



THE UNIVERSITY
of ADELAIDE

**Investigating the effects of mutations causative
for early-onset familial Alzheimer's disease
using zebrafish as a model organism**

Karissa Barthelson

A thesis submitted for the degree of Doctor of Philosophy

Discipline of Genetics

School of Biological Sciences

The University of Adelaide

March 2021

Table of Contents

<i>List of publications included in this thesis</i>	4
<i>Abstract</i>	5
<i>Thesis Declaration Statement</i>	7
<i>Ethics statement</i>	8
<i>Acknowledgements</i>	9
<i>Chapter 1: Literature Review</i>	10
1.1 Alzheimer’s disease is the most common form of dementia.....	11
1.2 Pathologies observed in Alzheimer’s disease.....	12
1.3 The genetics of early-onset familial Alzheimer’s disease	14
1.4 Risk factors for late-onset sporadic Alzheimer’s disease	22
1.5 Hypotheses of AD pathogenesis	25
1.6 Hypoxia, iron and HIF1- α in Alzheimer’s disease.....	32
1.7 Animal models of AD.....	36
1.8 Zebrafish as a model organism:.....	38
1.9 Behavioural characterisations of animal models	38
1.10 Transcriptome analysis using RNA-seq.....	40
1.11 Previous RNA-seq studies of AD.....	44
1.12 Previous work of the Alzheimer’s Disease Genetics Laboratory.....	47
1.13 Concluding remarks	50
1.14 References.....	51
<i>Chapter 2: Sorting out the role of SORTILIN-RELATED RECEPTOR 1 in Alzheimer’s disease</i>	66
<i>Chapter 3: No observed effect on brain vasculature of Alzheimer’s disease- related mutations in the zebrafish presenilin 1 gene</i>	67
<i>Chapter 4: Pilot behavioural characterisation of zebrafish early-onset familial Alzheimer’s disease models</i>	68
Introduction	69
Methods.....	71
Results.....	76
Discussion	90
Conclusion.....	94
References	95

Chapter 5: 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	101
Chapter 6: 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	102
Chapter 7: Frameshift and frame-preserving mutations in zebrafish presenilin 2 affect different cellular functions in young adult brains	103
Chapter 8: 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.....	104
Chapter 9: Comparative analysis of Alzheimer's disease knock-in model brain transcriptomes implies changes to energy metabolism as a causative pathogenic stress	105
Chapter 10: Final discussion and conclusion	106

List of publications included in this thesis

1. Barthelson K, Newman M, Lardelli M: **Sorting Out the Role of the Sortilin-Related Receptor 1 in Alzheimer's Disease.** *J Alzheimers Dis Rep* 2020, 4:123-140.
2. Barthelson K, Newman M, Nowell CJ, Lardelli M: **No observed effect on brain vasculature of Alzheimer's disease-related mutations in the zebrafish presenilin 1 gene.** *Molecular Brain* 2021, 14:22.
3. Barthelson K, Pederson SM, Newman M, Lardelli M: **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* 2020, 13:142.
4. Barthelson K, Pederson SM, Newman M, Lardelli M: **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* 2021, 79:1105-1119.
5. Barthelson K, Pederson SM, Newman M, Jiang H, Lardelli M: **Frameshift and frame-preserving mutations in zebrafish presenilin 2 affect different cellular functions in young adult brains.** *bioRxiv* 2020:2020.2011.2021.392761.
6. Barthelson K, Dong Y, Newman M, Lardelli M: **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.** *bioRxiv* 2021:2021.2001.2026.428321.
7. Barthelson K, Newman M, Lardelli M: **Comparative analysis of Alzheimer's disease knock-in model brain transcriptomes implies changes to energy metabolism as a causative pathogenic stress.** *bioRxiv* 2021:2021.2002.2016.431539.

Abstract

Development of an effective therapeutic for Alzheimer's disease (AD) is currently a global health priority. To prevent, or at least delay, the onset of AD, we must understand the initial cellular stresses/changes that drive the disease. These initiating changes are likely subtle and occur decades before symptom onset. We cannot easily investigate these changes in humans, as pre-symptomatic brain material from individuals genetically pre-disposed to AD is inaccessible for detailed molecular analyses. For this, we must utilise animal models.

The most frequently used animal models of AD are mice expressing one or more transgenes containing the sequences of human genes bearing mutations which cause AD. These transgenic models have been useful in elucidating some aspects of the pathogenic mechanisms of AD. However, they have not led to development of successful therapeutics.

Mutations in a small number of genes cause early-onset familial forms of AD (EOfAD). These mutations can be introduced into the orthologous, endogenous genes of an animal (i.e. knock-in models). However, relatively few papers describe research with knock-in models. Transcriptome analysis is currently the most detailed form of molecular phenotyping and can give a largely unbiased view of the molecular state of the brains of young knock-in models of EOfAD. Surprisingly, this has not previously been performed using knock-in models of EOfAD-like mutations.

To address this gap in our knowledge, the work presented in this thesis (along with previous work from the Alzheimer's Disease Genetics Laboratory (ADGL)), describes the generation, and/or characterisation of a collection of zebrafish knock-in models of EOfAD-like mutations. The power of zebrafish as a model organism lies in this

species' ability to generate large families of synchronous siblings which can be raised together in the same tank. This has allowed the assessment of the effects of heterozygosity for EOfAD-like mutations (closely mimicking the genetic state of human EOfAD) or the effects of non-EOfAD-like mutations (such as frameshift mutations in *presenilin* genes) on the brain transcriptome with minimal external sources of "noise."

These analyses have revealed that the only cellular process predicted to be affected by EOfAD-like mutations in the heterozygous state, and not by non-AD-related mutations, is oxidative phosphorylation. Comparison of these transcriptomes with recent, publicly available brain transcriptomes from two knock-in mouse models of late onset AD risk alleles revealed similar affected processes, thereby supporting the findings from the zebrafish models.

Preliminary non-transcriptomic characterisations of previously generated/novel zebrafish models were also performed. The effects of heterozygosity for EOfAD-like mutations on brain vasculature were assessed, as well as effects on spatial working memory. Only limited differences were observed in these studies. However, future work with greater statistical power and/or alternate study designs is recommended.

Overall, the research described in this thesis demonstrates the value of unbiased, transcriptome analyses of young, knock-in animals models for understanding the early stages of AD pathogenesis.

Thesis Declaration Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship

Karissa Barthelson

Date 24/03/2021

Ethics statement

The research described in this thesis was carried out under permits S-2017-089 and S-2017-073 from the Animal Ethics Committee of the University of Adelaide, as well as permit 15037 from the Institutional Biosafety Committee of the University of Adelaide.

Acknowledgements

First and foremost, I would like to thank my Principal Supervisor Associate Professor Michael Lardelli for giving me the opportunity to be able to join his group and perform this project. His seemingly endless knowledge, mentorship, advice and support over the last five years (including Honours) has shaped me into the researcher I am today. Thank you Michael for your guidance, and helping me develop my career as a researcher. I will be forever grateful.

Next, I would like to thank Dr. Morgan Newman. Morgan has played such an integral role in developing my skills in molecular biology. She has been an amazing role model for me. Thank you Morgan for your commitment to seeing this project out, despite moving onto your next chapter in life.

I also want to show immense gratitude to Dr. Steve Pederson. While he was not an official PhD supervisor, he has had a very significant impact on the direction this thesis took. In the beginning, I had absolutely no coding skills. Steve showed such kindness and patience with me while I have been learning the bioinformatics side of my research. Thank you Steve for providing virtually 24/7 assistance when I would ask for help, you have gone above and beyond for me.

I would also like to thank my fellow lab mates, notably Dr. Nhi Hin, Dr. Tanya Jayne, Dr. Yang Dong, Dr. Haowei Jiang, Dr. Seyyed Hani Moussavi Nik, Ewan Gerkin, Lachie Baer, Vanessa Chin, Georgie Burke and Jiayu Ruan. Thank you all so much for providing such a positive work environment. It has been so much fun working with all of you.

Next, I would like to thank my family, particularly my Mum, Teta, Baba and Dedo, as well as Graham and Tracy for supporting me during my time as a PhD Candidate.

Finally, I would like to thank my partner Ryan. Thank you for your love, encouragement, and endless support. Thank you asking how my day was every single day, and then sitting through what is probably hours of my venting over the last few years, and for sticking with me through the rollercoaster this PhD has been.

Chapter 1: Literature Review

The literature review is split into two chapters. **Chapter 1** consists of an overall introduction of Alzheimer's disease. Then, **Chapter 2** consists of a published review manuscript detailing the role of the sortilin-related receptor gene (*SORL1*) in Alzheimer's disease.

1.1 Alzheimer's disease is the most common form of dementia

Alzheimer's disease (AD) is a devastating neurodegenerative disease, and the most prevalent form of dementia [1]. Dementia affects approximately 472,000 people in Australia, and 50 million people worldwide [2]. Despite these disturbing statistics, no effective therapeutics exist to date.

AD is defined clinically by progressive cognitive decline and pathologically by the presence of neurofibrillary tangles (consisting of hyperphosphorylated tau protein) and senile neuritic plaques (consisting of a short peptide named amyloid beta [$A\beta$]) (**Figure 1**, reviewed in [3]).

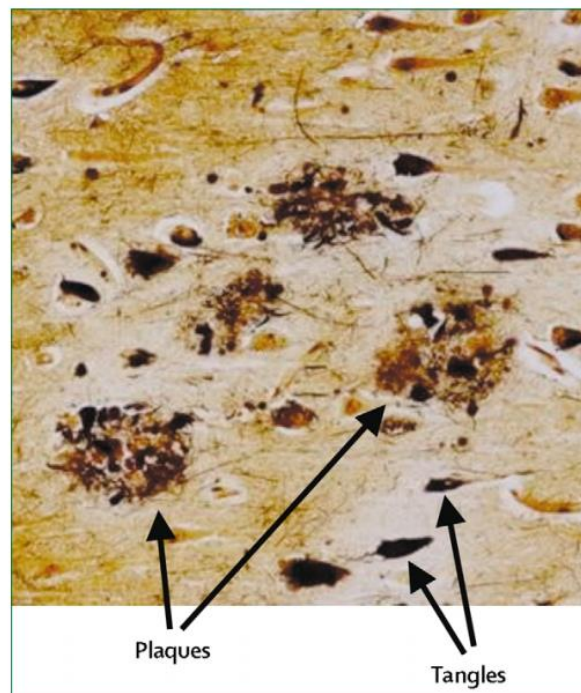


Figure 1: Neuritic plaques and tangles in the cerebral cortex found in AD patients. Neuritic plaques primarily extracellular deposits of a short peptide called amyloid β ($A\beta$), surrounded by other proteins and structures. Tangles are intracellular aggregates composed of a hyperphosphorylated form of the microtubule-associated protein tau. Image reproduced with permission from [3].

Cases of AD can be classified by the age of disease onset (early- or late-onset, where symptom onset occurs before or after an arbitrarily defined threshold of 65 years of age), and whether a family history of AD is present (familial or sporadic). Late-onset sporadic AD (LOAD) is the most common subtype of AD, accounting for the majority of AD cases [4]. The etiology of LOAD is a complex interaction between genetic and environmental risk factors and this is discussed in [1.4](#). In some rare cases (<1%), AD can be inherited due to autosomal dominant mutations in, mainly, three genes: amyloid β A4 precursor protein (*APP*), presenilin 1 (*PSEN1*) and presenilin 2 (*PSEN2*). This form of AD is known as early-onset familial AD (EOfAD). Early-onset sporadic cases, and late-onset familial cases also exist, although these are not as commonly discussed. The genetics of AD are discussed in more detail in [1.3](#) and [1.4](#).

1.2 Pathologies observed in Alzheimer's disease

In the AD brain, a wide range of pathologies are observed. These include, but are not limited to, oxidative stress [5], aberrant lipid homeostasis (reviewed in [6]), impaired glucose metabolism (reviewed in [7]), vascular irregularities (discussed in [1.5.3](#)), endo-lysosomal system irregularities [8, 9], gliosis [10] and inflammation (reviewed in [11]). A number of reviews describe these pathologies in detail (see [12-18]).

However, the literature regards the definitive pathological “hallmarks” of the disease to be amyloid plaques and neurofibrillary tangles (NFTs). These hallmarks are briefly described below.

1.2.1 Neurofibrillary tangles consist of the protein tau.

In humans, tau proteins are encoded by microtubule-associated protein tau (*MAPT*) on chromosome 17. They are normally bound to microtubules and promote

microtubule stabilisation and assembly from tubulin subunits [19]. In AD, tau is abnormally hyperphosphorylated (p-tau) and it disrupts microtubule formation by binding to normal tau and sequestering it away from its normal function. P-tau also self assembles into paired helical filaments which aggregate further and deposit as the neurofibrillary tangles observed in AD and other tauopathies such as fronto-temporal dementia and Pick's disease (reviewed in [19-21]).

1.2.2 Amyloid β is the main component of the senile neuritic plaques

Neuritic plaques primarily consist of aggregates of the short peptide $A\beta$. $A\beta$ is formed from intramembrane proteolysis of APP, a single transmembrane-spanning protein. The majority of APP is cleaved at the plasma membrane by α -secretase complexes to form a soluble sAPP α fragments. This pathway is non-amyloidogenic since α -secretase complexes cleave APP within the $A\beta$ peptide sequence. However, some APP may be internalised within endosomes by clathrin-mediated endocytosis. Internalisation occurs due to interaction of the NPxY motif in the cytoplasmic tail of APP and the clathrin adaptor AP2 [22, 23]. After internalisation, APP can move to the late endosomal compartments where $A\beta$ peptides of mainly 40-42 amino acids in length can be formed. This occurs by cleavage of APP by β -secretase, then by γ -secretase complexes (contains presenilin) at the C terminus (**Figure 2**, reviewed in [24, 25]). The longer of these peptides ($A\beta_{42}$) is prone to aggregation and eventually deposits as extracellular amyloid plaques [26]. $A\beta$ has been shown to form in other cellular locations [27] and this will be discussed in more detail in **Chapter 2**.

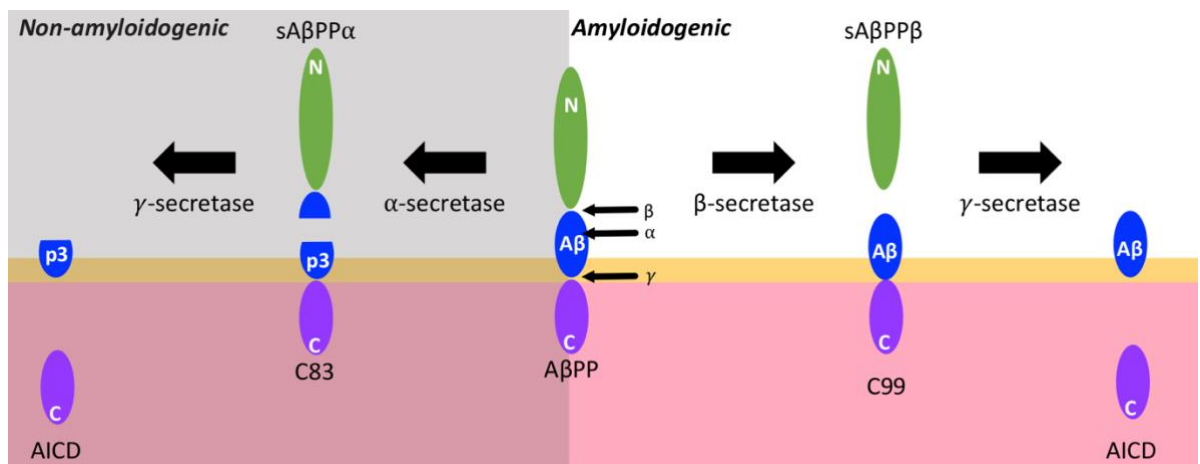


Figure 2: Proteolytic processing of APP. Membrane bound APP (AβPP) can be processed by a non-amyloidogenic pathway when it is proteolytically cleaved by α-secretase and then γ-secretase complexes. Soluble fragments are formed in this pathway. It also can be processed in an amyloidogenic pathway, where APP is sequentially cleaved by β-secretase and then γ-secretase to form the aggregation-prone amyloid β (Aβ) peptide. Aβ can be secreted into the extracellular space in which it can aggregate to contribute to the formation of amyloid plaques. Image reproduced from [28].

1.3 The genetics of early-onset familial Alzheimer's disease

Mutations in mainly three genes have been shown to cause EOfAD. The majority of mutations are found in *PSEN1*, followed by *APP*, then *PSEN2* (see [29] and <http://www.alzforum.org/mutations>). Sortilin-related receptor 1 (*SORL1*) was recently found to contain EOfAD mutations [30] and, remarkably, is also strongly associated with LOAD [31-39]. The role of *SORL1* in AD is discussed in detail in **Chapter 2**.

1.3.1 AMYLOID β A4 PRECURSOR PROTEIN

The *APP* gene is most well-known for encoding the protein precursor to $A\beta$.

However, the APP protein and its proteolytic processing products are also implicated in neurogenesis, cell adhesion, axonal transport and other neuro-developmental processes (reviewed in [40-42]).

EOfAD mutations in *APP* are generally missense mutations and therefore, maintain the generation of the full-length protein. The majority of the missense mutations are clustered around the sites for γ -secretase and β -secretase cleavage (**Figure 3**).

Additionally, *APP* gene duplication events, such as in trisomy 21 (Down's syndrome) are also associated with EOfAD [43] (reviewed in [44]). The general consensus in the field is that EOfAD mutations in *APP* result in elevated levels of total $A\beta$, or altered ratios of the longer $A\beta_{42}$ species (which are more prone to aggregation) to $A\beta_{40}$ species (reviewed in [45]). However, mutations in *APP* also can lead to increased levels of the β -secretase cleavage product of APP: C99 [46].

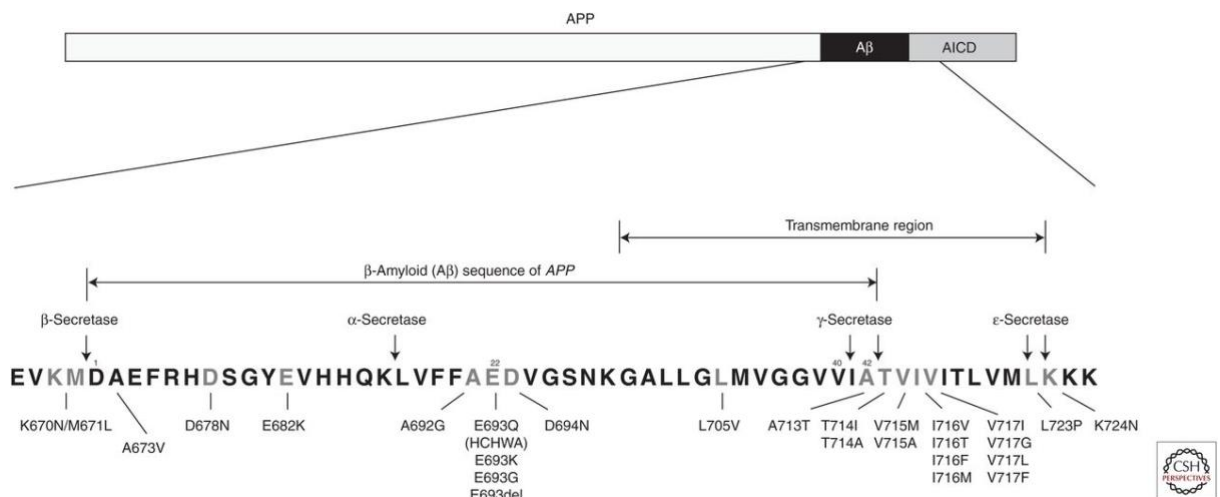


Figure 3: EOfAD mutations in the gene encoding the amyloid β precursor protein (APP). Schematic of the human APP protein, with the A β and APP intracellular domain (AICD) regions indicated. The amino acid sequence of the A β region, and the location of EOfAD mutations are shown below. Image reproduced with permission from [45]. Copyright © 2017 Cold Spring Harbor Laboratory Press; all rights reserved.

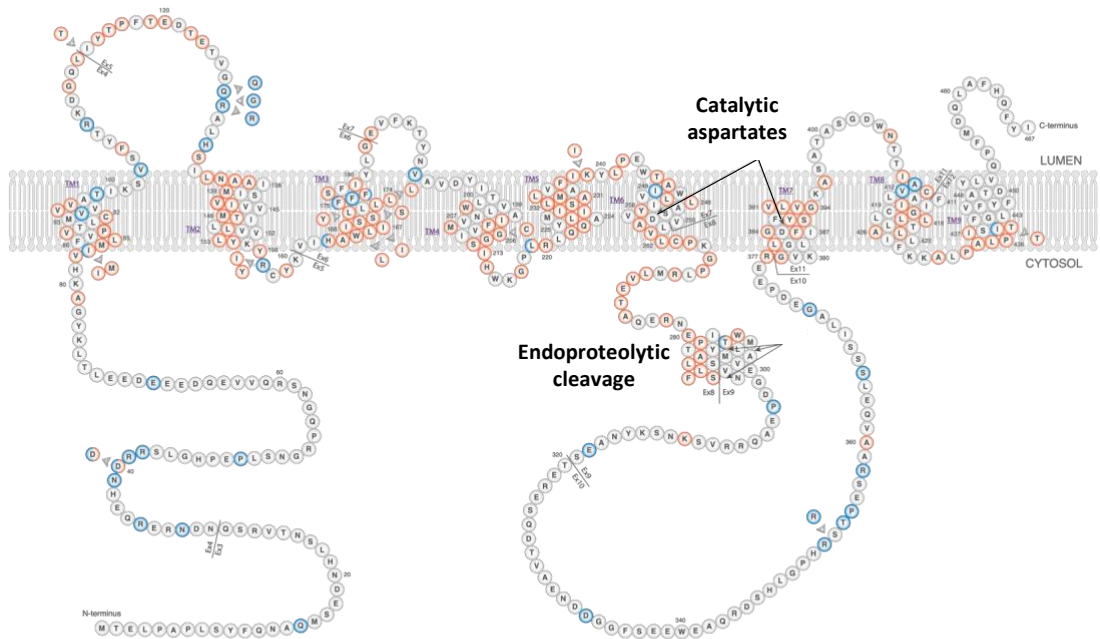
1.3.2 The *PRESENILINs*

The *PSEN1* and *PSEN2* genes encode for holoproteins of length 476 amino acids (52.6kDa) and 448 amino acids (50.1 kDa) respectively [47]. They share 67% of amino acid sequence identity. However, they are not fully redundant. *Psen1* homozygous knockout mice are lethal at early embryogenesis [48]. However, *Psen2* homozygous knockout mice (*Psen2*^{-/-}) are viable and fertile. *Psen2*^{-/-} mice do have a mild phenotype similar to pulmonary fibrosis, but show little to no effect on APP processing [49]. Intriguingly, rats lacking γ -secretase activity of *Psen1* are viable [50].

PSEN1 and PSEN2 proteins have similar structures, with 9 transmembrane domains connected by cytosolic or extracellular loops. A tenth transmembrane domain may exist when a PSEN protein is in the holoprotein state [51]. PSEN holoproteins

undergo auto-endoproteolytic cleavage in the cytosolic loop between transmembrane domains 6 and 7, and then the two fragments function together as an active heterodimer (**Figure 4**) [52-54]. The most characterised function of the active heterodimer is within the γ -secretase complex (discussed below).

A PSEN1



B PSEN2

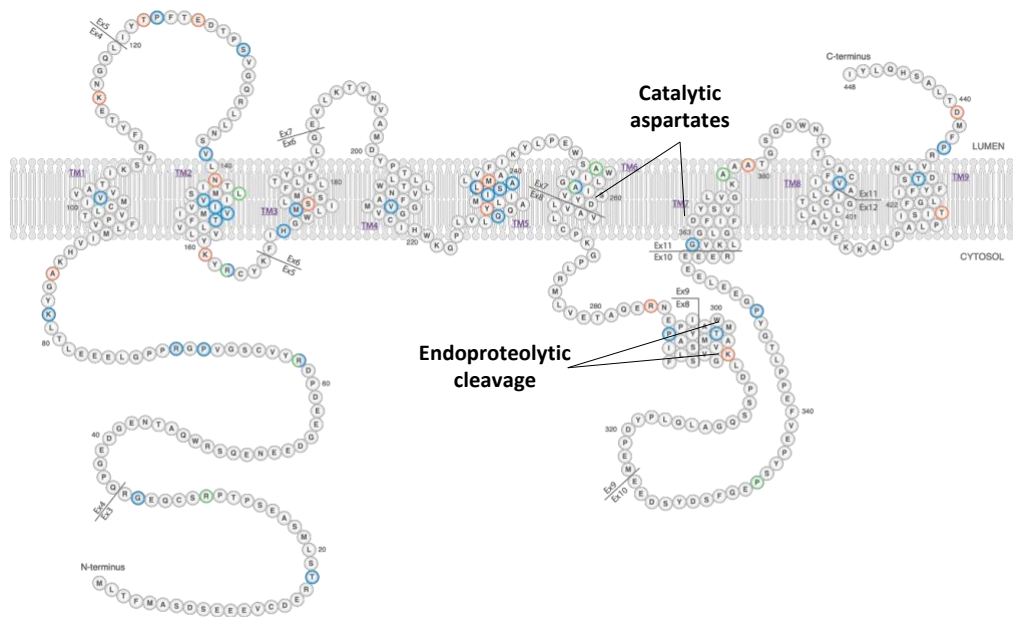


Figure 4: EOfAD mutations in the presenilins. A Schematic of the human presenilin 1 (PSEN1) and **B** PSEN2 proteins. The amino acid residues are coloured red if they are pathogenic for Alzheimer’s disease, blue if they have unclear pathogenicity, and green if they are non-pathogenic. The transmembrane domains (TM), catalytic aspartates required for γ -secretase activity, and the endoproteolytic cleavage sites are indicated. **A** was adapted with permission from www.alzforum.org/mutations/psen-1 Version 3.3 (2020). **B** was adapted with permission from www.alzforum.org/mutations/psen-2 Version 2.7 (2020). Copyright © 1996-2020 FBRI LLC All rights reserved.

Mutations which cause EOfAD in the *PSENs* follow a “reading frame preservation rule” [55]. Mutations which generate a mutant protein which still utilises the stop codon used to generate the wild type protein are causative for EOfAD. Examples are missense mutations which only alter a single codon (i.e. *PSEN1*^{E280A} [56]), in-frame insertion/deletion (indel) mutations (i.e. *PSEN1*^{T440del} [57]), and large deletions which preserve the reading frame (i.e. *PSEN1*^{ΔE9} [59]). Additional evidence supporting this rule comes from the fact that the only identified, disease-causing frameshift mutation of a *PSEN* gene: *PSEN1*^{P242LfsX11} causes an inherited skin condition (acne inversa) without EOfAD [60]. Other frameshift mutations in *PSEN1* have been observed in EOfAD patients [61, 62]. However, these mutations have not been functionally validated, and/or have not been demonstrated to cause EOfAD in large pedigrees.

A particularly interesting mutation of *PSEN2* is *PSEN2*^{K115Efs*} [63]. This mutation was originally thought to be the only frameshift mutation of a presenilin gene which generates a premature stop codon *and* is associated with EOfAD, and therefore was thought to be an exception to the reading-frame preservation rule. However, a more recent characterisation of this mutation in human *PSEN2*^{K115Efs*} mutation carrier tissue revealed the existence of a brain-specific alternative splicing product of *PSEN2* resulting in a partial intron retention, and restoration of the original, wild type

reading frame (after a short frameshift, **Figure 5**) [64]. Therefore, *PSEN2*^{K115Efs*} does indeed follow the reading frame preservation rule of EOfAD mutations in the presenilins, as a putatively translatable transcript is still generated that would utilise the wild type stop codon. Intriguingly, if translated, the *PSEN2*^{K115Efs*} transcripts that would be translated to form a truncated protein encode a protein similar to the PS2V isoform of human *PSEN2* which shows increased expression in LOAD (**Figure 5**) [65]. The expression of PS2V is induced by hypoxia and/or oxidative stress [65, 66], is associated with increased A β production [67, 68] and suppression of the unfolded protein response [67, 68].

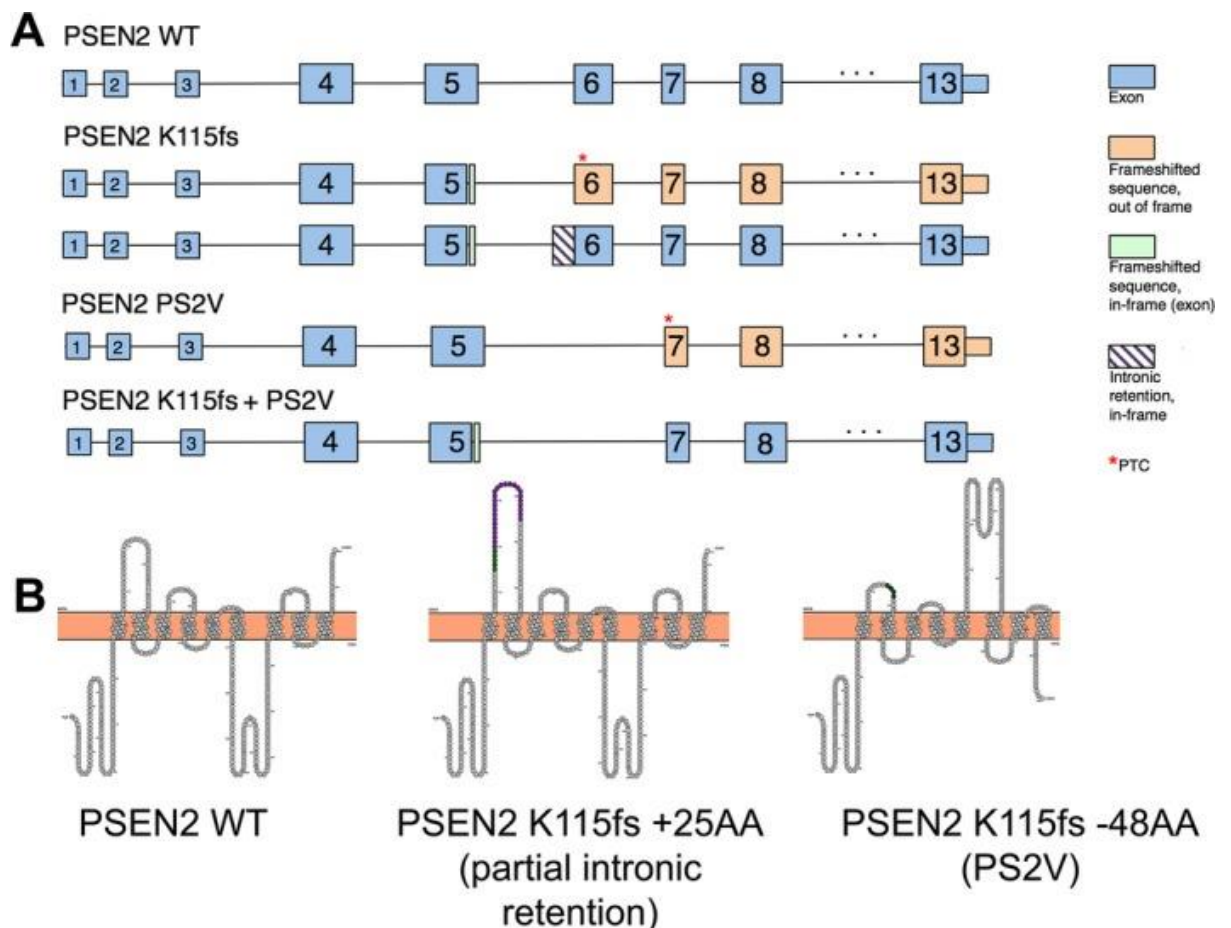


Figure 5: Alternative splicing of human *PSEN2*. **A** Schematic of the human WT *PSEN2* transcript. The two transcripts arising from the K115Efs (K115fs) allele of *PSEN2* either produce a premature termination codon (PTC, red asterisk), or a partial intron retention (striped box) and restoration of the wild type reading frame. The PS2V splice isoform results in a transcript which excludes *PSEN2* exon 6, leading to a frameshift and a PTC in exon 7. The PS2V splice isoform arising from the K115fs allele of *PSEN2* is predicted to encode a short frameshift, exclusion of exon 6 and restoration of the wild type reading frame **B** Schematic of predicted protein structures due to the presence of the K115fs mutation of *PSEN2*. Image reproduced with permission from [64].

