A., Botezatu, N. M., Wilson, K. A., Benson, A. E. and Langeland, J. A. (2014). Asynchronous evolutionary origins of Aβ and BACE1. *Mol Biol Evol* **31**, 696-702.

Morris, G. P., Clark, I. A. and Vissel, B. (2018). Questions concerning the role of amyloid-β in the definition, aetiology and diagnosis of Alzheimer's disease. *Acta Neuropathologica* **136**, 663-689.

Moussavi Nik, S. H., Newman, M., Wilson, L., Ebrahimie, E., Wells, S., Musgrave, I., Verdile, G., Martins, R. N. and Lardelli, M. (2015). Alzheimer's disease-related peptide PS2V plays ancient, conserved roles in suppression of the unfolded protein response under hypoxia and stimulation of γ-secretase activity. *Hum Mol Genet* **24**, 3662-3678.

Myers, A. and McGonigle, P. (2019). Overview of Transgenic Mouse Models for Alzheimer's Disease. *Curr Protoc Neurosci* **89**, e81.

Newman, M., Hin, N., Pederson, S. and Lardelli, M. (2019). Brain transcriptome analysis of a familial Alzheimer's disease-like mutation in the zebrafish presenilin 1 gene implies effects on energy production. *Molecular Brain* **12**.

Oblak, A. L., Forner, S., Territo, P. R., Sasner, M., Carter, G. W., Howell, G. R., Sukoff-Rizzo, S. J., Logsdon, B. A., Mangravite, L. M., Mortazavi, A. et al. (2020). Model organism development and evaluation for late-onset Alzheimer's disease: MODEL-AD. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* **6**, e12110.

Oddo, S., Caccamo, A., Shepherd, J. D., Murphy, M. P., Golde, T. E., Kayed, R., Metherate, R., Mattson, M. P., Akbari, Y. and LaFerla, F. M. (2003). Triple-transgenic model of Alzheimer's disease with plaques and tangles: intracellular Abeta and synaptic dysfunction. *Neuron* **39**, 409-21.

Oexle, H., Gnaiger, E. and Weiss, G. (1999). Iron-dependent changes in cellular energy metabolism: influence on citric acid cycle and oxidative phosphorylation. *Biochim Biophys Acta* **1413**, 99-107.

Ou, Y.-N., Xu, W., Li, J.-Q., Guo, Y., Cui, M., Chen, K.-L., Huang, Y.-Y., Dong, Q., Tan, L., Yu, J.-T. et al. (2019). FDG-PET as an independent biomarker for Alzheimer's biological diagnosis: a longitudinal study. *Alzheimer's Research & Therapy* **11**, 57.

Pasternak, S. H., Bagshaw, R. D., Guiral, M., Zhang, S., Ackerley, C. A., Pak, B. J., Callahan, J. W. and Mahuran, D. J. (2003). Presenilin-1, nicastrin, amyloid precursor protein, and gamma-secretase activity are co-localized in the lysosomal membrane. *Journal of Biological Chemistry* **278**, 26687-94.

Pellerin, L. and Magistretti, P. J. (1994). Glutamate uptake into astrocytes stimulates aerobic glycolysis: a mechanism coupling neuronal activity to glucose utilization. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 10625-10629.

Poisnel, G., Hérard, A.-S., El Tannir El Tayara, N., Bourrin, E., Volk, A., Kober, F., Delatour, B., Delzescaux, T., Debeir, T., Rooney, T. et al. (2012). Increased regional cerebral glucose uptake in an APP/PS1 model of Alzheimer's disease. *Neurobiology of aging* **33**, 1995-2005.

Pottier, C., Hannequin, D., Coutant, S., Rovelet-Lecrux, A., Wallon, D., Rousseau, S., Legallic, S., Paquet, C., Bombois, S., Pariente, J. et al. (2012). High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. *Mol Psychiatry* **17**, 875-879.

Prasad, H. and Rao, R. (2018). Amyloid clearance defect in ApoE4 astrocytes is reversed by epigenetic correction of endosomal pH. *Proceedings of the National Academy of Sciences* **115**, E6640.

Quiroz, Y. T., Schultz, A. P., Chen, K., Protas, H. D., Brickhouse, M., Fleisher, A. S., Langbaum, J. B., Thiyyagura, P., Fagan, A. M., Shah, A. R. et al. (2015). Brain Imaging and Blood Biomarker Abnormalities in Children With Autosomal Dominant Alzheimer Disease. **72**, 912.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47-e47.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2009). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* **26**, 139-140.

Rossor, M., Fox, N., Beck, J., Campbell, T. and Collinge, J. (1996). Incomplete penetrance of familial Alzheimer's disease in a pedigree with a novel presenilin-1 gene mutation. *The Lancet* **347**, 1560.

Saito, T., Matsuba, Y., Mihira, N., Takano, J., Nilsson, P., Itohara, S., Iwata, N. and Saido, T. C. (2014). Single App knock-in mouse models of Alzheimer's disease. *Nature Neuroscience* **17**, 661-663.

Saito, T., Matsuba, Y., Yamazaki, N., Hashimoto, S. and Saido, T. C. (2016). Calpain Activation in Alzheimer's Model Mice Is an Artifact of APP and Presenilin Overexpression. *The Journal of Neuroscience* **36**, 9933.

Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A. L., Nada, S. and Sabatini, D. M. (2010). Ragulator-Rag Complex Targets mTORC1 to the Lysosomal Surface and Is Necessary for Its Activation by Amino Acids. *Cell* **141**, 290-303.

Sannerud, R., Esselens, C., Ejsmont, P., Mattera, R., Rochin, L., Tharkeshwar, Arun K., De Baets, G., De Wever, V., Habets, R., Baert, V. et al. (2016). Restricted Location of PSEN2/γ-Secretase Determines Substrate Specificity and Generates an Intracellular Aβ Pool. *Cell* **166**, 193-208.

Sato, N., Hori, O., Yamaguchi, A., Lambert, J. C., Chartier-Harlin, M. C., Robinson, P. A., Delacourte, A., Schmidt, A. M., Furuyama, T., Imaizumi, K. et al. (1999). A novel presenilin-2 splice variant in human Alzheimer's disease brain tissue. *Journal of Neurochemistry* **72**, 2498-505.

Saunders, A. M., Schmader, K., Breitner, J. C., Benson, M. D., Brown, W. T., Goldfarb, L., Goldgaber, D., Manwaring, M. G., Szymanski, M. H., McCown, N. et al. (1993). Apolipoprotein E epsilon 4 allele distributions in late-onset Alzheimer's disease and in other amyloid-forming diseases. *Lancet* **342**, 710-1.

Saxton, R. A. and Sabatini, D. M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. *Cell* **168**, 960-976.

Schubert, M., Lindgreen, S. and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes* **9**, 88.

Sergushichev, A. A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv*, 060012.

Sherrington, R., Froelich, S., Sorbi, S., Campion, D., Chi, H., Rogaeva, E. A., Levesque, G., Rogaev, E. I., Lin, C., Liang, Y. et al. (1996). Alzheimer's Disease Associated with Mutations in Presenilin 2 is Rare and Variably Penetrant. *Hum Mol Genet* **5**, 985-988.

Simmen, T., Lynes, E. M., Gesson, K. and Thomas, G. (2010). Oxidative protein folding in the endoplasmic reticulum: Tight links to the mitochondria-associated membrane (MAM). *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1798**, 1465-1473.

Sims, R., Hill, M. and Williams, J. (2020). The multiplex model of the genetics of Alzheimer's disease. *Nature Neuroscience* **23**, 311-322.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545.

Sullivan, P. M., Mezdour, H., Aratani, Y., Knouff, C., Najib, J., Reddick, R. L., Quarfordt, S. H. and Maeda, N. (1997). Targeted Replacement of the Mouse Apolipoprotein E Gene with the Common Human APOE3 Allele Enhances Diet-induced Hypercholesterolemia and Atherosclerosis*. *Journal of Biological Chemistry* **272**, 17972-17980.

Sun, Y.-X., Ji, X., Mao, X., Xie, L., Jia, J., Galvan, V., Greenberg, D. A. and Jin, K. (2014). Differential Activation of mTOR Complex 1 Signaling in Human Brain with Mild to Severe Alzheimer's Disease. *Journal of Alzheimer's Disease* **38**, 437-444.

Tambini, M. D., Pera, M., Kanter, E., Yang, H., Guardia-Laguarta, C., Holtzman, D., Sulzer, D., Area-Gomez, E. and Schon, E. A. (2016). ApoE4 upregulates the activity of mitochondria-associated ER membranes. *EMBO Rep* **17**, 27-36.

R Core Team. (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.

Tellechea, P., Pujol, N., Esteve-Belloch, P., Echeveste, B., García-Eulate, M. R., Arbizu, J. and Riverol, M. (2018). Early- and late-onset Alzheimer disease: Are they the same entity? *Neurologia* **33**, 244-253.

Temitope, A., Ekaterina, R., Kurup, J. T., Beecham, G. and Christiane, R. (2021). Early-Onset Alzheimer's Disease: What Is Missing in Research? *Current Neurology and Neuroscience Reports* **21**.

Thonberg, H., Chiang, H.-H., Lilius, L., Forsell, C., Lindström, A.-K., Johansson, C., Björkström, J., Thordardottir, S., Sleegers, K., Van Broeckhoven, C. et al. (2017). Identification and description of three families with familial Alzheimer disease that segregate variants in the SORL1 gene. *Acta Neuropathologica Communications* **5**, 43.

Thordardottir, S., Rodriguez-Vieitez, E., Almkvist, O., Ferreira, D., Saint-Aubert, L., Kinhult-Ståhlbom, A., Thonberg, H., Schöll, M., Westman, E., Wall, A. et al. (2018). Reduced penetrance of the PSEN1 H163Y autosomal dominant Alzheimer mutation: a 22-year follow-up study. *Alzheimer's Research & Therapy* **10**, 45.

Tolomeo, D., Micotti, E., Serra, S. C., Chappell, M., Snellman, A. and Forloni, G. (2018). Chemical exchange saturation transfer MRI shows low cerebral 2-deoxy-D-glucose uptake in a model of Alzheimer's Disease. *Scientific Reports* **8**, 9576.

Tomiyama, T., Nagata, T., Shimada, H., Teraoka, R., Fukushima, A., Kanemitsu, H., Takuma, H., Kuwano, R., Imagawa, M., Ataka, S. et al. (2008). A new amyloid β variant favoring oligomerization in Alzheimer's-type dementia. *Annals of Neurology* **63**, 377-387.

Wang, Y.-T., Kang, M. S., Therriault, J., Pascoal, T. A., Lussier, F. Z., Savard, M., Benedet, A. L., Tissot, C., Arias, J. F., Gauthier, S. et al. (2020). APOE4 packs a punch in women: Sex-specific vulnerability for tau and neuroinflammation. *Alzheimer's & Dementia* **16**, e045098.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York.

Wilson, D. J. (2019). The harmonic mean p-value for combining dependent tests. *Proceedings of the National Academy of Sciences* **116**, 1195.

Wu, D., Lim, E., Vaillant, F., Asselin-Labat, M.-L., Visvader, J. E. and Smyth, G. K. (2010). ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics (Oxford, England)* **26**, 2176-2182.

Wu, D. and Smyth, G. K. (2012). Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Research* **40**, e133-e133.

Yambire, K. F., Rostosky, C., Watanabe, T., Pacheu-Grau, D., Torres-Odio, S., Sanchez-Guerrero, A., Senderovich, O., Meyron-Holtz, E. G., Milosevic, I., Frahm, J. et al. (2019). Impaired lysosomal acidification triggers iron deficiency and inflammation in vivo. *Elife* **8**.

Yim, W. W.-Y. and Mizushima, N. (2020). Lysosome biology in autophagy. *Cell Discovery* **6**, 6.

Young, M. D., Wakefield, M. J., Smyth, G. K. and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* **11**, R14.

Zhao, N., Ren, Y., Yamazaki, Y., Qiao, W., Li, F., Felton, L. M., Mahmoudiandehkordi, S., Kueider-Paisley, A., Sonoustoun, B., Arnold, M. et al. (2020). Alzheimer's Risk Factors Age, APOE Genotype, and Sex Drive Distinct Molecular Pathways. *Neuron* **106**, 727-742.e6.

Zhou, X., Liao, W.-J., Liao, J.-M., Liao, P. and Lu, H. (2015). Ribosomal proteins: functions beyond the ribosome. *Journal of Molecular Cell Biology* **7**, 92-104.

Zierler, K. (1999). Whole body glucose metabolism. *American Journal of Physiology-Endocrinology and Metabolism* **276**, E409-E426.

Zoncu, R., Bar-Peled, L., Efeyan, A., Wang, S., Sancak, Y. and Sabatini, D. M. (2011). mTORC1 Senses Lysosomal Amino Acids Through an Inside-Out Mechanism That Requires the Vacuolar H(+)-ATPase. *Science* **334**, 678.

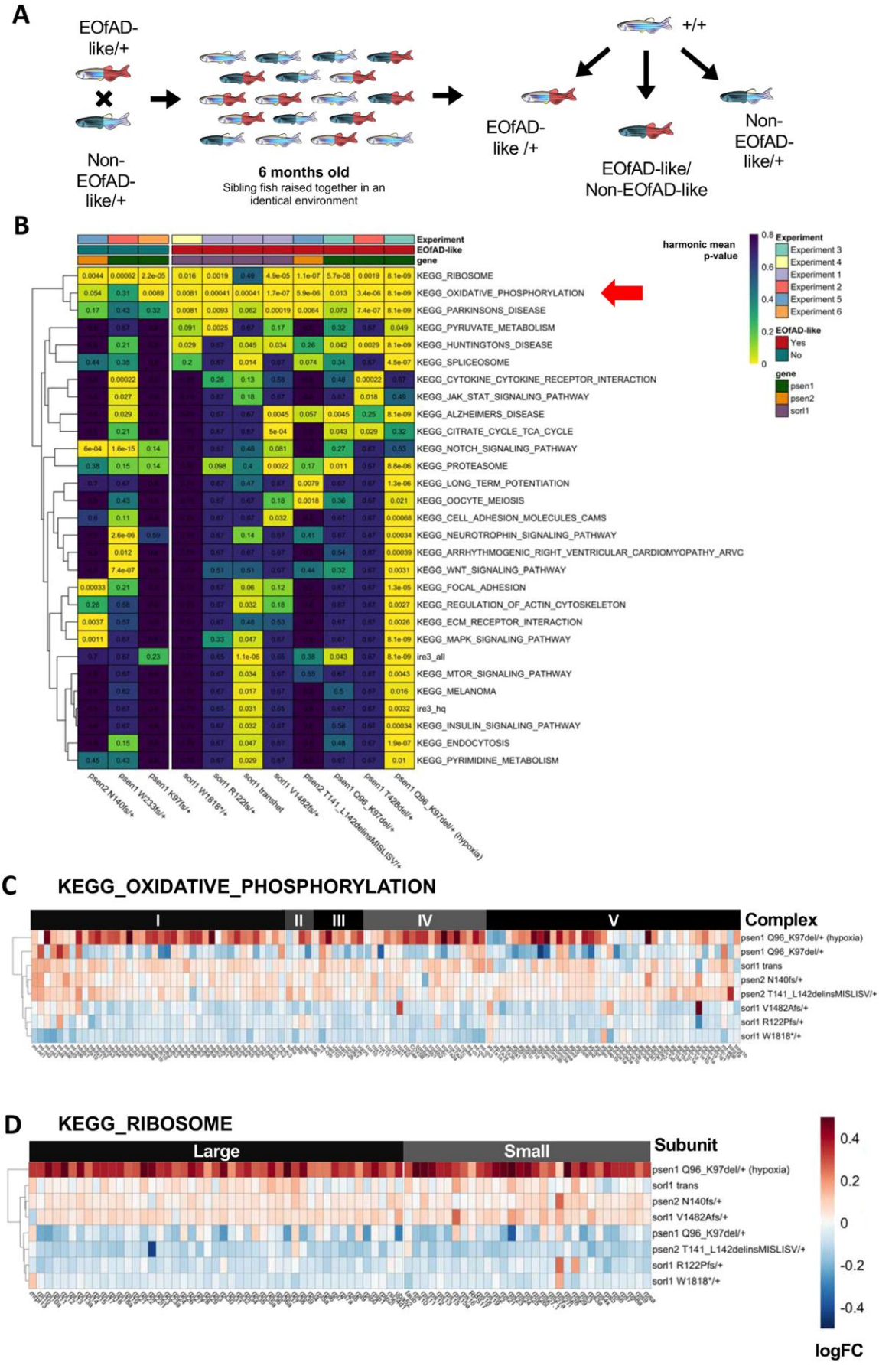


Fig 1: RNA-seq analysis of 6 month old zebrafish models of EOfAD. A. Schematic of a RNA-seq experiment using zebrafish. A single mating of a single pair of fish heterozygous for either an EOfAD-like or a non-EOfAD-like mutation results in a family heterozygous mutant, transheterozygous mutant, and wild type siblings. Comparisons made between genotypes in a RNA-seq experiment are depicted. **B.** Heatmap summary of significantly altered *KEGG* and *IRE* gene sets in zebrafish EOfAD genetic models at 6 months of age. Only gene sets significantly altered (FDR-adjusted harmonic mean p-value < 0.05) in at least two comparisons of mutant zebrafish to their corresponding wild type siblings are shown. Columns are grouped by whether or not the zebrafish genotype is EOfAD-like, while rows are clustered based on their Euclidean distance. The numbers are FDR-adjusted harmonic mean p-values. **C.** Heatmap indicating the logFC of genes in the *KEGG* gene sets for oxidative phosphorylation and, **D,** the ribosome in zebrafish mutants compared to their wild type siblings. Rows are clustered based on their Euclidean distance, while columns are grouped by the complex in the electron transport chain to which an encoded protein belongs (**C**), or whether an encoded protein forms part of the large or small ribosomal subunits (**D**). Only genes considered detectable in all RNA-seq experiments are depicted. See **Fig. S1-S6** and **Table 1** for more information on individual study designs.

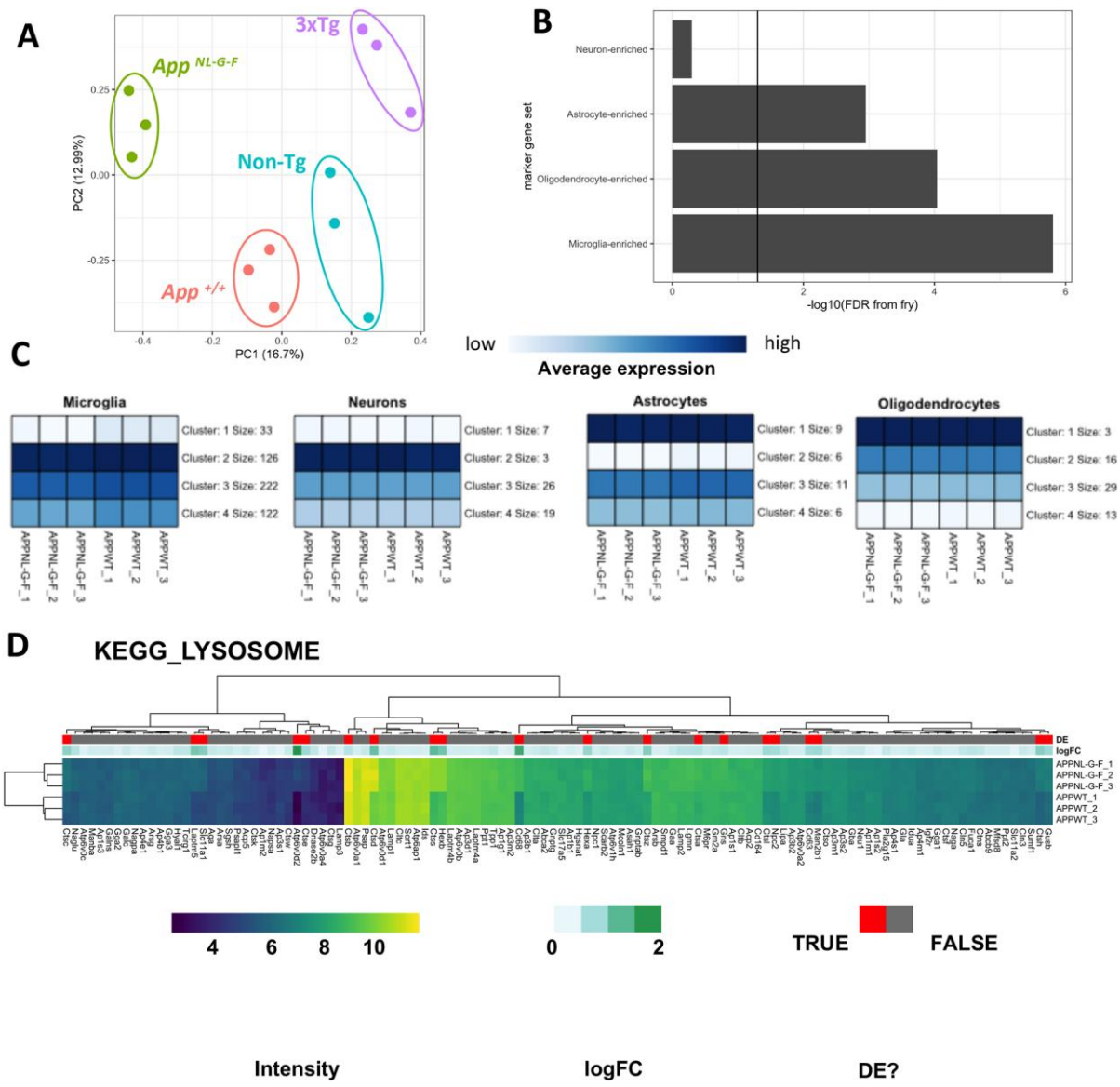


Fig 2: Microarray analysis of male, 12 month old homozygous *App*$^{NL\text{-}G\text{-}F}$ mice. **A.** Principal component (PC) analysis of brain transcriptome data from male, 12 month old homozygous *App*$^{NL\text{-}G\text{-}F}$ (n = 3), 3xTg (n = 3), *App* wild type (*App*$^{+/+}$, n = 3) and non-transgenic (non-Tg, n = 3) mice. The numbers between parentheses indicate the percentage of variation in the dataset explained by a PC. Each point represents a sample, which is coloured by genotype. **B.** Bar chart showing the FDR-adjusted p-value (directional hypothesis) from *fry* on marker genes of neurons, oligodendrocytes, astrocytes and microglia in *App*$^{NL\text{-}G\text{-}F}$ relative to wild type. **C.** Heatmap indicating the expression (intensity) of genes within these marker gene sets summarised using K-means (K = 4). **D.** Heatmap showing the expression of genes in the *KEGG_LYSOSOME* gene set, clustered by their Euclidean distance. Each gene is labelled in red if it was identified as differentially expressed, and the magnitude of the fold change (logFC) is shown in green.

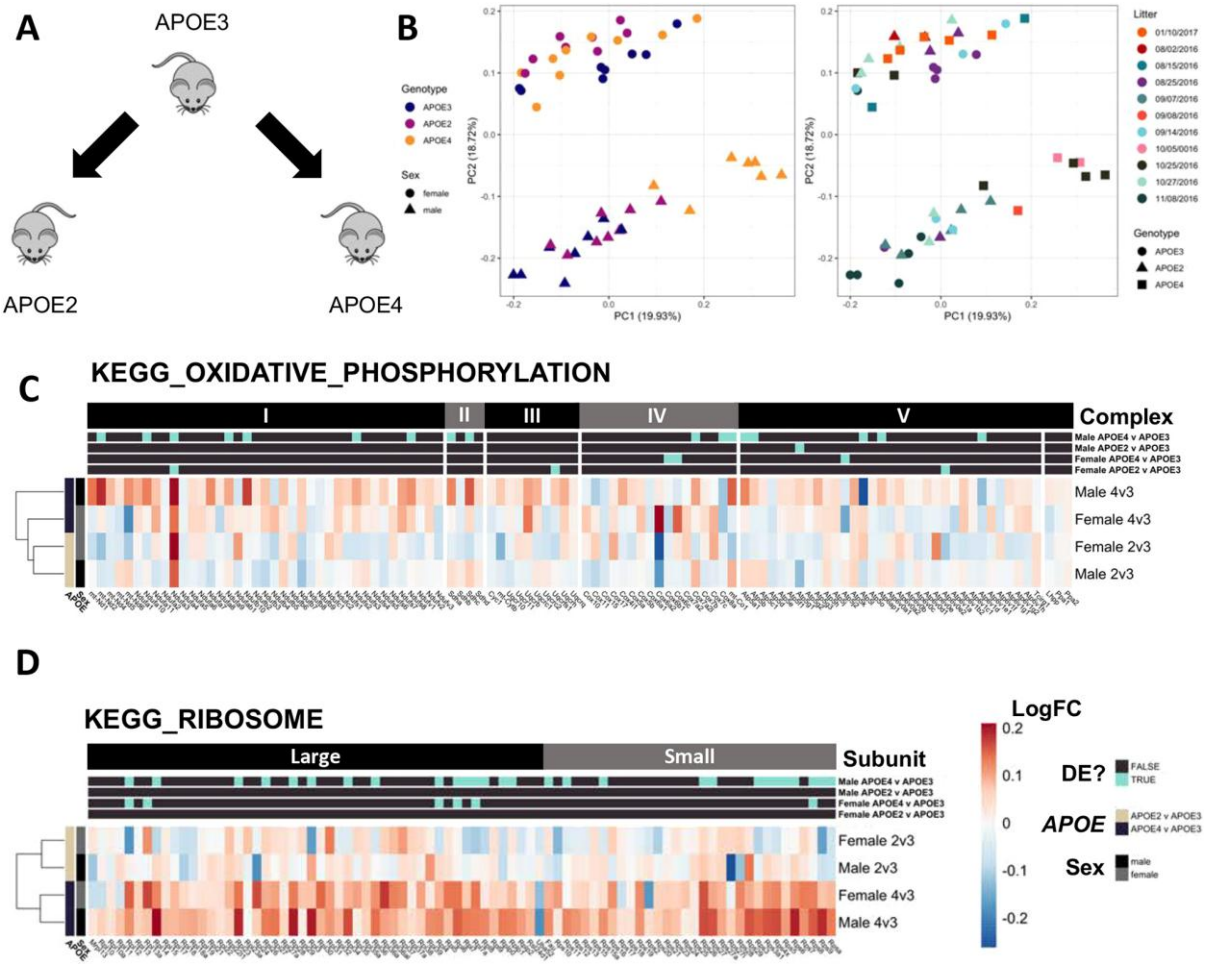


Fig 3: RNA-seq analysis of 3 month old APOE-TR mice. A. Visual representation of comparison of APOE4 (n = 7 males and 9 females) or APOE2 (n = 8 males and 8 females) mice to APOE3 (n = 8 males and 8 females). This was performed for both male and female mice separately. **B.** Principal component analysis (PCA) of three month old APOE-TR mice. Principal component 1 (PC1) is plotted against PC2. The numbers between parentheses indicate the percentage of variation in the dataset explained by a PC. In the left graph, each point represents a sample, which is coloured by *APOE* genotype, and shaped by sex. In the right plot, each point is coloured according to litter (implied from the date of birth of each mouse), and shaped by *APOE* genotype. **C** Heatmap showing the logFC of genes in the *KEGG_OXIDATIVE_PHOSPHORYLATION* and **D** *KEGG_RIBOSOME* gene sets in APOE-TR mice. Rows are clustered based on their Euclidean distance, while columns are grouped by the complex in the electron transport chain to which an encoded protein belongs (**C**), or whether an encoded protein forms part of the large or small ribosomal subunits (**D**). Genes are labelled in blue above whether they were classified as differentially expressed (FDR < 0.05) in the differential gene expression analysis in the listed comparisons.

Table 1: Summary of zebrafish RNA-seq experiments. For more detailed descriptions of study designs, see **Fig. S1-S6**. EOfAD: early-onset familial Alzheimer's disease. fAI: familial acne inversa.

Gene	Mutation(s)	Total number of zebrafish analysed	Sexes analysed	Comment	Accession	Ref
psen1	W233fs (fAI-like) and T428del (EOfAD-like)	24	Males and Females	High read depth (between 61 and 110 million reads per sample)	GSE164466	(Barthelson et al., 2021a)
psen1	Q96_K97del (EOfAD-like)	32	Males and females	Aged zebrafish and the effect of acute hypoxia treatment was also included	GSE149149	(Hin et al., 2020b; Newman et al., 2019)
psen1	K97fs	12	Females only	Aged zebrafish were also included	PRJEB24858	(Hin et al., 2020a)
psen2	T141_L142delinsMISLISV (EOfAD-like) and N140fs (not EOfAD-like)	15	Males and females	-	GSE158233	(Barthelson et al., 2021b)
sorl1	W1818*	12	Males and Females	-	GSE156167	(Barthelson et al., 2021c)
sorl1	V1482Afs (EOfAD-like) and R122Pfs (EOfAD-like)	24	Males and Females	The transheterozygous genotype was also analysed. Initially, R122Pfs was stated to be a putative null mutation. A recent case study now shows that this mutation is probably EOfAD-like.	GSE151999	(Barthelson et al., 2020b)

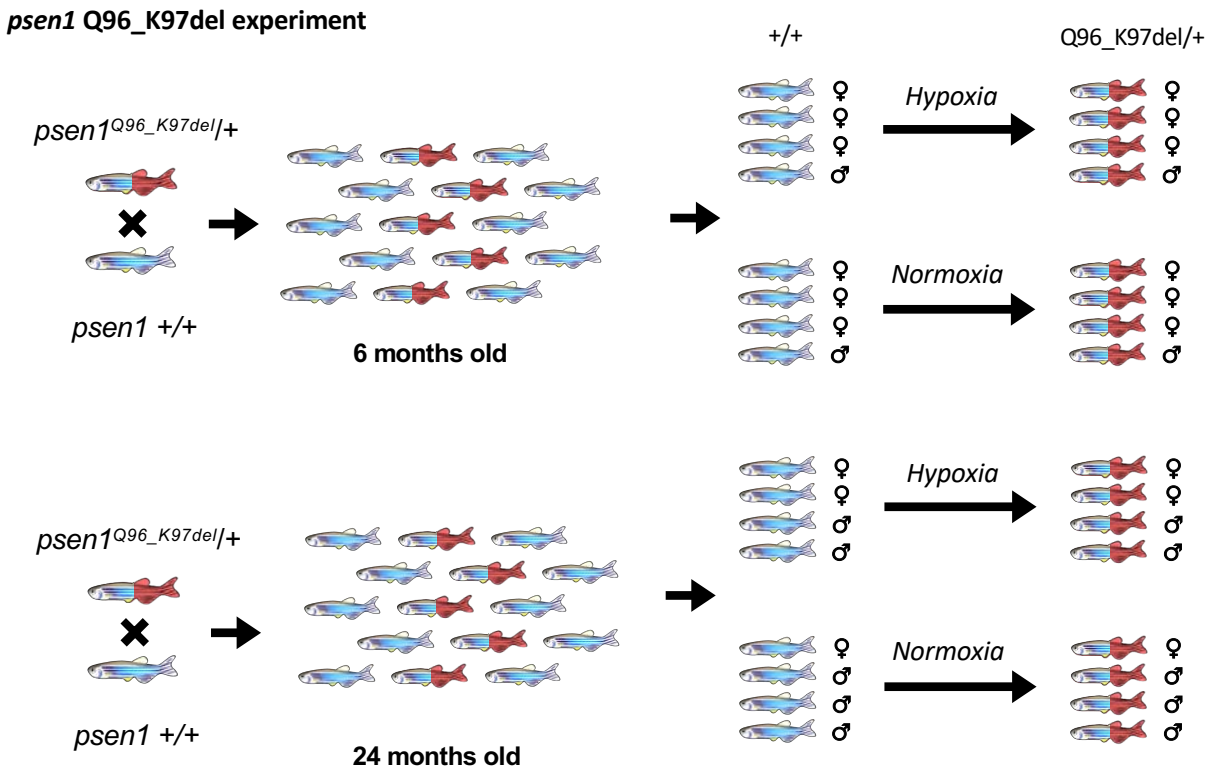


Fig. S1. Two families of zebrafish were generated by mating a wild type fish with a fish heterozygous for the Q96_K97del mutation of *psen1*, resulting in families of fish either heterozygous for the Q96_K97del mutation, or wild type. These families were raised together in single tanks until 6 or 24 months of age. Then, subsets of the families were genotyped using allele-specific polymerase chain reactions (PCRs), followed by hypoxia treatment. Then fish were sacrificed and n = 4 fish per genotype and treatment were subject to RNA-seq analysis.