The most characterised function of the presenilin proteins is their role as the catalytic core of the γ -secretase complex [69-71]. The γ -secretase complex is comprised of the active heterodimer of PSEN1 or PSEN2, nicastrin (NCSTN), presenilin enhancer 2 (PSENEN) and anterior pharynx defective 1 (APH1A or APH1B) in a 1:1:1:1 stoichiometry (**Figure 6**) [72]. This complex proteolytically cleaves at least 149 different substrates [73], with the most well studied substrates being the Notch receptor [74] and APP [69]. Many substrates cleaved by γ -secretase release an intracellular domain (ICD) which can move to the nucleus to regulate transcription of target genes, as well as generation of N-terminal fragments which have various biological functions. This is reviewed in detail in [73].

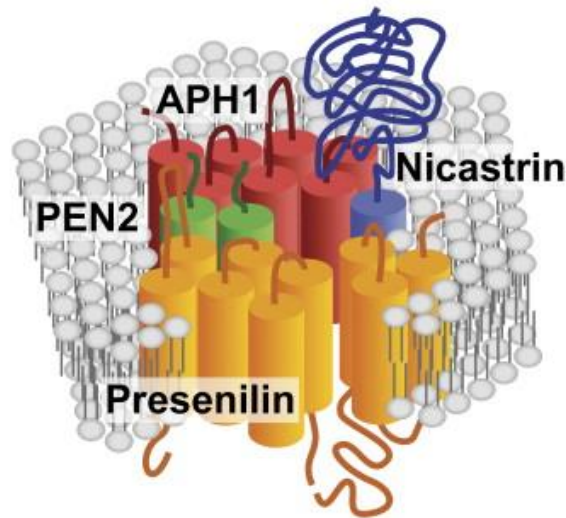


Figure 6: Representation of the γ -secretase multi-subunit complex. The complex is comprised of four core subunits, presenilin fragments (2) in yellow, nicastrin in blue, presenilin enhancer 2 (PEN2 or PSENEN) in green and anterior pharynx defective 1 (APH1A) or APH1B in red. PSEN is a nine-pass transmembrane aspartyl protease and provides the catalytic activity for the complex. It undergoes endoproteolysis and the N- and C-terminal fragments function together as a heterodimer. Image reproduced with permission from [75].

The presenilins also have γ -secretase independent functions. For example, the PSEN1 holoprotein is required for (macro-)autophagy by facilitating lysosomal acidification. PSEN1 holoprotein is required for N-glycosylation of the V0a1 subunit of v-ATPase, which is essential for its successful delivery from the endoplasmic reticulum to lysosomes [76]. The v-ATPase complex is a proton pump which acidifies components of the endolysosomal system such as newly created autolysosomes, which then can fuse with autophagosomes and late autophagic vacuoles (AVs) [77], to allow proteolysis of substrates and protein aggregates such as A β . This could explain the accumulation of autophagosomes and AVs in dystrophic dendrites in the AD brain [8].

In addition to lysosomal acidification, PSEN1 protein phosphorylated at amino acid residue S376 by casein kinase 1 γ isoform (CK1 γ), facilitates autophagosome-lysosome fusion by interacting with annexin A2, which then recruits other proteins to facilitate the fusion. Mutation of the S376 site resulted in decreased degradation of the C99 fragment of APP, consequently increasing A β levels [78, 79]. Taken together, these studies support that PSEN1 plays an important role in autophagy.

The PSEN1 holoprotein also plays a role in hypoxia. Psen1 and hypoxia inducible factor 1, α -subunit (Hif1- α , a master regulator of the cellular response to hypoxia), physically interact in immortalised mouse fibroblasts. The induction of Hif1- α was also shown to require Psen1 in a γ -secretase independent manner [80]. The role of hypoxia in AD is discussed further in section [1.6](#).

PSEN1 has been shown to play γ -secretase independent roles in other cellular processes such as Wnt/ β -catenin signalling [81] and regulation of calcium levels (reviewed in [82, 83]).

1.4 Risk factors for late-onset sporadic Alzheimer's disease

Compared to EO β AD, the development of LOAD is not the consequence of inheritance of a single variant in a single gene, but, rather, a complex interaction of both genetic and environmental factors. The heritability of LOAD is estimated to be between 56% and 79% [84], suggesting a strong genetic foundation. Indeed, genome-wide association studies (GWAS) have identified many loci associated with increased risk for LOAD. These are briefly introduced below.

1.4.1 APOLIPOPROTEIN E (*APOE*)

The apolipoprotein E (*APOE*) gene was the first gene to be associated with the development of LOAD [85]. There are three main alleles of *APOE*: $\epsilon 2$, $\epsilon 3$ and $\epsilon 4$, which have allele frequencies of approximately 8%, 75% and 13% respectively [86]. The $\epsilon 4$ allele is the strongest genetic risk factor for the development of LOAD [87, 88], and is associated with an increase in risk of approximately 3-fold when heterozygous, and 15-fold when homozygous [86]. Conversely, the $\epsilon 2$ allele is associated with a small protective effect [86], while the common $\epsilon 3$ allele is AD-risk neutral. The three alleles encode isoforms of *APOE* differing by only two amino acid residues (positions 112 and 158), which alter their secondary structure (reviewed in [89]).

APOE's main role is as a cholesterol and lipid transporter (reviewed in [90, 91]). However, *APOE* can also bind to $A\beta$ and facilitate its clearance (reviewed in [92, 93]). The $\epsilon 4$ isoform is least efficient at binding $A\beta$ [94], resulting in increased $A\beta$ levels in the brain [95]. This is consistent with the observations made in [96], where analysis of post mortem LOAD brains showed that patients who were homozygous for the $\epsilon 4$ allele of *APOE* had significantly greater plaque load than patients who were homozygous for the $\epsilon 3$ allele. Interestingly, the clearance of $A\beta$ via *APOE* receptors is mediated by *SORL1* [97].

1.4.2 Other loci associated with LOAD implicate particularly the endo-lysosomal system and inflammation in disease pathogenesis

GWAS have identified several other loci associated with increased risk for the development of LOAD (reviewed in [98, 99]). These loci include genes such as the ATP-binding cassette sub-family A member 7 (*ABCA7*), bridging integrator 1 (*B1N1*), triggering receptor expressed on myeloid cells 2 (*TREM2*), siglec-3 (*CD33*), clusterin

(*CLU*), complement receptor 1 (*CR1*), ephrin type-A receptor 1 (*EPHA1*), membrane-spanning 4-domains, subfamily A (*MS4A*) and phosphatidylinositol binding clathrin assembly protein (*PICALM*). Pathway analyses of LOAD risk genes have revealed their convergence on particular pathways in immune responses and the endo-lysosomal system (**Figure 7**) [87, 100]. This is consistent with other studies implicating these biological functions in LOAD pathogenesis (reviewed in [101, 102]).

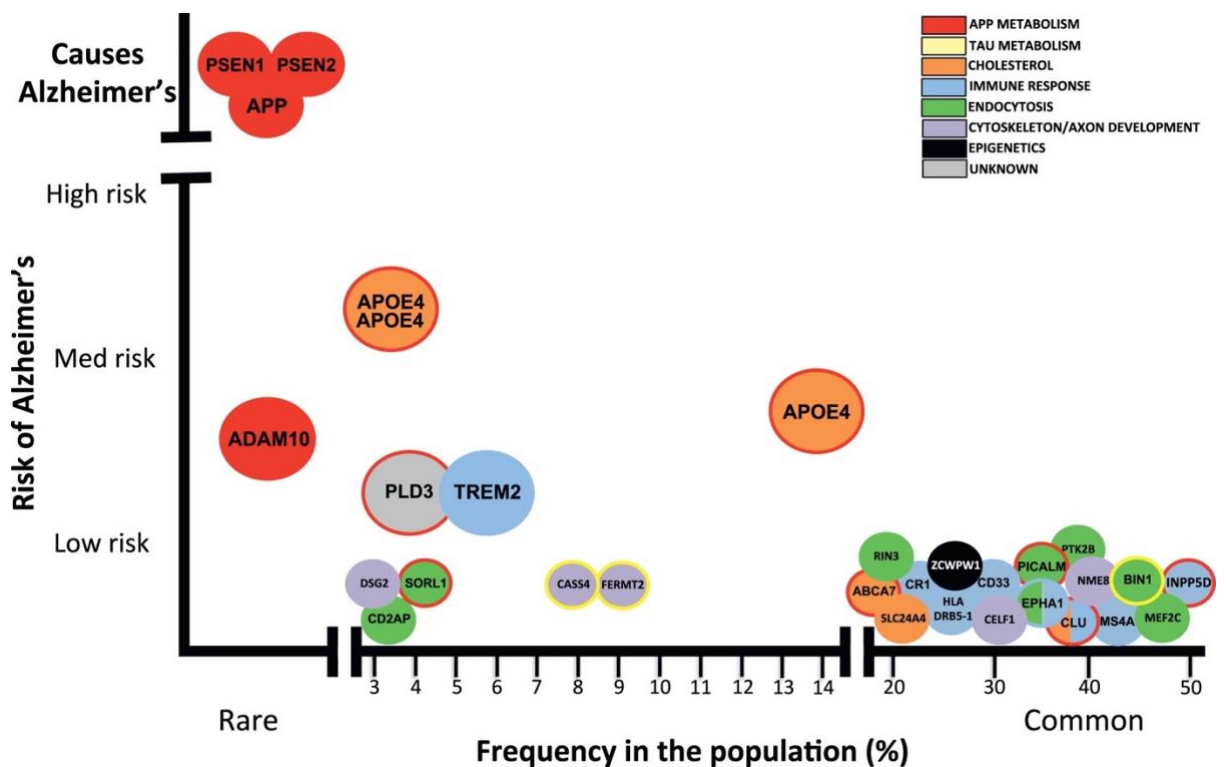


Figure 7: Genes associated with risk for late-onset Alzheimer's disease. The early onset AD genes *PSEN1*, *PSEN2* and *APP* are also shown. Image reproduced with permission from [98].

1.5 Hypotheses of AD pathogenesis

Many hypotheses have been proposed to explain the pathogenesis of AD (e.g. tau axis hypothesis [103], oxidative stress hypothesis, [104], the cholinergic hypothesis [105]). To review each hypothesis individually is beyond the scope of this thesis. However, three important hypotheses are described below.

1.5.1 The amyloid cascade hypothesis

The amyloid cascade hypothesis was formulated in 1992 and predicts that A β accumulation is the cause of AD pathology. The neurofibrillary tangles, cell loss, vascular damage and dementia follow as a result of this accumulation [106]. A flow chart of the hypothesis can be found in **Figure 8**.

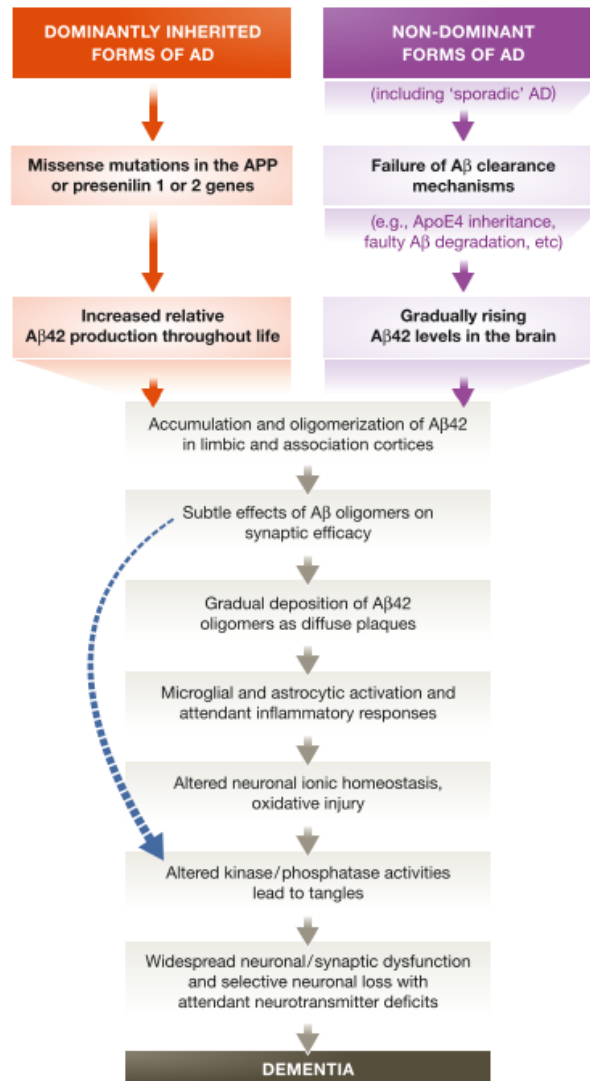


Figure 8: The amyloid cascade hypothesis. The flowchart summarises the steps that eventually result in dementia in dominantly or non-dominantly inherited forms of AD. Image reproduced from [107].

The amyloid cascade hypothesis has been very influential in the AD research field. However, it has recently been challenged. It was originally widely accepted since the senile plaques, a pathological hallmark of the disease, primarily consist of Aβ [108]. Also, the majority of mutations which cause EOfAD arise within genes directly involved in Aβ generation (*APP*, *PSEN1* and *PSEN2*) [47]. Recent evidence supporting the hypothesis was found by Leinenga and Götz [109], who removed Aβ from the brain of transgenic mice by scanning ultrasound and observed restoration of memory function.

There is also considerable evidence which does not support the amyloid cascade hypothesis. There is a poor correlation between the number of plaques and the severity of dementia in AD patients [110]. In fact, neurofibrillary tangle counts correlate better with the severity of dementia [111]. Also, amyloid plaques are found in cognitively normal individuals [112].

Perhaps the most important evidence against the validity of the amyloid cascade hypothesis is the repeated failure in clinical trials of drugs targeting A β production and clearance. Initial clinical trials showed that semagacestat (a γ -secretase inhibitor intended to reduce formation of A β), did reduce A β levels in the plasma of patients but not in the cerebrospinal fluid [113, 114]. In following trials, semagacestat did not improve cognitive status, and patients receiving the higher dose had significant worsening of functional ability. Consequently, the trial was terminated [115]. A more recent clinical trial was performed by the Dominantly Inherited Alzheimer Network (DIAN), a cohort of individuals who carry dominant mutations in *PSENs* and *APP*, and therefore are in all likelihood destined to develop AD. This cohort presents as an optimal cohort with which to perform prevention studies. The anti-amyloid drugs solanezumab and gantenerumab, two monoclonal anti-A β antibodies designed to facilitate A β clearance from the brain, were tested from a prodromal stage, but failed to slow cognitive decline (reviewed in [116]).

Together, these observations suggest A β may not be the causative agent in AD. Alternate views of AD pathogenic processes must be explored to develop therapeutics which will successfully slow or prevent cognitive decline.

1.5.2 The MAM hypothesis

Mitochondria-associated endoplasmic reticulum membranes (MAMs) are specialised, lipid-raft-like regions of the endoplasmic reticulum (ER) which are directly adjacent to mitochondria [117]. MAMs are enriched in lipids such as cholesterol [118] and the presenilin proteins [119]. The other components of γ -secretase and APP are present in high amounts [119]. Importantly, it appears that the majority of γ -secretase cleavage of APP in mouse brains occurs at MAMs [119]. MAMs are involved in lipid metabolism, calcium homeostasis, and in mitochondrial function (reviewed in [120-123]). The formation of autophagosomes occurs at MAM [124], implicating MAM in autophagy. Additionally, the NLRP3 inflammasome, a molecular complex which acts as a sensor for cellular stress, localises to MAM when reactive oxygen species (ROS) are generated (such as when mitochondria are dysfunctional and mitophagy is impaired) [125], implicating MAM in inflammation.

The MAM hypothesis of AD pathogenesis (**Figure 9**) posits that AD stems from upregulated MAM function [120-122, 126]. A very interesting study by Area-Gomez and colleagues [118], showed evidence for increased MAM function and ER-mitochondrial cross-talk in fibroblasts from EOfAD (various *PSEN1* mutations) and LOAD patients relative to controls. As a measure of MAM function, they measured the conversion of free cholesterol to cholesteryl esters, while the measure of cross-talk between the ER and mitochondria was assessed by rates of phospholipid synthesis and transfer [118]. Additionally, the physical area of MAM was observed to be increased in the AD cells, providing a possible explanation for the increased biochemical activity [118]. Another particularly interesting study from the same group explored the role of the C99 fragment of APP in MAM. Further investigation of the role of MAM in AD showed that the C99 fragment of APP was enriched in MAMs, and

that C99 (and not A β) appears to drive mitochondrial bioenergetic defects by upregulating sphingomyelin hydrolysis to form ceramide [127], an inhibitor of the oxidative phosphorylation pathway and a pro-apoptotic molecule (reviewed in [128]). Interesting, inhibition of C99 formation (by inhibiting β -secretase activity) in *PSEN* double knockout cells (which lack γ -secretase activity), rescued the mitochondrial respiration defects [127]. Together, these observations strongly implicate that C99, and not A β , is the pathological agent in AD pathogenesis.

In summary, the MAM hypothesis provides a mechanistic explanation for many pathologies observed in AD (reviewed in [126, 129]).

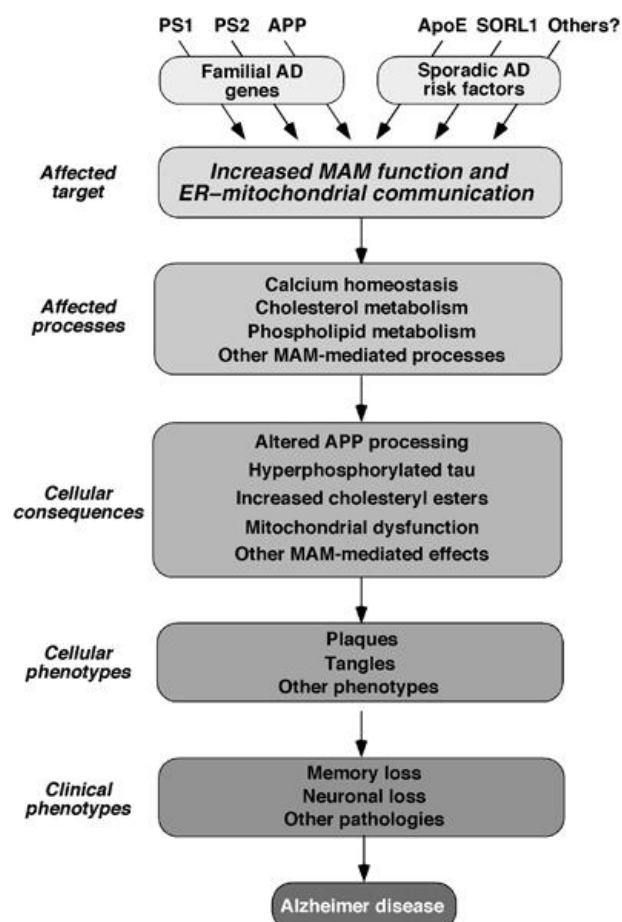


Figure 9: MAM hypothesis of AD pathogenesis. A flow diagram detailing the MAM hypothesis, described in a series of review articles from the group of Dr Estela Area-Gomez and Professor Eric Schon [120-122, 126, 129]. Image reproduced with permission from [118].

1.5.3 The vascular hypothesis

Another alternative hypothesis of AD pathogenesis is the vascular hypothesis. It was proposed by de la Torre JC and Mussivan T [130] and states that '*AD is caused by a cerebral microcirculatory impairment, probably evolving slowly and progressively over many years*' [130]. They suggest that aging brain capillaries are disrupted, and this causes blood flow problems which limit delivery of nutrients such as oxygen and glucose to the brain (**Figure 10**).

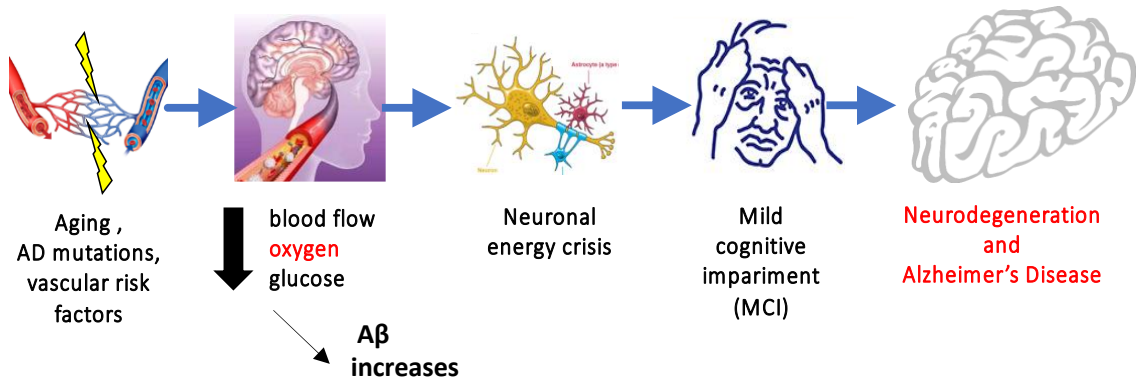


Figure 10: The vascular hypothesis of AD. The flow chart describes how aging, mutations and vascular risk factors eventually lead to dementia.

Magnetic resonance imaging (MRI) studies of AD patients have shown that, indeed, cerebral blood flow is reduced compared to age matched controls [131, 132].

Perhaps the most intriguing evidence supporting the vascular hypothesis of AD comes from a large scale (1,171 healthy and AD patients), multifactorial data-driven

analysis of various biomarkers of AD in the progression from healthy controls to LOAD. This displayed that vascular dysfunction was the earliest, and most prominent pathogenic factor in the progression from healthy controls, through the spectrum of mild cognitive impairment, to LOAD [133]. The vascular hypothesis of AD is further supported by the morphological changes to cerebral vasculature reported in post mortem AD brains (i.e. kinking and looping of vessels [134, 135], bumps and knobs forming on the surface of vessels [130] and capillary loss in certain regions [136]). Vascular abnormalities are also observed in animal models of AD. Gama Sosa et al. [137] observed that mice expressing human *PSEN1* EOfAD mutations in only neurons show morphological changes to brain vasculature. They observed kinked and looped vessels, 'string vessels', loss of vasculature density and thickening of the vascular basal laminae. Interestingly, these phenotypes were observed without A β accumulation. Morphological abnormalities in the vasculature were also observed in a transgenic mouse model of AD [138]. Scanning electron microscopy images of this transgenic mouse model (APP23) show regions in the left anterior cortex of the brain which lacked microvasculature. These regions were shown to contain amyloid plaques. Structures were also observed on the surface of the vessels, which the authors referred to as 'pompoms' and 'cubes' due to their appearance. These deposits were mainly composed of A β [138].

Other evidence for the vascular hypothesis of AD arises from a series of experiments performed by the group of de la Torre in the early 1990's. They investigated the effects of chronic brain hypoperfusion (CBH, i.e. reduced blood flow) on young, and aged rats and found that the aged animals were less effective at recovering from surgically induced CBH than the young animals. The aged rats showed impaired neuronal energy metabolism [139, 140], astrocytosis [141], oxidative stress [140], learning deficits [142], increased amyloid deposition in perivascular cells [143], brain

atrophy and death [139] after the CBH. These observations are reminiscent of AD pathology, and support the idea that impaired vasculature may play a role in AD.

1.6 Hypoxia, iron and HIF1- α in Alzheimer's disease

As we age, oxygen supply to the brain gradually becomes impaired as cerebral blood flow [131, 132] and angiogenesis [144] decrease with age. Unfortunately, the ability of cells in the brain to cope with hypoxic (low oxygen) conditions also decreases with age [145]. Longitudinal studies have shown that ischemic stroke increases an individual's risk of developing AD [146]. Also, markers of hypoxia in serum of patients with mild cognitive impairment (MCI, a precursor to AD) were found to predict whether the patients go onto develop AD or not [147]. Zetterberg et al. [148] observed that heart attacks (an extreme form of hypoxia) cause spikes in blood A β levels in humans and that the magnitude of the spike correlated with worsened clinical outcomes. Taken together, these results support that decreased cerebral oxygen levels may play a role in the pathogenesis of AD (which is supportive of the vascular hypothesis of AD).

The hypoxia inducible factor (HIF) complexes mediate the cellular response to changes in oxygen levels. The HIF complexes are part of the basic helix-loop-helix–per-arnt-sim (bHLH–PAS) protein family of transcription factors. The complex consists of two subunits, HIF- α and HIF- β . In humans, there are three homologues for the α subunit (HIF-1 α , HIF-2 α and HIF-3 α) and three homologues for the β subunit (HIF-1 β /aryl hydrocarbon receptor nuclear translocator [ARNT1], ARNT2, ARNT3). HIF-1 α and HIF-1 β are expressed in most tissue types while the other homologues are expressed predominantly in specific tissues (reviewed in [149]).

HIF-1 β is constitutively expressed, while HIF-1 α is degraded in an oxygen (and ferrous iron) dependent manner [150]. Under normal oxygen conditions (normoxia), HIF-1 α is labelled for degradation by prolyl hydroxylases (PHDs). However, under hypoxic (or ferrous iron depleted [151]) conditions, HIF-1 α is stabilised, allowing it to dimerise with HIF-1 β . This allows the complex to move into the nucleus to act as a transcription factor and regulate expression of target genes involved in cellular processes such as angiogenesis and metabolism [149, 150] (**Figure 11**).

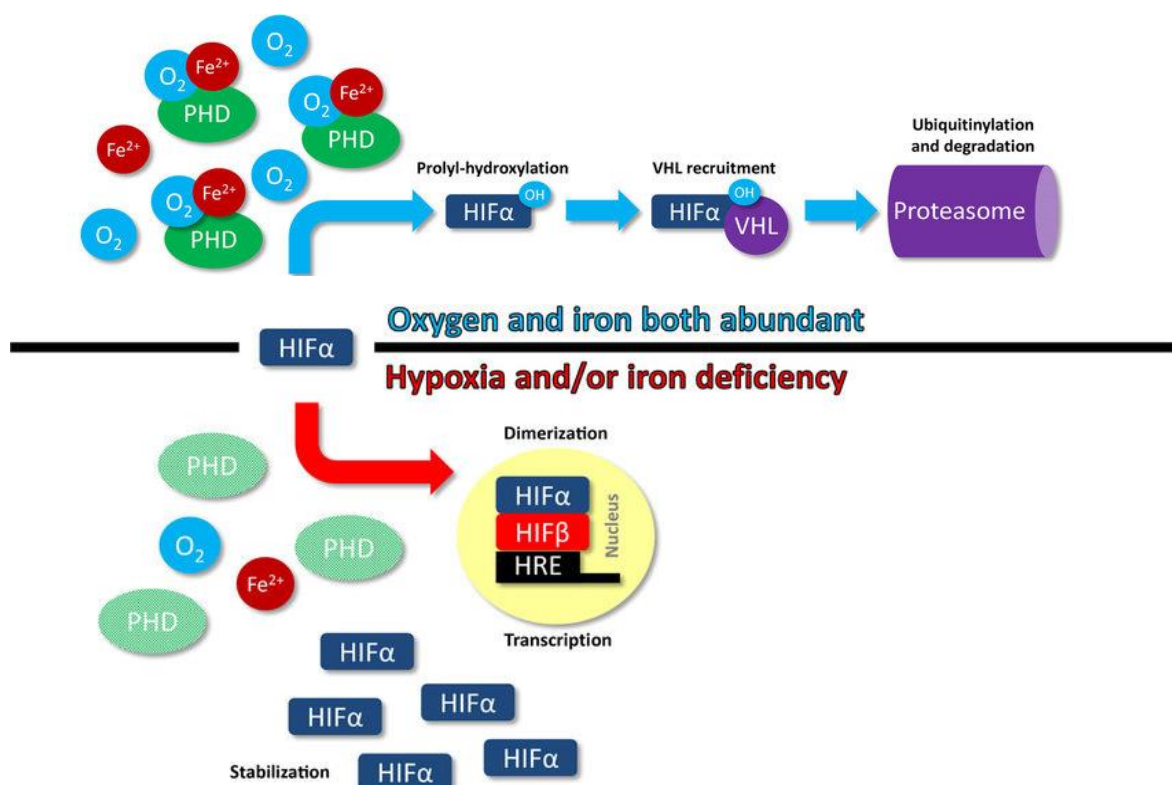


Figure 11: Schematic of the molecular consequences of HIF1- α in the presence of oxygen (O_2) and iron (Fe^{2+}), or under hypoxia and/or iron deficient conditions. Image is modified from [152].

An interesting experiment by Ndubuizu et al. [153] gave some insight into how the brain is more susceptible to hypoxia with age. They observed that, with advancing age in rat brains, came decreased HIF-1 α accumulation under hypoxia. They showed by Northern blot that *Hif-1 α* mRNA levels were similar in young and aged rat brains

when under normoxia and hypoxia, which suggested that the change in Hif-1 α levels is at the protein level. This was reflected in changes of transcript levels of HIF response genes such as vascular endothelial growth factor (*Vegf*). Transcriptional activation of this gene was significantly reduced in aging rat brains compared to young rat brains under hypoxic conditions. They also showed that this decreased expression of HIF target genes was not due to a global decrease in transcriptional activity, as the transcription of genes which are responsive to hypoxia independently of HIF showed a similar response in old and young rat brains. They showed that this age-dependent decrease of HIF-1 α under hypoxia is probably due to increased levels of PHD. This phenomenon was also observed by Frenkel-Denkberg et al. [154] in mouse brains, while Newman et al. [155] observed similar changes in zebrafish.

The stabilisation of HIF-1 α is dependent not only on oxygen but also iron. Iron (Fe) is essential for life and is involved in a number of processes in cells. Iron homeostasis is a complex phenomenon with many factors influencing the overall activity. For comprehensive reviews on this topic, see [156-158]). Briefly, iron is acquired through diet, and the majority of iron in circulation is in the ferric state (Fe³⁺), bound to a chaperone transferrin (TF) (reviewed in [157]). Ferric iron bound to TF is internalised into cells via TF-receptor-mediated import [159]. Acidification of the endosome releases Fe³⁺ from TF, then Fe³⁺ is reduced to more reactive ferrous iron (Fe²⁺) by endosomal ferrireductase (STEAP3) [160]. Fe²⁺ is then transported across the endosomal membrane by metal transporters such as divalent metal transporter 1 (DMT1) and then becomes part of the labile iron pool in the cytosol, which can be subsequently utilised by the cell (see **Figure 12** and reviewed in [161]). A few examples of iron-requiring cellular processes include incorporation into iron-containing proteins such as cytochromes and iron-sulfur cluster containing proteins, metabolism in the mitochondria and storage in the protein nanocage ferritin.

There is strong evidence that perturbation of iron metabolism may play a role in AD (a recent review from our laboratory details this evidence, see [162]). However, perhaps the most intriguing evidence that iron homeostasis is affected in AD is that the genes most commonly implicated in EOofAD: *PSEN1* and *APP*, encode proteins which affect acidification of the endo-lysosomal system. Evidence for whether *PSEN2* plays a similar role to *PSEN1* in iron homeostasis is still yet to be presented. Failure to acidify the endolysosomal system (such as in *PSEN1* EOofAD patient fibroblasts [76], or when levels of the C99 fragment of APP are increased [163]), was shown to result in a pseudo-hypoxic response (i.e. stabilisation of HIF1- α), inflammation and mitochondrial abnormalities *in vivo* [164], changes which are well accepted to occur in the AD brain.