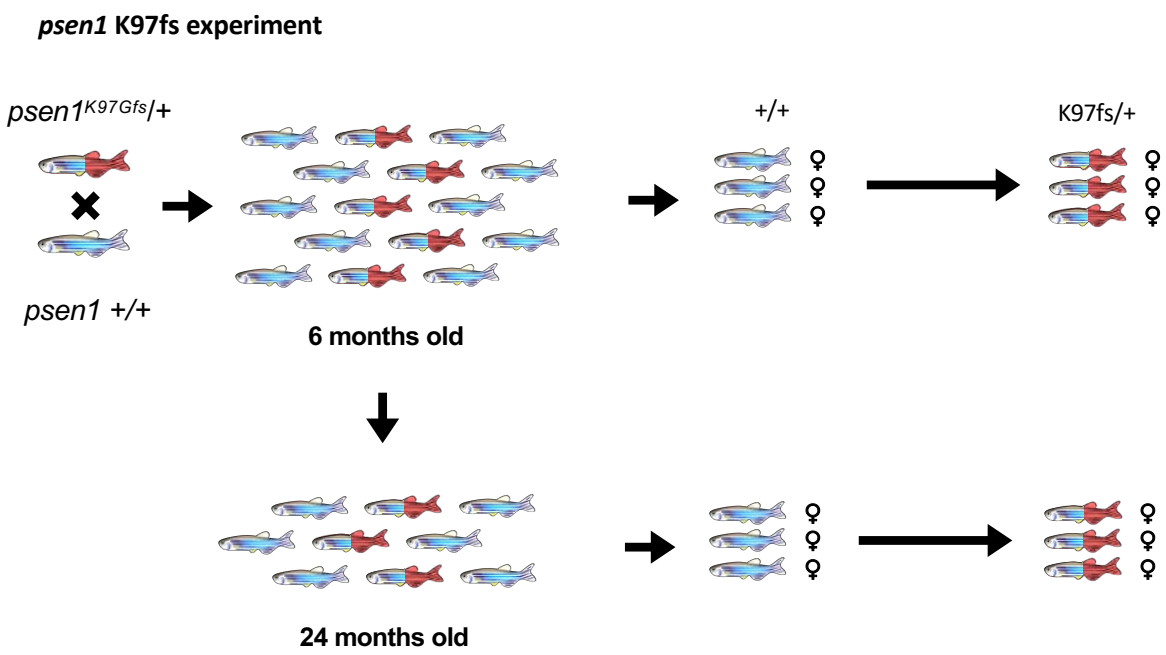


Fig. S2. A wild type fish was mated with a fish heterozygous for the K97Gfs mutation of *psen1*, resulting in a family fish either heterozygous for the K97Gfs mutation, or wild type. This family of fish was raised together in a single tank until 6 months of age. Then, a subset of the family were genotyped using allele-specific polymerase chain reactions (PCRs), then fish were sacrificed and brains were removed for RNA-seq. The remaining fish in the tank were allowed to develop until 24 months of age where this was repeated to generate the aged samples for RNA-seq.

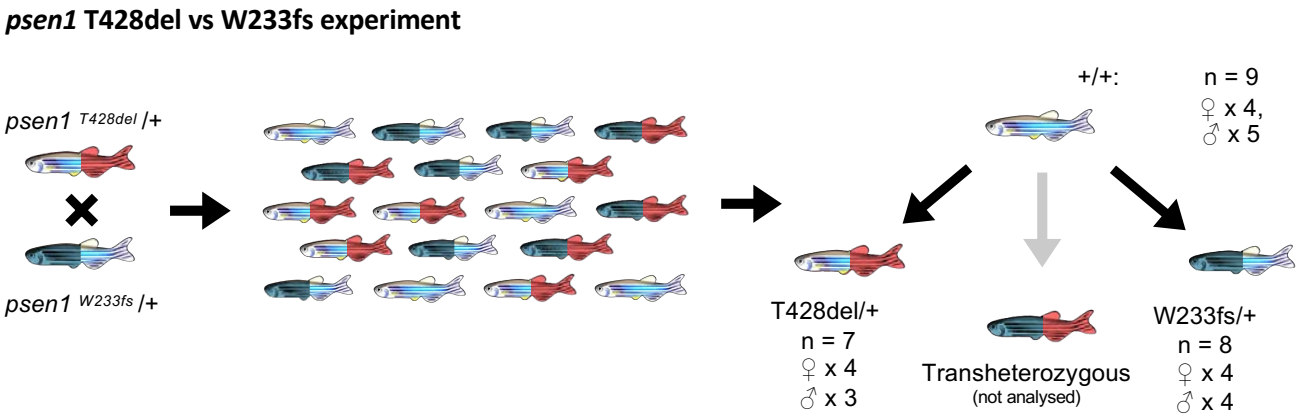


Fig. S3. A fish heterozygous for the T428del (EOfAD-like) mutation of *psen1* was mated with a fish heterozygous for the W233fs mutation (similar to the P242fs mutation of human *PSEN1* causative for familial acne inversa) to generate a family of sibling fish with four possible *psen1* genotypes. This family was raised together in the same tank until 6 months of age, at which time 50 fish were randomly selected and sacrificed in a loose ice slurry. Fish were genotyped after sacrifice by allele specific PCRs. Then n = 8 fish per genotype (4 females and 4 males) were subjected to RNA-seq analysis. During the RNA-seq analysis, one T428del/+ fish was identified to be incorrectly genotyped and was re-classified as wild type.

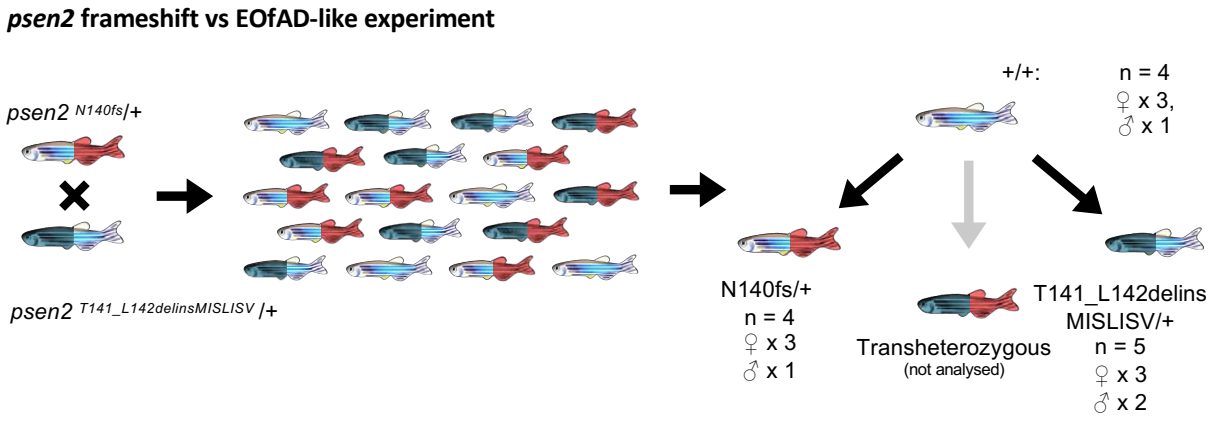


Fig. S4. A fish heterozygous for the N140fs (not EOfAD-like) mutation of *psen2* was mated with a fish heterozygous for the T141_L142delinsMISLISV (EOfAD-like) mutation of *psen2* to generate a family of sibling fish with four possible *psen2* genotypes. This family was raised together in the same tank until 6 months of age, at which time 24 fish were randomly selected and sacrificed in a loose ice slurry (to allow for n = 5 of each genotype in the RNA-seq analysis). Fish were genotyped after sacrifice by allele specific PCRs. Then n = 5 fish per genotype (3 females and 2 males) were subjected to RNA-seq analysis. During the RNA-seq analysis, one wild type fish was an obvious outlier and was omitted from the rest of the analysis, and one N140s/+ sample has been incorrectly genotyped and was also omitted.

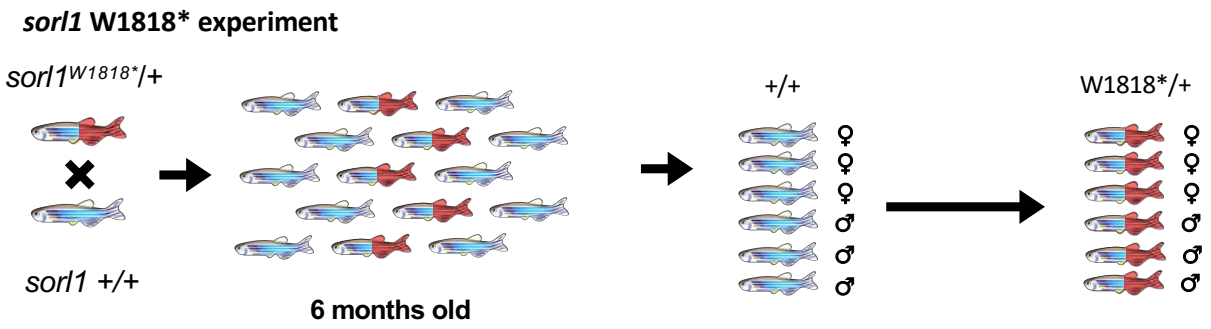


Fig. S5. A fish heterozygous for the W1818* (EOfAD-like) mutation of *sorl1* was mated with a wild type fish to generate a family of sibling fish with two possible *sorl1* genotypes. This family was raised together in the same tank until 6 months of age, at which time 20 fish were randomly selected and sacrificed in a loose ice slurry (to allow for n = 6 of each genotype in the RNA-seq analysis). Fish were genotyped after sacrifice by allele specific PCRs. Then n = 3 fish per genotype and sex were subjected to RNA-seq analysis.

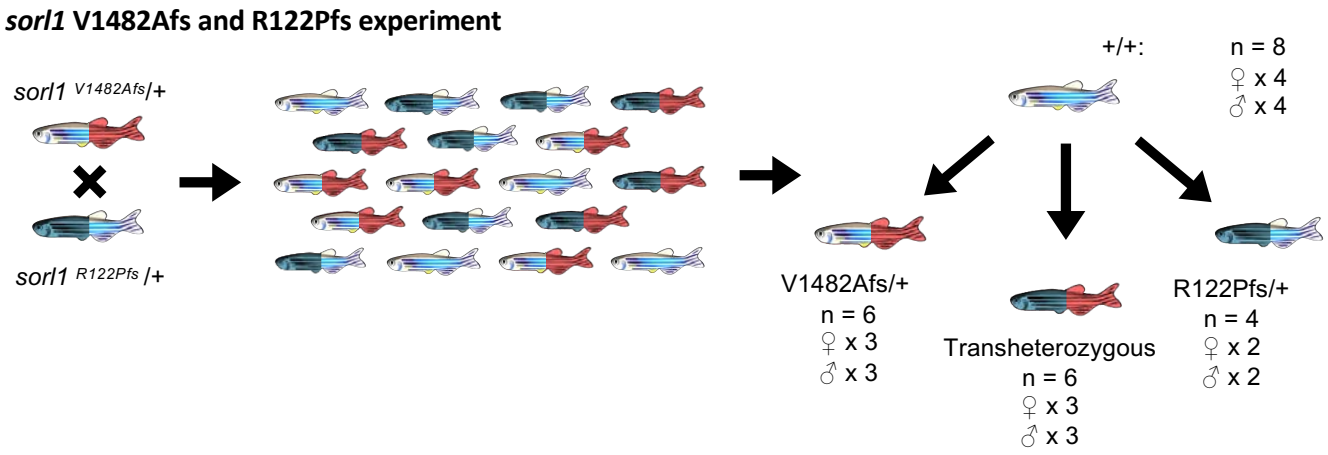


Fig. S6. A fish heterozygous for the V1482Afs (EOfAD-like) mutation of *sorl1* was mated with a fish heterozygous for the R122Pfs mutation to generate a family of sibling fish with four possible *sorl1* genotypes. This family was raised together in the same tank until 6 months of age, at which time 50 fish were randomly selected and sacrificed in a loose ice slurry (to allow for n = 6 of each genotype in the RNA-seq analysis). Fish were genotyped after sacrifice by allele specific PCRs. During the RNA-seq analysis, two R122Pfs/+ fish were identified to be incorrectly genotyped and were re-classified as wild type.

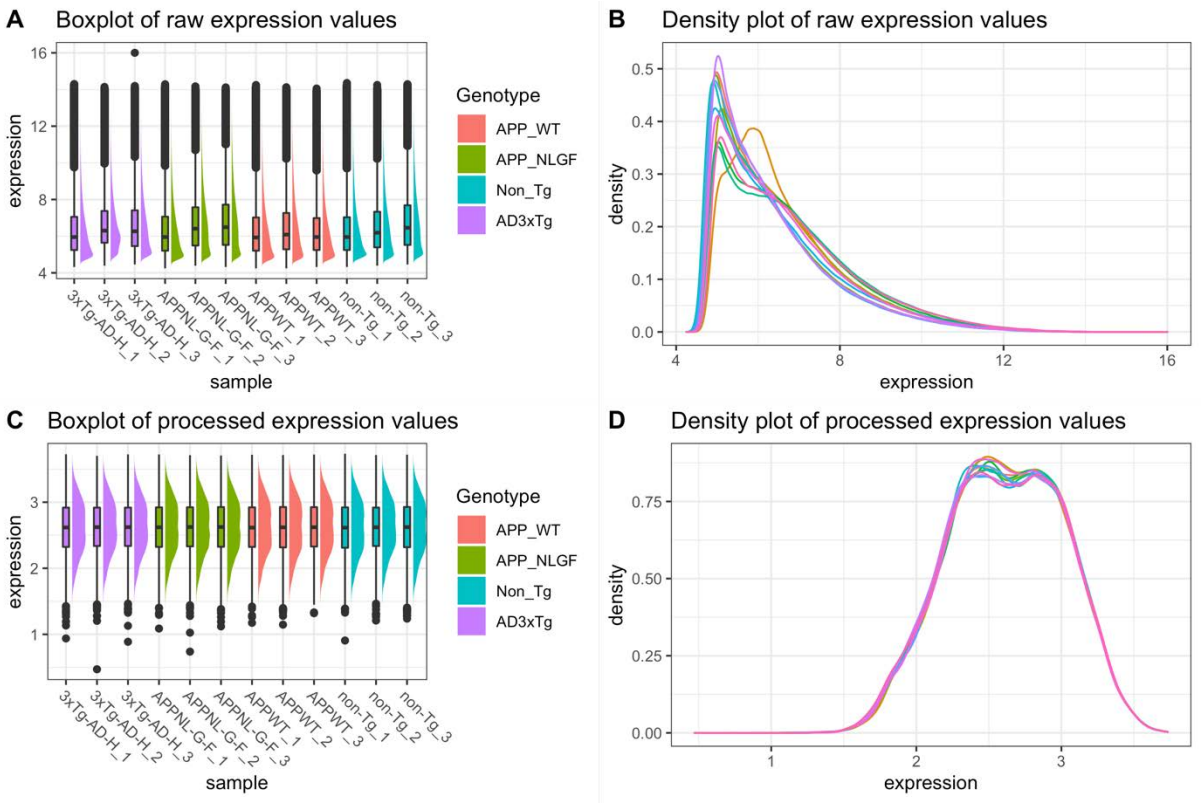


Fig. S7. A) Boxplot and **B)** density plot of the raw intensity data of GSE92926. **C)** Boxplot and **D)** density plot of the intensity data after rma normalisation and filtering for lowly expressed and multi-mapping probes.