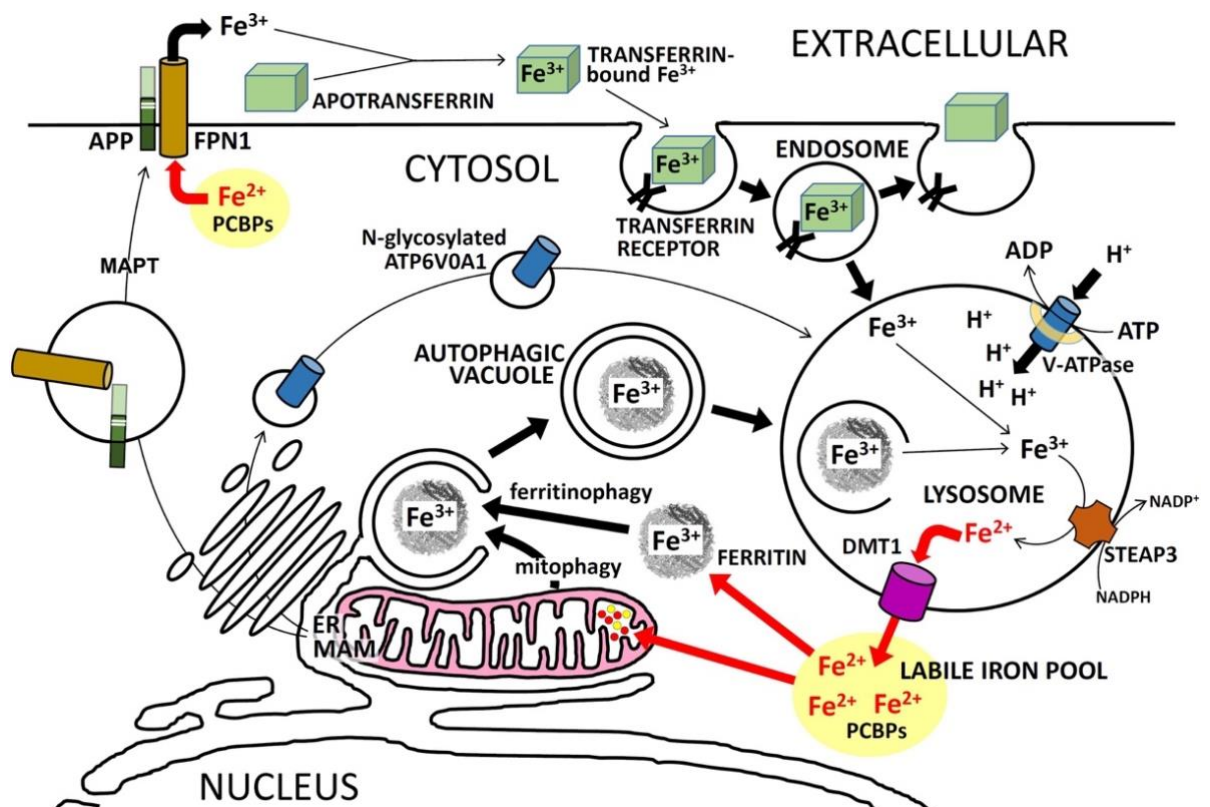


Figure 12: Schematic detailing transferrin-dependent ferrous iron (Fe^{3+}) import, conversion to Fe^{2+} via acidification of the lysosome (implicating PSEN1 and APP) and export to the labile iron pool in the cytoplasm of cells. Downstream cellular processes of Fe^{2+} from the labile iron pool are depicted. For a more complete description of this Figure, see [162] (from which this image is reproduced). Copyright © 2018 Lumsden, Rogers, Majd, Newman, Sutherland, Verdile and Lardelli

1.7 Animal models of AD

Animal models have been a vital tool to investigate the pathogenesis of human diseases. The normal function of genes and their products can be investigated through observing the phenotype of a knockout model. Knock-in models of human mutations can be engineered into the genome of the model organism to attempt to model the pathological effect(s) of that mutation.

Non-human primates share recent common ancestors with humans. Therefore, they are likely the most accurate models of AD pathology. Aged non-human primates do naturally show some pathologies reminiscent of AD (e.g. $\text{A}\beta$ deposition, reviewed in [165]). However, due to their long lifespan, and the cost associated with raising them until they are of an appropriate age, non-human primates are not the most practical model to study AD. Nevertheless, a *PSEN1*^{4E9} knock-in marmoset has been recently generated [166], presenting as the first non-human primate model of EOfAD. It will be some time until this recently generated model of will be old enough for investigations of AD pathogenesis. However, it will likely present as the most accurate animal model of AD to date.

Mus musculus (mouse) is by far the most popular organism for modelling AD. Many transgenic model lines have been developed which attempt to model AD pathology, the first being the PDAPP line. PDAPP mice express a transgene containing the

platelet-derived growth factor (*PDGF*)-b promotor (neuron-specific) driving the expression of a human *APP*^{V717F} (Indiana) mutant allele. Mice heterozygous for the transgene develop amyloid deposition at approximately 6-9 months old but do not develop NFT pathology [167]. The 5xFAD mouse model is another commonly used transgenic AD model. It overexpresses five EOfAD mutations in two genes (*APP*^{K670N/M671L}, *APP*^{I716V}, *APP*^{V717I}, *PSEN1*^{M146L} and *PSEN1*^{L286V}) driven by a thymocyte antigen 1 promotor (*Thy1*, neuronal expression). This mouse line exhibits high levels of A β ₄₂, aggressive plaque formation, gliosis, degeneration of synapses, neuron loss and memory impairment [168]. Although this mouse model shows many of the pathologies observed in AD, it does not show NFT pathology. The 3xTg mouse model contains the *PSEN1*^{M146V} mutation in the endogenous *Psen1* gene, and also expresses the human *APP*^{K670N/M671L}, and *MAPT*^{P301L} mutant alleles under the control of a *Thy1.2* promotor. This line of mouse may possibly represent a more complete model of AD as it exhibits plaque formation, NFTs, cognitive defects and neuron loss [169].

Although these mouse models capture a range of the pathologies observed in AD, whether or not they reflect the true pathology of AD is debateable. This is because human AD patients generally only carry one EOfAD mutation. Also, most mouse models overexpress the human mutant genes from novel promotors which does not reflect the physiology of AD in the human brain. Knock-in models of EOfAD exist in the literature (e.g. [170, 171]). However, they are not as well characterised and are infrequently exploited in comparison to the transgenic models, likely due to their lack of discernible phenotypes.

1.8 Zebrafish as a model organism:

Zebrafish (*Danio rerio*) can be a useful tool for modelling AD pathology. Most importantly this species has genes orthologous to the human genes thought to play essential roles in AD such as *APP*, *PSEN1*, *PSEN2* and *SORL1* (reviewed in [51]). Zebrafish embryos develop *ex utero* and are particularly manipulable due to their large size. Therefore, targeted mutagenesis (i.e. using the CRISPR-Cas9 system) is relatively straightforward in zebrafish (compared to mice, where the embryo develops inside the mother). Large scale mutagenesis and screening strategies are feasible in zebrafish due to their larger progeny numbers. However, one of the most important properties of zebrafish in the context of the work presented in this thesis is that genetic and environmental variation can be reduced between samples by raising the progeny of a single pair mating event together in the same tank.

1.9 Behavioural characterisations of animal models

The most characteristic pathology observed in AD is memory impairment. Many types of memory are proposed for humans and detailing all of the proposed types is beyond the scope of this thesis. However, impairment in working memory (which can be thought of as short term memory) is an early pathological change observed in AD (and MCI) (reviewed in [172]). Working memory is often tested in animal models by assessment of spatial working memory using various shaped mazes, including the Y-maze (reviewed in [173]). Impairment of working memory has been observed in transgenic models of AD mutations (reviewed in [174]), and in the *APP* knock-in models of Saito et al. [171]. Variable changes to working memory have been observed in *PSEN1*^{M146V} knock in mice [175, 176].

To assess working memory in zebrafish, the free-movement pattern (FMP) Y-maze task was developed [177]. This method involves continuous tracking of zebrafish while they freely explore a Y-maze for a set amount of time, and is based on the intrinsic behaviour of animals to explore a new area, and alternate their choices [178]. The sequences of choices the fish makes of turning left or right when it reaches the centre of the Y-maze is noted, and a series of frequencies of overlapping tetragrams (i.e. four consecutive arm entries) is calculated. There are 16 possible tetragrams (ranging from LLLL to RRRR), and the alternation tetragrams (LRLR and RLRL) are proposed to be the measure of working memory (**Figure 13**). This method is verified by the reduced number of alternations zebrafish perform when treated with small molecules known to be inhibitors of memory [177]. Intriguingly, both mice and humans behave similarly to zebrafish during the FMP Y-maze task [177], supporting the translatability of this behavioural task.

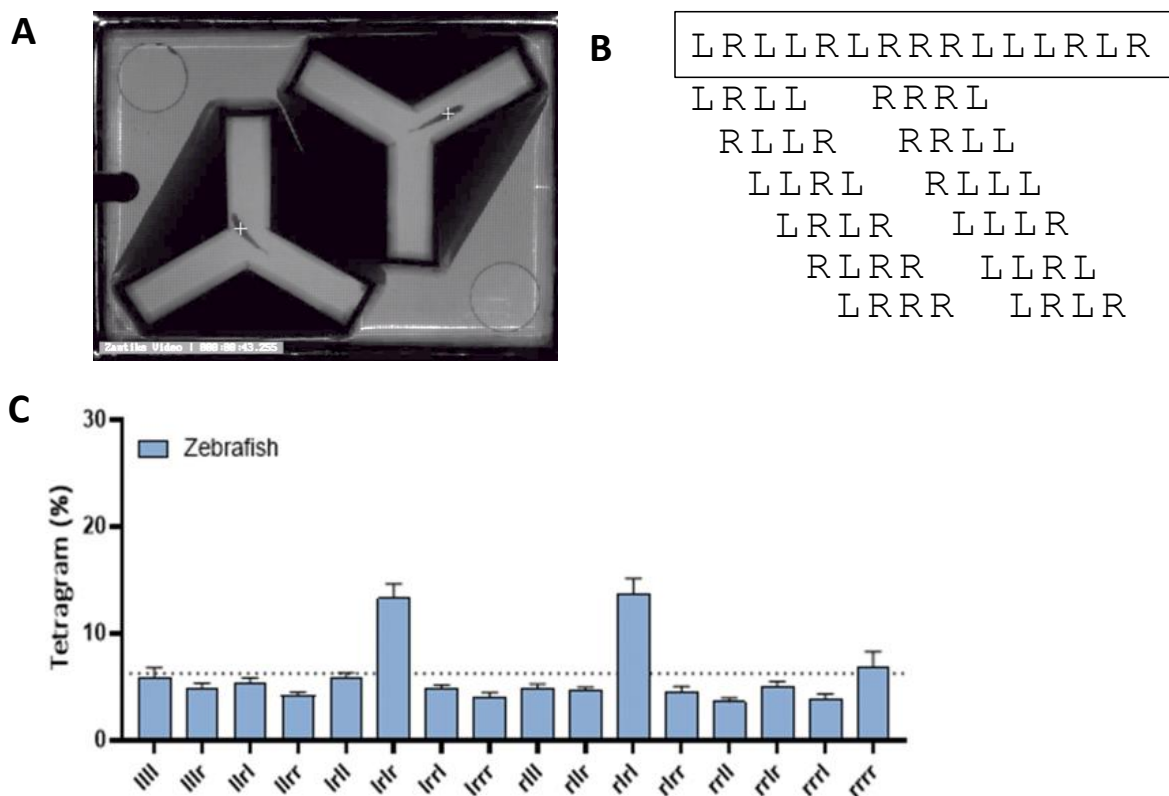


Figure 13: The Free Movement Pattern (FMP) Y-Maze task. **A** Image of zebrafish inside Y-mazes. A zebrafish is allowed to swim freely in a maze for 1 hour, and the choices of whether it turns left (L) or right (R) when it reaches the centre of the maze is recorded. **B** Identification of overlapping tetragrams in the behaviour data. **C** Zebrafish naturally show greater proportions of the LRLR and RLRL (i.e. alternation) tetragrams. **C** is reproduced from [177].

1.10 Transcriptome analysis using RNA-seq

The transcriptome is the global set of RNA molecules transcribed in a cell or a population of cells. Comparisons of transcriptomes between a condition of interest (i.e. mutation, treatment) and a control allows changes to gene expression to be assessed, and changes to cellular function can be predicted. Initially, transcriptomes were assessed using DNA-microarrays. However, with the emergence of next generation sequencing technologies, RNA-sequencing (RNA-seq) has become the most popular method for transcriptome analysis. Unlike microarrays, RNA-seq does not require any prior knowledge of the genome of the species from which the RNA was derived, as the transcriptome can be generated *de novo*. Additionally, RNA-seq has greater sensitivity for genes which are expressed at very high or low levels. It also overcomes the problem of cross hybridisation which can occur in microarray experiments between highly similar sequences. Additionally, sequence-level information is readily available in RNA-seq data (reviewed in [179]).

The methods typically used to perform RNA-seq experiments first involve extracting high quality total RNA from a cell or population of cells. The most abundant type of RNA in total RNA is ribosomal RNA (rRNA), which is mostly uninformative for the assessment of the transcriptome state. Two popular methods for removing rRNA from a total RNA sample include rRNA depletion, or selection of mRNA by enriching for polyadenylated (polyA) tails (reviewed in [180]). Then, the remaining RNA is

fragmented, reverse-transcribed into cDNA libraries, and sequencing adaptors are ligated. Finally, the cDNA libraries are sequenced using a high-throughput platform such as Illumina to generate, typically, millions of short reads per sample. After assessment of the quality of the resulting data, and if the reference genome is known, reads can then be mapped back to the genome. However, this is not necessary and the reads can be aligned into contigs, which then can be used to generate transcript sequences. The number of reads per transcript is proportional to the level of gene expression (**Figure 14**, for a review on the principles and methods of RNA-seq, see [181]).

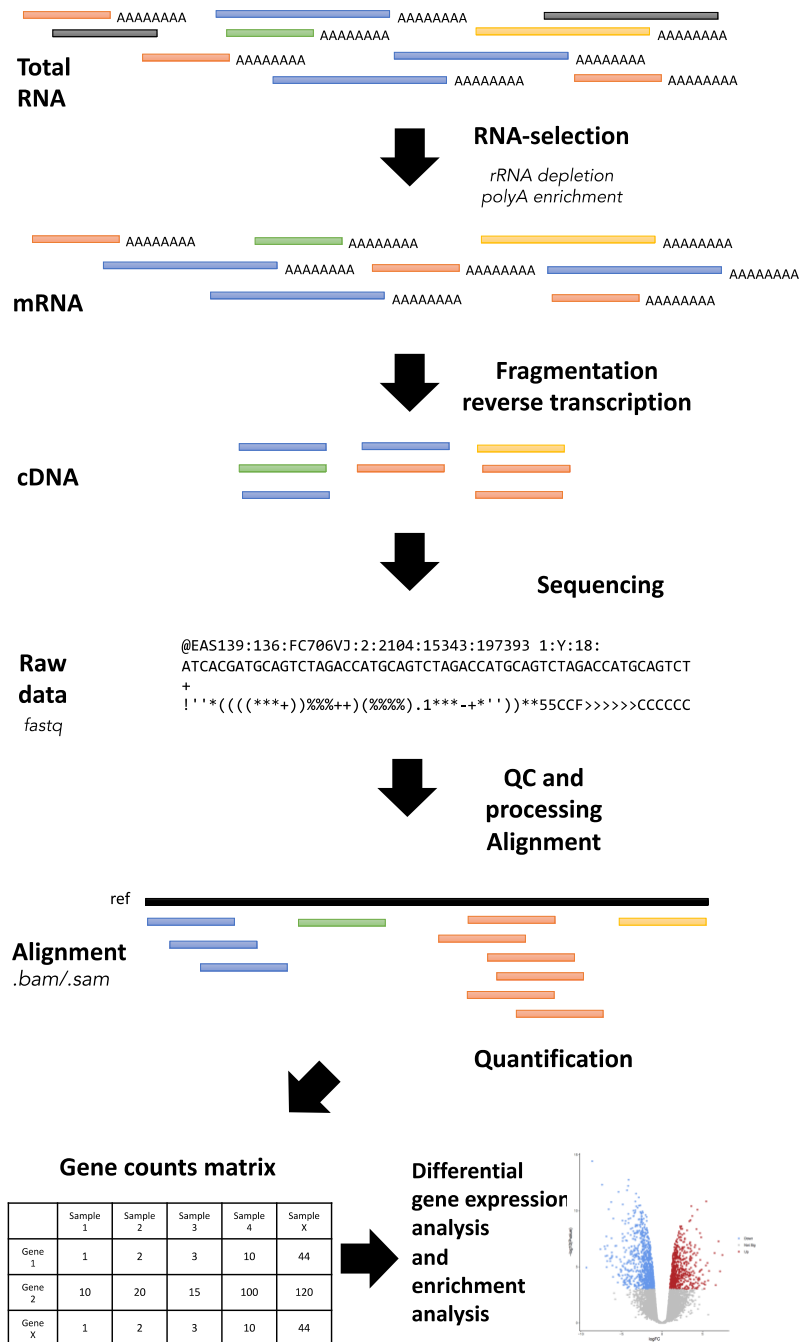


Figure 14: A simplified, typical, RNA-seq experiment workflow. Total RNA is extracted from cells, ribosomal rRNA is depleted (grey boxes), then the remaining RNA is converted to cDNA libraries. cDNA is sequenced to produce fastq files containing typically millions of reads. After quality control and processing, reads can be aligned back to the reference genome/transcriptome. The number of reads per transcript/gene is proportional to the level of gene expression and is summarised in a counts matrix. This counts matrix is used for differential gene expression analysis and enrichment analysis.

The most common goal of an RNA-seq experiment is to perform differential gene expression (DGE) analysis to obtain a list of differentially expressed genes. After normalising for differences between library sizes between samples, and filtering genes which only display a few counts (which are statistically uninformative), a differential expression test can be performed. Three main algorithms are typically used to perform DGE analysis (see **Table 1**).

Table 1: Summary of commonly used differential gene expression tests			
Name	Type of distribution	Type of differential expression test	Reference
<i>limma</i>	Voom* transformation to approximate normality	Moderated t-tests with empirical Bayes methods	[182]
<i>edgeR</i>	Negative binomial distribution	Exact test, quasi-likelihood F-test or likelihood ratio test	[183]
<i>DESeq2</i>	Negative binomial distribution	Wald test or likelihood ratio test	[184]
* Variance modelling at the observational level [185]			

To put a list of differentially expressed genes into biological context, enrichment analysis of pre-defined gene sets can be performed. Enrichment analysis will test whether *groups* of genes together show significant differential expression, and can give insight into whether a particular biological process is altered. Various enrichment analysis methods have been developed (see [186] and **Table 2**). The gene sets used in enrichment analysis are pre-defined, based on prior knowledge of genes involved in particular biological processes. Common gene sets used in enrichment analysis include gene ontologies (GO) [187], and the HALLMARK [188] and Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets [189].

Table 2: Summary of commonly used gene set enrichment analysis methods				
Name	Type of hypothesis *	Statistic	Accounts for inter-gene correlations?	Reference
<i>goseq</i>	Competitive	Overlap of differentially expressed genes and gene set, correcting for sampling bias (typically gene length or GC content)	No	[190]
<i>fry/roast</i>	Self-contained	Weighted mean	Yes	[191]
<i>camera</i>	Competitive	t-statistic accounting for inter-gene correlation	Yes	[192]
<i>GSEA</i>	Self-contained	Kolmogorov–Smirnov	No	[193]
<p>* a self-contained null hypothesis is that no genes in the gene set are differentially expressed in the condition of interest (i.e. genotype). A competitive null hypothesis is genes in the gene set are not more differentially expressed in the condition of interest than genes not in the gene set</p>				

1.11 Previous RNA-seq studies of AD

RNA-seq analysis on post-mortem brain tissue from AD patients has shown that a wide range of biological processes are disrupted (see **Table 3** for a summary of selected RNA-seq studies). The expression of genes involved in neuronal processes and inflammation is commonly observed. However, post-mortem AD brain tissues represent the end stages of AD pathology, where damage to the brain is considerable and the causative molecular event may not be discernible.

Reference	Brain region	Samples	Main findings
Twine et al. 2011 [194]	Whole brain, frontal and temporal lobes	13 male and 10 female whole brain, 5 male frontal lobe and 5 male temporal lobe control samples, 1 male AD total brain, 1 male frontal lobe and 1 male temporal lobe AD samples.	Common and region specific changes are observed in all brain samples. Stronger changes to synaptic processes are observed in the frontal lobe. Differential usage of <i>APOE</i> isoforms.
Mills et al. 2013 [195]	Parietal lobe	1 pooled RNA sample from 5 male LOAD samples 1 pooled RNA sample from 5 control samples	Alterations to developmental processes, endocytosis, ATPases, protein degradation, nucleotide metabolism, synaptic transmission, lipid metabolism and inflammation.
Bennet and Keeney 2018 [196]	Frontal cortex (ventral midbrain for Parkinson's disease)	10 AD patients and 10 controls, 14 PD patients and 10 controls	Synthesis and salvage of nicotinamide adenine dinucleotide (NAD) appears to a common pathway altered in AD and Parkinson's disease.
Neff et al. 2021 [197]	frontal pole, superior temporal gyrus, parahippocampal gyrus, and inferior frontal gyrus	1,543 AD and control transcriptomes	Three major sub-types of LOAD are revealed, classified by differential expression of genes involved in: responses to tau and A β , inflammation, synaptic signalling, mitochondria, and myelination.

An individual is said to show mild cognitive impairment (MCI) if cognitive decline is present without meeting the criteria for dementia (reviewed in [198]). MCI is often a precursor to AD, so the changes to gene expression observed in MCI could give insight to the biological processes eventually leading to AD. Berchtold et al. [199] investigated the gene expression patterns of MCI compared to AD and non-

demented controls using microarrays. They observed that synaptic and mitochondrial bioenergetic genes do not show a progressive linear relationship. Rather, they show upregulation in MCI and downregulation in AD, suggestive of that mitochondrial changes occur early in AD progression, and that changes observed late in AD disease progression are different to the early changes occurring before dementia sets in.

A general consensus in the AD-research field is that early detection and treatment is necessary to slow, or even prevent cognitive decline. To prevent cognitive decline, we must understand the initiating molecular/cellular stresses which eventually lead to AD. As highlighted beforehand, this likely cannot be achieved using post-mortem AD brain tissue. Analysis of the molecular changes occurring in the brains of young individuals who are predisposed to AD (i.e. who carry mutations in *PSEN1/2* or *APP*) would likely reveal these changes. However, the living brains of such individuals are inaccessible for detailed molecular characterisations. Therefore, we must utilise animal models to investigate these changes.

The brain transcriptomes of commonly used transgenic mouse models of AD show low concordance with human AD, and with each other [200]. This questions the validity of these models in interpretation of transcriptomic data relevant to human AD. These discrepancies are likely due to the different and complex genetic manipulations performed to generate these mice, which deviate significantly from the human disease genetic state. Knock-in genetic models which mimic the genetic state of the human disease may paint a more accurate picture of how the disease manifests. Since mutations in the presenilin genes, *APP* and, potentially, *SORL1*, are causative for EOfAD, analysis of the effects of these mutations in young animal models may illuminate the early cellular processes which eventually lead to AD.

1.12 Previous work of the Alzheimer's Disease Genetics Laboratory

The brain transcriptomes of any young, knock-in EOfAD animal models have not been analysed. Filling this gap in knowledge has been the subject of a long-term research program of the Alzheimer's Disease Genetics Laboratory (ADGL). This program has aimed to generate a suite of zebrafish genetic models that possess similar mutations to those that cause EOfAD in humans. We predict that transcriptomic analysis of young zebrafish brains will reveal an "early EOfAD transcriptomic signature". Comparisons with mutations which do not cause EOfAD (i.e. frameshift mutations in the presenilin genes) can reveal, by exclusion, the specific effects of the EOfAD-like mutations in zebrafish. As shown in **Figure 15A**, we employ a breeding program that results in a family of sibling zebrafish containing the genotypes of interest that are raised together in a shared environment. This minimises unwanted sources of genetic and environmental variation, and, theoretically, should maximise statistical power to detect subtle changes to the transcriptome due to the mutation. The cellular processes predicted to be affected in common by the different mutations in different genes in the zebrafish brain transcriptomes may illuminate the early pathological changes which would eventually lead to EOfAD in humans (**Figure 15B**).

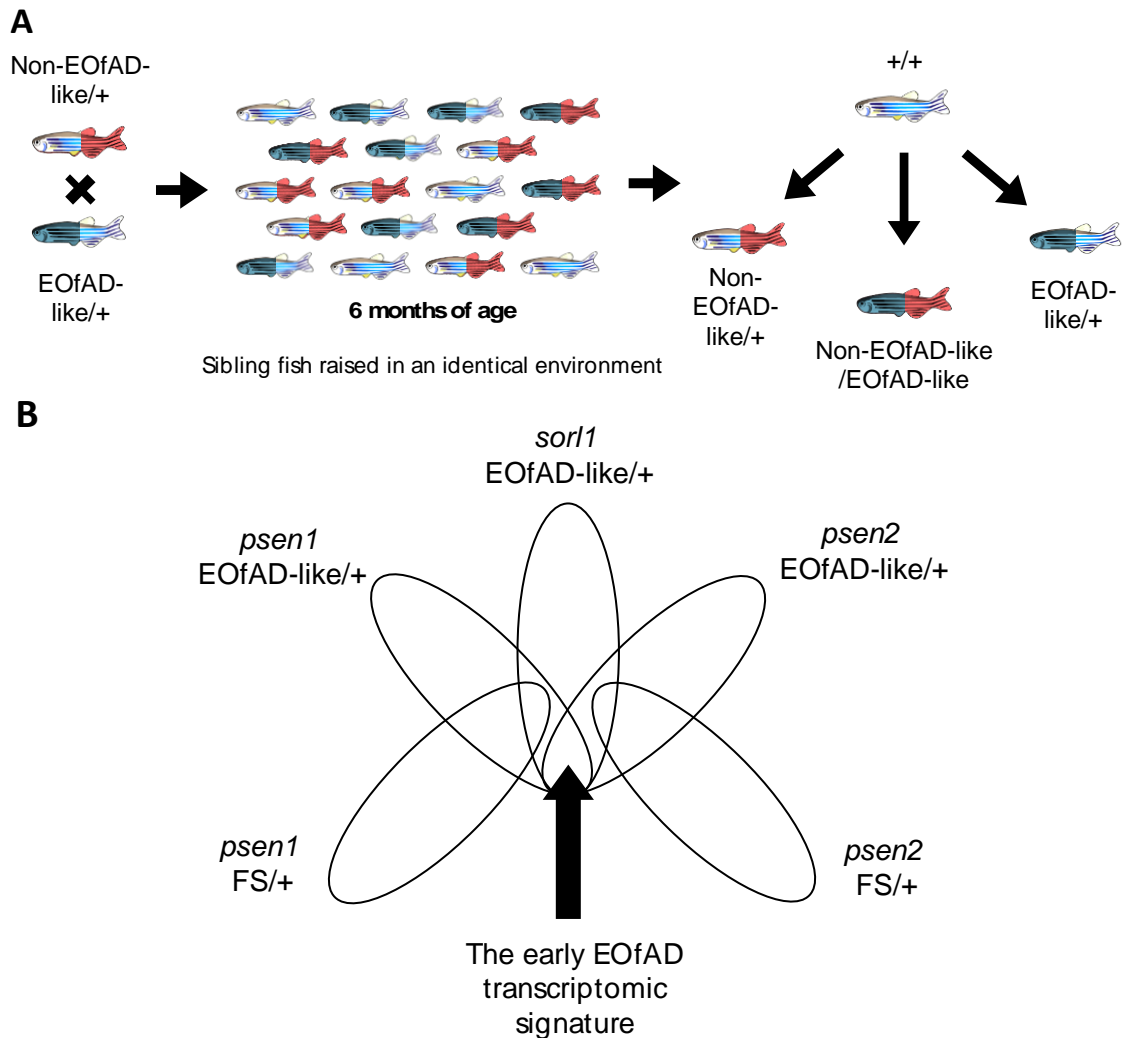


Figure 15: **A** Typical breeding strategy for a zebrafish RNA-seq experiment. A fish heterozygous for an EOfAD-like mutation (EOfAD-like/+) is pair-mated with a fish heterozygous for a non-EOfAD-like mutation (Non-EOfAD-like/+), such as a frameshift mutation in a presenilin gene. This results in a family comprised of sibling fish (i.e. with limited genetic variation), each with one of four genotypes: wild type (+/+), EOfAD-like/+, Non-EOfAD-like/+ or transheterozygous for these mutations (Non-EOfAD-like/EOfAD-like). These sibling fish are raised together in a single tank, reducing environmental variation. Pairwise comparisons between each mutant genotype relative to their wild type siblings as can be performed in a transcriptome analysis are depicted. **B** Simplified Venn diagram depicting the early EOfAD transcriptomic signature. Each oval represents genes or gene sets altered in the brain transcriptome(s) of a particular zebrafish mutant. The transcriptomic changes which occur in EOfAD-like transcriptomes, but not in non-EOfAD-like transcriptomes, define the signature.

Previously, the AGDL has generated a zebrafish knock-in model of the frameshift mutation *PSEN2*^{K115Efs} (this mutation was previously described in section [1.3.2](#)) [201]. As described earlier, the truncated transcript of *PSEN2*^{K115Efs*} is predicted to encode a protein similar to the PS2V isoform of human *PSEN2* which shows increased expression in LOAD [65]. The PS2V-like isoform in zebrafish is expressed from their endogenous *psen1* gene, i.e. PS1IV, rather than their orthologous *psen2* gene [68]. Therefore, the *psen1*^{K97fs} mutant line of zebrafish was established to model the molecular effects of PS2V. During the isolation of the *psen1*^{K97fs} mutant line of zebrafish, an in-frame deletion of two-codons at the K97 site of *psen1* was also identified: *psen1*^{Q96_K97del}. This in-frame deletion follows the reading-frame preservation rule of EOfAD mutations and therefore, models a typical EOfAD mutation in a presenilin gene. Heterozygosity for the *psen1*^{Q96_K97del} or the *psen1*^{K97fs} mutation results in apparent hypoxic stress under normoxia in young adult brains, and age-dependent changes in transcriptional responses to hypoxia [155].

RNA-seq analysis has previously been performed to address the effects of these mutations on the young zebrafish brain transcriptome [201-203]. The effect of age was also examined in both mutant analyses, while the effect of both age and acute hypoxia treatment was examined for the *psen1*^{Q96_K97del} mutant. In these analyses, the brain transcriptomes of 6 month old (i.e. young adult) heterozygous mutant zebrafish are compared to their wild type siblings. The *psen1*^{K97fs} mutation results in changes to gene expression suggestive of accelerated aging, and increased glucocorticoid signalling [201], while the *psen1*^{Q96_K97del} mutation appears to have effects on mitochondrial function, lysosomal acidification and iron homeostasis [202, 203]. These transcriptomic changes occurring in young *psen1* mutant zebrafish may give insight into the early molecular stresses which eventually lead to AD.

1.13 Concluding remarks

The overall aim of the manuscripts presented in this thesis contribute to the ADGL's goal of understanding the changes occurring in young-adult, EOfAD mutation carrier brains using zebrafish *knock-in* models. First, a detailed review on the role of *SORL1* in AD is presented in **Chapter 2**. **Chapter 3** consists of an investigation of the hypothesis that young, heterozygous *psen1*^{Q96_K97del} and *psen1*^{K97fs} mutant zebrafish show an apparent hypoxic response at normal oxygen levels in their brains due to changes in brain vasculature. **Chapter 4** consists of a preliminary investigation as to whether the previously established *psen1*^{Q96_K97del} zebrafish, and novel *sorl1* mutant zebrafish (generation of which will be described in **Chapters 5** and **6**) show any age-related changes to spatial working memory by assessment of their performance in a Y-maze.