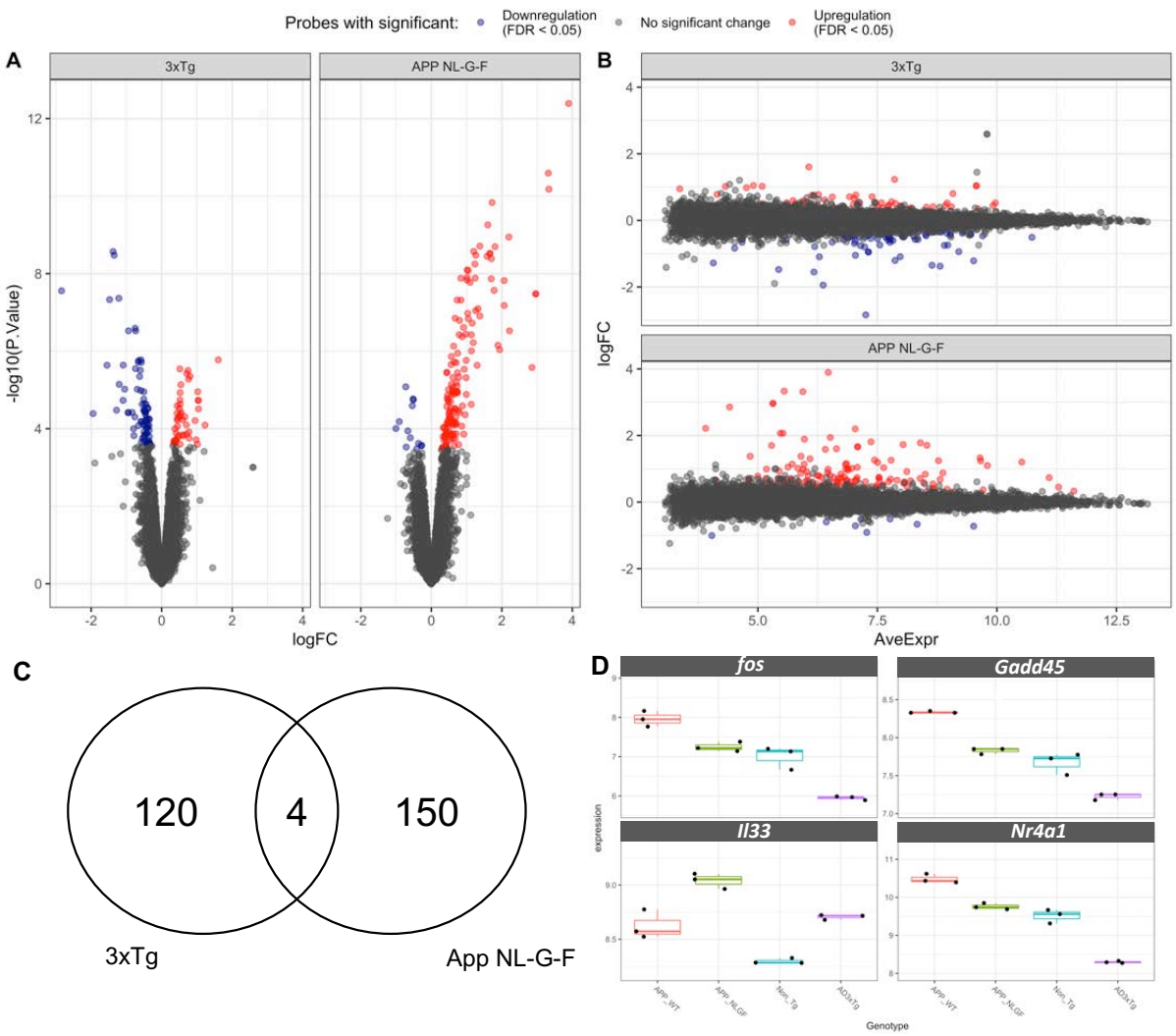
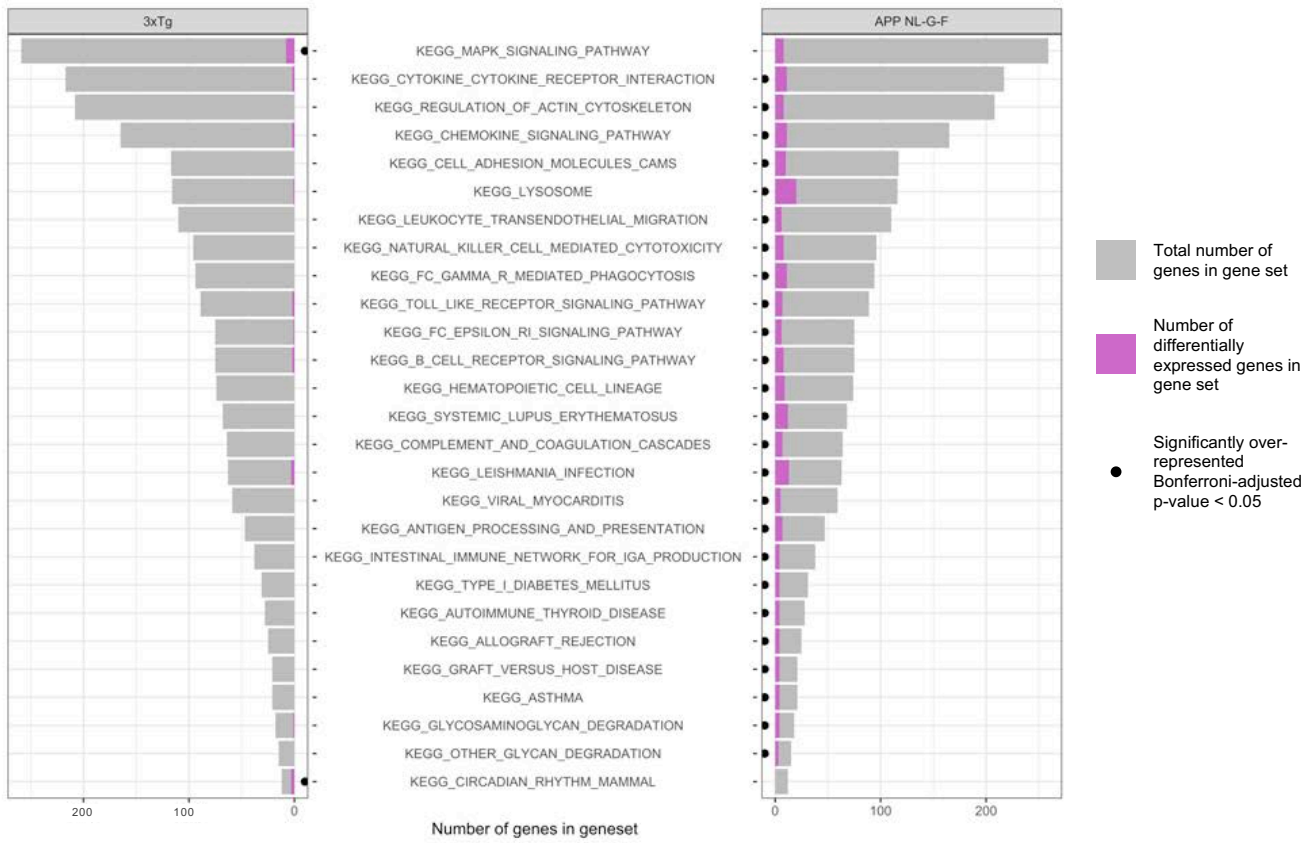


Fig. S8. Differential gene expression analysis of *APP* mutation mouse models. **A)** Volcano plot and **B)** MD plot of the changes to gene expression in $App^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ and 3xTg mice relative to controls. **C)** Venn diagram showing four genes are identified as differentially expressed (DE) in both comparisons. **D)** Boxplot of expression values of the four shared DE genes. The FDR-adjusted p-values as calculated from the differential gene expression analysis using *limma* are indicated.

A


3xTg
APP NL-G-F
KEGG_MAPK_SIGNALING_PATHWAY
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION
KEGG_REGULATION_OF_ACTIN_CYTOSKELETON
KEGG_CHEMOKINE_SIGNALING_PATHWAY
KEGG_CELL_ADHESION_MOLECULES_CAMS
KEGG_LYSOSOME
KEGG_LEUKOCYTE_TRANSENDOTHELIAL_MIGRATION
KEGG_NATURAL_KILLER_CELL_MEDIATED_CYTOTOXICITY
KEGG_FC_GAMMA_R_MEDIATED_PHAGOCYTOSIS
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY
KEGG_FC_EPSILON_RI_SIGNALING_PATHWAY
KEGG_B_CELL_RECEPTOR_SIGNALING_PATHWAY
KEGG_HEMATOPOIETIC_CELL_LINEAGE
KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES
KEGG_LEISHMANIA_INFECTION
KEGG_VIRAL_MYOCARDITIS
KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION
KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION
KEGG_TYPE_I_DIABETES_MELLITUS
KEGG_AUTOIMMUNE_THYROID_DISEASE
KEGG_ALLOGRAFT_REJECTION
KEGG_GRAFT_VERSUS_HOST_DISEASE
KEGG_ASTHMA
KEGG_GLYCOSAMINOGLYCAN_DEGRADATION
KEGG_OTHER_GLYCAN_DEGRADATION
KEGG_CIRCADIAN_RHYTHM_MAMMAL
200
100
0
0
100
200
Number of genes in geneset
Total number of genes in gene set
Number of differentially expressed genes in gene set
Significantly over-represented Bonferroni-adjusted p-value < 0.05

B

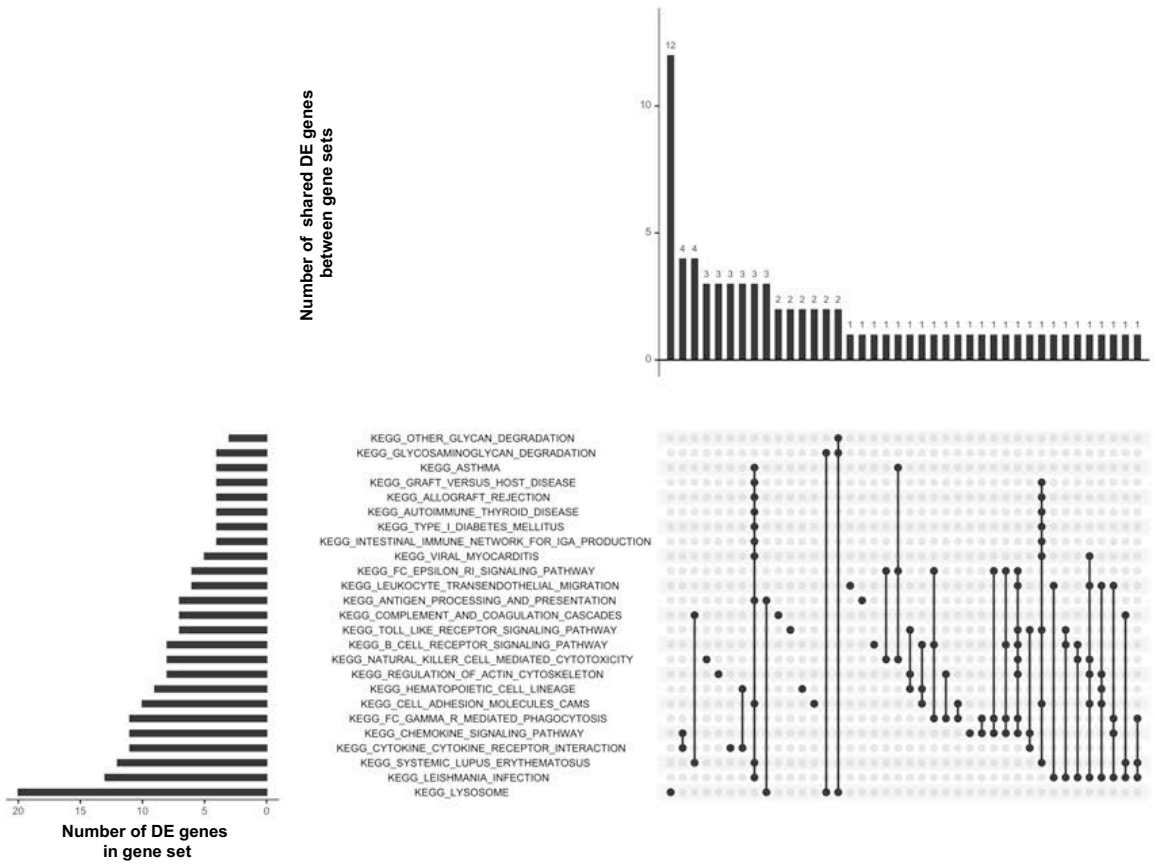

Number of shared DE genes between gene sets
12
10
5
0
4 4
3 3 3 3 3 3
2 2 2 2 2 2
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
KEGG_OTHER_GLYCAN_DEGRADATION
KEGG_GLYCOSAMINOGLYCAN_DEGRADATION
KEGG_ASTHMA
KEGG_GRAFT_VERSUS_HOST_DISEASE
KEGG_ALLOGRAFT_REJECTION
KEGG_AUTOIMMUNE_THYROID_DISEASE
KEGG_TYPE_I_DIABETES_MELLITUS
KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION
KEGG_VIRAL_MYOCARDITIS
KEGG_FC_EPSILON_RI_SIGNALING_PATHWAY
KEGG_LEUKOCYTE_TRANSENDOTHELIAL_MIGRATION
KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY
KEGG_B_CELL_RECEPTOR_SIGNALING_PATHWAY
KEGG_NATURAL_KILLER_CELL_MEDIATED_CYTOTOXICITY
KEGG_REGULATION_OF_ACTIN_CYTOSKELETON
KEGG_HEMATOPOIETIC_CELL_LINEAGE
KEGG_CELL_ADHESION_MOLECULES_CAMS
KEGG_FC_GAMMA_R_MEDIATED_PHAGOCYTOSIS
KEGG_CHEMOKINE_SIGNALING_PATHWAY
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION
KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS
KEGG_LEISHMANIA_INFECTION
KEGG_LYSOSOME
20
15
10
5
0
Number of DE genes in gene set

Fig. S9. Over-representation analysis of *APP* mutation mouse models. **A)** Pyramid bar plot indicating the number of genes in the significantly enriched KEGG and IRE gene sets in $App^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ and 3xTg mice. Only gene sets with a Bonferroni adjusted p-value from *kegga* are shown (and are indicated by a black dot). The total numbers of genes in these gene sets are shown by grey bars, while the numbers of significantly differentially expressed (DE) genes in these gene sets are shown in magenta. **B)** Upset plot indicating the high degree of overlap of DE genes across the significantly enriched gene sets in $App^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ mice.

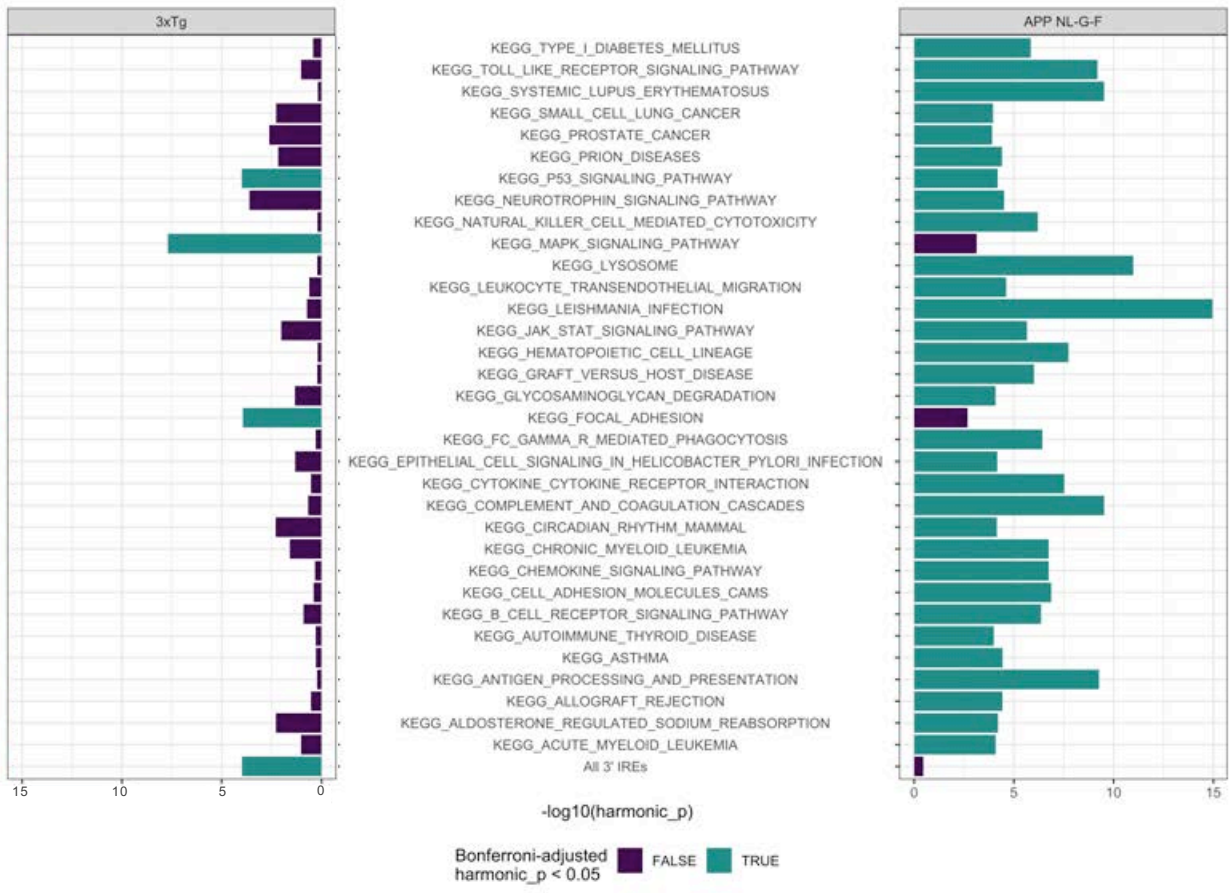


Fig. S10. Ranked-list enrichment testing. Summary of significantly enriched *KEGG* and *IRE* gene sets in *App*$^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ and 3xTg mice. Gene sets are coloured according to whether they were below the threshold of a Bonferroni-adjusted harmonic mean p-value of < 0.05