Following these Chapters, this thesis presents a series of manuscripts describing the generation and/or brain transcriptome analyses of various knock-in models of EOfAD mutations in zebrafish. It concludes with a summary of all of our analyses of zebrafish EOfAD genetic models to define the "early EOfAD transcriptome signature." A common theme in these manuscripts is the analysis of zebrafish mutants compared to their wild type siblings which are raised together in a single tank (**Figure 15A**). **Chapters 5** and **6** detail the generation and brain transcriptome analysis of the first knock-in animal models of EOfAD-like mutations in *SORL1 in vivo*. **Chapter 7** consists of a transcriptome analysis of an EOfAD-like, and a frameshift mutation, of *psen2*. **Chapter 8** performs a similar experiment for *psen1*. It compares a frameshift mutation (similar to that of the *PSEN1*^{P242fs} mutation causing the skin disease familial acne inversa), with an exact equivalent of the human EOfAD mutation *PSEN1*^{T440del} [57]. **Chapter 9** consists of a hybrid Results and Discussion chapter presenting, as a

manuscript, a summary of the results from **Chapters 5-8**, together with those from previously analysed transcriptomes of *psen1^{Q96_K97del}* and *psen1^{K97fs}* mutant zebrafish. These transcriptomes are subsequently compared with publicly available brain transcriptome data from the *App^{NL-G-F}* knock-in mouse model of EOfAD, and knock-in mouse models of the *APOE* alleles implicated in LOAD. In summary, this thesis adopts mainly as a data-driven approach to illuminating the early cellular stresses that, eventually, would lead to EOfAD in humans.

1.14 References

1. Lobo A, Launer LJ, Fratiglioni L, Andersen K, Di Carlo A, Breteler MM, Copeland JR, Dartigues JF, Jagger C, Martinez-Lage J, et al: **Prevalence of dementia and major subtypes in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group.** *Neurology* 2000, **54**:S4-9.
2. **Dementia Prevalence Data 2018-2058**
[<https://www.dementia.org.au/statistics>]
3. Blennow K, de Leon MJ, Zetterberg H: **Alzheimer's disease.** *The Lancet* 2006, **368**:387-403.
4. Association As: **2018 Alzheimer's disease facts and figures.** *Alzheimer's & Dementia* 2018, **14**:367-429.
5. Huang W-J, Zhang XIA, Chen W-W: **Role of oxidative stress in Alzheimer's disease.** *Biomedical Reports* 2016, **4**:519-522.
6. Di Paolo G, Kim T-W: **Linking lipids to Alzheimer's disease: cholesterol and beyond.** 2011, **12**:284-296.
7. Calsolaro V, Edison P: **Alterations in Glucose Metabolism in Alzheimer's Disease.** *Recent patents on endocrine, metabolic & immune drug discovery* 2016, **10**:31-39.
8. Yu WH, Cuervo AM, Kumar A, Peterhoff CM, Schmidt SD, Lee J-H, Mohan PS, Mercken M, Farmery MR, Tjernberg LO, et al: **Macroautophagy—a novel β -amyloid peptide-generating pathway activated in Alzheimer's disease.** *The Journal of Cell Biology* 2005, **171**:87-98.
9. Cataldo AM, Peterhoff CM, Troncoso JC, Gomez-Isla T, Hyman BT, Nixon RA: **Endocytic pathway abnormalities precede amyloid beta deposition in sporadic Alzheimer's disease and Down syndrome: differential effects of APOE genotype and presenilin mutations.** *The American journal of pathology* 2000, **157**:277-286.
10. Beach TG, Walker R, McGeer EG: **Patterns of gliosis in alzheimer's disease and aging cerebrum.** *Glia* 1989, **2**:420-436.

11. Crews L, Masliah E: **Molecular mechanisms of neurodegeneration in Alzheimer's disease.** *Human molecular genetics* 2010, **19**:R12-R20.
12. Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT: **Neuropathological Alterations in Alzheimer Disease.** *Cold Spring Harb Perspect Med* 2011, **1**.
13. Selkoe DJ: **The molecular pathology of Alzheimer's disease.** *Neuron* 1991, **6**:487-498.
14. Pimplikar SW, Nixon RA, Robakis NK, Shen J, Tsai L-H: **Amyloid-Independent Mechanisms in Alzheimer's Disease Pathogenesis.** *The Journal of Neuroscience* 2010, **30**:14946.
15. Hane FT, Lee BY, Leonenko Z: **Recent Progress in Alzheimer's Disease Research, Part 1: Pathology.** *Journal of Alzheimer's Disease* 2017, **57**:1-28.
16. de la Monte SM, Tong M: **Brain metabolic dysfunction at the core of Alzheimer's disease.** *Biochemical Pharmacology* 2014, **88**:548-559.
17. Chew H, Solomon VA, Fonteh AN: **Involvement of Lipids in Alzheimer's Disease Pathology and Potential Therapies.** *Frontiers in Physiology* 2020, **11**:598.
18. Maruszak A, Żekanowski C: **Mitochondrial dysfunction and Alzheimer's disease.** *Progress in neuro-psychopharmacology & biological psychiatry* 2011, **35**:320-330.
19. Drubin DG, Kirschner MW: **Tau protein function in living cells.** *The Journal of Cell Biology* 1986, **103**:2739.
20. Iqbal K, Alonso Adel C, Chen S, Chohan MO, El-Akkad E, Gong CX, Khatoon S, Li B, Liu F, Rahman A, et al: **Tau pathology in Alzheimer disease and other tauopathies.** *Biochimica et biophysica acta* 2005, **1739**:198-210.
21. Iqbal K, Liu F, Gong C-X: **Tau and neurodegenerative disease: the story so far.** *Nat Rev Neurol* 2016, **12**:15-27.
22. Lai A, Sisodia SS, Trowbridge IS: **Characterization of Sorting Signals in the β -Amyloid Precursor Protein Cytoplasmic Domain.** *Journal of Biological Chemistry* 1995, **270**:3565-3573.
23. Boll W, Rapoport I, Brunner C, Modis Y, Prehn S, Kirchhausen T: **The mu 2 subunit of the clathrin adaptor AP-2 binds to FDNPVY and YppO sorting signals at distinct sites.** *Traffic* 2002, **3**:590-600.
24. Yin RH, Yu JT, Tan L: **The Role of SORL1 in Alzheimer's Disease.** *Mol Neurobiol* 2015, **51**:909-918.
25. Willnow TE, Andersen OM: **Sorting receptor SORLA--a trafficking path to avoid Alzheimer disease.** *J Cell Sci* 2013, **126**:2751-2760.
26. Iwatsubo T, Odaka A, Suzuki N, Mizusawa H, Nukina N, Ihara Y: **Visualization of A beta 42(43) and A beta 40 in senile plaques with end-specific A beta monoclonals: evidence that an initially deposited species is A beta 42(43).** *Neuron* 1994, **13**:45-53.
27. Hartmann T, Bieger SC, Bruhl B, Tienari PJ, Ida N, Allsop D, Roberts GW, Masters CL, Dotti CG, Unsicker K, Beyreuther K: **Distinct sites of intracellular production for Alzheimer's disease A[beta]40/42 amyloid peptides.** *Nature Medicine* 1997, **3**:1016-1020.

28. Barthelson K, Newman M, Lardelli M: **Sorting Out the Role of the Sortilin-Related Receptor 1 in Alzheimer's Disease.** *J Alzheimers Dis Rep* 2020, **4**:123-140.
29. Ryman DC, Acosta-Baena N, Aisen PS, Bird T, Danek A, Fox NC, Goate A, Frommelt P, Ghetti B, Langbaum JBS, et al: **Symptom onset in autosomal dominant Alzheimer disease: A systematic review and meta-analysis.** *Neurology* 2014, **83**:253-260.
30. Pottier C, Hannequin D, Coutant S, Rovelet-Lecrux A, Wallon D, Rousseau S, Legallic S, Paquet C, Bombois S, Pariente J, et al: **High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease.** *Mol Psychiatry* 2012, **17**:875-879.
31. Bellenguez C, Charbonnier C, Grenier-Boley B, Quenez O, Le Guennec K, Nicolas G, Chauhan G, Wallon D, Rousseau S, Richard AC, et al: **Contribution to Alzheimer's disease risk of rare variants in TREM2, SORL1, and ABCA7 in 1779 cases and 1273 controls.** *Neurobiol Aging* 2017.
32. Lee JH, Cheng R, Honig LS, Vonsattel JP, Clark L, Mayeux R: **Association between genetic variants in SORL1 and autopsy-confirmed Alzheimer disease.** *Neurology* 2008, **70**:887-889.
33. Lee JH, Cheng R, Schupf N, Manly J, Lantigua R, Stern Y, Rogaeva E, Wakutani Y, Farrer L, St George-Hyslop P, Mayeux R: **The association between genetic variants in SORL1 and Alzheimer disease in an urban, multiethnic, community-based cohort.** *Archives of neurology* 2007, **64**:501-506.
34. Liu G, Sun JY, Xu M, Yang XY, Sun BL: **SORL1 Variants Show Different Association with Early-Onset and Late-Onset Alzheimer's Disease Risk.** *Journal of Alzheimer's disease : JAD* 2017, **58**:1121-1128.
35. Louwersheimer E, Cohn-Hokke PE, Pijnenburg YA, Weiss MM, Sistermans EA, Rozemuller AJ, Hulsman M, van Swieten JC, van Duijn CM, Barkhof F, et al: **Rare Genetic Variant in SORL1 May Increase Penetrance of Alzheimer's Disease in a Family with Several Generations of APOE-varepsilon4 Homozygosity.** *Journal of Alzheimer's disease : JAD* 2017, **56**:63-74.
36. Meng Y, Lee JH, Cheng R, George-Hyslop PS, Mayeux R, Farrer LA: **Association between SORL1 and Alzheimer disease in a genome-wide study.** *Neuroreport* 2007, **18**:1761-1764.
37. Miyashita A, Koike A, Jun G, Wang L-S, Takahashi S, Matsubara E, Kawarabayashi T, Shoji M, Tomita N, Arai H, et al: **SORL1 Is Genetically Associated with Late-Onset Alzheimer's Disease in Japanese, Koreans and Caucasians.** *PLoS ONE* 2013, **8**:e58618.
38. Rogaeva E, Meng Y, Lee JH, Gu Y, Kawarai T, Zou F, Katayama T, Baldwin CT, Cheng R, Hasegawa H, et al: **The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease.** *Nat Genet* 2007, **39**:168-177.
39. Verheijen J, Van den Bossche T, van der Zee J, Engelborghs S, Sanchez-Valle R, Lladó A, Graff C, Thonberg H, Pastor P, Ortega-Cubero S, et al: **A comprehensive study of the genetic impact of rare variants in SORL1 in**

- European early-onset Alzheimer's disease.** *Acta Neuropathologica* 2016, **132**:213-224.
40. Nicolas M, Hassan BA: **Amyloid precursor protein and neural development.** *Development* 2014, **141**:2543.
 41. Soldano A, Hassan BA: **Beyond pathology: APP, brain development and Alzheimer's disease.** *Current Opinion in Neurobiology* 2014, **27**:61-67.
 42. Müller UC, Zheng H: **Physiological functions of APP family proteins.** *Cold Spring Harb Perspect Med* 2012, **2**:a006288-a006288.
 43. Burger PC, Vogel FS: **The development of the pathologic changes of Alzheimer's disease and senile dementia in patients with Down's syndrome.** *Am J Pathol* 1973, **73**:457-476.
 44. Mann DMA: **Alzheimer's disease and Down's syndrome.** *Histopathology* 1988, **13**:125-137.
 45. Tcw J, Goate AM: **Genetics of β -Amyloid Precursor Protein in Alzheimer's Disease.** *Cold Spring Harb Perspect Med* 2017, **7**:a024539.
 46. De Jonghe C, Esselens C, Kumar-Singh S, Craessaerts K, Serneels S, Checler F, Annaert W, Van Broeckhoven C, De Strooper B: **Pathogenic APP mutations near the gamma-secretase cleavage site differentially affect Abeta secretion and APP C-terminal fragment stability.** *Human molecular genetics* 2001, **10**:1665-1671.
 47. Cruts M, Theuns J, Van Broeckhoven C: **Locus-specific mutation databases for neurodegenerative brain diseases.** *Human mutation* 2012, **33**:1340-1344.
 48. Shen J, Bronson RT, Chen DF, Xia W, Selkoe DJ, Tonegawa S: **Skeletal and CNS defects in Presenilin-1-deficient mice.** *Cell* 1997, **89**:629-639.
 49. Herreman A, Hartmann D, Annaert W, Saftig P, Craessaerts K, Serneels L, Umans L, Schrijvers V, Checler F, Vanderstichele H, et al: **Presenilin 2 deficiency causes a mild pulmonary phenotype and no changes in amyloid precursor protein processing but enhances the embryonic lethal phenotype of presenilin 1 deficiency.** *Proceedings of the National Academy of Sciences* 1999, **96**:11872-11877.
 50. Tambini MD, D'Adamio L: **Knock-in rats with homozygous PSEN1(L435F) Alzheimer mutation are viable and show selective γ -secretase activity loss causing low A β 40/42 and high A β 43.** *The Journal of biological chemistry* 2020, **295**:7442-7451.
 51. Laudon H, Hansson EM, Melén K, Bergman A, Farmery MR, Winblad B, Lendahl U, von Heijne G, Näslund J: **A nine-transmembrane domain topology for presenilin 1.** *The Journal of biological chemistry* 2005, **280**:35352-35360.
 52. Thinakaran G, Borchelt DR, Lee MK, Slunt HH, Spitzer L, Kim G, Ratovitsky T, Davenport F, Nordstedt C, Seeger M, et al: **Endoproteolysis of presenilin 1 and accumulation of processed derivatives in vivo.** *Neuron* 1996, **17**:181-190.
 53. Kim TW, Pettingell WH, Hallmark OG, Moir RD, Wasco W, Tanzi RE: **Endoproteolytic cleavage and proteasomal degradation of presenilin 2 in transfected cells.** *Journal of Biological Chemistry* 1997, **272**:11006-11010.

54. Dries DR, Yu G: **Assembly, maturation, and trafficking of the gamma-secretase complex in Alzheimer's disease.** *Current Alzheimer's Research* 2008, **5**:132-146.
55. Jayne T, Newman M, Verdile G, Sutherland G, Munch G, Musgrave I, Moussavi Nik SH, Lardelli M: **Evidence for and Against a Pathogenic Role of Reduced gamma-Secretase Activity in Familial Alzheimer's Disease.** *Journal of Alzheimer's Disease* 2016.
56. Group AsDC: **The structure of the presenilin 1 (S182) gene and identification of six novel mutations in early onset AD families.** *Nat Genet* 1995, **11**:219-222.
57. Ishikawa A, Piao Y-S, Miyashita A, Kuwano R, Onodera O, Ohtake H, Suzuki M, Nishizawa M, Takahashi H: **A mutant PSEN1 causes dementia with lewy bodies and variant Alzheimer's disease.** *Annals of Neurology* 2005, **57**:429-434.
58. Guo J, Wei J, Liao S, Wang L, Jiang H, Tang B: **A novel presenilin 1 mutation (Ser169del) in a Chinese family with early-onset Alzheimer's disease.** *Neurosci Lett* 2010, **468**:34-37.
59. Sato S, Kamino K, Miki T, Doi A, li K, St George-Hyslop PH, Ogihara T, Sakaki Y: **Splicing mutation of presenilin-1 gene for early-onset familial Alzheimer's disease.** *Human mutation* 1998, **11**:S91-S94.
60. Wang B, Yang W, Wen W, Sun J, Su B, Liu B, Ma D, Lv D, Wen Y, Qu T, et al: **-Secretase Gene Mutations in Familial Acne Inversa.** 2010, **330**:1065-1065.
61. Wong TH, Seelaar H, Melhem S, Rozemuller AJM, van Swieten JC: **Genetic screening in early-onset Alzheimer's disease identified three novel presenilin mutations.** *Neurobiology of aging* 2020, **86**:201.e209-201.e214.
62. El Kadmiri N, Zaid N, Zaid Y, Tadevosyan A, Hachem A, Dubé MP, Hamzi K, El Moutawakil B, Slassi I, Nadifi S: **Novel presenilin mutations within Moroccan patients with Early-Onset Alzheimer's Disease.** *Neuroscience* 2014, **269**:215-222.
63. Jayadev S, Leverenz JB, Steinbart E, Stahl J, Klunk W, Yu CE, Bird TD: **Alzheimer's disease phenotypes and genotypes associated with mutations in presenilin 2.** *Brain* 2010, **133**:1143-1154.
64. Braggin JE, Bucks SA, Course MM, Smith CL, Sopher B, Osnis L, Shuey KD, Domoto-Reilly K, Caso C, Kinoshita C, et al: **Alternative splicing in a presenilin 2 variant associated with Alzheimer disease.** *Annals of clinical and translational neurology* 2019, **6**:762-777.
65. Sato N, Hori O, Yamaguchi A, Lambert JC, Chartier-Harlin MC, Robinson PA, Delacourte A, Schmidt AM, Furuyama T, Imaizumi K, et al: **A novel presenilin-2 splice variant in human Alzheimer's disease brain tissue.** *Journal of Neurochemistry* 1999, **72**:2498-2505.
66. Manabe T, Katayama T, Sato N, Gomi F, Hitomi J, Yanagita T, Kudo T, Honda A, Mori Y, Matsuzaki S, et al: **Induced HMGA1a expression causes aberrant splicing of Presenilin-2 pre-mRNA in sporadic Alzheimer's disease.** *Cell Death & Differentiation* 2003, **10**:698-708.
67. Sato N, Imaizumi K, Manabe T, Taniguchi M, Hitomi J, Katayama T, Yoneda T, Morihara T, Yasuda Y, Takagi T, et al: **Increased Production of β -**

Amyloid and Vulnerability to Endoplasmic Reticulum Stress by an Aberrant Spliced Form of Presenilin 2 *. *Journal of Biological Chemistry* 2001, **276**:2108-2114.

68. Moussavi Nik SH, Newman M, Wilson L, Ebrahimie E, Wells S, Musgrave I, Verdile G, Martins RN, Lardelli M: **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**. *Human molecular genetics* 2015, **24**:3662-3678.
69. De Strooper B, Saftig P, Craessaerts K, Vanderstichele H, Guhde G, Annaert W, Von Figura K, Van Leuven F: **Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein**. *Nature* 1998, **391**:387-390.
70. Wolfe MS, Xia W, Ostaszewski BL, Diehl TS, Kimberly WT, Selkoe DJ: **Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and $[\gamma]$ -secretase activity**. *Nature* 1999, **398**:513-517.
71. Ahn K, Shelton CC, Tian Y, Zhang X, Gilchrist ML, Sisodia SS, Li Y-M: **Activation and intrinsic γ -secretase activity of presenilin 1**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**:21435-21440.
72. Sato T, Diehl TS, Narayanan S, Funamoto S, Ihara Y, De Strooper B, Steiner H, Haass C, Wolfe MS: **Active γ -secretase complexes contain only one of each component**. *The Journal of biological chemistry* 2007, **282**:33985-33993.
73. Güner G, Lichtenthaler SF: **The substrate repertoire of γ -secretase/presenilin**. *Seminars in Cell & Developmental Biology* 2020, **105**:27-42.
74. De Strooper B, Annaert W, Cupers P, Saftig P, Craessaerts K, Mumm JS, Schroeter EH, Schrijvers V, Wolfe MS, Ray WJ, et al: **A presenilin-1-dependent γ -secretase-like protease mediates release of Notch intracellular domain**. *Nature* 1999, **398**:518.
75. Parks AL, Curtis D: **Presenilin diversifies its portfolio**. *Trends in Genetics* 2007, **23**:140-150.
76. Lee JH, Yu WH, Kumar A, Lee S, Mohan PS, Peterhoff CM, Wolfe DM, Martinez-Vicente M, Massey AC, Sovak G, et al: **Lysosomal proteolysis and autophagy require presenilin 1 and are disrupted by Alzheimer-related PS1 mutations**. *Cell* 2010, **141**:1146-1158.
77. Yamamoto A, Tagawa Y, Yoshimori T, Moriyama Y, Masaki R, Tashiro Y: **Bafilomycin A1 prevents maturation of autophagic vacuoles by inhibiting fusion between autophagosomes and lysosomes in rat hepatoma cell line, H-4-II-E cells**. *Cell structure and function* 1998, **23**:33-42.
78. Bustos V, Pulina MV, Bispo A, Lam A, Flajolet M, Gorelick FS, Greengard P: **Phosphorylated Presenilin 1 decreases beta-amyloid by facilitating autophagosome-lysosome fusion**. *Proc Natl Acad Sci U S A* 2017.
79. Bustos V, Pulina MV, Kelahmetoglu Y, Sinha SC, Gorelick FS, Flajolet M, Greengard P: **Bidirectional regulation of A β levels by Presenilin 1**. *Proceedings of the National Academy of Sciences* 2017, **114**:7142.

80. De Gasperi R, Sosa MAG, Dracheva S, Elder GA: **Presenilin-1 regulates induction of hypoxia inducible factor-1 α : altered activation by a mutation associated with familial Alzheimer's disease.** *Molecular Neurodegeneration* 2010, **5**:38-38.
81. Kang DE, Soriano S, Frosch MP, Collins T, Naruse S, Sisodia SS, Leibowitz G, Levine F, Koo EH: **Presenilin 1 Facilitates the Constitutive Turnover of β -Catenin: Differential Activity of Alzheimer's Disease-Linked PS1 Mutants in the β -Catenin-Signaling Pathway.** *The Journal of Neuroscience* 1999, **19**:4229.
82. Zhang S, Zhang M, Cai F, Song W: **Biological function of Presenilin and its role in AD pathogenesis.** *Translational neurodegeneration* 2013, **2**:15-15.
83. Kabir MT, Uddin MS, Setu JR, Ashraf GM, Bin-Jumah MN, Abdel-Daim MM: **Exploring the Role of PSEN Mutations in the Pathogenesis of Alzheimer's Disease.** *Neurotoxicity Research* 2020, **38**:833-849.
84. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL: **Role of Genes and Environments for Explaining Alzheimer Disease.** 2006, **63**:168.
85. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, Roses AD: **Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease.** *Proceedings of the National Academy of Sciences of the United States of America* 1993, **90**:1977-1981.
86. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, Myers RH, Pericak-Vance MA, Risch N, van Duijn CM: **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* 1997, **278**:1349-1356.
87. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, Boland A, Vronskaya M, Van Der Lee SJ, Amlie-Wolf A, et al: **Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing.** *Nature Genetics* 2019, **51**:414-430.
88. Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, DeStefano AL, Bis JC, Beecham GW, et al: **Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease.** *Nature Genetics* 2013, **45**:1452.
89. Frieden C, Garai K: **Concerning the structure of apoE.** *Protein Science* 2013, **22**:1820-1825.
90. Li Z, Shue F, Zhao N, Shinohara M, Bu G: **APOE2: protective mechanism and therapeutic implications for Alzheimer's disease.** *Molecular Neurodegeneration* 2020, **15**:63.
91. Yamazaki Y, Zhao N, Caulfield TR, Liu C-C, Bu G: **Apolipoprotein E and Alzheimer disease: pathobiology and targeting strategies.** *Nature Reviews Neurology* 2019, **15**:501-518.
92. Mahley RW: **Apolipoprotein E: cholesterol transport protein with expanding role in cell biology.** *Science* 1988, **240**:622.

93. Mahley RW, Stanley C, Rall J: **Apolipoprotein E: Far More Than a Lipid Transport Protein.** *Annual Review of Genomics and Human Genetics* 2000, **1**:507-537.
94. LaDu MJ, Falduto MT, Manelli AM, Reardon CA, Getz GS, Frail DE: **Isoform-specific binding of apolipoprotein E to beta-amyloid.** *Journal of Biological Chemistry* 1994, **269**:23403-23406.
95. Holtzman DM, Bales KR, Tenkova T, Fagan AM, Parsadanian M, Sartorius LJ, Mackey B, Olney J, McKeel D, Wozniak D, Paul SM: **Apolipoprotein E isoform-dependent amyloid deposition and neuritic degeneration in a mouse model of Alzheimer's disease.** *Proc Natl Acad Sci U S A* 2000, **97**:2892-2897.
96. Schmechel DE, Saunders AM, Strittmatter WJ, Crain BJ, Hulette CM, Joo SH, Pericak-Vance MA, Goldgaber D, Roses AD: **Increased amyloid beta-peptide deposition in cerebral cortex as a consequence of apolipoprotein E genotype in late-onset Alzheimer disease.** *Proceedings of the National Academy of Sciences* 1993, **90**:9649-9653.
97. Yajima R, Tokutake T, Koyama A, Kasuga K, Tezuka T, Nishizawa M, Ikeuchi T: **ApoE-isoform-dependent cellular uptake of amyloid-beta is mediated by lipoprotein receptor LR11/SorLA.** *Biochemical and biophysical research communications* 2015, **456**:482-488.
98. Karch CM, Goate AM: **Alzheimer's Disease Risk Genes and Mechanisms of Disease Pathogenesis.** *Biological Psychiatry* 2015, **77**:43-51.
99. Tábuas-Pereira M, Santana I, Guerreiro R, Brás J: **Alzheimer's Disease Genetics: Review of Novel Loci Associated with Disease.** *Current Genetic Medicine Reports* 2020, **8**:1-16.
100. International Genomics of Alzheimer's Disease C: **Convergent genetic and expression data implicate immunity in Alzheimer's disease.** *Alzheimer's & dementia : the journal of the Alzheimer's Association* 2015, **11**:658-671.
101. Van Acker ZP, Bretou M, Annaert W: **Endo-lysosomal dysregulations and late-onset Alzheimer's disease: impact of genetic risk factors.** *Molecular Neurodegeneration* 2019, **14**:20.
102. Newcombe EA, Camats-Perna J, Silva ML, Valmas N, Huat TJ, Medeiros R: **Inflammation: the link between comorbidities, genetics, and Alzheimer's disease.** *Journal of Neuroinflammation* 2018, **15**:276.
103. Ittner LM, Götz J: **Amyloid- β and tau — a toxic pas de deux in Alzheimer's disease.** *Nature Reviews Neuroscience* 2011, **12**:67-72.
104. Markesbery WR: **Oxidative Stress Hypothesis in Alzheimer's Disease.** *Free Radical Biology and Medicine* 1997, **23**:134-147.
105. Francis PT, Palmer AM, Snape M, Wilcock GK: **The cholinergic hypothesis of Alzheimer's disease: a review of progress.** *Journal of Neurology, Neurosurgery & Psychiatry* 1999, **66**:137.
106. Hardy JA, Higgins GA: **Alzheimer's Disease: The Amyloid Cascade Hypothesis.** *Science* 1992, **256**:184-185.
107. Selkoe DJ, Hardy J: **The amyloid hypothesis of Alzheimer's disease at 25 years.** *EMBO Mol Med* 2016, **8**:595-608.

108. Glenner GG, Wong CW: **Alzheimer's disease and Down's syndrome: sharing of a unique cerebrovascular amyloid fibril protein.** *Biochem Biophys Res Commun* 1984, **122**:1131-1135.
109. Leinenga G, Götz J: **Scanning ultrasound removes amyloid- β and restores memory in an Alzheimer's disease mouse model.** *Science Translational Medicine* 2015, **7**:278ra233-278ra233.
110. Terry RD, Masliah E, Salmon DP, Butters N, DeTeresa R, Hill R, Hansen LA, Katzman R: **Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment.** *Ann Neurol* 1991, **30**:572-580.
111. Tomlinson BE, Blessed G, Roth M: **Observations on the brains of demented old people.** *J Neurol Sci* 1970, **11**:205-242.
112. Nordberg A: **Amyloid imaging in Alzheimer's disease.** *Neuropsychologia* 2008, **46**:1636-1641.
113. Fleisher AS, Raman R, Siemers ER, Becerra L, Clark CM, Dean RA, Farlow MR, Galvin JE, Peskind ER, Quinn JF, et al: **Phase 2 safety trial targeting amyloid beta production with a gamma-secretase inhibitor in Alzheimer disease.** *Archives of neurology* 2008, **65**:1031-1038.
114. Siemers ER, Dean RA, Friedrich S, Ferguson-Sells L, Gonzales C, Farlow MR, May PC: **Safety, tolerability, and effects on plasma and cerebrospinal fluid amyloid-beta after inhibition of gamma-secretase.** *Clinical Neuropharmacology* 2007, **30**:317-325.
115. Doody RS, Raman R, Farlow M, Iwatsubo T, Vellas B, Joffe S, Kieburtz K, He F, Sun X, Thomas RG, et al: **A phase 3 trial of semagacestat for treatment of Alzheimer's disease.** *The New England Journal of Medicine* 2013, **369**:341-350.
116. Imbimbo BP, Lucca U, Watling M: **Can Anti- β -amyloid Monoclonal Antibodies Work in Autosomal Dominant Alzheimer Disease?** *Neurology Genetics* 2021, **7**:e535.
117. de Brito OM, Scorrano L: **Mitofusin 2 tethers endoplasmic reticulum to mitochondria.** *Nature* 2008, **456**:605-610.
118. Area-Gomez E, del Carmen Lara Castillo M, Tambini MD, Guardia-Laguarta C, de Groof AJC, Madra M, Ikenouchi J, Umeda M, Bird TD, Sturley SL, Schon EA: **Upregulated function of mitochondria-associated ER membranes in Alzheimer disease.** *The EMBO Journal* 2012, **31**:4106-4123.
119. Area-Gomez E, de Groof AJ, Boldogh I, Bird TD, Gibson GE, Koehler CM, Yu WH, Duff KE, Yaffe MP, Pon LA, Schon EA: **Presenilins are enriched in endoplasmic reticulum membranes associated with mitochondria.** *American Journal of Pathology* 2009, **175**:1810-1816.
120. Schon EA, Area-Gomez E: **Is Alzheimer's disease a disorder of mitochondria-associated membranes?** *Journal of Alzheimer's disease : JAD* 2010, **20 Suppl 2**:S281-292.
121. Schon EA, Area-Gomez E: **Mitochondria-associated ER membranes in Alzheimer disease.** 2013, **55**:26-36.
122. Area-Gomez E, Schon EA: **On the Pathogenesis of Alzheimer's Disease: The MAM Hypothesis.** *FASEB J* 2017, **31**:864-867.

123. Vance JE: **MAM (mitochondria-associated membranes) in mammalian cells: Lipids and beyond.** *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 2014, **1841**:595-609.
124. Hamasaki M, Furuta N, Matsuda A, Nezu A, Yamamoto A, Fujita N, Oomori H, Noda T, Haraguchi T, Hiraoka Y, et al: **Autophagosomes form at ER-mitochondria contact sites.** *Nature* 2013, **495**:389-393.
125. Zhou R, Yazdi AS, Menu P, Tschopp J: **A role for mitochondria in NLRP3 inflammasome activation.** *Nature* 2011, **469**:221-225.
126. Area-Gomez E, de Groof A, Bonilla E, Montesinos J, Tanji K, Boldogh I, Pon L, Schon EA: **A key role for MAM in mediating mitochondrial dysfunction in Alzheimer disease.** *Cell Death & Disease* 2018, **9**:335.
127. Pera M, Larrea D, Guardia-Laguarta C, Montesinos J, Velasco KR, Agrawal RR, Xu Y, Chan RB, Di Paolo G, Mehler MF, et al: **Increased localization of APP-C99 in mitochondria-associated ER membranes causes mitochondrial dysfunction in Alzheimer disease.** *The EMBO journal* 2017, **36**:3356-3371.
128. Kogot-Levin A, Saada A: **Ceramide and the mitochondrial respiratory chain.** *Biochimie* 2014, **100**:88-94.
129. Agrawal RR, Montesinos J, Larrea D, Area-Gomez E, Pera M: **The silence of the fats: A MAM's story about Alzheimer.** *Neurobiology of Disease* 2020, **145**:105062.
130. de la Torre JC, Mussivan T: **Can disturbed brain microcirculation cause Alzheimer's disease?** *Neurological Research* 1993, **15**:146-153.
131. Roher AE, Debbins JP, Malek-Ahmadi M, Chen K, Pipe JG, Maze S, Belden C, Maarouf CL, Thiyyagura P, Mo H, et al: **Cerebral blood flow in Alzheimer's disease.** *Vascular Health and Risk Management* 2012, **8**:599-611.
132. Rivera-Rivera LA, Schubert T, Turski P, Johnson KM, Berman SE, Rowley HA, Carlsson CM, Johnson SC, Wieben O: **Changes in intracranial venous blood flow and pulsatility in Alzheimer's disease: A 4D flow MRI study.** *J Cereb Blood Flow Metab* 2017, **37**:2149-2158.
133. Iturria-Medina Y, Sotero R, Toussaint P, Mateos-Pérez J, Evans A, Initiative AsDN: **Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis.** *Nature Communications* 2016, **7**.
134. Fischer VW, Siddiqi A, Yusufaly Y: **Altered angioarchitecture in selected areas of brains with Alzheimer's disease.** *Acta Neuropathol* 1990, **79**:672-679.
135. Hashimura T, Kimura T, Miyakawa T: **Morphological changes of blood vessels in the brain with Alzheimer's disease.** *Jpn J Psychiatry Neurol* 1991, **45**:661-665.
136. Nielsen RB, Egefjord L, Angleys H, Mouridsen K, Gejl M, Moller A, Brock B, Braendgaard H, Gottrup H, Rungby J, et al: **Capillary dysfunction is associated with symptom severity and neurodegeneration in Alzheimer's disease.** *Alzheimers Dement* 2017.