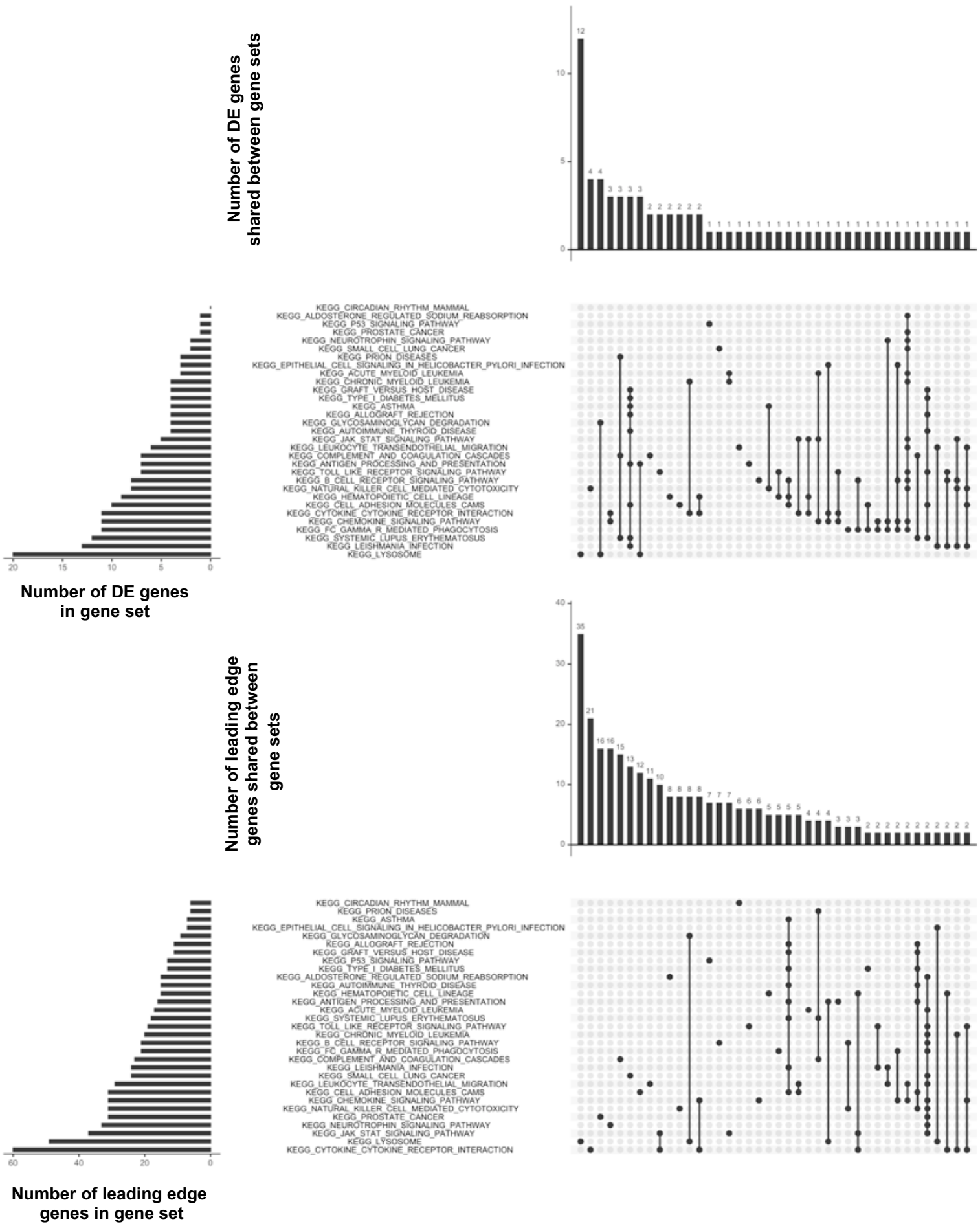


Fig. S11. Upset plots indicating the overlap of differentially expressed (DE, upper) and leading edge (lower) genes across the significantly altered gene sets in $^{AppNL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ mice. DE genes were identified using *limma,* and the leading edge genes were obtained from the *GSEA* algorithm.

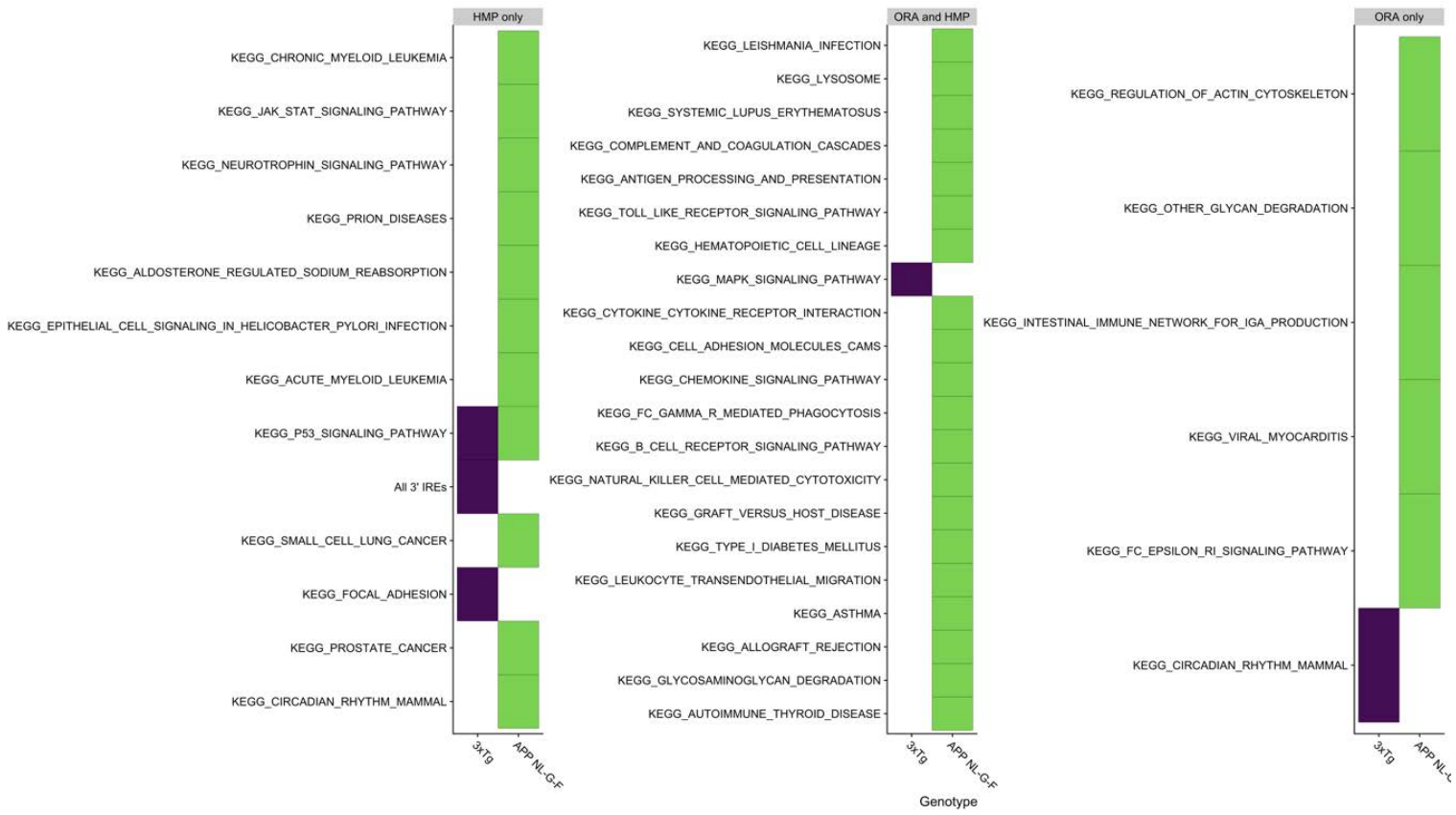
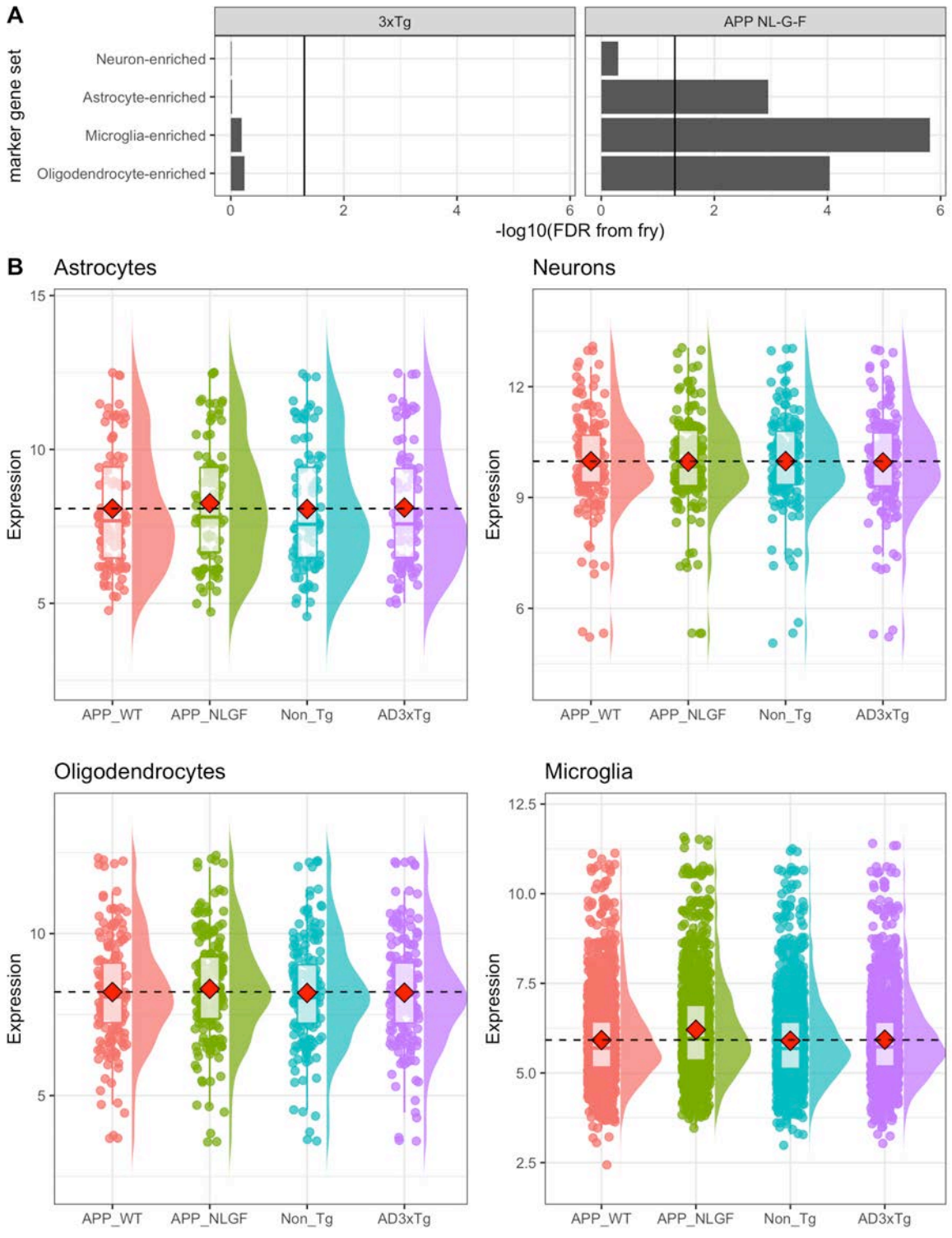


Fig. S12. Summary of significantly altered gene sets in *App*$^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ and 3xTg mice. Gene sets only found to be altered by calculation of the harmonic mean p-value (HMP) are shown on the left. Gene sets only found to be altered by over-representation analysis (ORA) using *kegga* are shown on the right. Gene sets found to be altered in both types of enrichment are shown in the middle. Gene sets are coloured according to the comparison in which they are significantly altered.



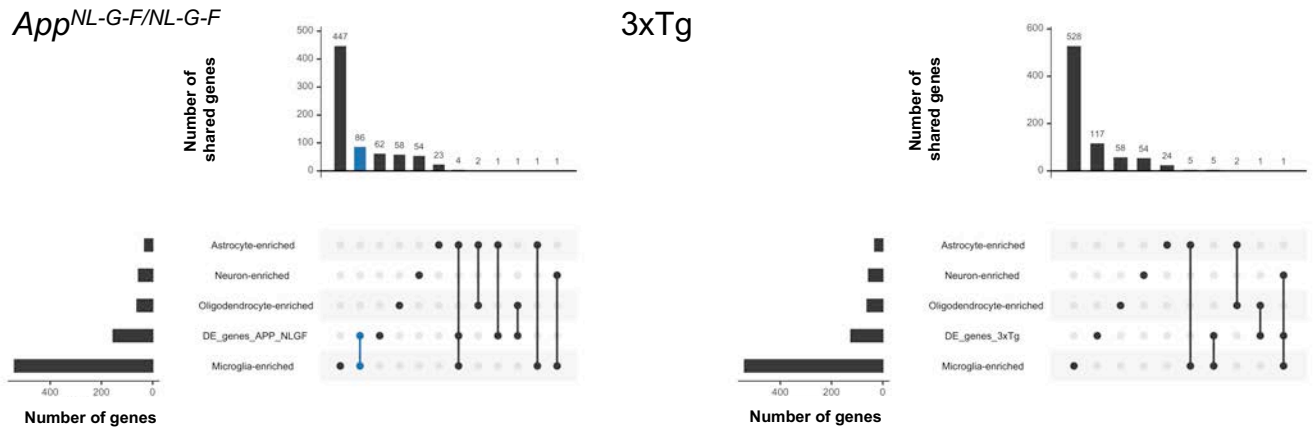


Fig. S13. Proportions of cell types in *App*$^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ mice are altered. A) Gene set testing using *fry* on marker gene sets of neurons, astrocytes, oligodendrocytes and microglia from. The black line indicates an FDR-adjusted p-value of 0.05. **B)** Distribution of intensities of the marker genes across genotypes. The boxplots indicate summary statistics. The mean intensity value for each genotype is indicated by the red diamonds. To assist with visualisation of the increased expression of marker genes in *App*$^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ mice, the mean expression in *App*$^{+/+}$ mice is also shown as a black dashed line. **C)** Upset plot showing the overlap of DE genes found in *App*$^{NL\text{-}G\text{-}F/NL\text{-}G\text{-}F}$ (left) and 3xTg (right) mice with the cell type marker gene sets.