137. Gama Sosa MA, Gasperi RD, Rocher AB, Wang AC, Janssen WG, Flores T, Perez GM, Schmeidler J, Dickstein DL, Hof PR, Elder GA: **Age-related vascular pathology in transgenic mice expressing presenilin 1-associated familial Alzheimer's disease mutations.** *Am J Pathol* 2010, **176**:353-368.
138. Meyer EP, Ulmann-Schuler A, Staufenbiel M, Krucker T: **Altered morphology and 3D architecture of brain vasculature in a mouse model for Alzheimer's disease.** *Proceedings of the National Academy of Sciences* 2008, **105**:3587-3592.
139. de la Torre JC, Butler K, Kozlowski P, Fortin T, Saunders JK: **Correlates between nuclear magnetic resonance spectroscopy, diffusion weighted imaging and CA1 morphometry following chronic brain ischemia.** *Journal of Neuroscience Research* 1995, **41**:238-245.
140. Cada A, de la Torre JC, Gonzalez-Lima F: **Chronic cerebrovascular ischemia in aged rats: effects on brain metabolic capacity and behavior.** *Neurobiol Aging* 2000, **21**:225-233.
141. de la Torre JC, Fortin T, Park GA, Saunders JK, Kozlowski P, Butler K, de Socarraz H, Pappas B, Richard M: **Aged but not young rats develop metabolic, memory deficits after chronic brain ischaemia.** *Neurol Res* 1992, **14**:177-180.
142. de la Torre JC, Fortin T, Park GA, Butler KS, Kozlowski P, Pappas BA, de Socarraz H, Saunders JK, Richard MT: **Chronic cerebrovascular insufficiency induces dementia-like deficits in aged rats.** *Brain Res* 1992, **582**:186-195.
143. de la Torre JC, Aliev G: **Inhibition of Vascular Nitric Oxide after Rat Chronic Brain Hypoperfusion: Spatial Memory and Immunocytochemical Changes.** *Journal of Cerebral Blood Flow & Metabolism* 2005, **25**:663-672.
144. Rivard A, Fabre J-E, Silver M, Chen D, Murohara T, Kearney M, Magner M, Asahara T, Isner JM: **Age-Dependent Impairment of Angiogenesis.** *Circulation* 1999, **99**:111.
145. Yager JY, Wright S, Armstrong EA, Jahraus CM, Saucier DM: **The influence of aging on recovery following ischemic brain damage.** *Behavioural Brain Research* 2006, **173**:171-180.
146. Desmond DW, Moroney JT, Sano M, Stern Y: **Incidence of dementia after ischemic stroke: results of a longitudinal study.** *Stroke* 2002, **33**:2254-2260.
147. Oresic M, Hyotylainen T, Herukka SK, Sysi-Aho M, Mattila I, Seppanan-Laakso T, Julkunen V, Gopalacharyulu PV, Hallikainen M, Koikkalainen J, et al: **Metabolome in progression to Alzheimer's disease.** *Transl Psychiatry* 2011, **1**:e57.
148. Zetterberg H, Mörtberg E, Song L, Chang L, Provuncher GK, Patel PP, Ferrell E, Fournier DR, Kan CW, Campbell TG, et al: **Hypoxia Due to Cardiac Arrest Induces a Time-Dependent Increase in Serum Amyloid β Levels in Humans.** *PLOS ONE* 2011, **6**:e28263.
149. Ke Q, Costa M: **Hypoxia-inducible factor-1 (HIF-1).** *Mol Pharmacol* 2006, **70**:1469-1480.

150. Wang GL, Jiang BH, Rue EA, Semenza GL: **Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O₂ tension.** *Proceedings of the National Academy of Sciences* 1995, **92**:5510-5514.
151. Wang GL, Semenza GL: **Desferrioxamine induces erythropoietin gene expression and hypoxia-inducible factor 1 DNA-binding activity: implications for models of hypoxia signal transduction.** *Blood* 1993, **82**:3610-3615.
152. Frise MC, Robbins PA: **Iron, oxygen, and the pulmonary circulation.** *Journal of Applied Physiology* 2015, **119**:1421-1431.
153. Ndubuizu OI, Chavez JC, LaManna JC: **Increased Prolyl 4-Hydroxylase (PHD) Expression and Differential Regulation of Hypoxia Inducible Factors (HIFs) in the Aged Rat Brain.** *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology* 2009.
154. Frenkel-Denkberg G, Gershon D, Levy AP: **The function of hypoxia-inducible factor 1 (HIF-1) is impaired in senescent mice.** *FEBS Letters* 1999, **462**:341-344.
155. Newman M, Nik HM, Sutherland GT, Hin N, Kim WS, Halliday GM, Jayadev S, Smith C, Laird AS, Lucas CW, et al: **Accelerated loss of hypoxia response in zebrafish with familial Alzheimer's disease-like mutation of presenilin 1.** *Human molecular genetics* 2020, **29**:2379-2394.
156. Waldvogel-Abramowski S, Waeber G, Gassner C, Buser A, Frey BM, Favrat B, Tissot J-D: **Physiology of iron metabolism.** *Transfus Med Hemother* 2014, **41**:213-221.
157. Wang J, Pantopoulos K: **Regulation of cellular iron metabolism.** *Biochemical Journal* 2011, **434**:365-381.
158. Pantopoulos K, Porwal SK, Tartakoff A, Devireddy L: **Mechanisms of mammalian iron homeostasis.** *Biochemistry* 2012, **51**:5705-5724.
159. Bien-Ly N, Yu YJ, Bumbaca D, Elstrott J, Boswell CA, Zhang Y, Luk W, Lu Y, Dennis MS, Weimer RM, et al: **Transferrin receptor (TfR) trafficking determines brain uptake of TfR antibody affinity variants.** *The Journal of experimental medicine* 2014, **211**:233-244.
160. Ohgami RS, Campagna DR, Greer EL, Antiochos B, McDonald A, Chen J, Sharp JJ, Fujiwara Y, Barker JE, Fleming MD: **Identification of a ferrireductase required for efficient transferrin-dependent iron uptake in erythroid cells.** *Nat Genet* 2005, **37**:1264-1269.
161. Sargent PJ, Farnaud S, Evans RW: **Structure/function overview of proteins involved in iron storage and transport.** *Curr Med Chem* 2005, **12**:2683-2693.
162. Lumsden AL, Rogers JT, Majd S, Newman M, Sutherland GT, Verdile G, Lardelli M: **Dysregulation of Neuronal Iron Homeostasis as an Alternative Unifying Effect of Mutations Causing Familial Alzheimer's Disease.** *Frontiers in neuroscience* 2018, **12**:533-533.
163. Jiang Y, Sato Y, Im E, Berg M, Bordi M, Darji S, Kumar A, Mohan PS, Bandyopadhyay U, Diaz A, et al: **Lysosomal Dysfunction in Down Syndrome Is APP-Dependent and Mediated by APP- β CTF (C99).** *The Journal of Neuroscience* 2019, **39**:5255.

164. Yambire KF, Rostosky C, Watanabe T, Pacheu-Grau D, Torres-Odio S, Sanchez-Guerrero A, Senderovich O, Meyron-Holtz EG, Milosevic I, Frahm J, et al: **Impaired lysosomal acidification triggers iron deficiency and inflammation in vivo.** *Elife* 2019, **8**.
165. Toledano A, Alvarez MI, López-Rodríguez AB, Toledano-Díaz A, Fernández-Verdecia CI: **[Does Alzheimer's disease exist in all primates? Alzheimer pathology in non-human primates and its pathophysiological implications (I)].** *Neurologia (Barcelona, Spain)* 2012, **27**:354-369.
166. Sato K, Sasaguri H, Kumita W, Inoue T, Kurotaki Y, Nagata K, Mihira N, Sato K, Sakuma T, Yamamoto T, et al: **A non-human primate model of familial Alzheimer's disease.** *bioRxiv* 2020:2020.2008.2024.264259.
167. Games D, Adams D, Alessandrini R, Barbour R, Borthellette P, Blackwell C, Carr T, Clemens J, Donaldson T, Gillespie F, et al: **Alzheimer-type neuropathology in transgenic mice overexpressing V717F [beta]-amyloid precursor protein.** *Nature* 1995, **373**:523-527.
168. Oakley H, Cole SL, Logan S, Maus E, Shao P, Craft J, Guillozet-Bongaarts A, Ohno M, Disterhoft J, Van Eldik L, et al: **Intraneuronal β -Amyloid Aggregates, Neurodegeneration, and Neuron Loss in Transgenic Mice with Five Familial Alzheimer's Disease Mutations: Potential Factors in Amyloid Plaque Formation.** *The Journal of Neuroscience* 2006, **26**:10129.
169. Oddo S, Caccamo A, Shepherd JD, Murphy MP, Golde TE, Kaye R, Metherate R, Mattson MP, Akbari Y, LaFerla FM: **Triple-transgenic model of Alzheimer's disease with plaques and tangles: intracellular Abeta and synaptic dysfunction.** *Neuron* 2003, **39**:409-421.
170. Guo Q, Fu W, Sopher BL, Miller MW, Ware CB, Martin GM, Mattson MP: **Increased vulnerability of hippocampal neurons to excitotoxic necrosis in presenilin-1 mutant knock-in mice.** *Nat Med* 1999, **5**:101-106.
171. Saito T, Matsuba Y, Mihira N, Takano J, Nilsson P, Itohara S, Iwata N, Saido TC: **Single App knock-in mouse models of Alzheimer's disease.** *Nature Neuroscience* 2014, **17**:661-663.
172. Kirova A-M, Bays RB, Lagalwar S: **Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimer's disease.** *BioMed research international* 2015, **2015**:748212-748212.
173. Sharma S, Rakoczy S, Brown-Borg H: **Assessment of spatial memory in mice.** *Life Sci* 2010, **87**:521-536.
174. Janus C, Westaway D: **Transgenic mouse models of Alzheimer's disease.** *Physiology & Behavior* 2001, **73**:873-886.
175. Sun X, Beglopoulos V, Mattson MP, Shen J: **Hippocampal Spatial Memory Impairments Caused by the Familial Alzheimer's Disease-Linked Presenilin 1 M146V Mutation.** *Neurodegenerative Diseases* 2005, **2**:6-15.
176. Pak K, Chan SL, Mattson MP: **Presenilin-1 mutation sensitizes oligodendrocytes to glutamate and amyloid toxicities, and exacerbates white matter damage and memory impairment in mice.** *Neuromolecular Med* 2003, **3**:53-64.
177. Cleal M, Fontana BD, Ranson DC, McBride SD, Swinny JD, Redhead ES, Parker MO: **The Free-movement pattern Y-maze: A cross-species**

- measure of working memory and executive function.** *Behavior Research Methods* 2020.
178. Tolman EC: **Purpose and cognition: the determiners of animal learning.** *Psychological Review* 1925, **32**:285-297.
 179. Kukurba KR, Montgomery SB: **RNA Sequencing and Analysis.** *Cold Spring Harbor Protocols* 2015.
 180. Wilhelm BT, Landry J-R: **RNA-Seq—quantitative measurement of expression through massively parallel RNA-sequencing.** *Methods* 2009, **48**:249-257.
 181. Wolf JBW: **Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial.** *Molecular Ecology Resources* 2013, **13**:559-572.
 182. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: **limma powers differential expression analyses for RNA-sequencing and microarray studies.** *Nucleic Acids Research* 2015, **43**:e47-e47.
 183. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics* 2009, **26**:139-140.
 184. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.** *Genome Biology* 2014, **15**:550.
 185. Law CW, Chen Y, Shi W, Smyth GK: **voom: precision weights unlock linear model analysis tools for RNA-seq read counts.** *Genome Biology* 2014, **15**:R29.
 186. Geistlinger L, Csaba G, Santarelli M, Ramos M, Schiffer L, Turaga N, Law C, Davis S, Carey V, Morgan M, et al: **Toward a gold standard for benchmarking gene set enrichment analysis.** *Briefings in bioinformatics* 2021, **22**:545-556.
 187. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al: **Gene Ontology: tool for the unification of biology.** *Nature genetics* 2000, **25**:25-29.
 188. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov Jill P, Tamayo P: **The Molecular Signatures Database Hallmark Gene Set Collection.** *Cell Systems* 2015, **1**:417-425.
 189. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.** *Nucleic acids research* 2000, **28**:27-30.
 190. Young MD, Wakefield MJ, Smyth GK, Oshlack A: **Gene ontology analysis for RNA-seq: accounting for selection bias.** *Genome Biology* 2010, **11**:R14.
 191. Wu D, Lim E, Vaillant F, Asselin-Labat M-L, Visvader JE, Smyth GK: **ROAST: rotation gene set tests for complex microarray experiments.** *Bioinformatics* 2010, **26**:2176-2182.
 192. Wu D, Smyth GK: **Camera: a competitive gene set test accounting for inter-gene correlation.** *Nucleic acids research* 2012, **40**:e133-e133.
 193. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: **Gene set**

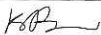
- enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles.** *Proceedings of the National Academy of Sciences* 2005, **102**:15545.
194. Twine NA, Janitz K, Wilkins MR, Janitz M: **Whole Transcriptome Sequencing Reveals Gene Expression and Splicing Differences in Brain Regions Affected by Alzheimer's Disease.** *PLOS ONE* 2011, **6**:e16266.
 195. Mills JD, Nalpathamkalam T, Jacobs HIL, Janitz C, Merico D, Hu P, Janitz M: **RNA-Seq analysis of the parietal cortex in Alzheimer's disease reveals alternatively spliced isoforms related to lipid metabolism.** *Neuroscience Letters* 2013, **536**:90-95.
 196. Bennett JP, Keeney PM: **RNA-Sequencing Reveals Similarities and Differences in Gene Expression in Vulnerable Brain Tissues of Alzheimer's and Parkinson's Diseases.** *Journal of Alzheimer's Disease Reports* 2018, **2**:129-137.
 197. Neff RA, Wang M, Vatansever S, Guo L, Ming C, Wang Q, Wang E, Horgusluoglu-Moloch E, Song W-m, Li A, et al: **Molecular subtyping of Alzheimer's disease using RNA sequencing data reveals novel mechanisms and targets.** *Science Advances* 2021, **7**:eabb5398.
 198. Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, Belleville S, Brodaty H, Bennett D, Chertkow H, et al: **Mild cognitive impairment.** *The Lancet* 2006, **367**:1262-1270.
 199. Berchtold NC, Sabbagh MN, Beach TG, Kim RC, Cribbs DH, Cotman CW: **Brain gene expression patterns differentiate mild cognitive impairment from normal aged and Alzheimer's disease.** *Neurobiology of Aging* 2014, **35**:1961-1972.
 200. Hargis KE, Blalock EM: **Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us?** *Behavioural Brain Research* 2017, **322**:311-328.
 201. Hin N, Newman M, Kaslin J, Douek AM, Lumsden A, Nik SHM, Dong Y, Zhou X-F, Mañucat-Tan NB, Ludington A, et al: **Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2.** *PLOS ONE* 2020, **15**:e0227258.
 202. Hin N, Newman M, Pederson SM, Lardelli MM: **Iron Responsive Element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease.** *bioRxiv* 2020:2020.2005.2001.071498.
 203. Newman M, Hin N, Pederson S, Lardelli M: **Brain transcriptome analysis of a familial Alzheimer's disease-like mutation in the zebrafish presenilin 1 gene implies effects on energy production.** *Molecular Brain* 2019, **12**.

Chapter 2: Sorting out the role of *SORTILIN-RELATED RECEPTOR 1* in Alzheimer's disease

Statement of Authorship

Title of Paper	Sorting Out the Role of the Sortilin-Related Receptor 1 in Alzheimer's Disease		
Publication Status	<input checked="" type="checkbox"/> Published	<input type="checkbox"/> Accepted for Publication	<input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
	<input type="checkbox"/> Submitted for Publication		
Publication Details	Published in the Journal of Alzheimer's disease Reports doi: 10.3233/ADR-200177		

Principal Author

Name of Principal Author (Candidate)	Karissa Barthelson		
Contribution to the Paper	Reviewed the literature, drafted the manuscript		
Overall percentage (%)	95%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	11/2/21

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Morgan Newman		
Contribution to the Paper	Editing of manuscript		
Signature		Date	15/02/2021

Name of Co-Author	Michael Lardelli		
Contribution to the Paper	Editing of manuscript		
Signature		Date	19/02/2021

Please cut and paste additional co-author panels here as required.

Review

Sorting Out the Role of the *Sortilin-Related Receptor 1* in Alzheimer's Disease

Karissa Barthelson, Morgan Newman and Michael Lardelli*
School of Biological Sciences, University of Adelaide, Adelaide, SA, Australia

Accepted 26 March 2020

Abstract. *Sortilin-related receptor 1 (SORL1)* encodes a large, multi-domain containing, membrane-bound receptor involved in endosomal sorting of proteins between the trans-Golgi network, endosomes and the plasma membrane. It is genetically associated with Alzheimer's disease (AD), the most common form of dementia. *SORL1* is a unique gene in AD, as it appears to show strong associations with the common, late-onset, sporadic form of AD and the rare, early-onset familial form of AD. Here, we review the genetics of *SORL1* in AD and discuss potential roles it could play in AD pathogenesis.

Keywords: Alzheimer's disease, amyloid, amyloid-beta protein precursor, endocytosis, endosomes, protein transport

INTRODUCTION

Sortilin-related receptor 1 (SORL1), also known as *LR11* or *SORLA* encodes a hybrid receptor with multiple domains involved in intracellular sorting and trafficking of proteins into their respective subcellular compartments. The sorting of proteins is essential for normal cell function and defects in these pathways are thought to be important factors in the pathogenesis of Alzheimer's disease (AD).

AD, a progressive neurodegenerative disorder, is the most common form of dementia. It is defined by progressive cognitive decline, widespread neurodegeneration and two proteinopathies: senile neuritic plaques [consisting of the product of the proteolytic processing of the amyloid- β protein precursor (A β PP), amyloid- β (A β)], and neurofibrillary tangles (consisting of hyperphosphorylated tau protein) (reviewed in [1]). AD can be classified into two broad

subtypes defined by the age of disease onset: early-onset AD (<65 years of age, EOAD) and late onset AD (>65 years of age, LOAD). There can be both sporadic and familial varieties of both EOAD and LOAD. The majority of AD cases (~95%) belong to the LOAD subgroup, and these mainly arise sporadically. Conversely, ~5% of cases belong to the EOAD subgroup [2]. A portion of EOAD cases are caused by autosomal dominant variants in *A β PP* [3–7], *presenilin 1 (PSEN1)* [8–13], and *presenilin 2 (PSEN2)* [14–17] (EOAD). Recently, variants in *SORL1* have been recognized as likely causing EOAD [18, 19]. Intriguingly, *SORL1* is also recognized as a genetic risk factor for development of late-onset, sporadic AD (sAD) [20–22]. The exact role of *SORL1* in the pathogenesis of AD is unclear since there are multiple pathways which link it to AD pathological mechanisms.

This review will discuss the normal expression of *SORL1* and the genetic evidence which supports a role in LOAD and EOAD. We then discuss the possible roles which *SORL1* may play in different pathologies observed in AD. *SORL1* appears to be strongly associated with both EOAD and LOAD and

*Correspondence to: Michael Lardelli, Department of Molecular & Biomedical Science, School of Biological Sciences, The University of Adelaide, SA 5005, Australia. Tel.: +61 8 83033212; Fax: +61 8 83034362; E-mail: michael.lardelli@adelaide.edu.au.

provides a unique opportunity to understand the similarities and differences between these two subtypes of AD.

SORL1 HAS MULTIPLE TRANSCRIPT SPLICING PRODUCTS

SORL1 is located on chromosome 11q23.2-24.2 in humans and is widely expressed in the brain. It is particularly highly expressed in neurons of the hippocampus and some nuclei of the brainstem and Purkinje cells, and has slightly weaker expression in neurons of the thalamus and the hypothalamus [23]. It is also expressed in other tissue types such as the testes, ovaries, thyroid, and lymph nodes [24].

Alternative splicing occurs for *SORL1*, with at least 5 protein-coding transcripts arising from the gene in humans according to the ENSEMBL database [25] (Fig. 1). The shortest alternative splice product, *SORL1-206*, has a transcription initiation site in intron 30 and is expressed highly in the hippocampus and temporal lobe, moderately in the entorhinal cortex and frontal cortex and is undetectable in the testes or kidney [26]. Expression of this particular splice product is unchanged in the cerebellum of AD patients. However, total *SORL1* mRNA levels are relatively constant in the cerebellum of AD brains [27], so this is not surprising. Gear et al. [28] reported two *SORL1* mRNA isoforms expressed in the temporal lobe in both AD and non-AD samples: *SORL1-Δ2* and *SORL1-Δ9*, which lack exon 2 and exon 19 coding sequences respectively. *SORL1-Δ2* comprises of up to 5% of total *SORL1* expression and *SORL1-Δ19* comprises of less than 1% (and was not investigated further). The *SORL1-Δ2* isoform lacks sequences coding for amino acid residues V96 to D134, the N-terminal proximal region of the VPS10 domain of SORL1 protein. It has increased abundance in white matter compared to grey matter, which suggests varying expression across cell types (as white and grey matter contain different ratios of cell types). However, expression levels of the *SORL1-Δ2* isoform did not correlate with Braak staging of AD progression (while the expression of the longest *SORL1* transcript splice form did show correlation). Finally, a non-coding RNA (ncRNA) termed 51A maps in anti-sense configuration to intron 1 of the *SORL1* gene and drives the alternative splicing of *SORL1* towards two splice products: 1) splice product B (*SORL1-B*), with a coding sequence beginning in exon 23 and ending with the stop codon used for translation of the

full length protein and 2) splice product F (*SORL1-F*), with a coding sequence beginning with the start codon used in translation of the full length protein within exon 1 and ending in exon 14. Overexpression of the 51A ncRNA in a neuroblastoma cell line resulted in increased levels of Aβ peptides. Interestingly, 51A is upregulated in AD brains (although the authors observed substantial variation among individuals) [29].

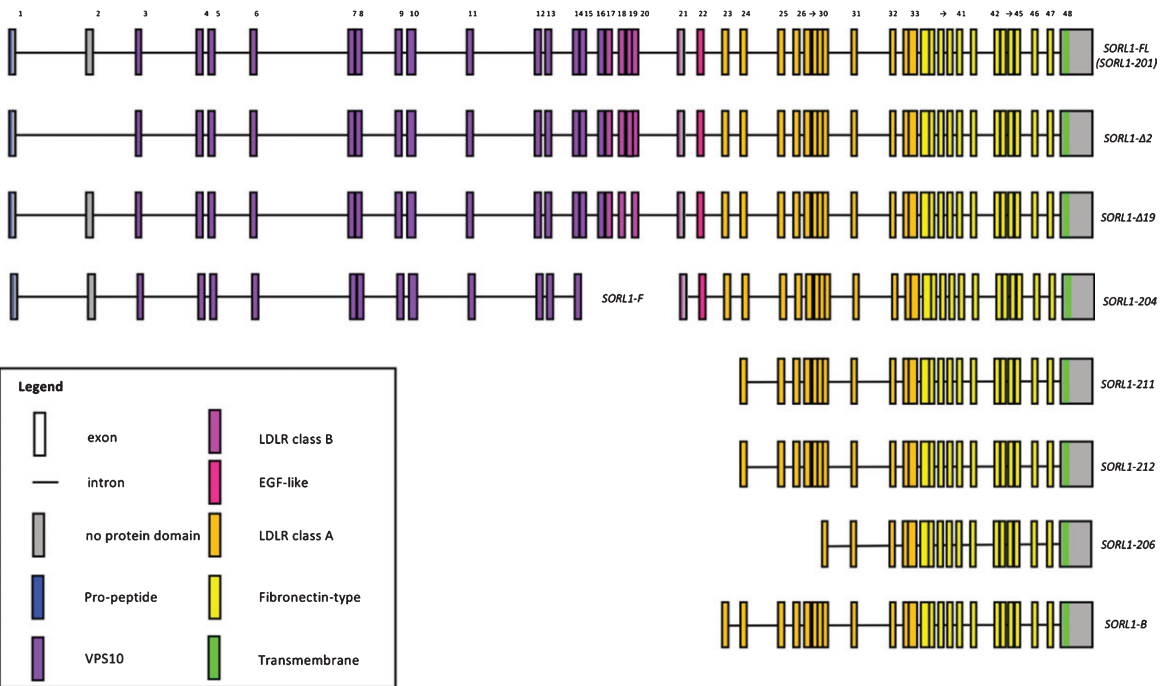
Together, these studies show the complexity of alternative splicing of *SORL1* transcripts. However, the scientific literature lacks functional studies investigating the roles of these alternative splice products.

SORL1 ENCODES A MEMBRANE-BOUND RECEPTOR WITH MULTIPLE FUNCTIONAL DOMAINS

The full-length transcript of *SORL1* encodes a ~250 kDa, membrane-bound protein which primarily localizes in the endosomal and Golgi compartments [30]. The protein is comprised of multiple functional domains. These include a vacuolar protein sorting 10 (VPS10) domain, five low-density lipoprotein receptor (LDLR) class B repeats, an epidermal growth factor-like (EGF-like) domain, eleven LDLR class A repeats, six fibronectin-type (FN-type) repeats, a transmembrane domain and a cytosolic domain containing recognition motifs for cytosolic adaptors [31–34] (Fig. 1B).

Nascent SORL1 proteins are generated in endoplasmic reticulum (ER) and are then transported through the trans-Golgi network (TGN) and eventually to the cell surface. They are initially inactive due to the pro-peptide at the NH₂-terminal which is removed by furin-mediated cleavage in the TGN due to the RRKR furin recognition sequence [35]. This allows the receptor to be directed to the cell surface. Once at the cell surface, it appears that SORL1 can be utilized either in a signaling pathway or a trafficking pathway. In the signaling pathway, SORL1 is subject to ectodomain shedding by the α-secretase tumor necrosis factor-A converting enzyme (TACE/ADAM-17) [36–38]. The exact site at which TACE cleaves SORL1 is unclear. SORL1 is then processed further by γ-secretase at the plasma membrane, releasing a fragment of SORL1 into the extracellular space and a SORL1 intracellular domain (SORL1-ICD) into the cytosol [39, 40]. The SORL1-ICD contains a nuclear localization motif KHRR. It was demonstrated in a reporter assay that the SORL1-

A



B

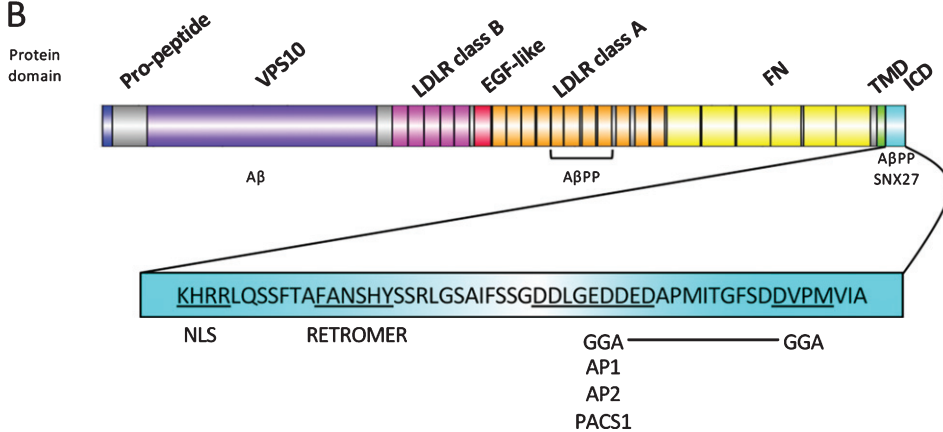


Fig. 1. *SORL1* encodes a multi-domain containing protein and its transcripts are subject to alternative splicing. (A) depicts gene models for alternative *SORL1* splice products from the ENSEMBL database (ENSG00000137642) and published literature. Exons are numbered and color-coded to indicate which protein domains they encode. Protein-coding domains are as given for the human *SORL1* protein (Uniprot ID: Q92673) with SMART database annotations. VPS10, vacuolar protein sorting 10; LDLR, low density lipoprotein receptor; EGF, epidermal growth factor. (B) depicts a schematic of the full length *SORL1* protein consisting of a pro-peptide sequence, a VPS10 domain, five LDLR class B repeats, an EGF-like domain, eleven LDLR class A repeats, six fibronectin-type (FN) repeats, a transmembrane domain (TMD) and a cytosolic intracellular domain (ICD) containing recognition motifs for cytosolic adaptors. Binding sites of amyloid- β protein precursor (A β PP), amyloid- β (A β), sorting nexin 27 (SNX27), Golgi-localizing, γ -adaptin ear homology domain ARF-interaction (GGA), clathrin adaptor protein 1/2 (AP1/2), and phosphofurin acidic cluster sorting protein 1 (PACS1) and the nuclear localization signal (NLS) are indicated.

ICD can enter the nucleus to regulate transcription [39]. Whether the ICD has this function in reality and what genes it might regulate are yet to be determined.

As an alternative to processing at the cell surface for signaling, *SORL1* can also enter a trafficking pathway if it is internalized by clathrin-mediated

endocytosis employing the chaperone clathrin adaptor protein 2 (AP2) [34]. Subsequently, internalized SORL1 receptors generally shuttle between the TGN and endosomes, guided by cytosolic adaptor proteins which will be discussed later in this review in the context of A β PP trafficking (Fig. 2) (reviewed in [41, 42]).

GENETIC EVIDENCE FOR A ROLE OF *SORL1* IN AD

The first characterization of genetic variation of *SORL1* in AD came from a candidate gene approach in 2007. Rogaeva et al. [21] investigated 29 single nucleotide polymorphisms (SNPs) associated with AD throughout the *SORL1* locus in 6 cohorts of familial and sporadic forms of LOAD from different ethnicities. Associations of these individual SNPs with AD were modest, with odds ratios (ORs) ranging from 1.4–2.6 (compared to 14.9 for homozygosity

for the $\epsilon 4$ allele of *APOE* [43]). Haplotype analysis using a sliding window covering 3 SNPs confirmed these associations by demonstrating that two clusters of SNPs in *SORL1* were independently associated with AD: the 5' cluster and the 3' cluster. The 5' cluster consists of SNPs 8, 9 and 10 which are located within intron 6 of *SORL1*. Possession of the T – A – T haplotype was associated with decreased risk of AD in Caribbean-Hispanic families ($p = 0.0086$) and in Israeli-Arab ($p = 0.0037$) and North European ($p = 0.068$) case-control cohorts. Conversely, the C – G – C haplotype for these SNPs was associated with increased risk for developing AD. The 3' cluster consists of SNPs 22 – 25 and two overlapping haplotypes were associated with increased risk of AD: the C – T – T alleles at SNPs 22–24 and the T – T – C alleles at SNPs 23–25 in North European families and North European case-control cohorts. This region was also associated with AD in an African-American cohort. However, it was the A – C – T alleles for

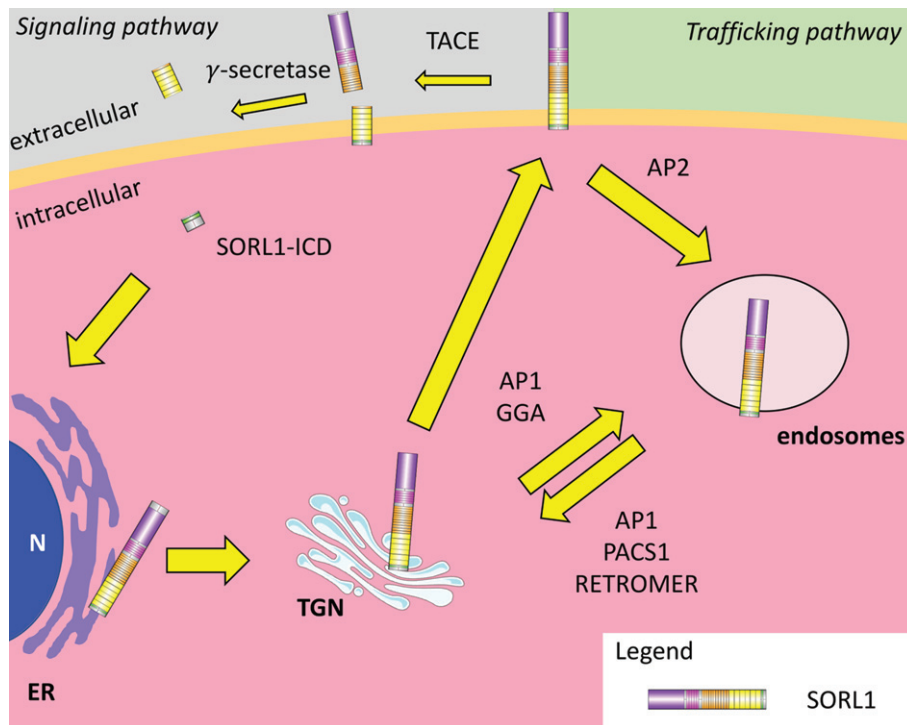


Fig. 2. SORL1 trafficking pathways. Nascent SORL1 peptides are generated in the endoplasmic reticulum (ER) and follow the constituent secretory pathway to the trans-Golgi network (TGN) where the pro-peptide is removed by furin-mediated cleavage. This allows the receptor to move to the plasma membrane where it can follow a signaling pathway (left) or a trafficking pathway (right). In the signaling pathway, SORL1 is cleaved by tumor necrosis factor- α converting enzyme (TACE) and then by γ -secretase, releasing luminal fragments of SORL1 and a cytosolic SORL1 intracellular domain (SORL1-ICD). SORL1-ICD can move to the nucleus (N) and regulate transcription of as yet unknown genes. In the trafficking pathway, SORL1 can be internalized via clathrin-mediated endocytosis utilizing the chaperone clathrin adaptor protein 2 (AP2). Internalized SORL1 receptors then shuttle between the TGN and the endosomes, guided by cytosolic adaptor proteins such as adaptor protein 1 (AP1), phosphofurin acidic cluster sorting protein (PACS1), Golgi-localizing, γ -adaptn ear homology domain ARF-interaction (GGA) and the retromer complex.

SNPs 23 – 24 ($p=0.0025$) which showed the latter association.

Further evidence for the association of the *SORL1* locus with LOAD comes from a large-scale meta-analysis of four published genome-wide association studies (GWAS) with a combined sample size of 74,046 individuals (25,580 LOAD cases and 48,466 controls). The *SORL1* locus was one of the 14 genomic regions to be associated with LOAD ($p < 5 \times 10^{-8}$). The strongest signal in the *SORL1* region was rs11218343, an intronic variant which was observed to be protective with an odds ratio of 0.77 (95% CI=0.72 – 0.82) [44]. This study was extended to include 35,274 cases and 59,163 controls and *SORL1* remained as one of the 25 genes reaching the threshold for genome-wide significance ($p < 5 \times 10^{-8}$) [45].

Due to the nature of GWAS, only common variants can be identified. With the advancement of next generation sequencing, whole exome sequencing (WES) experiments have identified rare variants in *SORL1* associated with AD. Intriguingly, an early WES study performed on families showing autosomal dominant inheritance of EOAD but without pathological variants in *AβPP*, *PSEN1*, or *PSEN2* found that 7/29 families carried putatively pathological variants in *SORL1*. This supported the idea that *SORL1* could be a fourth autosomal dominant EOfAD gene alongside the presenilins and AβPP. These variants were found throughout the coding sequence of the gene and were either missense (Gly511Arg, Asn924Ser, Asn1358Ser, Gly1681Asp, and Tyr141Cys) or nonsense (Cys1478X and Trp1821X) [18] (Fig. 3). However, Campion et al. [46] noted that some of the affected individuals in these families possessed vari-

ants in *SORL1* and other AD risk loci (e.g., *APOE* ε4, *TREM2*, and *ABCA7*). It is also worthy of note that the ages of disease onset of some of these reported patients were close to the conventional threshold of 65 years of age. Therefore, it is still somewhat uncertain whether *SORL1* variants can cause EOfAD.

Despite the observations of Campion et al. [46], numerous other studies have found that variants in *SORL1* are associated with EOAD. An independent study found 3 additional variants in *SORL1* segregating with EOAD in families (Arg416X, Gly1017-Glu1074del, and Arg1303Cys; Fig. 3), although some of the affected individuals had *APOE* genotype ε3/ε4 and had a later age of disease onset [19]. Nicolas et al. [47] found an enrichment of rare variants of *SORL1* in AD cases compared to controls in a WES study in a French cohort (OR = 5.03, 95% CI = 2.02–14.99, $p = 7.49 \times 10^{-5}$). When the authors restricted the analysis to cases showing a family history of AD, the OR increased to 8.86 (95% CI = 3.35–27.31, $p = 3.82 \times 10^{-7}$). Other studies have shown variants in *SORL1* to be associated with EOAD in different cohorts [48, 49]. However, these studies do not describe family histories of the EOAD variants, so no conclusions can be made about Mendelian inheritance.

SORL1 expression levels are also reduced in the sporadic form of AD. Gene expression profiling of lymphoblasts from LOAD patients and age-matched controls found that *SORL1* was down-regulated approximately 2-fold at the mRNA level, and 2.5-fold at the protein level [50]. Sager et al. [51] quantified SORL1 protein levels in frontal cortex of AD patients and found that approximately 30% of AD cases had reduced SORL1 protein levels. However,

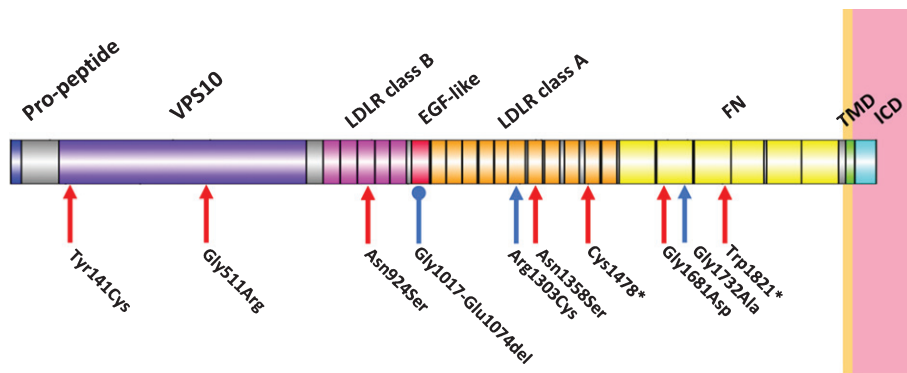


Fig. 3. EOfAD variants in *SORL1*. Figure 3 depicts a schematic of the *SORL1* full length protein, indicating the sites for early-onset, familial Alzheimer's disease (EOfAD) variants for which pedigrees have been published to date. Red arrows indicate variants published in [18] and blue arrows indicate variants published from [19]. VPS10, vacuolar protein sorting 10; LDLR, low density lipoprotein receptor; EGF, epidermal growth factor; FN, fibronectin-type; TMD, transmembrane domain; ICD, intracellular domain.

they did not find any significant differences between AD and control brains overall. Another independent study found that *SORL1* protein levels were decreased in post mortem, sporadic AD brains but not in EOfAD brains [52]. These results suggest that *SORL1* plays different roles in EOfAD and the sporadic forms of AD. Thus, the relationship between *SORL1* and the different subtypes of AD is still unclear.

Taken together, the above genetic studies support that *SORL1* is a unique gene in AD as, currently, no other genes are strongly implicated in both the early and late onset forms of the disease. Understanding the molecular and cellular changes occurring due to variation in *SORL1* could help us illuminate the differences between EOfAD and LOAD.

FUNCTIONAL STUDIES OF *SORL1* VARIANTS

Functional studies characterizing the effects of genetic variation in *SORL1* have largely been based on variants' effects on A β PP processing (A β PP processing and trafficking are discussed in detail later in this review). Cuccaro et al. [53] showed that the EOfAD variants T588I and T2134M of *SORL1* reduce the binding affinity of *SORL1* protein for A β PP, resulting in altered A β PP trafficking in HEK293 cell lines. Similarly, Vardarajan et al. [54] showed that the E270K, A528K, and T947M variants found in LOAD patients also have reduced binding affinities for A β PP and alter A β PP trafficking in HEK293 cell lines.

Young et al. [55] used human induced pluripotent stem cell (hiPSC)-derived neurons from LOAD patients to investigate the cellular effects of the risk (R) or protective (P) haplotype at the 5' SNP cluster in *SORL1*. They found expression of *SORL1* was variable among their hiPSC-derived neurons, and there were no significant differences between levels of *SORL1* mRNAs due to *SORL1* haplotype or disease state. This supported that the effect of this haplotype was likely not due to altered basal *SORL1* expression levels. However, they did find that patients carrying the R/R haplotype generally do not increase *SORL1* expression levels in response to brain-derived neurotrophic factor (BDNF, a known inducer of *SORL1* expression [56, 57] that is also implicated in AD [58, 59]). This led to increased A β levels in R/R haplotype-containing hiPSC-derived neurons relative to hiPSC-derived neurons carrying the R/P or

P/P haplotypes and, therefore, can partially explain the increased risk associated with the R/R haplotype. Having one copy of the protective haplotype appears to be sufficient to increase *SORL1* expression in response to BDNF. Therefore, increasing *SORL1* expression in response to BDNF is likely a protective mechanism.

SORL1 AND A β PP

The most studied role of *SORL1* in AD is its role in A β PP trafficking. A β PP is a single transmembrane-spanning protein which can be subjected to an amyloidogenic, or non-amyloidogenic proteolytic processing pathway. In the non-amyloidogenic pathway, A β PP is cleaved by α -secretase, giving two fragments: a secreted form of A β PP, sA β PP α , and C-terminal fragment C83. The ADAM (A disintegrin and metalloprotease domain) family of proteolytic enzymes are responsible for α -secretase cleavage of A β PP. ADAM-9 [60], ADAM-10 [60, 61], ADAM-17 [60, 62], and ADAM-19 [63] have been shown to contain α -secretase activity. However, ADAM-10 appears to be the main α -secretase [64]. Further processing of sA β PP α by γ -secretase (of which the PSENs are catalytic subunits (reviewed in [65])) gives a p3 fragment and an A β PP intracellular domain (AICD). This pathway is known as non-amyloidogenic as α -secretase cleaves within, and prevents the formation of, the A β peptide. In the amyloidogenic pathway, A β PP is cleaved by β -secretase (β -site A β PP cleaving enzyme 1, BACE1), giving sA β PP β and C99 fragments. Then C99 is further processed by γ -secretase to give A β peptides of mainly 40-42 amino acids in length (Fig. 4) (reviewed in [66]).

A β PP undergoes a series of trafficking steps within cells (Fig. 5). Nascent A β PP holoprotein is generated in the ER, before further processing in the Golgi. Some A β PP molecules can be transported through secretory vesicles to the plasma membrane, where the majority are cleaved by α -secretase and so are non-amyloidogenically processed [67]. Some surface A β PP may be internalized within endosomes by clathrin-mediated endocytosis. Internalization occurs due to interaction of the NPxY motif in the cytoplasmic tail of A β PP and the clathrin adaptor AP2 [68, 69]. Amyloidogenic processing of A β PP has been shown to proceed in many membranous organelles in the cell including the early and late-endosomal compartments [70, 71], the plasma membrane [71, 72],

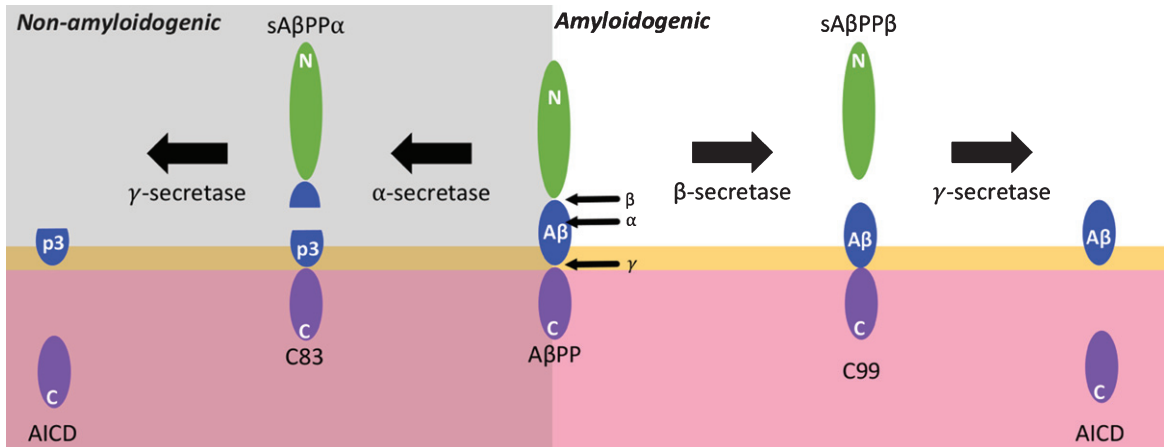


Fig. 4. Proteolytic processing of A β PP. Membrane bound amyloid- β protein precursor (A β PP) can be subjected to non-amyloidogenic or amyloidogenic processing. In the non-amyloidogenic pathway, A β PP is first cleaved by α -secretase within the amyloid- β (A β) sequence, producing a soluble sA β PP α fragment and a membrane-bound C83 fragment. C83 can be processed further by the γ -secretase complex to give a p3 fragment and an A β PP intracellular domain (AICD). In the amyloidogenic pathway, A β PP can be cleaved by β -secretase, giving the C99 and sA β PP β fragments. Then C99 is cleaved by γ -secretase to give A β peptides and an AICD. β -secretase and γ -secretase may function together as a supramolecular complex. Yellow depicts a lipid bi-layer and pink depicts a cytosolic region.

the TGN [73] (non-amyloidogenic processing is also thought to occur in the TGN [74]), the lysosomes [75, 76], and the ER [77, 78]. In regards to amyloidogenic processing at the ER, Area-Gomez et al. [79] showed that PSENs (i.e., γ -secretase) are particularly enriched in the mitochondrial associated membranes (MAMs) of the ER, which are lipid raft-like regions of the ER that form close associations with mitochondria and regulate their activity (reviewed in [80, 81]). BACE1 also is enriched in lipid rafts, and has been specifically detected in MAMs [82]. Therefore, the proteolytic cleavage of A β PP to form A β , which appears to occur in the ER, likely occurs in MAMs. Intriguingly, MAMs have been implicated in AD (see [83, 84] for excellent reviews on this topic).

An interesting observation is that SORL1 forms a complex with BACE1 in perinuclear regions in neurons [85]. BACE1 and γ -secretase have also been shown to form a complex, which sequentially cleaves A β PP holoprotein to form A β and this also occurs in perinuclear regions of the cell [86]. MAMs also have a perinuclear subcellular distribution [79, 87] and, as mentioned above, are enriched in PSENs and contain the other components of γ -secretase [79]. We have also detected SORL1 protein in the MAM [88]. Taken together, these results support the idea that SORL1, BACE1, and γ -secretase may form a complex in MAMs which proteolytically cleaves A β PP holoprotein to form A β . However, direct evidence for this SORL1/A β PP/BACE1/ γ -secretase complex has not been observed. As mentioned previously, SORL1

is also cleaved by γ -secretase [39, 40] and this could also occur in this complex. This idea is intriguing, as it links, functionally, all four EOfAD loci (if *SORL1* indeed is an EOfAD locus). Further characterization of the interactions between these AD-related proteins is required.

SORL1 can act to modulate the distribution of A β PP within cells. It has been shown that overexpression of *SORL1* results in an accumulation of A β PP in the Golgi and prevents it from being sorted to the late endosomal membranes where some A β is produced [27, 30, 89]. Conversely, Sorl1 deficiency results in increased A β levels compared to controls [30, 90]. SORL1 interacts with the carbohydrate-linked domain of A β PP through its LDLR class A domain (repeats 5-8 appear to be crucial region for binding) [30, 91, 92], as well as with the cytoplasmic domain of A β PP through its intracellular domain [93] (Fig. 1B). There is evidence supporting that this interaction between SORL1 and A β PP prevents oligomerization of A β PP, which is required for the γ -secretase-cleavage of A β PP [94]. It has also been shown that SORL1 acts as an adaptor between A β PP and other intracellular sorting proteins (including the retromer complex, AP1, GGA, and PACS1) which guide A β PP through the endosomal/secretory pathways, keeping it away from subcellular compartments where A β is thought to form (such as in endosomes). These adaptors are discussed below and are summarized in Fig. 5. Binding sites of these adaptors in the SORL1 cytoplasmic tail are shown in Fig. 1B.

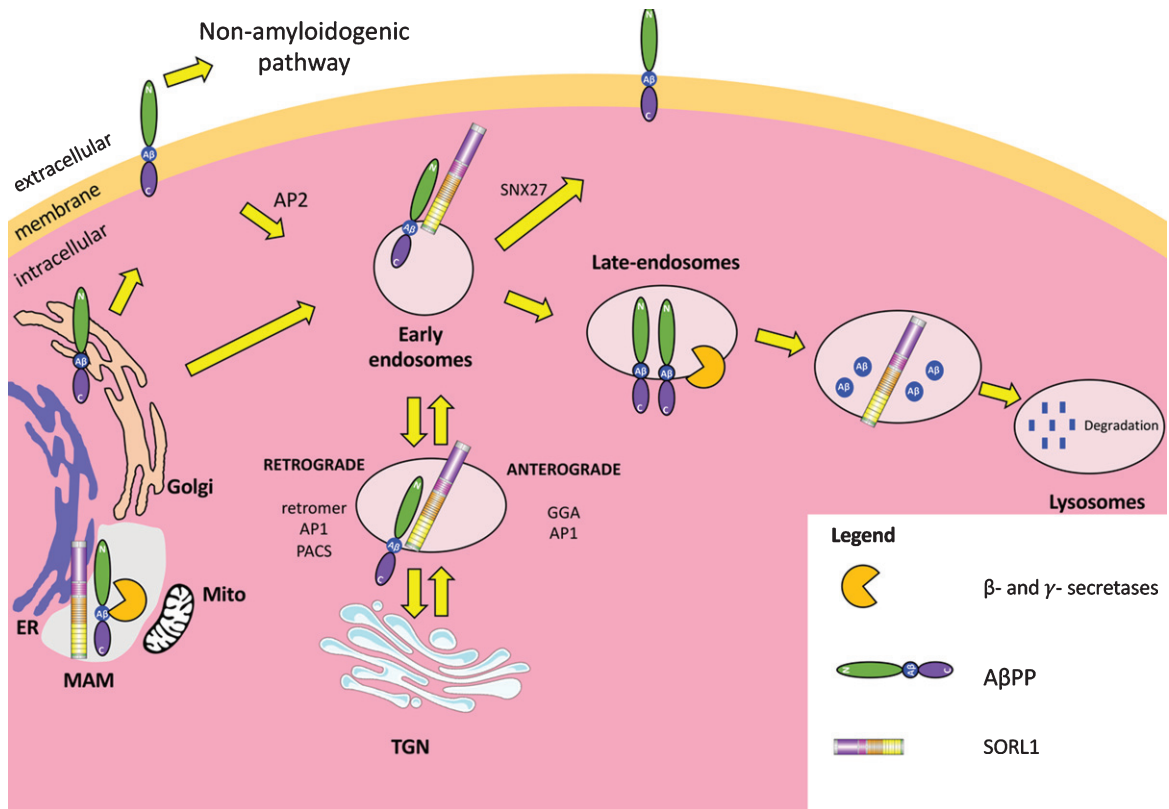


Fig. 5. SORL1 dependent trafficking of A β PP. A β PP is translated at the endoplasmic reticulum (ER) and is processed in the Golgi for direction to the cell surface. Some A β PP is cleaved in the non-amyloidogenic pathway by α -secretase and some A β PP is internalized by clathrin-mediated endocytosis via the chaperone clathrin adaptor protein 2 (AP2). SORL1 is present in early endosomes and can guide A β PP throughout different pathways in the cell by interacting with different adaptor proteins. SORL1 and A β PP can move directly back to the plasma membrane (orange) mediated by sorting nexin 27 (SNX27). They can also move retrogradely to the trans-Golgi network (TGN) mediated by the retromer complex, clathrin adaptor protein 1 (AP1), and/or phosphofurin acidic cluster sorting protein (PACS1). They can also move anterogradely from the TGN to the early endosomes mediated by Golgi-localizing, γ -adaptor ear homology domain ARF-interaction (GGA) proteins. Without SORL1, A β PP can move to late endosomal compartments where some β - and γ -secretase activities are thought to be located and can be proteolytically cleaved to form A β . SORL1 can also bind newly-formed A β and direct it to the lysosome for degradation. SORL1, A β PP, β - and γ -secretases are also present in the mitochondrial associated membranes (MAMs) of the ER.

The retromer complex is part of the endocytic machinery in cells and is responsible for retrograde transport of cargo from endosomes to the TGN as well as recycling of cargo from endosomes to the plasma membrane. It is composed of two sub-complexes of proteins encoded by genes from the vacuolar protein sorting (VPS) family. The subcomplex responsible for cargo selection consists of the proteins VPS35, VPS29 and VPS26 while the subcomplex responsible for vesicle formation consists of VPS5 and VPS17 (reviewed in [95, 96]). SORL1 interacts with the VPS26 subunit of the retromer complex through a FANSHY motif (Fig. 1B) in SORL1's cytoplasmic domain and loss of this binding motif in SORL1 resulted in aberrant subcellular localization of SORL1 and an accumulation of A β PP in

late endosomes [31]. Loss of retrograde transport of SORL1 back to the TGN would result in retention of SORL1 in endosomes. This would, in turn, increase retention of A β PP in endosomes where some A β is formed. This is supported by results from a cell culture study, where siRNA knockdown of the retromer complex caused impaired A β PP trafficking to the TGN [97]. Interestingly, gene expression profiling of the entorhinal cortex and the dentate gyrus of AD brains found that expression levels of genes encoding retromer components *VPS35* and *VPS26* were decreased in AD patients (a mix of EOAD and LOAD) compared to age-matched controls [98]. Also a GWAS showed that SNPs in retromer-associated genes were associated with sporadic AD in a Caucasian cohort [99]. In animal

models, retromer activity knockdown in the mouse brain resulted in memory defects, synaptic dysfunction and increased A β levels, which is reminiscent of AD pathology [100, 101]. Taken together, these results indicate the possibility that loss of retromer function could also play a role in sporadic forms of AD.

Golgi-localizing, γ -adaptin ear homology domain ARF-interaction (GGA) proteins are responsible for anterograde transport from the TGN to endosomes. GGA binds to SORL1 via recognition of a DVPM motif (Fig. 1B) and an acidic cluster of amino acid residues in the cytoplasmic domain of SORL1 [32, 33]. Loss of the recognition motif for GGA proteins results in an inability of SORL1 to return to the TGN so that it accumulates in early endosomes and at the plasma membrane. This also results in increased sA β PP α products as A β PP also accumulates in the organelles in which α -secretase resides [33, 102, 103].

Phosphofurin acidic cluster sorting protein 1 (PACS1) is another sorting protein which mediates both retrograde and anterograde transport between the Golgi and endosomes [104]. The furin binding region of PACS1 interacts with SORL1 by recognizing the acid cluster motif DDLGEDDED in the SORL1 cytoplasmic domain [33, 34] (Fig. 1B). Loss of PACS1 activity has a similar outcome to loss of GGA proteins, where A β PP accumulates in early endosomes, likely due to aberrant subcellular localization of SORL1 [105].

Clathrin adaptor protein 1 (AP1) binds to an overlapping region in the cytoplasmic domain of SORL1 (DDLGEDDE) (Fig. 1B) and knockdown of AP1 also results in aberrant subcellular localization of SORL1 [34].

Sorting nexin 27 (SNX27) is a sorting protein involved in retrograde transport from the endosomes to the plasma membrane. SNX27 mediates the transport of A β PP via binding to the membrane-proximal region of the SORL1 cytoplasmic tail. Knockdown of SNX27 resulted in a reduced amount of cell surface SORL1 and A β PP protein levels. Conversely, overexpression of SNX27 enhanced distribution of A β PP and SORL1 to the cell surface. An increase in A β PP protein levels at the cell surface also resulted in increased α -secretase cleavage of A β PP to give sA β PP α [106]. However, SNX27 has been shown to interact with PSEN1 protein and this appears to reduce γ -secretase activity (i.e., reduced cleavage of A β PP and Notch) [107] so this could also explain the increased levels of sA β PP α .

SORL1 also plays a role in trafficking of A β . It interacts with A β through its VPS10 domain as shown by fluorescence polarization assays. Overexpression of *SORL1* in a neuronal cell line also expressing a human A β PP isoform resulted in faster turnover of intracellular A β and accumulation of A β in lysosomes. Furthermore, an EOfAD variant in the VPS10 domain of SORL1 disrupts the ability for SORL1 to bind A β [108, 109]. This process also appears to be mediated by GGA proteins [103]. These results indicate that SORL1 directs A β to lysosomes for degradation.

In summary, SORL1 plays an important and central role in the complex trafficking pathways of A β PP within cells. Therefore, it is no surprise that the loss of *SORL1* activity in sporadic and some familial forms of AD leads to A β PP-related pathologies.

OTHER POTENTIAL ROLES OF *SORL1* IN AD

SORL1 and lipoprotein metabolism

When AD was first described by Alois Alzheimer over a century ago, he observed a third pathological hallmark of ‘adipose inclusions’ and ‘lipid granules’ in the glia of his deceased patient Auguste Deter [110]. Subsequently, these observations were largely overlooked, but it is now well established that aberrant lipid metabolism occurs in AD brains. Lipids in the brain include glycerophospholipids, sphingolipids and cholesterol. These lipids play important roles such as in the structure of the myelin sheath which insulates axons and in formation of the plasma membrane. Cholesterol is one of the most well studied lipids and its homeostasis is implicated in AD. Indeed, there is an enrichment of cholesterol in cell membranes in AD patients compared to controls and there appears to be a positive correlation between cholesterol level and disease severity [111]. It is thought that cholesterol in the CNS is synthesized by glial cells and then imported into neurons as APOE-containing lipoproteins. However, neurons have also been shown to express genes involved in cholesterol synthesis [112] so the reality of brain cholesterol metabolism may be more complex. SORL1 mediates cholesterol intake into neurons by acting as a receptor for APOE. SORL1 has preferential binding affinity for each of the isoforms of APOE which corresponds with the relative risk that each APOE allele presents for development of AD. SORL1’s highest affinity is with the ϵ 4 isoform, followed by ϵ 3 then ϵ 2 [113]. Each isoform of APOE differs by cysteine (Cys) and

arginine (Arg) content at positions 112 and 158: $\epsilon 4$ (Arg, Arg), $\epsilon 3$ (Cys, Arg), and $\epsilon 2$ (Cys, Cys). These amino acid residue changes result in conformational changes to the protein structure of APOE (reviewed in [114]) and could explain the differential affinities between each of the APOE isoforms and SORL1.

Overexpression of *SORL1* increases cellular uptake of APOE $\epsilon 4$ and $\epsilon 3$ (but not $\epsilon 2$, which is protective for AD (reviewed in [115])) as shown by a cell culture study [113]. Additionally, neural stem cells isolated from an AD patient with genotype $\epsilon 4/\epsilon 4$ had lower *SORL1* expression after 5 weeks in culture than neural stem cells from AD and control subjects with the other possible *APOE* genotypes [116]. However, this study only analyzed one homozygous $\epsilon 4$ patient. Collectively, these results indicate that SORL1 and APOE are functionally related.

SORL1 and insulin signaling

It has been proposed that AD may represent a ‘type 3 diabetes’ (reviewed in [117]) and that insulin resistance/deficiency underlies AD pathology. The insulin signaling pathway has been shown to be perturbed in AD brains [118, 119]. Since insulin signaling stimulates glucose metabolism, this could partly explain the aberrant glucose metabolism observed in AD brains in FDG-PET studies (reviewed in [120]).

SORL1 has been shown to act as a sorting receptor for the insulin receptor in adipocytes. It directs internalized insulin receptor back to the cell surface, preventing lysosomal catabolism and, thereby amplifying insulin signals [121]. Loss of *SORL1* activity could decrease the amount of insulin receptor at the cell surface and perturb the insulin signaling pathway. It is unknown whether this also occurs in the CNS and further investigation is required.

Is SORL1 involved in iron homeostasis?

There is strong evidence that perturbation of iron homeostasis may play a role in AD (reviewed in [122]). To our knowledge, *SORL1* has not been linked directly to iron homeostasis. However, A β PP is thought to play a role in maintaining iron homeostasis and so *SORL1* may have indirect effects.

Iron content is increased in both living [123] and post-mortem [124, 125] AD brains relative to controls. Iron dyshomeostasis may also be linked to many of the pathologies observed in AD such as mitochondrial dysfunction, vascular pathologies, changes in

energy metabolism, and inflammation (reviewed in [122]).

A recent, elegant study by Yambire et al. [126] showed that acidification of the endo-lysosomal system was critical for ferric iron (Fe^{3+}) to be reduced to the more reactive ferrous (Fe^{2+}) form and released into the cytosol. Inhibition of endo-lysosomal acidification was shown to lead to a cellular ferrous iron deficiency and a pseudo-hypoxic response (as HIF-1 α , a master regulator of cellular responses to hypoxia, is normally inhibited from acting under normoxia by a degradative mechanism requiring Fe^{2+}). Mitochondrial biogenesis and function were also disturbed and inflammation markers were increased. It has been shown that the C99 fragment of A β PP (and interestingly, also the PSEN1 holoprotein [127]) is required for acidification of the endo-lysosomal system [128]. Also, *A β PP*^{-/-} mice are observed to have accumulation of intraneuronal iron [129, 130]. This demonstrates that interference with A β PP activity affects cellular iron homeostasis. If SORL1 is required for the proper subcellular localization of A β PP, then the loss of *SORL1* activity observed in sporadic AD and some EOfAD cases may also indirectly affect iron homeostasis via A β PP. Further investigation of the role of *SORL1* in iron homeostasis is required.

SORL1 and estrogens

One hypothesis to explain the female bias observed in AD incidence [141, 142], is the reduced estrogen levels in females due to menopause. This may increase the susceptibility of post-menopausal females to AD. Indeed, estrogen levels have been shown to be reduced in female AD brains relative to healthy, age-matched, female controls [143].

Estrogens have been shown to have neuroprotective effects in the brain. Estrogen administration reduces the severity of lesions in the brain after either permanent or transient ischemia [144]. This is relevant to AD as hypoxia/reduced blood flow is evident in AD brains [145, 146]. It has also been shown that estrogens are protective against A β toxicity *in vitro* [147, 148].

SORL1 expression appears to be stimulated by estrogens. Ratnakumar et al. [149] found that *SORL1* was one of the 504 genes significantly differentially expressed in laser-dissected, serotonergic neurons from female, adult, ovariectomized Rhesus macaques treated with estrogen. Interestingly, *A β PP* and *APOE*

were also among these genes significantly differentially expressed ($p < 0.05$, log fold change > 2). *SORL1* has also been shown to be differentially expressed in various cancer cell lines in response to estrogen treatments [150, 151]. Intriguingly, some studies have found an association between *SORL1* variants and AD in a sex-specific manner in various cohorts [152, 153]. However, a detailed investigation of the relationship between estrogens and *SORL1* in the context of AD is yet to be performed.

CONCLUSION

It is no surprise that variation at the *SORL1* locus is associated with AD, as *SORL1* plays roles in many of the cellular processes that have been linked to AD pathologies. *SORL1* shows potential to illuminate mechanistic similarities and differences between rare EOfAD and the much more common sporadic forms of LOAD. However, more family-based studies analyzing the segregation of variants in *SORL1* with EOAD are required to confirm whether *SORL1* can be regarded as a causative EOfAD locus alongside the *PSENs* and *AβPP*.