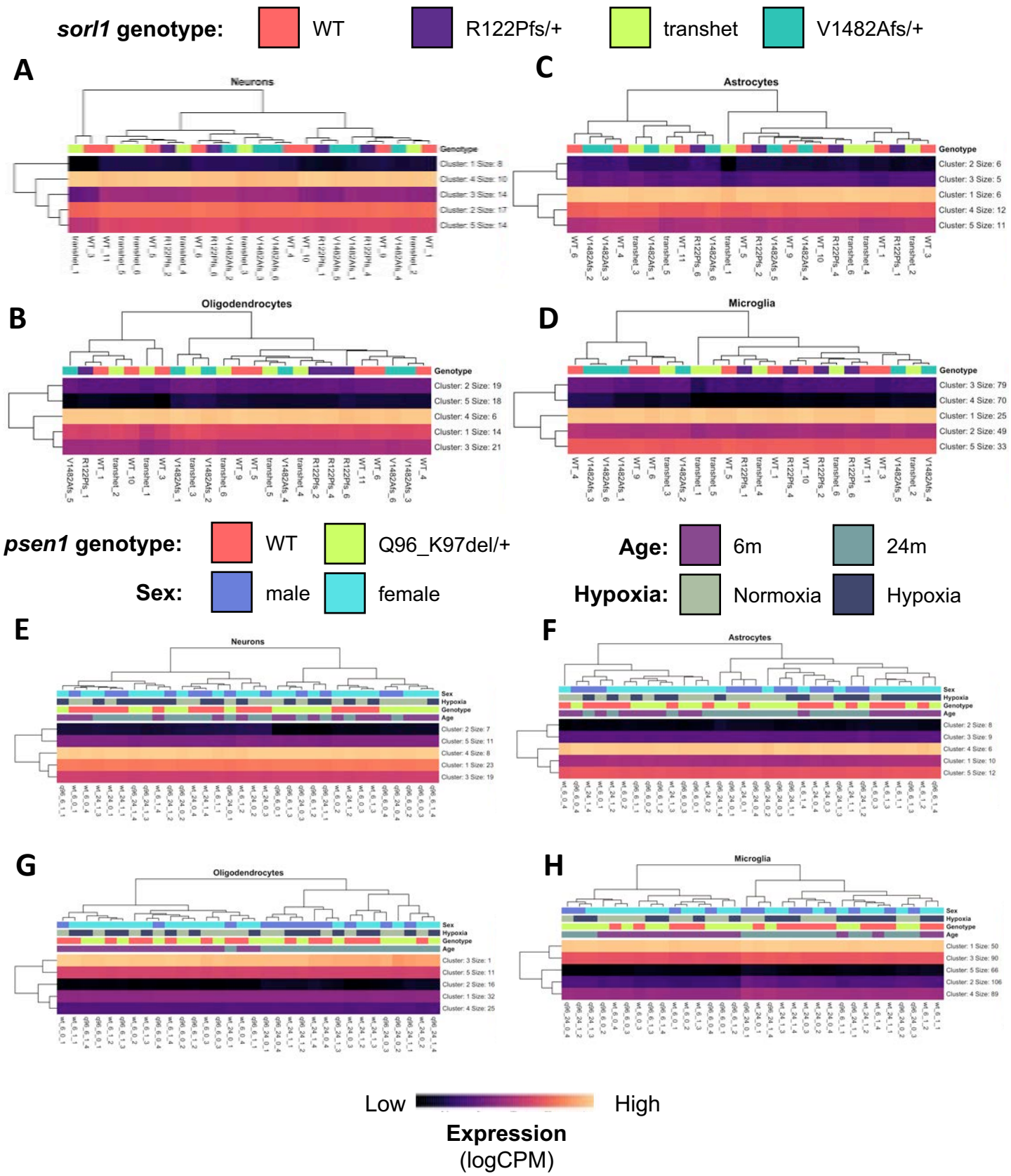


Fig. S14: Changes to cell type proportions are not observed in the brains of young, knock-in zebrafish models of EOfAD mutations. Expression values (log2 counts per million, logCPM) of marker genes of **A)** neurons (63 genes), **B)** astrocytes (40 genes), **C)** oligodendrocytes (78 genes) and **D)** microglia (256 genes) in *sorl1* knock-in mutant zebrafish. LogCPM values are also shown for *psen1* Q96_K97del mutant zebrafish in **E-H**. Rows represent clusters of genes with similar gene expression values summarised with k-means (k = 5). Columns represent samples, and are labelled with metadata of each experiment (see legends). Rows and columns are clustered based on their Euclidean distance. Samples do not cluster by genotype (or by genotype within age or treatment groups) in either experiment, meaning that expression of cell-type specific marker genes are consistent across genotypes.

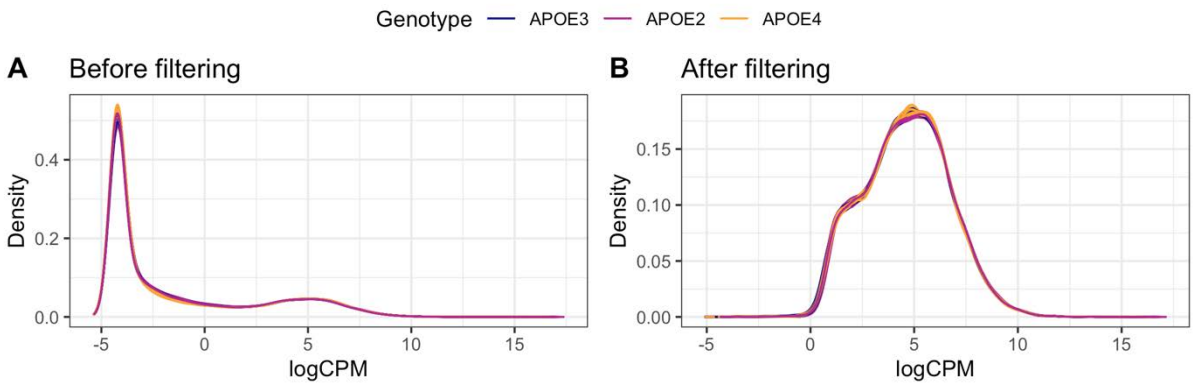


Fig. S15. The density of the log2 counts per million (logCPM) values detected in 3 month old APOE-TRmouse brain samples is shown before filtering in **A),** then after omitting samples with a logCPM of < 2 in at least one third of the samples in **B)**.

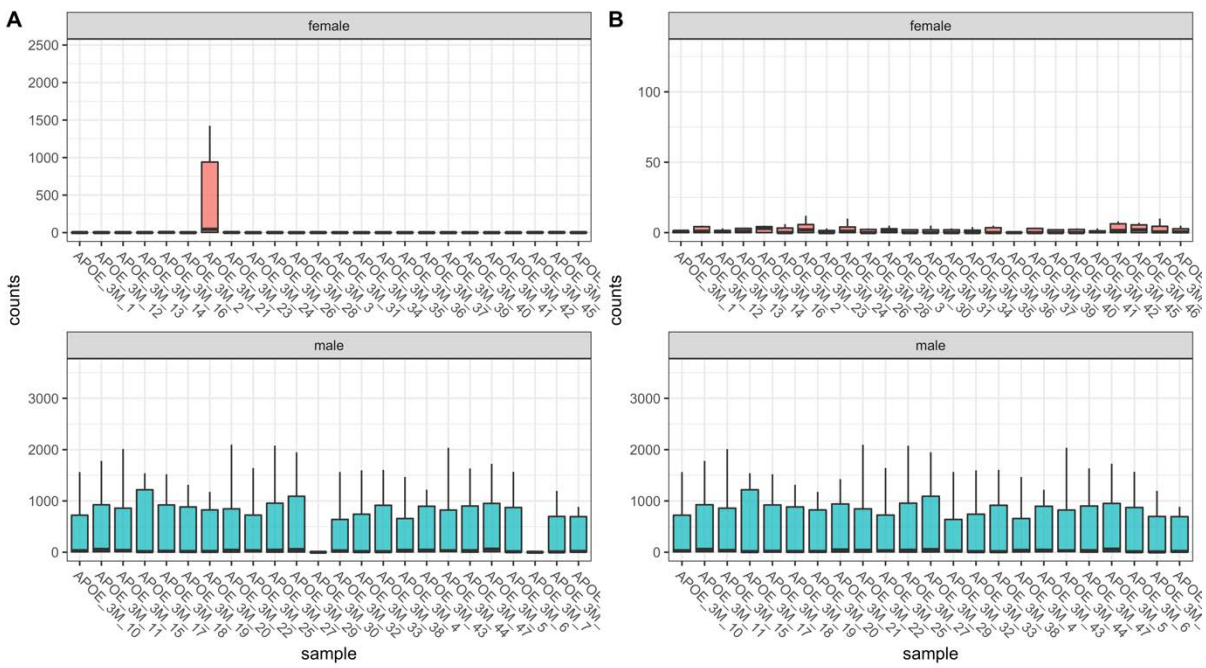


Fig. S16. Assessment of expression of genes from the Y-chromosome. A) Boxplots showing the summary statistics of the number of reads aligning to male-specific genes (located on the Y-chromosome) in the cortex of 3 month old APOE-TR mice, grouped by the metadata obtained from the AD Knowledge Portal database. Sample APOE_3M_21 appears to be a male sample as it expresses genes from the Y chromosome. Samples APOE_3M_30 and APOE_3M_7 appear to be female as they do not express genes from the Y-chromosome. **B)** Number of reads aligning to male-specific genes after correcting the sex of the samples.

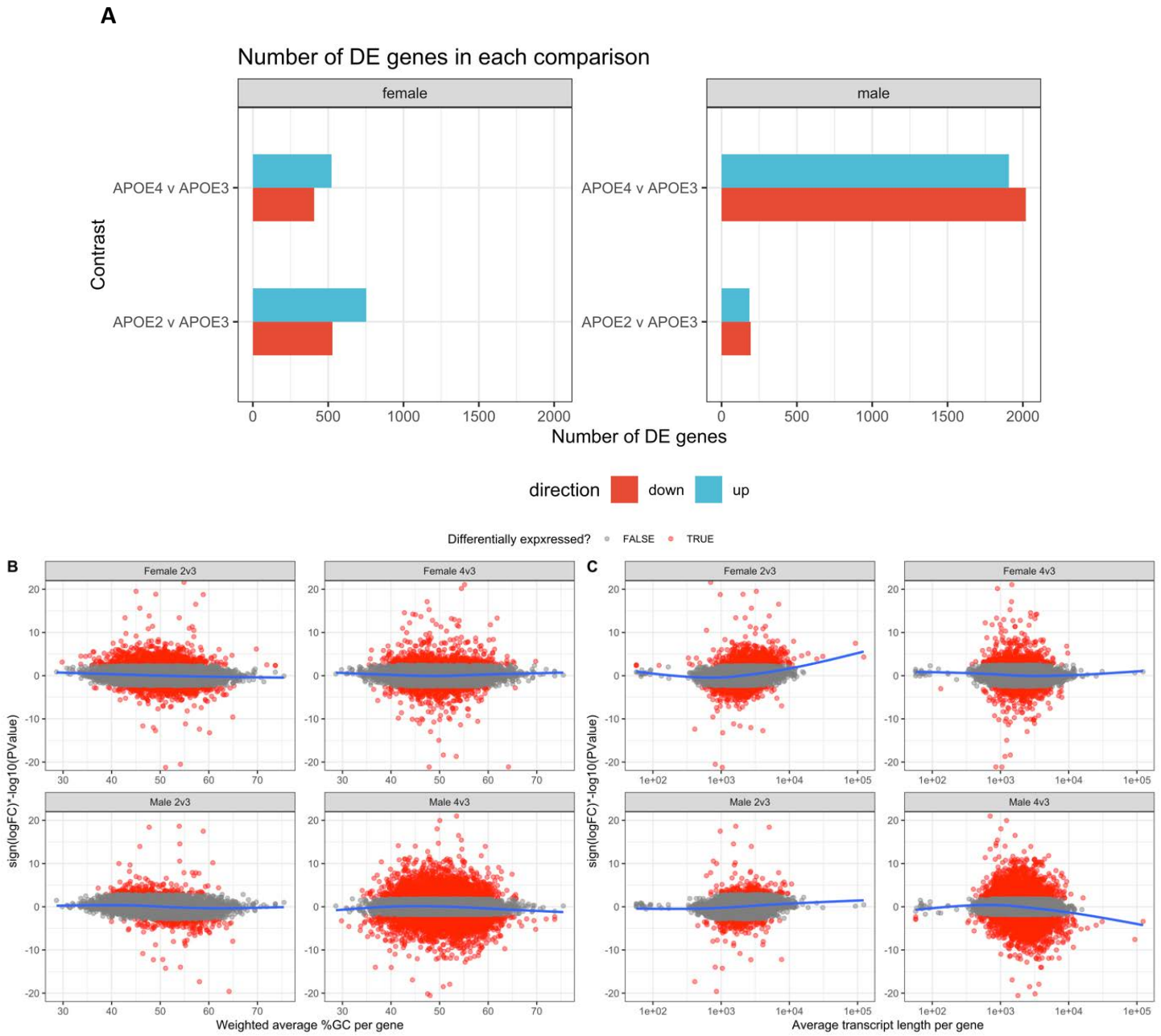


Fig. S17. 4: Initial differential expression analysis. A) Bar chart showing the number of differentially expressed genes (DE) in each comparison of the APOE4 or APOE2 genotype to APOE3. **B)** A ranking statistic per gene was calculated as the sign of the logFC multiplied by the negative log10 of the p-value from the likelihood ratio tests in *edgeR*. This was plotted against a weighted (by transcript length) average %GC content per gene and **C)** average transcript length. The blue generalised additive model fit (gam) lines are not centred on 0, indicating a bias.