ACKNOWLEDGMENTS

The authors declare no conflicts of interest, financial or otherwise, in publication of this paper. KB is supported by an Australian Government Research Training Program Scholarship. MN is supported by National Health and Medical Research Council (NHMRC) Project Grant APP1061006. ML is an academic employee of the University of Adelaide. ML is also supported by the NHMRC (APP1045507 and APP1061006).

CONFLICT OF INTEREST

The authors have no conflicts of interest to report.

REFERENCES

- [1] Blennow K, de Leon MJ, Zetterberg H (2006) Alzheimer's disease. *Lancet* **368**, 387-403.
- [2] Masters CL, Bateman R, Blennow K, Rowe CC, Sperling RA, Cummings JL (2015) Alzheimer's disease. *Nat Rev Dis Primers* **1**, 15056.
- [3] Mullan M, Crawford F, Axelman K, Houlden H, Lilius L, Winblad B, Lannfelt L (1992) A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of β -amyloid. *Nat Genet* **1**, 345-347.
- [4] Kamino K, Orr HT, Payami H, Wijsman EM, Alonso ME, Pulst SM, Anderson L, O'Dahl S, Nemens E, White JA, et al. (1992) Linkage and mutational analysis of familial Alzheimer disease kindreds for the APP gene region. *Am J Hum Genet* **51**, 998-1014.
- [5] Kumar-Singh S, De Jonghe C, Cruts M, Kleinert R, Wang R, Mercken M, De Strooper B, Vanderstichele H, Lofgren A, Vanderhoeven I, Backhovens H, Vanmechelen E, Kroisel PM, Van Broeckhoven C (2000) Nonfibrillar diffuse amyloid deposition due to a gamma(42)-secretase site mutation points to an essential role for N-truncated A beta(42) in Alzheimer's disease. *Hum Mol Genet* **9**, 2589-2598.
- [6] Murrell J, Farlow M, Ghetti B, Benson MD (1991) A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease. *Science* **254**, 97-99.
- [7] Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes A, Irving N, James L, et al. (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* **349**, 704-706.
- [8] Campion D, Flaman JM, Brice A, Hannequin D, Dubois B, Martin C, Moreau V, Charbonnier F, Didierjean O, Tardieu S, et al. (1995) Mutations of the presenilin I gene in families with early-onset Alzheimer's disease. *Hum Mol Genet* **4**, 2373-2377.
- [9] Hutton M, Busfield F, Wrang M, Crook R, Perez-Tur J, Clark RF, Prihar G, Talbot C, Phillips H, Wright K, Baker M, Lendon C, Duff K, Martinez A, Houlden H, Nichols A, Karran E, Roberts G, Roques P, Rossor M, Venter JC, Adams MD, Cline RT, Phillips CA, Goate A, et al. (1996) Complete analysis of the presenilin 1 gene in early onset Alzheimer's disease. *Neuroreport* **7**, 801-805.
- [10] Crook R, Ellis R, Shanks M, Thal LJ, Perez-Tur J, Baker M, Hutton M, Haltia T, Hardy J, Galasko D (1997) Early-onset Alzheimer's disease with a presenilin-1 mutation at the site corresponding to the Volga German presenilin-2 mutation. *Ann Neurol* **42**, 124-128.
- [11] Lanoiselee HM, Nicolas G, Wallon D, Rovelet-Lecrux A, Lacour M, Rousseau S, Richard AC, Pasquier F, Rollin-Sillaire A, Martinaud O, Quillard-Muraine M, de la Sayette V, Boutoleau-Bretonniere C, Etcharry-Bouyx F, Chauvire V, Sarazin M, le Ber I, Epelbaum S, Jonveaux T, Rouaud O, Ceccaldi M, Felician O, Godefroy O, Formaglio M, Croisile B, Auriaacombes S, Chamard L, Vincent JL, Sauvee M, Marelli-Tosi C, Gabelle A, Ozsancak C, Pariente J, Paquet C, Hannequin D, Campion D, collaborators of the CNR-MAJ project (2017) APP, PSEN1, and PSEN2 mutations in early-onset Alzheimer disease: A genetic screening study of familial and sporadic cases. *PLoS Med* **14**, e1002270.
- [12] Lemere CA, Lopera F, Kosik KS, Lendon CL, Ossa J, Saito TC, Yamaguchi H, Ruiz A, Martinez A, Madrigal L, Hincapie L, Arango JC, Anthony DC, Koo EH, Goate AM, Selkoe DJ, Arango JC (1996) The E280A presenilin 1 Alzheimer mutation produces increased A beta 42 deposition and severe cerebellar pathology. *Nat Med* **2**, 1146-1150.
- [13] Ishikawa A, Piao Y-S, Miyashita A, Kuwano R, Onodera O, Ohtake H, Suzuki M, Nishizawa M, Takahashi H (2005) A mutant PSEN1 causes dementia with Lewy bodies and variant Alzheimer's disease. *Ann Neurol* **57**, 429-434.
- [14] Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, Yu CE, Jondro PD, Schmidt SD, Wang K, et al. (1995) Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* **269**, 973-977.

- [15] Finckh U, Alberici A, Antoniazzi M, Benussi L, Fedi V, Giannini C, Gal A, Nitsch RM, Binetti G (2000) Variable expression of familial Alzheimer disease associated with presenilin 2 mutation M239I. *Neurology* **54**, 2006-2008.
- [16] Jayadev S, Leverenz JB, Steinbart E, Stahl J, Klunk W, Yu CE, Bird TD (2010) Alzheimer's disease phenotypes and genotypes associated with mutations in presenilin 2. *Brain* **133**, 1143-1154.
- [17] Rogaev EI, Sherrington R, Rogaeva EA, Levesque G, Ikeda M, Liang Y, Chi H, Lin C, Holman K, Tsuda T, et al. (1995) Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature* **376**, 775-778.
- [18] Pottier C, Hannequin D, Coutant S, Rovelet-Lecrux A, Wallon D, Rousseau S, Legallic S, Paquet C, Bombois S, Pariente J, Thomas-Anterion C, Michon A, Croisile B, Etchary-Bouyx F, Berr C, Dartigues JF, Amouyel P, Dauchel H, Boutoleau-Bretonniere C, Thauvin C, Frebourg T, Lambert JC, Campion D, PHRC GMAJ Collaborators (2012) High frequency of potentially pathogenic *SORL1* mutations in autosomal dominant early-onset Alzheimer disease. *Mol Psychiatry* **17**, 875-879.
- [19] Thonberg H, Chiang H-H, Lilius L, Forsell C, Lindström A-K, Johansson C, Björkström J, Thordardottir S, Slegers K, Van Broeckhoven C, Rönnbäck A, Graff C (2017) Identification and description of three families with familial Alzheimer disease that segregate variants in the *SORL1* gene. *Acta Neuropathol Commun* **5**, 43.
- [20] Wen Y, Miyashita A, Kitamura N, Tsukie T, Saito Y, Hattuta H, Murayama S, Kakita A, Takahashi H, Akatsu H, Yamamoto T, Kosaka K, Yamaguchi H, Akazawa K, Ihara Y, Kuwano R (2013) *SORL1* is genetically associated with neuropathologically characterized late-onset Alzheimer's disease. *J Alzheimers Dis* **35**, 387-394.
- [21] Rogaeva E, Meng Y, Lee JH, Gu Y, Kawarai T, Zou F, Katayama T, Baldwin CT, Cheng R, Hasegawa H, Chen F, Shibata N, Lunetta KL, Pardossi-Piquard R, Bohm C, Wakutani Y, Cupples LA, Cuenco KT, Green RC, Pinessi L, Rainero I, Sorbi S, Bruni A, Duara R, Friedland RP, Inzelberg R, Hampe W, Bujo H, Song Y-Q, Andersen OM, Willnow TE, Graff-Radford N, Petersen RC, Dickson D, Der SD, Fraser PE, Schmitt-Ulms G, Younkin S, Mayeux R, Farrer LA, St George-Hyslop P (2007) The neuronal sortilin-related receptor *SORL1* is genetically associated with Alzheimer disease. *Nat Genet* **39**, 168-177.
- [22] Liu G, Sun JY, Xu M, Yang XY, Sun BL (2017) *SORL1* variants show different association with early-onset and late-onset Alzheimer's disease risk. *J Alzheimers Dis* **58**, 1121-1128.
- [23] Motoi Y, Aizawa T, Haga S, Nakamura S, Namba Y, Ikeda K (1999) Neuronal localization of a novel mosaic apolipoprotein E receptor, LR11, in rat and human brain. *Brain Res* **833**, 209-215.
- [24] Jacobsen L, Madsen P, Moestrup SK, Lund AH, Tommerup N, Nykjaer A, Sottrup-Jensen L, Gliemann J, Petersen CM (1996) Molecular characterization of a novel human hybrid-type receptor that binds the alpha2-macroglobulin receptor-associated protein. *J Biol Chem* **271**, 31379-31383.
- [25] Frankish A, Vullo A, Zadissa A, Yates A, Thormann A, Parker A, Gall A, Moore B, Walts B, Aken BL, Cummins C, Girón CG, Ong CK, Sheppard D, Staines DM, Murphy DN, Zerbino DR, Ogeh D, Perry E, Haskell E, Martin FJ, Cunningham F, Riat HS, Schuilenburg H, Sparrow H, Lavidas I, Loveland JE, To JK, Mudge J, Bhai J, Taylor K, Billis K, Gil L, Haggerty L, Gordon L, Amode MR, Ruffier M, Patricio M, Laird MR, Muffato M, Nuhn M, Kostadima M, Langridge N, Izuogu OG, Achuthan P, Hunt SE, Janacek SH, Trevanion SJ, Hourlier T, Juettemann T, Maurel T, Newman V, Akanni W, McLaren W, Liu Z, Barrell D, Flicek P (2017) Ensembl 2018. *Nucleic Acids Res* **46**, D754-D761.
- [26] Blechingsberg J, Poulsen ASA, Kjølbj M, Monti G, Allen M, Ivarsen AK, Lincoln SJ, Thotakura G, Vægter CB, Ertekin-Taner N, Nykjær A, Andersen OM (2018) An alternative transcript of the Alzheimer's disease risk gene *SORL1* encodes a truncated receptor. *Neurobiol Aging* **71**, 266.e211-266.e224.
- [27] Offe K, Dodson SE, Shoemaker JT, Fritz JJ, Gearing M, Levey AI, Lah JJ (2006) The lipoprotein receptor LR11 regulates amyloid beta production and amyloid precursor protein traffic in endosomal compartments. *J Neurosci* **26**, 1596-1603.
- [28] Grear KE, Ling IF, Simpson JF, Furman JL, Simmons CR, Peterson SL, Schmitt FA, Markesbery WR, Liu Q, Crook JE, Younkin SG, Bu G, Estus S (2009) Expression of *SORL1* and a novel *SORL1* splice variant in normal and Alzheimer's disease brain. *Mol Neurodegener* **4**, 46.
- [29] Ciarlo E, Massone S, Penna I, Nizzari M, Gigoni A, Dieci G, Russo C, Florio T, Cancedda R, Pagano A (2013) An intronic ncRNA-dependent regulation of *SORL1* expression affecting Aβ formation is upregulated in post-mortem Alzheimer's disease brain samples. *Dis Model Mech* **6**, 424.
- [30] Andersen OM, Reiche J, Schmidt V, Gotthardt M, Spoelgen R, Behlke J, von Arnim CA, Breiderhoff T, Jansen P, Wu X, Bales KR, Cappai R, Masters CL, Gliemann J, Mufson EJ, Hyman BT, Paul SM, Nykjær A, Willnow TE (2005) Neuronal sorting protein-related receptor sorLA/LR11 regulates processing of the amyloid precursor protein. *Proc Natl Acad Sci U S A* **102**, 13461-13466.
- [31] Fjorback AW, Seaman M, Gustafsen C, Mehmedbasic A, Gokool S, Wu C, Militz D, Schmidt V, Madsen P, Nyengaard JR, Willnow TE, Christensen EI, Mobley WB, Nykjær A, Andersen OM (2012) Retromer binds the FANSHY sorting motif in SorLA to regulate amyloid precursor protein sorting and processing. *J Neurosci* **32**, 1467-1480.
- [32] Jacobsen L, Madsen P, Nielsen MS, Geraerts WPM, Gliemann J, Smit AB, Petersen CM (2002) The sorLA cytoplasmic domain interacts with GGA1 and -2 and defines minimum requirements for GGA binding. *FEBS Lett* **511**, 155-158.
- [33] Schmidt V, Sporbert A, Rohe M, Reimer T, Rehm A, Andersen OM, Willnow TE (2007) SorLA/LR11 regulates processing of amyloid precursor protein via interaction with adaptors GGA and PACS-1. *J Biol Chem* **282**, 32956-32964.
- [34] Nielsen MS, Gustafsen C, Madsen P, Nyengaard JR, Hermey G, Bakke O, Mari M, Schu P, Pohlmann R, Dennes A, Petersen CM (2007) Sorting by the cytoplasmic domain of the amyloid precursor protein binding receptor SorLA. *Mol Cell Biol* **27**, 6842-6851.
- [35] Jacobsen L, Madsen P, Jacobsen C, Nielsen MS, Gliemann J, Petersen CM (2001) Activation and functional characterization of the mosaic receptor SorLA/LR11. *J Biol Chem* **276**, 22788-22796.
- [36] Guo L (2002) A proteomic approach for the identification of cell-surface proteins shed by metalloproteases. *Mol Cell Proteomics* **1**, 30-36.

- [37] Hampe W, Riedel IB, Lintzel J, Bader CO, Franke I, Schaller HC (2000) Ectodomain shedding, translocation and synthesis of SorLA are stimulated by its ligand head activator. *J Cell Sci* **113 Pt 24**, 4475-4485.
- [38] Hermeij G, Sjøgaard SS, Petersen CM, Nykjaer A, Gliemann J (2006) Tumour necrosis factor alpha-converting enzyme mediates ectodomain shedding of Vps10p-domain receptor family members. *Biochem J* **395**, 285-293.
- [39] Böhm C, Seibel NM, Henkel B, Steiner H, Haass C, Hampe W (2006) SorLA signaling by regulated intramembrane proteolysis. *J Biol Chem* **281**, 14547-14553.
- [40] Nyborg AC, Ladd TB, Zwizinski CW, Lah JJ, Golde TE (2006) Sortilin, SorCS1b, and SorLA Vps10p sorting receptors, are novel γ -secretase substrates. *Mol Neurodegener* **1**, 3.
- [41] Willnow TE, Andersen OM (2013) Sorting receptor SORLA—a trafficking path to avoid Alzheimer disease. *J Cell Sci* **126**, 2751-2760.
- [42] Schmidt V, Subkhanulova A, Willnow TE (2017) Sorting receptor SORLA: cellular mechanisms and implications for disease. *Cell Mol Life Sci* **74**, 1475-1483.
- [43] Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, Myers RH, Pericak-Vance MA, Risch N, van Duijn CM (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**, 1349-1356.
- [44] Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, DeStefano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thornton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin C-F, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau M-T, Choi S-H, Reitz C, Pasquier F, Hollingworth P, Ramirez A, Hanon O, Fitzpatrick AL, Buxbaum JD, Campion D, Crane PK, Baldwin C, Becker T, Gudnason V, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Morón FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fiévet N, Huettelmann MJ, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuinness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, St George-Hyslop P, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossù P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Naranjo MCD, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, European Alzheimer's Disease Initiative (EADI); Genetic and Environmental Risk in Alzheimer's Disease; Alzheimer's Disease Genetic Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannfelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley Jr TH, Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RFAG, Passmore P, Montine TJ, Bettens K, Rotter JI, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JSK, Boerwinkle E, Riemenschneider M, Boada M, Hiltunen M, Martin ER, Schmidt R, Rujescu D, Wang L-S, Dartigues J-F, Mayeux R, Tzourio C, Hofman A, Nöthen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM, Van Broeckhoven C, Moskvina V, Seshadri S, Williams J, Schellenberg GD, Amouyel P (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452.
- [45] Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, Boland A, Vronskaya M, Van Der Lee SJ, Amlie-Wolf A, Bellenguez C, Frizzati A, Chouraki V, Martin ER, Sleegers K, Badarinarayan N, Jakobsdottir J, Hamilton-Nelson KL, Moreno-Grau S, O'Laso R, Raybould R, Chen Y, Kuzma AB, Hiltunen M, Morgan T, Ahmad S, Vardarajan BN, Epelbaum J, Hoffmann P, Boada M, Beecham GW, Garnier J-G, Harold D, Fitzpatrick AL, Valladares O, Moutet M-L, Gerrish A, Smith AV, Qu L, Bacq D, Denning N, Jian X, Zhao Y, Del Zompo M, Fox NC, Choi S-H, Mateo I, Hughes JT, Adams HH, Malamon J, Sanchez-Garcia F, Patel Y, Brody JA, Dombroski BA, Naranjo MCD, Daniilidou M, Eiriksdottir G, Mukherjee S, Wallon D, Uphill J, Aspelund T, Cantwell LB, Garzia F, Galimberti D, Hofer E, Butkiewicz M, Fin B, Scarpini E, Sarnowski C, Bush WS, Meslage S, Kornhuber J, White CC, Song Y, Barber RC, Engelborghs S, Sordon S, Voijnovic D, Adams PM, Vandenberghe R, Mayhaus M, Cupples LA, Albert MS, De Deyn PP, Gu W, Himali JJ, Beekly D, Squassina A, Hartmann AM, Orellana A, Blacker D, Rodriguez-Rodriguez E, Lovestone S, Garcia ME, Doody RS, Munoz-Fernandez C, Sussars R, Lin H, Fairchild TJ, Benito YA, Holmes C, Karamujic-Comić H, Frosch MP, Thonberg H, Maier W, Roshchupkin G, Ghatti B, Giedraitis V, Kawalia A, Li S, Huebinger RM, Kilander L, Moebus S, Hernández I, Kamboh MI, Brundin R, Turton J, Yang Q, Katz MJ, Concari L, Lord J, Beiser AS, Keene CD, Helisalmi S, Kloszewska I, Kukull WA, Koivisto AM, Lynch A, Tarraga L, Larson EB, Haapasalo A, Lawlor B, Mosley TH, Lipton RB, Solfrizzi V, Gill M, Longstreth WT, Montine TJ, Frisardi V, Diez-Fairen M, Rivadeneira F, Petersen RC, Deramecourt V, Alvarez I, Salani F, Ciarrella A, Boerwinkle E, Reiman EM, Fievet N, Rotter JI, Reisch JS, Hanon O, Cupidi C, Andre Uitterlinden AG, Royall DR, Dufouil C, Maletta RG, De Rojas I, Sano M, Brice A, Cecchetti R, George-Hyslop PS, Ritchie K, Tsolaki M, Tsuang DW, Dubois B, Craig D, Wu C-K, Soininen H, Avramidou D, Albin RL, Fratiglioni L, Germanou A, Apostolova LG, Keller L, Koutroumani M, Arnold SE, Panza F, Gkatzima O, Asthana S, Hannequin D, Whitehead P, Atwood CS, Caffarra P, Hampel H, Quintela I, Carracedo Á, Lannfelt L, Rubinsztein DC, Barnes LL, Pasquier F, Frölich L, Barral S, McGuinness B, Beach TG, Johnston JA, Becker JT, Passmore P, Bigio EH, Schott JM, Bird TD, Warren JD, Boeve BF, Lupton MK, Bowen JD, Proitsi P, Boxer A, Powell JF, Burke JR, Kauwe JSK, Burns JM, Mancuso M, Buxbaum JD, Bonuccelli U, Cairns NJ, McQuillin A, Cao C, Livingston G, Carlson CS, Bass NJ, Carlsson CM, Hardy J, Carney RM, Bras J, Carrasquillo MM, Guerreiro R, Allen M, Chui HC, Fisher E, Masullo C, Crocco EA, Decarli C, Bisceglia G, Dick M, Ma L, Duara R, Graff-Radford NR, Evans DA, Hodges A, Faber KM, Scherer M, Fallon KB, Riemenschneider M, Fardo DW, Heun R, Farlow MR, Kölsch H, Ferris S, Leber M, Foroud TM, Heuser I, Galasko DR, Giegling I, Gearing

- M, Hüll M, Geschwind DH, Gilbert JR, Morris J, Green RC, Mayo K, Growdon JH, Feulner T, Hamilton RL, Harrell LE, Drichel D, Honig LS, Cushion TD, Huentelman MJ, Hollingworth P, Hulette CM, Hyman BT, Marshall R, Jarvik GP, Meggy A, Abner E, Menzies GE, Jin L-W, Leonenko G, Real LM, Jun GR, Baldwin CT, Grozeva D, Karydas A, Russo G, Kaye JA, Kim R, Jessen F, Kowall NW, Vellas B, Kramer JH, Vardy E, Laferla FM, Jöckel K-H, Lah JJ, Dichgans M, Leverenz JB, Mann D, Levey AI, Pickering-Brown S, Lieberman AP, Klopp N, Lunetta KL, Wichmann HE, Lyketsos CG, Morgan K, Marson DC, Brown K, Martiniuk F, Medway C, Mash DC, Nöthen MM, Masliah E, Hooper NM, McCormick WC, Daniele A, McCurry SM, Bayer A, McDavid AN, Gallacher J, McKee AC, Van Den Bussche H, Mesulam M, Brayne C, Miller BL, Riedel-Heller S, Miller CA, Miller JW, Al-Chalabi A, Morris JC, Shaw CE, Myers AJ, Wiltfang J, O'Bryant S, Olichney JM, Alvarez V, Parisi JE, Singleton AB, Paulson HL, Collinge J, Perry WR, Mead S, Peskind E, Cribbs DH, Rossor M, Pierce A, Ryan NS, Poon WW, Nacmias B, Potter H, Sorbi S, Quinn JF, Sacchinelli E, Raj A, Spalletta G, Raskind M, Caltagirone C, Bossù P, Orfei MD, Reisberg B, Clarke R, Reitz C, Smith AD, Ringman JM, Warden D, Roberson ED, Wilcock G, Rogaeva E, Bruni AC, Rosen HJ, Gallo M, Rosenberg RN, Ben-Shlomo Y, Sager MA, Mecocci P, Saykin AJ, Pastor P, Cuccaro ML, Vance JM, Schneider JA, Schneider LS, Slifer S, Seeley WW, Smith AG, Sonnen JA, Spina S, Stern RA, Swerdlow RH, Tang M, Tanzi RE, Trojanowski JQ, Troncoso JC, Van Deerlin VM, Van Eldik LJ, Vinters HV, Vonsattel JP, Weintraub S, Welsh-Bohmer KA, Wilhelmsen KC, Williamson J, Wingo TS, Woltjer RL, Wright CB, Yu C-E, Yu L, Saba Y, Pilotto A, Bullido MJ, Peters O, Crane PK, Bennett D, Bosco P, Coto E, Boccardi V, De Jager PL, Lleó A, Warner N, Lopez OL, Ingelsson M, Deloukas P, Cruchaga C, Graff C, Gwilliam R, Fornage M, Goate AM, Sanchez-Juan P, Kehoe PG, Amin N, Ertekin-Taner N, Berr C, Debette S, Love S, Launer LJ, Younkin SG, Dartigues J-F, Corcoran C, Ikram MA, Dickson DW, Nicolas G, Campion D, Tschanz J, Schmidt H, Hakonarson H, Clarimon J, Munger R, Schmidt R, Farrer LA, Van Broeckhoven C, C. O'Donovan M, Destefano AL, Jones L, Haines JL, Deleuze J-F, Owen MJ, Gudnason V, Mayeux R, Escott-Price V, Psaty BM, Ramirez A, Wang L-S, Ruiz A, Van Duijn CM, Holmans PA, Seshadri S, Williams J, Amouyel P, Schellenberg GD, Lambert J-C, Pericak-Vance MA (2019) Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat Genet* **51**, 414-430.
- [46] Campion D, Charbonnier C, Nicolas G (2019) *SORL1* genetic variants and Alzheimer disease risk: a literature review and meta-analysis of sequencing data. *Acta Neuropathol* **138**, 173-186.
- [47] Nicolas G, Charbonnier C, Wallon D, Quenez O, Bellenguez C, Grenier-Boley B, Rousseau S, Richard AC, Rovelet-Lecruex A, Le Guennec K, Bacq D, Garnier JG, Olaso R, Boland A, Meyer V, Deleuze JF, Amouyel P, Munter HM, Bourque G, Lathrop M, Frebourg T, Redon R, Letenneur L, Dartigues JF, Génin E, Lambert JC, Hannequin D, Campion D (2016) *SORL1* rare variants: a major risk factor for familial early-onset Alzheimer's disease. *Mol Psychiatry* **21**, 831-836.
- [48] Verheijen J, Van den Bossche T, van der Zee J, Engelborghs S, Sanchez-Valle R, Lladó A, Graff C, Thonberg H, Pastor P, Ortega-Cubero S, Pastor MA, Benussi L, Ghidoni R, Binetti G, Clarimon J, Lleó A, Fortea J, de Mendonça A, Martins M, Grau-Rivera O, Gelpi E, Bettens K, Mateiu L, Dillen L, Cras P, De Deyn PP, Van Broeckhoven C, Sleegers K (2016) A comprehensive study of the genetic impact of rare variants in *SORL1* in European early-onset Alzheimer's disease. *Acta Neuropathol* **132**, 213-224.
- [49] Bellenguez C, Charbonnier C, Grenier-Boley B, Quenez O, Le Guennec K, Nicolas G, Chauhan G, Wallon D, Rousseau S, Richard AC, Boland A, Bourque G, Munter HM, Olaso R, Meyer V, Rollin-Sillaire A, Pasquier F, Letenneur L, Redon R, Dartigues JF, Tzourio C, Frebourg T, Lathrop M, Deleuze JF, Hannequin D, Genin E, Amouyel P, Debette S, Lambert JC, Campion D; CNR MAJ collaborators (2017) Contribution to Alzheimer's disease risk of rare variants in *TREM2*, *SORL1*, and *ABCA7* in 1779 cases and 1273 controls. *Neurobiol Aging* **59**, 220.e1-220.e9.
- [50] Scherzer CR, Offe K, Gearing M, Rees HD, Fang G, Heilman CJ, Schaller C, Bujo H, Levey AI, Lah JJ (2004) Loss of apolipoprotein e receptor *Irl1* in Alzheimer disease. *Arch Neurol* **61**, 1200-1205.
- [51] Sager KL, Wu J, Herskowitz JH, Mufson EJ, Levey AI, Lah JJ (2012) Neuronal *LR11* expression does not differentiate between clinically-defined Alzheimer's disease and control brains. *PLoS One* **7**, e40527.
- [52] Dodson SE, Gearing M, Lippa CF, Montine TJ, Levey AI, Lah JJ (2006) *LR11/SorLA* expression is reduced in sporadic Alzheimer disease but not in familial Alzheimer disease. *J Neuropathol Exp Neurol* **65**, 866-872.
- [53] Cuccaro ML, Carney RM, Zhang Y, Bohm C, Kunkle BW, Vardarajan BN, Whitehead PL, Cukier HN, Mayeux R, St George-Hyslop P, Pericak-Vance MA (2016) *SORL1* mutations in early- and late-onset Alzheimer disease. *Neurol Genet* **2**, e116.
- [54] Vardarajan BN, Zhang Y, Lee JH, Cheng R, Bohm C, Ghani M, Reitz C, Reyes-Dumeyer D, Shen Y, Rogaeva E, St George-Hyslop P, Mayeux R (2015) Coding mutations in *SORL1* and Alzheimer disease. *Ann Neurol* **77**, 215-227.
- [55] Young JE, Boulanger-Weill J, Williams DA, Woodruff G, Buen F, Revilla AC, Herrera C, Israel MA, Yuan SH, Edland SD, Goldstein LS (2015) Elucidating molecular phenotypes caused by the *SORL1* Alzheimer's disease genetic risk factor using human induced pluripotent stem cells. *Cell Stem Cell* **16**, 373-385.
- [56] Rohe M, Carlo A-S, Breyhan H, Sporbert A, Militz D, Schmidt V, Wozny C, Harmeyer A, Erdmann B, Bales KR, Wolf S, Kempermann G, Paul SM, Schmitz D, Bayer TA, Willnow TE, Andersen OM (2008) Sortilin-related Receptor with A-type Repeats (*SORLA*) affects the amyloid precursor protein-dependent stimulation of ERK signaling and adult neurogenesis. *J Biol Chem* **283**, 14826-14834.
- [57] Rohe M, Synowitz M, Glass R, Paul SM, Nykjaer A, Willnow TE (2009) Brain-derived neurotrophic factor reduces amyloidogenic processing through control of *SORLA* gene expression. *J Neurosci* **29**, 15472-15478.
- [58] Ng TKS, Ho CSH, Tam WWS, Kua EH, Ho RC-M (2019) Decreased serum brain-derived neurotrophic factor (BDNF) levels in patients with Alzheimer's disease (AD): a systematic review and meta-analysis. *Int J Mol Sci* **20**, E257.
- [59] Weinstein G, Beiser AS, Choi SH, Preis SR, Chen TC, Vorgas D, Au R, Pikula A, Wolf PA, DeStefano AL, Vasan RS, Seshadri S (2014) Serum brain-derived neurotrophic