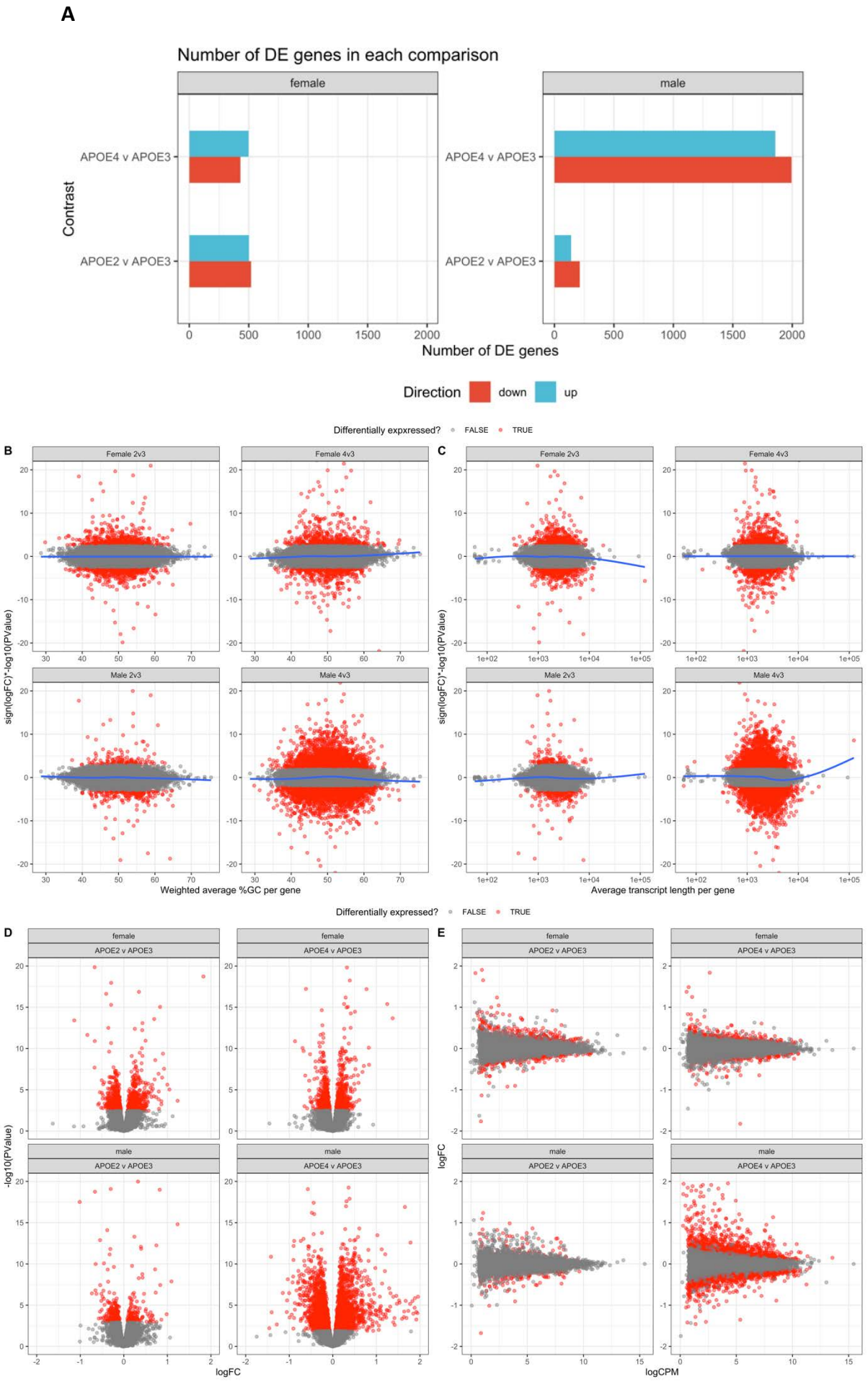
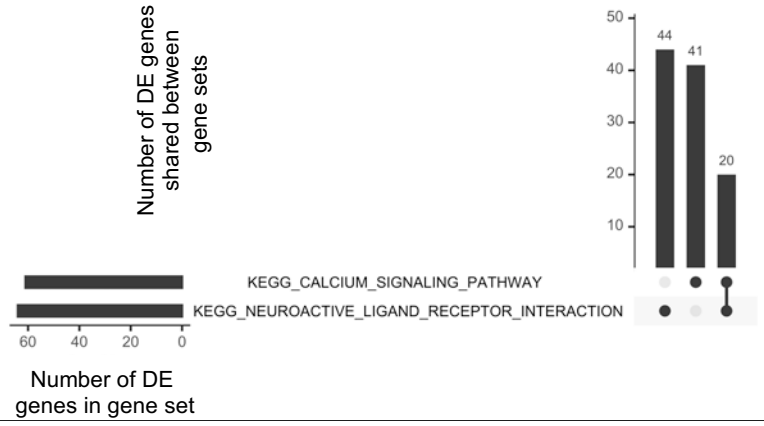
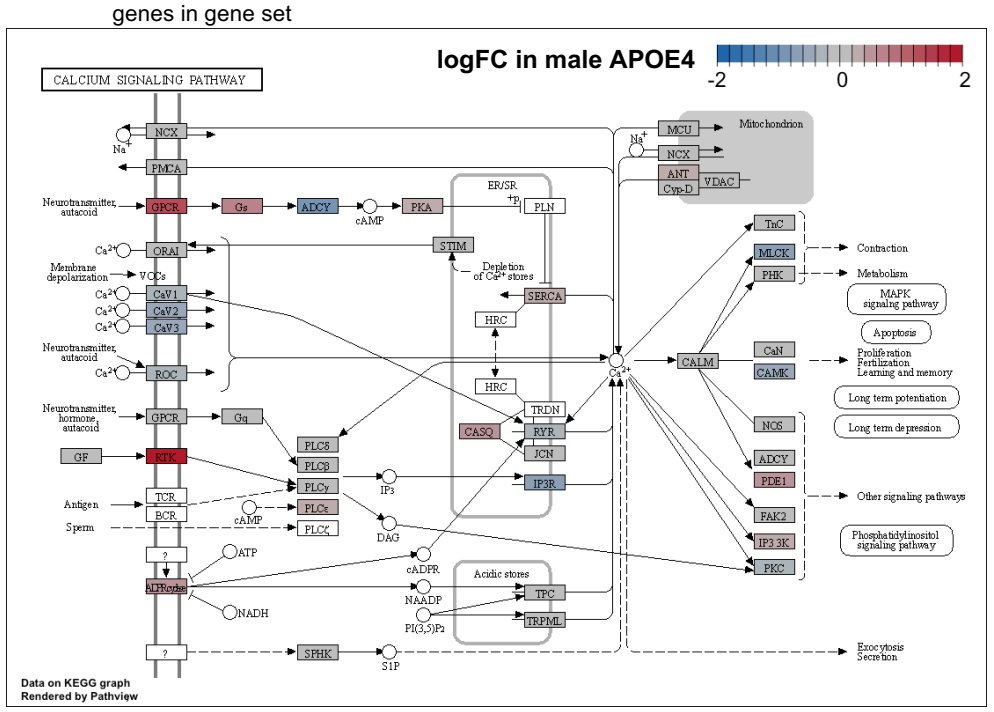


Fig. S18. Differential gene expression analysis after *cqn*. A) Number of genes identified to be differentially expressed (DE) after *cqn*. **B)** Improvement of observed bias between %GC content and **C)** gene length for differential expression after *cqn*. The remaining bias for transcript length in the female APOE2 and male APOE4 comparisons appear to be only driven by a small number of genes and can be ignored. **D)** Volcano plots and **E)** mean difference (MD) plots of the changes to gene expression observed due to homozygosity for the APOE4 or APOE2 alleles relative to APOE3 in male and female mice. The limits of the x-axis in **D)** and the y-axis in **E)** are constrained to -2 and 2, and of the y-axis in **D)** to between 0 and 20, for visualisation purposes.

A



B



C

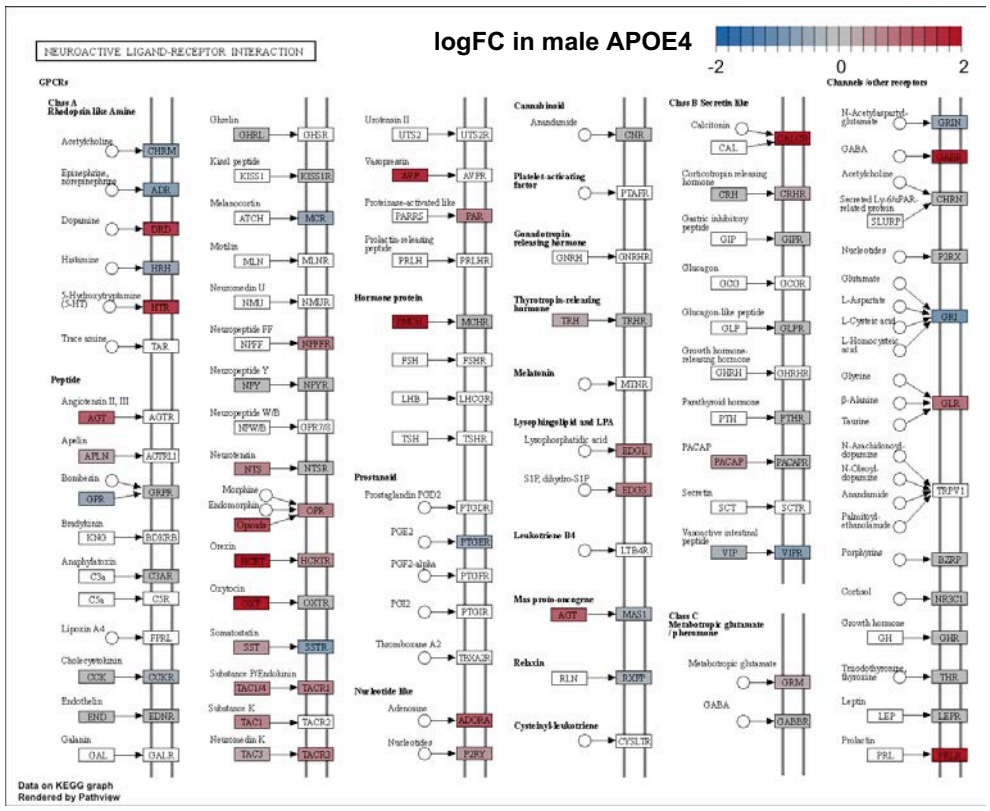


Fig. S19. Enrichment analysis within the lists of differentially expressed (DE) genes in APOE4 mice. **A)** Upset plot indicating the overlap of DE genes in male APOE4 samples for the two significantly enriched gene sets. **B)** Pathview visualisation of the logFC in male APOE4 samples for the *KEGG_CALCIUM_SIGNALING_PATHWAY* gene set and **C)** *KEGG_NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION* gene set.

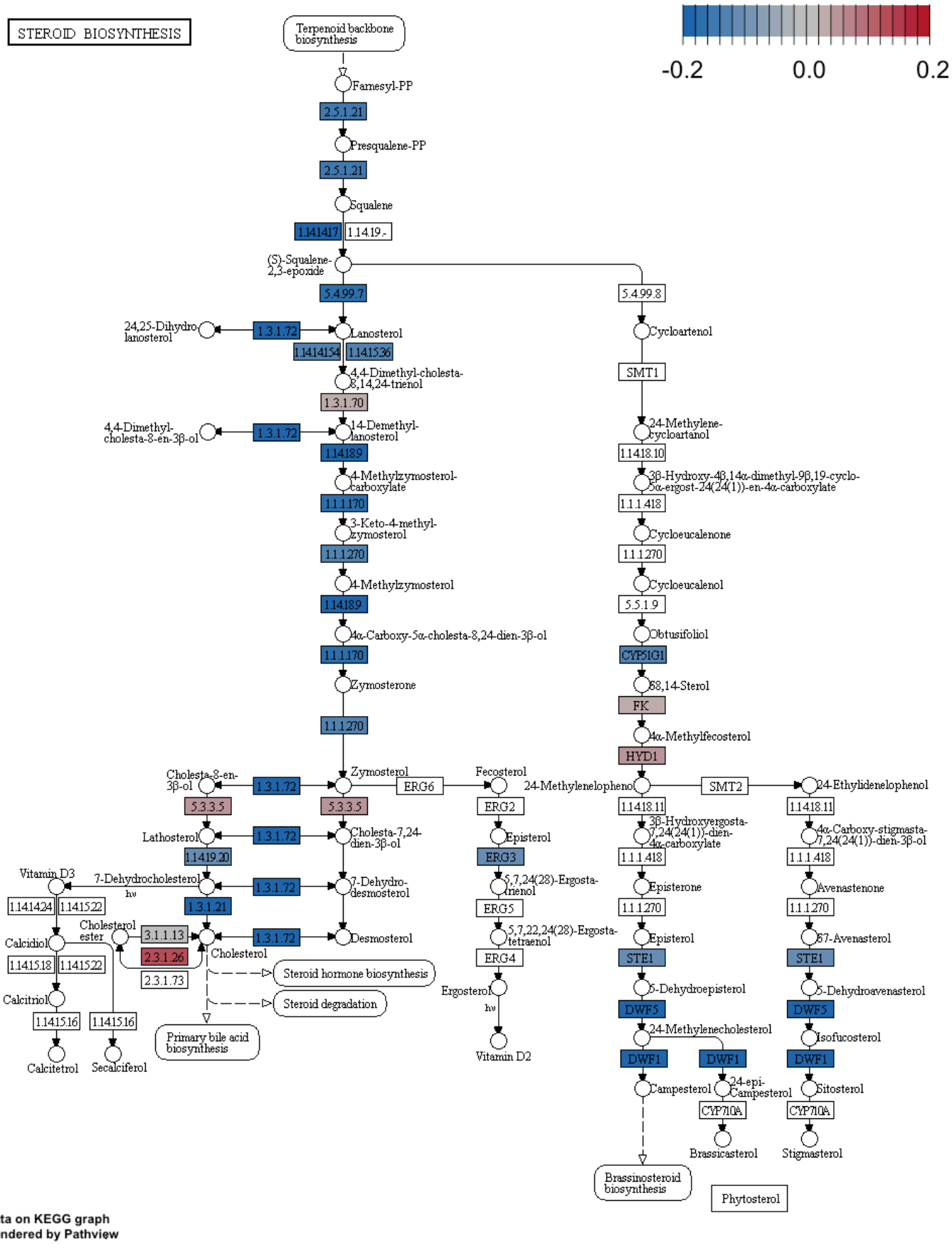


Fig. S20. Pathview visualisation indicating the logFC of genes in the *KEGG_STEROID_BIOSYNTHESIS* gene set in female APOE2 mice.

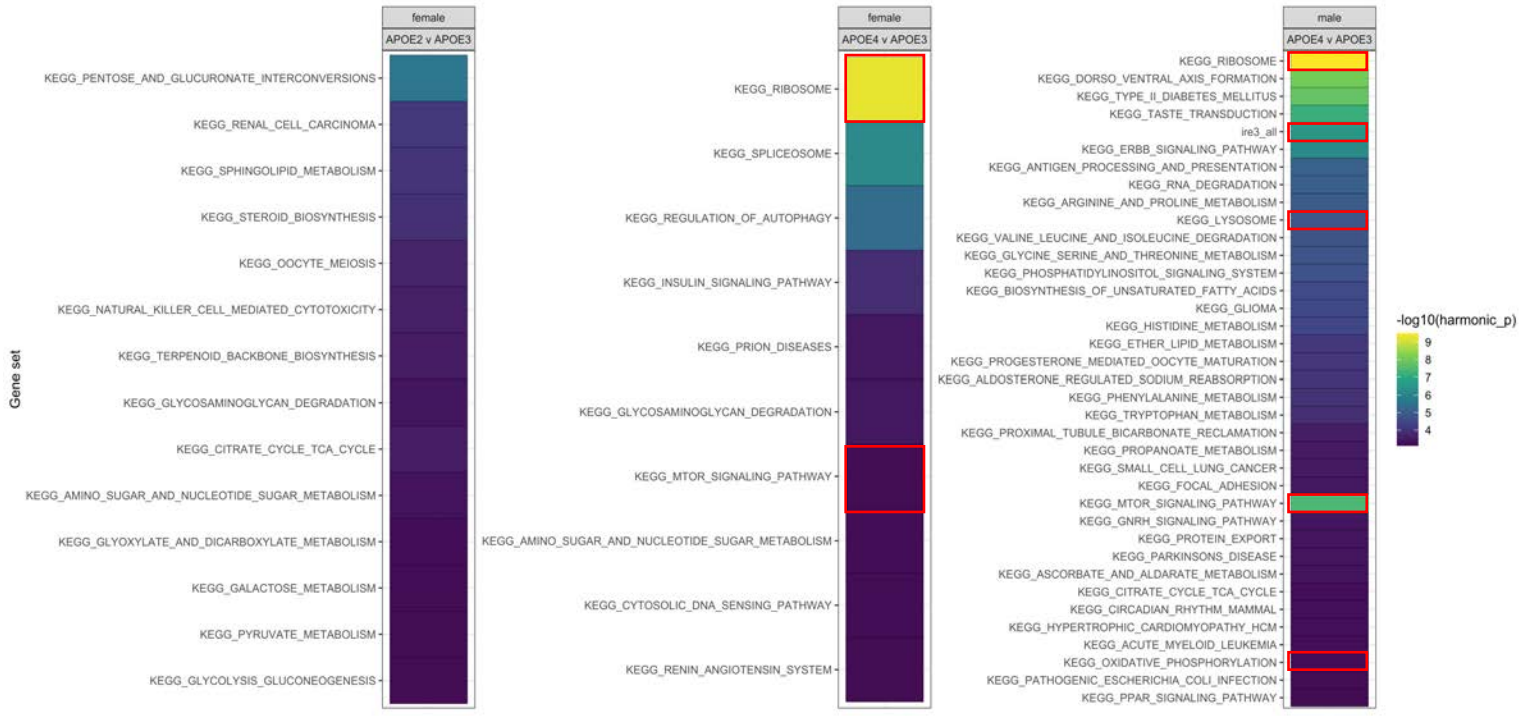


Fig. S21. Heatmap indicating gene sets with a FDR-adjusted harmonic mean p-value < 0.01 in APOE-TR mice at 3 months of age. Gene sets of interest are highlighted with a red box. Note that no gene sets were found to contain an FDR-adjusted harmonic mean p-value of < 0.01 in male APOE2 mice

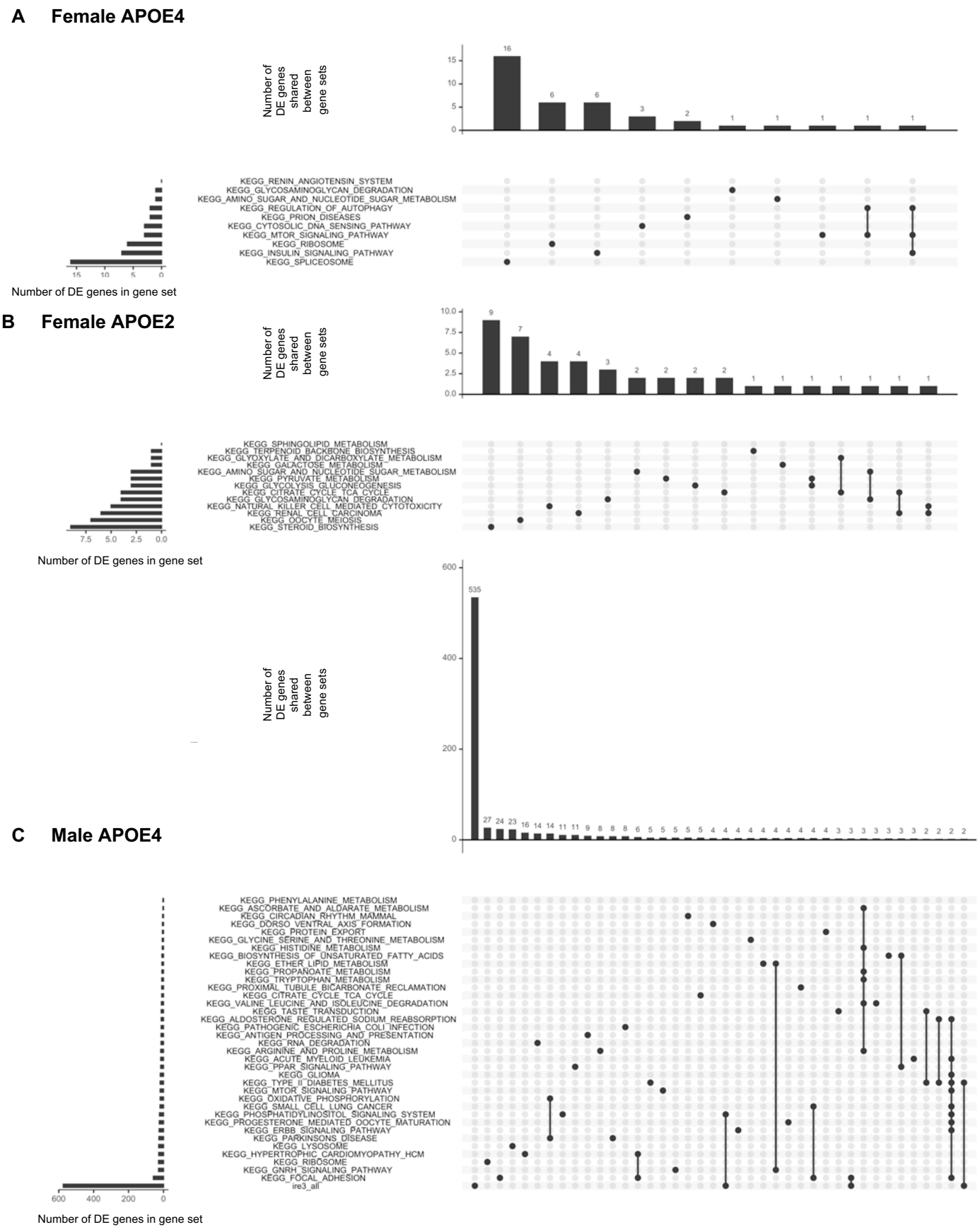


Fig. S22. Upset plots indicating the overlap of DE genes across the gene sets which were calculated to have a FDR-adjusted harmonic mean p-value < 0.01 in **A)** female APOE4 mice, **B)** female APOE2 mice and **C)** male APOE4 mice. Note that the *KEGG* gene sets for oxidative phosphorylation and Parkinson's disease share 14 DE genes.

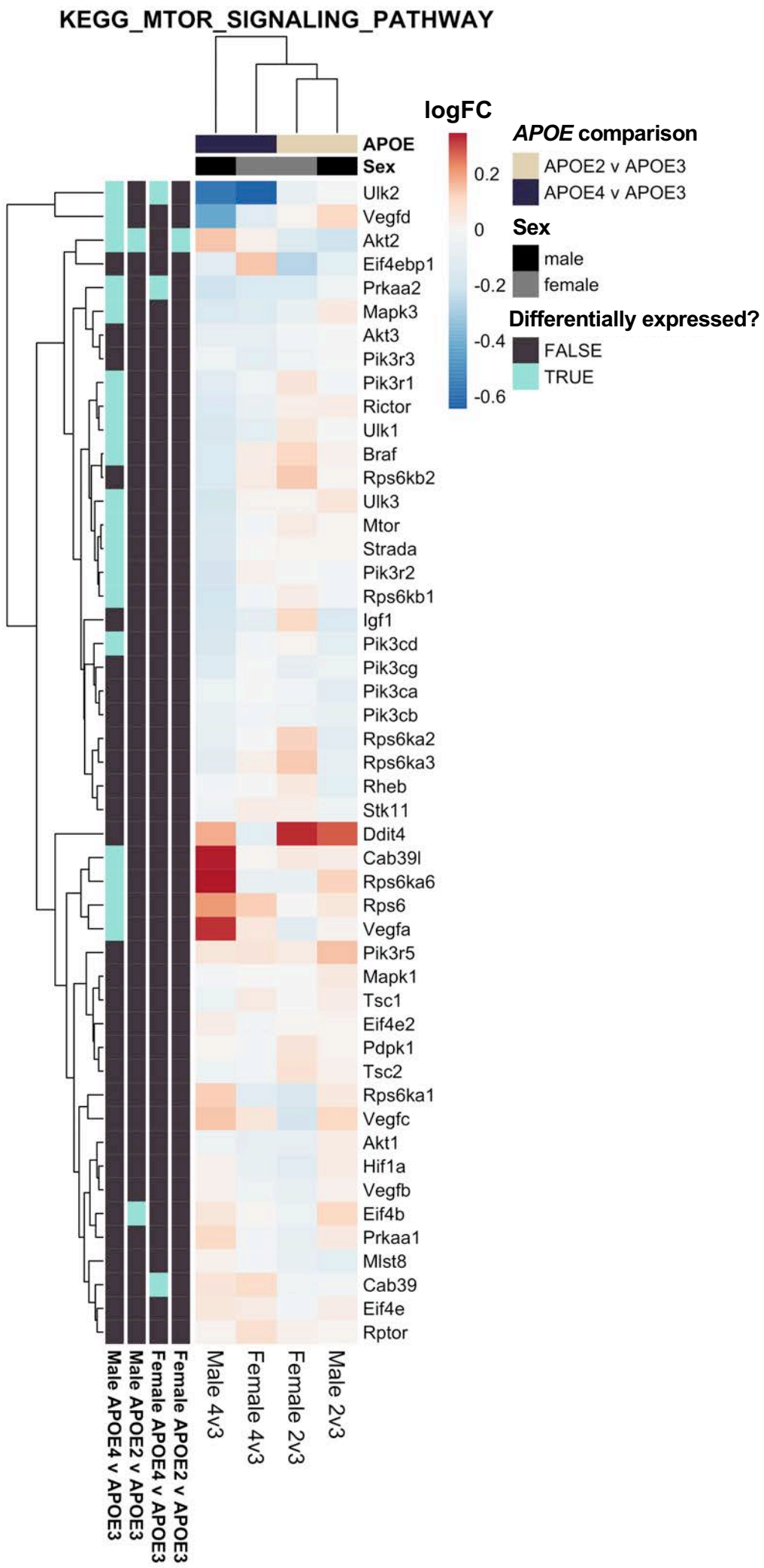


Fig. S23. Heatmap showing the log2 fold change (logFC) of genes detected in APOE-TR mice in the *KEGG_MTOR_SIGNALING_PATHWAY* gene set. Rows (genes) and columns (comparisons) are clustered based on their Euclidean distance. Genes are labelled whether they were found to be significantly differentially expressed (DE) in the DE analysis with *edgeR*. Columns are labelled with the *APOE* genotype and sex comparison.

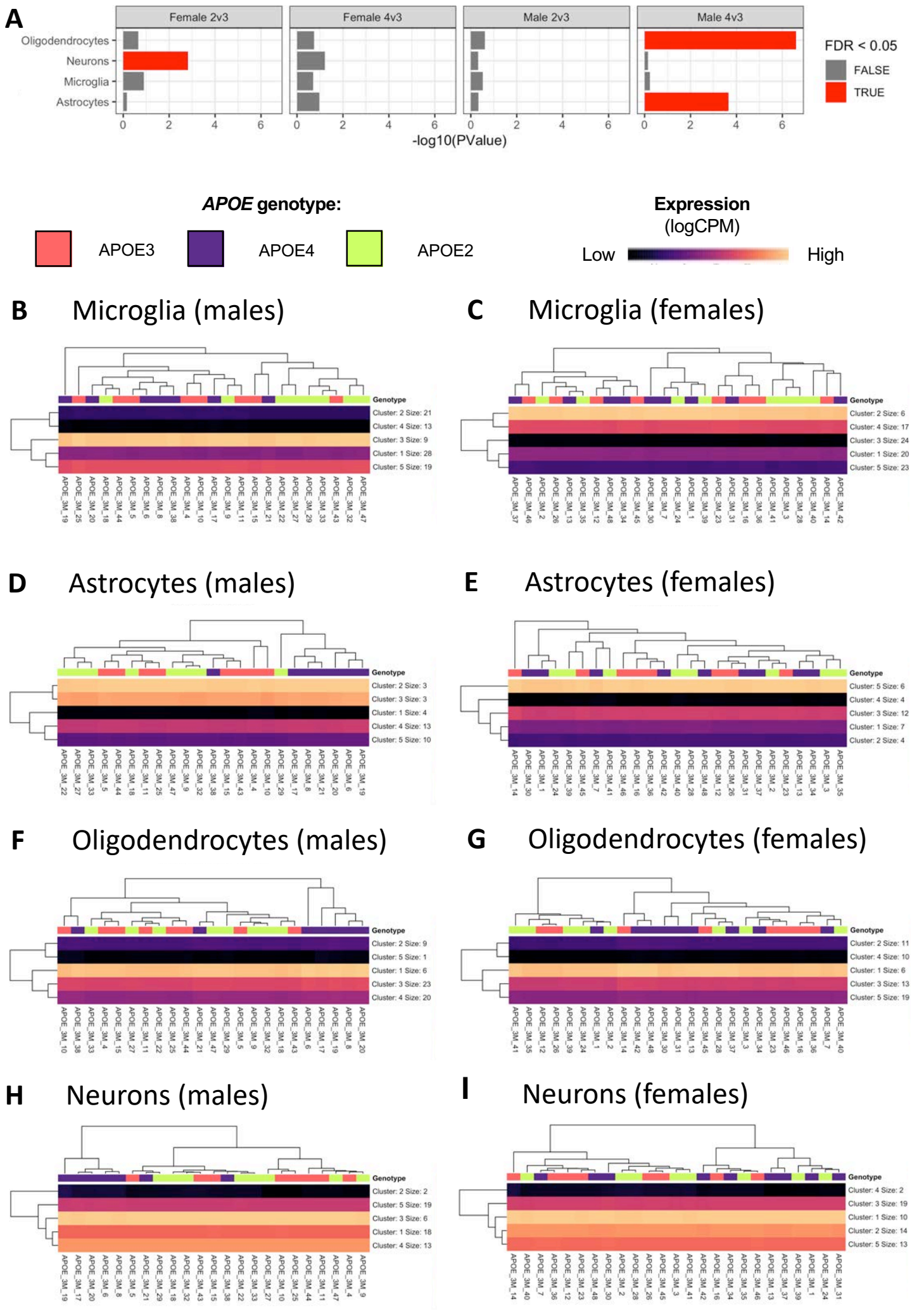


Fig. S24: Changes to cell type proportions in APOE-TR mice. A) Significance of gene set testing from *fry* with a directional hypothesis of gene sets consisting of marker genes of neurons, astrocytes, oligodendrocytes and microglia. **B)** Expression (logCPM) of these cell type marker genes in APOE-TR mice. Rows represent clusters of genes with similar gene expression values summarised with k-means (k = 5). Columns represent samples, and are labelled with *APOE* genotype (see legend). Rows and columns are clustered based on their Euclidean distance. Male APOE4 samples mostly form distinct clusters in **D)** and **F)**, indicating that expression of cell-type specific marker genes are distinct in these genotypes.

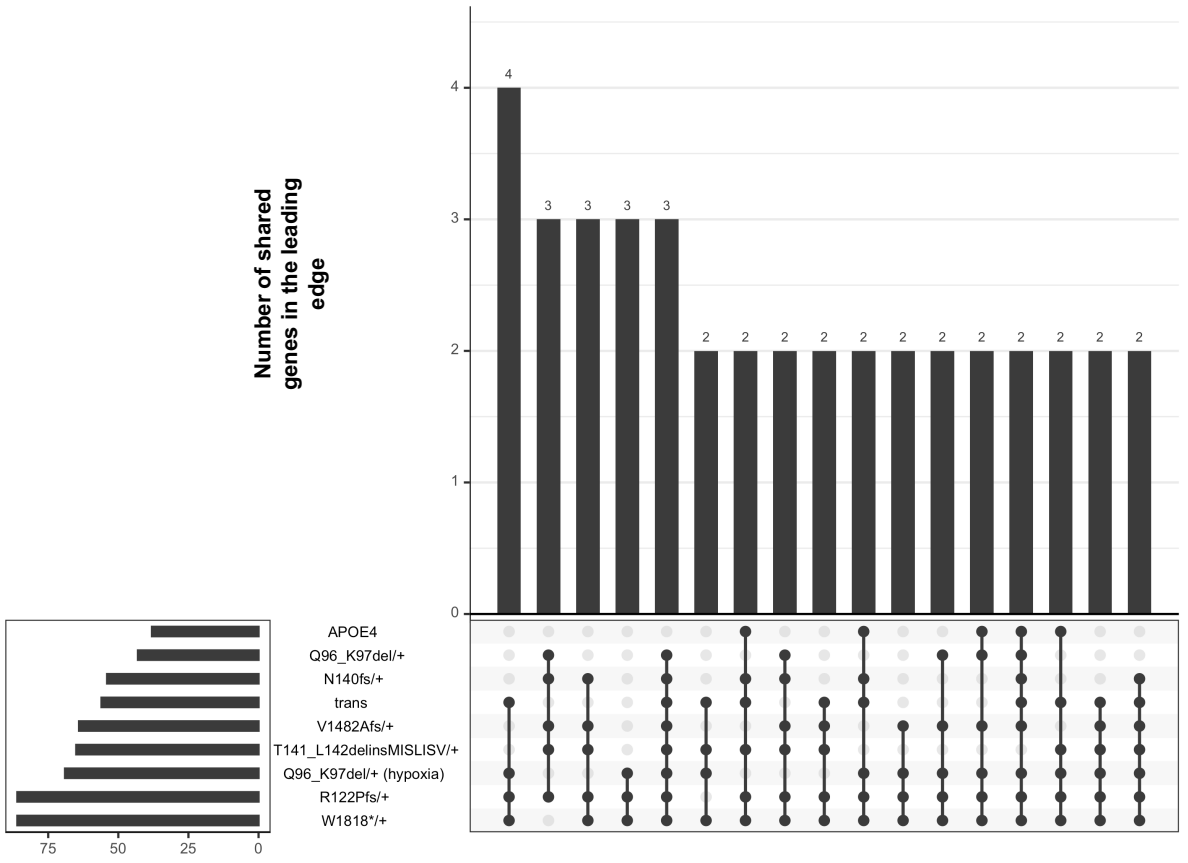


Fig. S25. The genes which drive the statistical significance of the *KEGG_OXIDATIVE_PHOSPHORYALTION* gene set are frequently different in male APOE4 mice and EOfAD model zebrafish. The upset plot shows the overlap of genes in the leading edge of the *KEGG_OXIDATIVE_PHOSPHORYALTION* gene set from the *fgsea* algorithm. Only genes which were found in the leading edge for at least two mutations are displayed.

Table S1. Gene sets significantly enriched with differentially expressed genes in APOE-TR mice.		
Gene set	**FDR**	**Coef**
KEGG_CALCIUM_SIGNALING_PATHWAY	0.03	APOE4 male
KEGG_NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION	0.04	APOE4 male
KEGG_STEROID_BIOSYNTHESIS	1e-5	APOE2 female

Table S2. Significance of the *KEGG_OXIDATIVE_PHOSPHORYLATION* gene set in young APOE-TR mice.

Sex	***APOE***	**FDR-adjusted harmonic mean p-value**
Male	APOE4	0.00948
	APOE2	0.794
Female	APOE4	0.794
	APOE2	0.248

Table S3: Significance of the *KEGG_RIBOSOME* gene set in young APOE-TR mice.

Sex	***APOE***	**FDR-adjusted harmonic mean p-value**
Male	APOE4	0.0000000782
	APOE2	0.691
Female	APOE4	0.000000101
	APOE2	0.793