- factor and the risk for dementia: the Framingham Heart Study. *JAMA Neurol* **71**, 55-61.
- [60] Asai M, Hattori C, Szabó B, Sasagawa N, Maruyama K, Tanuma S-i, Ishiura S (2003) Putative function of ADAM9, ADAM10, and ADAM17 as APP α -secretase. *Biochem Biophys Res Commun* **301**, 231-235.
- [61] Lammich S, Kojro E, Postina R, Gilbert S, Pfeiffer R, Jasionowski M, Haass C, Fahrenholz F (1999) Constitutive and regulated alpha-secretase cleavage of Alzheimer's amyloid precursor protein by a disintegrin metalloprotease. *Proc Natl Acad Sci U S A* **96**, 3922-3927.
- [62] Buxbaum JD, Liu KN, Luo Y, Slack JL, Stocking KL, Peschon JJ, Johnson RS, Castner BJ, Cerretti DP, Black RA (1998) Evidence that tumor necrosis factor alpha converting enzyme is involved in regulated alpha-secretase cleavage of the Alzheimer amyloid protein precursor. *J Biol Chem* **273**, 27765-27767.
- [63] Tanabe C, Hotoda N, Sasagawa N, Sehara-Fujisawa A, Maruyama K, Ishiura S (2007) ADAM19 is tightly associated with constitutive Alzheimer's disease APP α -secretase in A172 cells. *Biochem Biophys Res Commun* **352**, 111-117.
- [64] uhn P-H, Wang H, Dislich B, Colombo A, Zeitschel U, Ellwart JW, Kremmer E, Roßner S, Lichtenthaler SF (2010) ADAM10 is the physiologically relevant, constitutive α -secretase of the amyloid precursor protein in primary neurons. *EMBO J* **29**, 3020-3032.
- [65] Tolia A, De Strooper B (2009) Structure and function of γ -secretase. *Semin Cell Dev Biol* **20**, 211-218.
- [66] Haass C, Kaether C, Thinakaran G, Sisodia S (2012) Trafficking and proteolytic processing of APP. *Cold Spring Harb Perspect Med* **2**, a006270.
- [67] Sisodia SS (1992) Beta-amyloid precursor protein cleavage by a membrane-bound protease. *Proc Natl Acad Sci U S A* **89**, 6075-6079.
- [68] Lai A, Sisodia SS, Trowbridge IS (1995) Characterization of sorting signals in the β -amyloid precursor protein cytoplasmic domain. *J Biol Chem* **270**, 3565-3573.
- [69] Boll W, Rapoport I, Brunner C, Modis Y, Prehn S, Kirchhausen T (2002) The μ 2 subunit of the clathrin adaptor AP-2 binds to FDNPVY and YppO sorting signals at distinct sites. *Traffic* **3**, 590-600.
- [70] Vassar R, Bennett BD, Babu-Khan S, Kahn S, Mendiaz EA, Denis P, Teplow DB, Ross S, Amarante P, Loeloff R, Luo Y, Fisher S, Fuller J, Edenson S, Lile J, Jarosinski MA, Biere AL, Curran E, Burgess T, Louis JC, Collins F, Treanor J, Rogers G, Citron M (1999) Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* **286**, 735-741.
- [71] Frykman S, Hur J-Y, Frånberg J, Aoki M, Winblad B, Nahalkova J, Behbahani H, Tjernberg LO (2010) Synaptic and endosomal localization of active γ -secretase in rat brain. *PLoS One* **5**, e8948.
- [72] Chyung JH, Raper DM, Selkoe DJ (2005) Gamma-secretase exists on the plasma membrane as an intact complex that accepts substrates and effects intramembrane cleavage. *J Biol Chem* **280**, 4383-4892.
- [73] Choy RW-Y, Cheng Z, Schekman R (2012) Amyloid precursor protein (APP) traffics from the cell surface via endosomes for amyloid β ($A\beta$) production in the trans-Golgi network. *Proc Natl Acad Sci U S A* **109**, E2077-E2082.
- [74] Tan JZA, Gleeson PA (2019) The trans-Golgi network is a major site for α -secretase processing of amyloid precursor protein in primary neurons. *J Biol Chem* **294**, 1618-1631.
- [75] Pasternak SH, Bagshaw RD, Guiral M, Zhang S, Acklerley CA, Pak BJ, Callahan JW, Mahuran DJ (2003) Presenilin-1, nicastrin, amyloid precursor protein, and gamma-secretase activity are co-localized in the lysosomal membrane. *J Biol Chem* **278**, 26687-26694.
- [76] Sannerud R, Esselens C, Ejsmont P, Mattera R, Rochin L, Tharkeshwar Arun K, De Baets G, De Wever V, Habets R, Baert V, Vermeire W, Michiels C, Groot Arjan J, Wouters R, Dillen K, Vints K, Baatsen P, Munck S, Derua R, Waelkens E, Basi Guriqbal S, Mercken M, Vooijs M, Bollen M, Schymkowitz J, Rousseau F, Bonifacino Juan S, Van Niel G, De Strooper B, Annaert W (2016) Restricted location of PSEN2/ γ -secretase determines substrate specificity and generates an intracellular $A\beta$ pool. *Cell* **166**, 193-208.
- [77] Cook DG, Forman MS, Sung JC, Leight S, Kolson DL, Iwatsubo T, Lee VMY, Doms RW (1997) Alzheimer's $A\beta$ (1-42) is generated in the endoplasmic reticulum/intermediate compartment of NT2N cells. *Nat Med* **3**, 1021-1023.
- [78] Hartmann T, Bieger SC, Bruhl B, Tienari PJ, Ida N, Allsop D, Roberts GW, Masters CL, Dotti CG, Unsicker K, Beyreuther K (1997) Distinct sites of intracellular production for Alzheimer's disease A[beta]40/42 amyloid peptides. *Nat Med* **3**, 1016-1020.
- [79] Area-Gomez E, de Groof AJ, Boldogh I, Bird TD, Gibson GE, Koehler CM, Yu WH, Duff KE, Yaffe MP, Pon LA, Schon EA (2009) Presenilins are enriched in endoplasmic reticulum membranes associated with mitochondria. *Am J Pathol* **175**, 1810-1816.
- [80] Vance JE (2014) MAM (mitochondria-associated membranes) in mammalian cells: Lipids and beyond. *Biochim Biophys Acta* **1841**, 595-609.
- [81] Raturi A, Simmen T (2013) Where the endoplasmic reticulum and the mitochondrion tie the knot: The mitochondria-associated membrane (MAM). *Biochim Biophys Acta* **1833**, 213-224.
- [82] Del Prete D, Suski JM, Oulès B, Debayle D, Gay AS, Lacas-Gervais S, Bussiere R, Bauer C, Pinton P, Paterlini-Bréchet P, Wiecekowski MR, Checler F, Chami M (2016) Localization and processing of the amyloid- β protein precursor in mitochondria-associated membranes. *J Alzheimers Dis* **55**, 1549-1570.
- [83] Paillusson S, Stoica R, Gomez-Suaga P, Lau DHW, Mueller S, Miller T, Miller CCJ (2016) There's something wrong with my MAM; the ER-mitochondria axis and neurodegenerative diseases. *Trends Neurosci* **39**, 146-157.
- [84] Schon EA, Area-Gomez E (2010) Is Alzheimer's disease a disorder of mitochondria-associated membranes? *J Alzheimers Dis* **20** Suppl 2, S281-292.
- [85] Spoelgen R, von Arnim CA, Thomas AV, Peltan ID, Koker M, Deng A, Irizarry MC, Thomsen OM, Willnow TE, Hyman BT (2006) Interaction of the cytosolic domains of sorLA/LR11 with the amyloid precursor protein (APP) and beta-secretase beta-site APP-cleaving enzyme. *J Neurosci* **26**, 418-428.
- [86] Liu L, Ding L, Rovere M, Wolfe MS, Selkoe DJ (2019) A cellular complex of BACE1 and γ -secretase sequentially generates $A\beta$ from its full-length precursor. *J Cell Biol* **218**, 644-663.
- [87] Newman M, Wilson L, Verdile G, Lim A, Khan I, Mousavi Nik SH, Pursglove S, Chapman G, Martins RN, Lardelli M (2014) Differential, dominant activation and inhibition of Notch signalling and APP cleavage by trun-

- cations of PSEN1 in human disease. *Hum Mol Genet* **23**, 602-617.
- [88] Lim AHL (2015) *Analysis of the subcellular localization of proteins implicated in Alzheimer's disease* (thesis). Department of Genetics and Evolution, University of Adelaide, p. 235.
- [89] Herskowitz JH, Offe K, Deshpande A, Kahn RA, Levey AI, Lah JJ (2012) GGA1-mediated endocytic traffic of LR11/SorLA alters APP intracellular distribution and amyloid- β production. *Mol Biol Cell* **23**, 2645-2657.
- [90] Dodson SE, Andersen OM, Karmali V, Fritz JJ, Cheng D, Peng J, Levey AI, Willnow TE, Lah JJ (2008) Loss of LR11/SORLA enhances early pathology in a mouse model of amyloidosis: evidence for a proximal role in Alzheimer's disease. *J Neurosci* **28**, 12877-12886.
- [91] Mehmedbasic A, Christensen SK, Nilsson J, Rüetschi U, Gustafsen C, Poulsen ASA, Rasmussen RW, Fjorback AN, Larson G, Andersen OM (2015) SorLA complement-type repeat domains protect the amyloid precursor protein against processing. *J Biol Chem* **290**, 3359-3376.
- [92] Andersen OM, Schmidt V, Spoelgen R, Gliemann J, Behlke J, Galatis D, McKinstry WJ, Parker MW, Masters CL, Hyman BT, Cappai R, Willnow TE (2006) Molecular dissection of the interaction between amyloid precursor protein and its neuronal trafficking receptor SorLA/LR11. *Biochemistry* **45**, 2618-2628.
- [93] Spoelgen R, von Arnim CAF, Thomas AV, Peltan ID, Koker M, Deng A, Irizarry MC, Andersen OM, Willnow TE, Hyman BT (2006) Interaction of the cytosolic domains of sorLA/LR11 with the amyloid precursor protein (APP) and β -secretase β -site APP-cleaving enzyme. *J Neurosci* **26**, 418.
- [94] Schmidt V, Baum K, Lao A, Rateitschak K, Schmitz Y, Teichmann A, Wiesner B, Petersen CM, Nykjaer A, Wolf J, Wolkenhauer O, Willnow TE (2012) Quantitative modelling of amyloidogenic processing and its influence by SORLA in Alzheimer's disease. *EMBO J* **31**, 187-200.
- [95] Seaman MN (2012) The retromer complex - endosomal protein recycling and beyond. *J Cell Sci* **125**, 4693-4702.
- [96] Zhang H, Huang T, Hong Y, Yang W, Zhang X, Luo H, Xu H, Wang X (2018) The retromer complex and sorting nexins in neurodegenerative diseases. *Front Aging Neurosci* **10**, 79.
- [97] Vieira SI, Rebelo S, Esselmann H, Wiltfang J, Lah J, Lane R, Small SA, Gandy S, Da Cruz E Silva EF, Da Cruz E Silva OA (2010) Retrieval of the Alzheimer's amyloid precursor protein from the endosome to the TGN is S655 phosphorylation state-dependent and retromer-mediated. *Mol Neurodegener* **5**, 40.
- [98] Small SA, Kent K, Pierce A, Leung C, Kang MS, Okada H, Honig L, Vonsattel J-P, Kim T-W (2005) Model-guided microarray implicates the retromer complex in Alzheimer's disease. *Ann Neurol* **58**, 909-919.
- [99] Vardarajan BN, Bruesegem SY, Harbour ME, George-Hyslop PS, Seaman MNJ, Farrer LA (2012) Identification of Alzheimer disease-associated variants in genes that regulate retromer function. *Neurobiol Aging* **33**, 2231.e2215-2231.e2233.
- [100] Muhammad A, Flores I, Zhang H, Yu R, Staniszewski A, Planel E, Herman M, Ho L, Kreber R, Honig LS, Ganetzky B, Duff K, Arancio O, Small SA (2008) Retromer deficiency observed in Alzheimer's disease causes hippocampal dysfunction, neurodegeneration, and Abeta accumulation. *Proc Natl Acad Sci U S A* **105**, 7327-7332.
- [101] Wen L, Tang F-L, Hong Y, Luo S-W, Wang C-L, He W, Shen C, Jung J-U, Xiong F, Lee D-H, Zhang Q-G, Brann D, Kim T-W, Yan R, Mei L, Xiong W-C (2011) VPS35 haploinsufficiency increases Alzheimer's disease neuropathology. *J Cell Biol* **195**, 765-779.
- [102] Von Arnim CAF, Spoelgen R, Peltan ID, Deng M, Courchesne S, Koker M, Matsui T, Kowa H, Lichtenthaler SF, Irizarry MC, Hyman BT (2006) GGA1 acts as a spatial switch altering amyloid precursor protein trafficking and processing. *J Neurosci* **26**, 9913-9922.
- [103] Dumanis SB, Burgert T, Caglayan S, Fuchtbauer A, Fuchtbauer EM, Schmidt V, Willnow TE (2015) Distinct functions for anterograde and retrograde sorting of SORLA in amyloidogenic processes in the brain. *J Neurosci* **35**, 12703-12713.
- [104] Wan L, Molloy SS, Thomas L, Liu G, Xiang Y, Rybak SL, Thomas G (1998) PACS-1 defines a novel gene family of cytosolic sorting proteins required for trans-Golgi network localization. *Cell* **94**, 205-216.
- [105] Burgert T, Schmidt V, Caglayan S, Lin F, Fuchtbauer A, Fuchtbauer EM, Nykjaer A, Carlo AS, Willnow TE (2013) SORLA-dependent and -independent functions for PACS1 in control of amyloidogenic processes. *Mol Cell Biol* **33**, 4308-4320.
- [106] Huang TY, Zhao Y, Li X, Wang X, Tseng IC, Thompson R, Tu S, Willnow TE, Zhang YW, Xu H (2016) SNX27 and SORLA interact to reduce amyloidogenic subcellular distribution and processing of amyloid precursor protein. *J Neurosci* **36**, 7996-8011.
- [107] Wang X, Huang T, Zhao Y, Zheng Q, Robert, Bu G, Zhang Y-W, Hong W, Xu H (2014) Sorting nexin 27 regulates A β production through modulating γ -secretase activity. *Cell Rep* **9**, 1023-1033.
- [108] Caglayan S, Takagi-Niidome S, Liao F, Carlo A-S, Schmidt V, Burgert T, Kitago Y, Fuchtbauer E-M, Fuchtbauer A, Holtzman DM, Takagi J, Willnow TE (2014) Lysosomal sorting of amyloid- β by the SORLA receptor is impaired by a familial Alzheimer's disease mutation. *Sci Transl Med* **6**, 223ra220.
- [109] Kitago Y, Nagae M, Nakata Z, Yagi-Utsumi M, Takagi-Niidome S, Mihara E, Nogi T, Kato K, Takagi J (2015) Structural basis for amyloidogenic peptide recognition by sorLA. *Nat Struct Mol Biol* **22**, 199-206.
- [110] Foley P (2010) Lipids in Alzheimer's disease: A century-old story. *Biochim Biophys Acta* **1801**, 750-753.
- [111] Cutler RG, Kelly J, Storie K, Pedersen WA, Tammaru A, Hatanpaa K, Troncoso JC, Mattson MP (2004) Involvement of oxidative stress-induced abnormalities in ceramide and cholesterol metabolism in brain aging and Alzheimer's disease. *Proc Natl Acad Sci U S A* **101**, 2070-2075.
- [112] Valdez CM, Smith MA, Perry G, Phelix CF, Santamaria F (2010) Cholesterol homeostasis markers are localized to mouse hippocampal pyramidal and granule layers. *Hippocampus* **20**, 902-905.
- [113] Yajima R, Tokutake T, Koyama A, Kasuga K, Tezuka T, Nishizawa M, Ikeuchi T (2015) ApoE-isoform-dependent cellular uptake of amyloid-beta is mediated by lipoprotein receptor LR11/SorLA. *Biochem Biophys Res Commun* **456**, 482-488.
- [114] Frieden C, Garai K (2012) Structural differences between apoE3 and apoE4 may be useful in developing therapeutic agents for Alzheimer's disease. *Proc Natl Acad Sci U S A* **109**, 8913-8918.

- [115] Suri S, Heise V, Trachtenberg AJ, Mackay CE (2013) The forgotten APOE allele: a review of the evidence and suggested mechanisms for the protective effect of APOE varepsilon2. *Neurosci Biobehav Rev* **37**, 2878-2886.
- [116] Zollo A, Allen Z, Rasmussen HF, Iannuzzi F, Shi Y, Larsen A, Maier TJ, Matrone C (2017) Sortilin-related receptor expression in human neural stem cells derived from Alzheimer's disease patients carrying the APOE epsilon 4 allele. *Neural Plast* **2017**, 1892612.
- [117] de la Monte SM (2014) Type 3 diabetes is sporadic Alzheimer's disease: mini-review. *Eur Neuropsychopharmacol* **24**, 1954-1960.
- [118] Lee S, Tong M, Hang S, Deochand C, de la Monte S (2013) CSF and brain indices of insulin resistance, oxidative stress and neuro-inflammation in early versus late Alzheimer's disease. *J Alzheimers Dis Parkinsonism* **3**, 128.
- [119] Rivera EJ, Goldin A, Fulmer N, Tavares R, Wands JR, de la Monte SM (2005) Insulin and insulin-like growth factor expression and function deteriorate with progression of Alzheimer's disease: link to brain reductions in acetylcholine. *J Alzheimers Dis* **8**, 247-268.
- [120] Mosconi L (2005) Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease. FDG-PET studies in MCI and AD. *Eur J Nucl Med Mol Imaging* **32**, 486-510.
- [121] Schmidt V, Schulz N, Yan X, Schurmann A, Kempa S, Kern M, Blüher M, Poy MN, Olivecrona G, Willnow TE (2016) SORLA facilitates insulin receptor signaling in adipocytes and exacerbates obesity. *J Clin Invest* **126**, 2706-2720.
- [122] Lumsden AL, Rogers JT, Majd S, Newman M, Sutherland GT, Verdile G, Lardelli M (2018) Dysregulation of neuronal iron homeostasis as an alternative unifying effect of mutations causing familial Alzheimer's disease. *Front Neurosci* **12**, 533-533.
- [123] Du L, Zhao Z, Cui A, Zhu Y, Zhang L, Liu J, Shi S, Fu C, Han X, Gao W, Song T, Xie L, Wang L, Sun S, Guo R, Ma G (2018) Increased iron deposition on brain quantitative susceptibility mapping correlates with decreased cognitive function in Alzheimer's disease. *ACS Chem Neurosci* **9**, 1849-1857.
- [124] Lovell MA, Robertson JD, Teesdale WJ, Campbell JL, Markesbery WR (1998) Copper, iron and zinc in Alzheimer's disease senile plaques. *J Neurol Sci* **158**, 47-52.
- [125] Samudralwar DL, Diprete CC, Ni BF, Ehmann WD, Markesbery WR (1995) Elemental imbalances in the olfactory pathway in Alzheimer's disease. *J Neurol Sci* **130**, 139-145.
- [126] Yambire KF, Rostosky C, Watanabe T, Pacheu-Grau D, Torres-Odio S, Sanchez-Guerrero A, Senderovich O, Meyron-Holtz EG, Milosevic I, Frahm J, West AP, Raimundo N (2019) Impaired lysosomal acidification triggers iron deficiency and inflammation *in vivo*. *Elife* **8**, e51031.
- [127] Lee JH, McBrayer MK, Wolfe DM, Haslett LJ, Kumar A, Sato Y, Lie PP, Mohan P, Coffey EE, Kompella U, Mitchell CH, Lloyd-Evans E, Nixon RA (2015) Presenilin 1 maintains lysosomal Ca(2+) homeostasis via TRPML1 by regulating vATPase-mediated lysosome acidification. *Cell Rep* **12**, 1430-1444.
- [128] Jiang Y, Sato Y, Im E, Berg M, Bordi M, Darji S, Kumar A, Mohan PS, Bandyopadhyay U, Diaz A, Cuervo AM, Nixon RA (2019) Lysosomal dysfunction in Down syndrome is APP-dependent and mediated by APP-βCTF (C99). *J Neurosci* **39**, 5255-5268.
- [129] Duce JA, Tsatsanis A, Cater MA, James SA, Robb E, Wikke K, Leong SL, Perez K, Johanssen T, Greenough MA, Cho H-H, Galatis D, Moir RD, Masters CL, McLean C, Tanzi RE, Cappai R, Barnham KJ, Ciccotosto GD, Rogers JT, Bush AI (2010) Iron-export ferroxidase activity of β-amyloid precursor protein is inhibited by zinc in Alzheimer's disease. *Cell* **142**, 857-867.
- [130] Needham BE, Ciccotosto GD, Cappai R (2014) Combined deletions of amyloid precursor protein and amyloid precursor-like protein 2 reveal different effects on mouse brain metal homeostasis. *Metallomics* **6**, 598-603.
- [131] Waldvogel-Abramowski S, Waeber G, Gassner C, Buser A, Frey BM, Favrat B, Tissot J-D (2014) Physiology of iron metabolism. *Transfus Med Hemother* **41**, 213-221.
- [132] Gunshin H, Mackenzie B, Berger UV, Gunshin Y, Romero MF, Boron WF, Nussberger S, Gollan JL, Hediger MA (1997) Cloning and characterization of a mammalian proton-coupled metal-ion transporter. *Nature* **388**, 482-488.
- [133] Jenkitkasemwong S, Wang C-Y, Mackenzie B, Knutson MD (2012) Physiologic implications of metal-ion transport by ZIP14 and ZIP8. *Biomaterials* **25**, 643-655.
- [134] Richardson DR, Ponka P (1997) The molecular mechanisms of the metabolism and transport of iron in normal and neoplastic cells. *Biochim Biophys Acta* **1331**, 1-40.
- [135] Sargent PJ, Farnaud S, Evans RW (2005) Structure/function overview of proteins involved in iron storage and transport. *Curr Med Chem* **12**, 2683-2693.
- [136] Arosio P, Levi S (2010) Cytosolic and mitochondrial ferritins in the regulation of cellular iron homeostasis and oxidative damage. *Biochim Biophys Acta* **1800**, 783-792.
- [137] Wong BX, Tsatsanis A, Lim LQ, Adlard PA, Bush AI, Duce JA (2014) β-amyloid precursor protein does not possess ferroxidase activity but does stabilize the cell surface ferrous iron exporter ferroportin. *PLoS One* **9**, e114174.
- [138] Dlouhy AC, Bailey DK, Steimle BL, Parker HV, Kosman DJ (2019) Fluorescence resonance energy transfer links membrane ferroportin, hephaestin but not ferroportin, amyloid precursor protein complex with iron efflux. *J Biol Chem* **294**, 4202-4214.
- [139] Ji C, Steimle BL, Bailey DK, Kosman DJ (2018) The ferroxidase hephaestin but not amyloid precursor protein is required for ferroportin-supported iron efflux in primary hippocampal neurons. *Cell Mol Neurobiol* **38**, 941-954.
- [140] Lane DJR, Merlot AM, Huang MLH, Bae DH, Jansson PJ, Sahni S, Kalinowski DS, Richardson DR (2015) Cellular iron uptake, trafficking and metabolism: Key molecules and mechanisms and their roles in disease. *Biochim Biophys Acta* **1853**, 1130-1144.
- [141] Gao S, Hendrie HC, Hall KS, Hui S (1998) The relationships between age, sex, and the incidence of dementia and Alzheimer disease: a meta-analysis. *Arch Gen Psychiatry* **55**, 809-815.
- [142] Andersen K, Launer LJ, Dewey ME, Letenneur L, Ott A, Copeland JR, Dartigues JF, Kragh-Sorensen P, Baldereschi M, Brayne C, Lobo A, Martinez-Lage JM, Stijnen T, Hofman A (1999) Gender differences in the incidence of AD and vascular dementia: The EURODEM Studies. EURODEM Incidence Research Group. *Neurology* **53**, 1992-1997.
- [143] Yue X, Lu M, Lancaster T, Cao P, Honda SI, Staufenbiel M, Harada N, Zhong Z, Shen Y, Li R (2005) Brain estrogen deficiency accelerates Abeta plaque formation in an

- Alzheimer's disease animal model. *Proc Natl Acad Sci U S A* **102**, 19198-19203.
- [144] Gibson CL, Gray LJ, Murphy SP, Bath PM (2006) Estrogens and experimental ischemic stroke: a systematic review. *J Cereb Blood Flow Metab* **26**, 1103-1113.
- [145] Iturria-Medina Y, Sotero R, Toussaint P, Mateos-Pérez J, Evans A, Alzheimer's Disease Neuroimaging Initiative (2016) Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nat Commun* **7**, 11934.
- [146] Rivera-Rivera LA, Schubert T, Turski P, Johnson KM, Berman SE, Rowley HA, Carlsson CM, Johnson SC, Wieben O (2017) Changes in intracranial venous blood flow and pulsatility in Alzheimer's disease: A 4D flow MRI study. *J Cereb Blood Flow Metab* **37**, 2149-2158.
- [147] Marin R, Guerra B, Hernández-Jiménez JG, Kang XL, Fraser JD, López FJ, Alonso R (2003) Estradiol prevents amyloid- β peptide-induced cell death in a cholinergic cell line via modulation of a classical estrogen receptor. *Neuroscience* **121**, 917-926.
- [148] Nilsen J, Chen S, Irwin RW, Iwamoto S, Brinton RD (2006) Estrogen protects neuronal cells from amyloid beta-induced apoptosis via regulation of mitochondrial proteins and function. *BMC Neurosci* **7**, 74.
- [149] Ratnakumar A, Zimmerman SE, Jordan BA, Mar JC (2019) Estrogen activates Alzheimer's disease genes. *Alzheimers Dement (N Y)* **5**, 906-917.
- [150] Zhang S-t, Zuo C, Li W-n, Fu X-q, Xing S, Zhang X-p (2016) Identification of key genes associated with the effect of estrogen on ovarian cancer using microarray analysis. *Arch Gynecol Obstet* **293**, 421-427.
- [151] Shen M, Cao J, Shi H (2018) Effects of estrogen and estrogen receptors on transcriptomes of HepG2 cells: a preliminary study using RNA sequencing. *Int J Endocrinol* **2018**, 5789127.
- [152] Cellini E, Tedde A, Bagnoli S, Pradella S, Piacentini S, Sorbi S, Nacmias B (2009) Implication of sex and *SORL1* variants in Italian patients with Alzheimer disease. *Arch Neurol* **66**, 1260-1266.
- [153] Casingal CR, Daroy MLG, Mapua CA, Florendo DJA, Natividad FF, Dominguez JC (2014) Polymorphisms in the Sortilin-Related Receptor 1 gene are associated with cognitive impairment in Filipinos. *Asian J Neurosci* **2014**, 1-7.

**Chapter 3: No observed effect on brain vasculature of
Alzheimer's disease-related mutations in the zebrafish
presenilin 1 gene**

Statement of Authorship

Title of Paper	No observed effect on brain vasculature of Alzheimer's disease-related mutations in the zebrafish presenilin 1 gene
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Published in Molecular brain https://doi.org/10.1186/s13041-021-00734-5

Principal Author

Name of Principal Author (Candidate)	Karissa Barthelson		
Contribution to the Paper	Sample preparation and imaging, image analysis. Drafting the manuscript		
Overall percentage (%)	95%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	14/2/21

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Morgan Newman		
Contribution to the Paper	Generated the zebrafish used in this study. Supervision		
Signature		Date	15/02/2021

Name of Co-Author	Cameron Nowell		
Contribution to the Paper	Assistance with image analysis, editing of manuscript		
Signature		Date	16/2/21

Please cut and paste additional co-author panels here as required.


Name of Co-Author	Michael Lardelli		
Contribution to the Paper	Supervision over entire project, editing of manuscript		
Signature		Date	19/02/2021

MICRO REPORT

Open Access



No observed effect on brain vasculature of Alzheimer's disease-related mutations in the zebrafish presenilin 1 gene

Karissa Barthelson^{1*} , Morgan Newman¹, Cameron J. Nowell² and Michael Lardelli¹

Abstract

Previously, we found that brains of adult zebrafish heterozygous for Alzheimer's disease-related mutations in their presenilin 1 gene (*psen1*, orthologous to human *PSEN1*) show greater basal expression levels of hypoxia responsive genes relative to their wild type siblings under normoxia, suggesting hypoxic stress. In this study, we investigated whether this might be due to changes in brain vasculature. We generated and compared 3D reconstructions of GFP-labelled blood vessels of the zebrafish forebrain from heterozygous *psen1* mutant zebrafish and their wild type siblings. We observed no statistically significant differences in vessel density, surface area, overall mean diameter, overall straightness, or total vessel length normalised to the volume of the telencephalon. Our findings do not support that changes in vascular morphology are responsible for the increased basal expression of hypoxia responsive genes in *psen1* heterozygous mutant brains.

Keywords: Zebrafish, Vasculature, Confocal laser scanning microscopy, 3D reconstruction

Introduction

The dominant hypothesis of Alzheimer's disease (AD) pathogenesis is the amyloid cascade hypothesis (ACH) [1], which postulates the amyloid β peptide ($A\beta$) as initiating a pathological process resulting in neurodegeneration and dementia (reviewed in [2]). An alternative to the ACH is the vascular hypothesis [3], asserting that age-related cerebral vascular abnormalities induce AD pathologies by limiting nutrient and oxygen delivery to produce hypoxic stress, a neural energy crisis and, consequently, neurodegeneration. Significant evidence supports the vascular hypothesis of AD (reviewed in [4]).

Rare, inherited forms of AD are caused by dominant mutations in a small number of genes (early-onset familial AD, EOfAD). Most EOfAD cases are due to heterozygous mutations in the gene presenilin 1 (*PSEN1*) that

obey a "reading-frame preservation rule" [5]. Mutations allowing production of a transcript(s) with an altered coding sequence but, nevertheless, utilising the original stop codon cause EOfAD while mutant alleles coding only for truncated proteins do not. We previously generated knock-in models in zebrafish with each of these types of mutant *psen1* allele: K97Gfs, a frameshift mutation encoding a truncated protein similar to the human PS2V isoform that is increased in sporadic, late onset AD [6], and Q96_K97del: an EOfAD-like, reading-frame-preserving deletion of two codons [7].

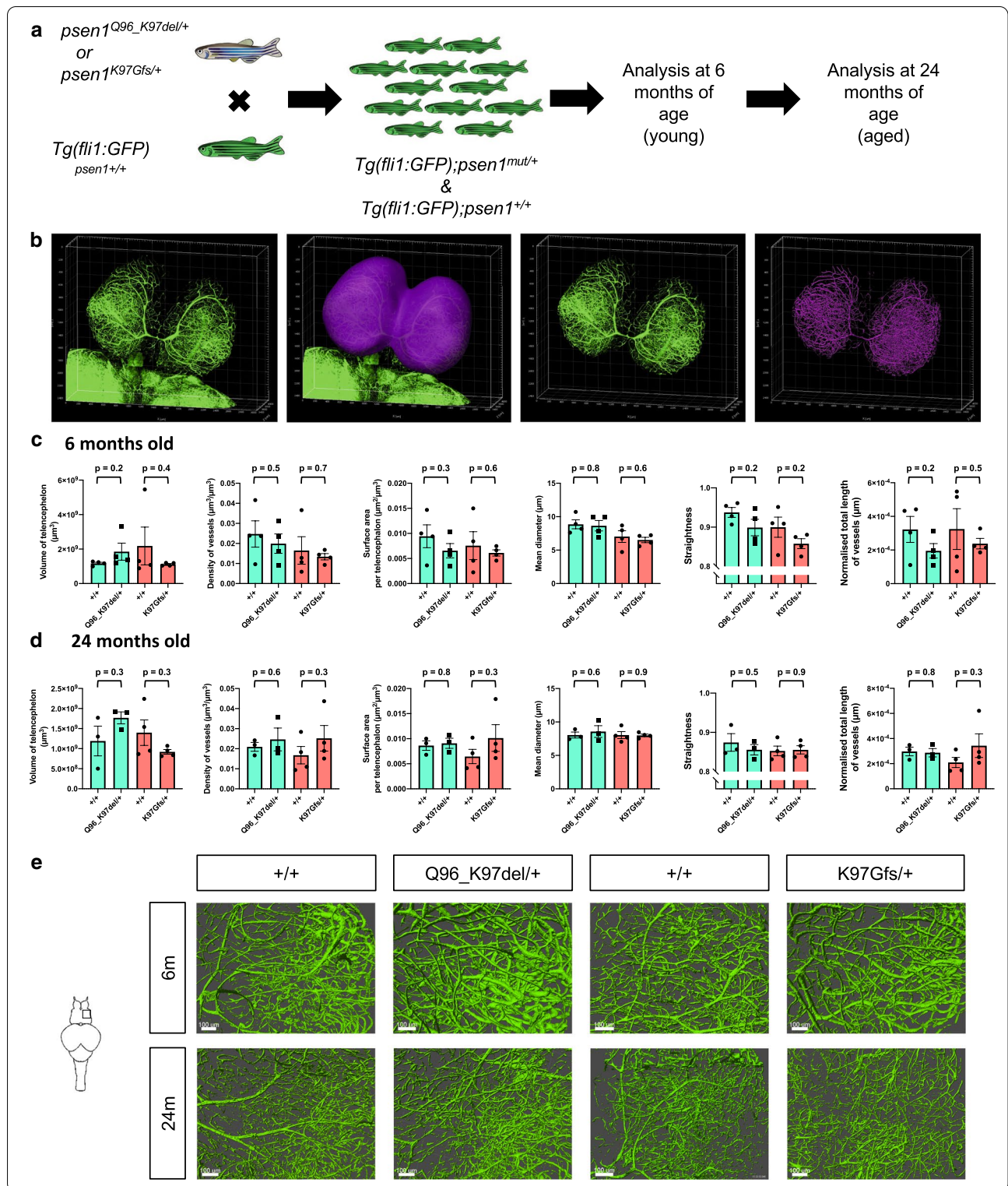
We recently observed in normoxic adult zebrafish brains that heterozygosity for either of the above two mutations causes increased basal expression levels of hypoxia responsive genes (HRGs, genes with expression regulated by a master regulator of the transcriptional response to hypoxia: hypoxia-inducible factor 1 (HIF1)). This implied that the heterozygous *psen1* mutant fish brains were already under some form of hypoxic stress [8], possibly due to changes in vasculature, as have been observed in

*Correspondence: karissa.barthelson@adelaide.edu.au

¹ Alzheimer's Disease Genetics Laboratory, School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, SA 5005, Australia
Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.



transgenic mice expressing human *PSEN1* EoFAD mutation-bearing transgenes in neurons [9]. Therefore, we examined the effects on forebrain vasculature with age of

heterozygosity for the K97Gfs and Q96_K97del mutations of *psen1* by exploiting the *fli1:GFP* transgene that labels zebrafish endothelial cells [10].

(See figure on previous page.)

Fig. 1 No statistically significant changes to brain vascular network parameters due to heterozygosity for the Q96_K97del or K97Gfs mutations of *psen1*. **a** Experimental design flow diagram. Genome-edited *psen1* heterozygous mutant fish were pair-mated with transgenic zebrafish expressing green fluorescent protein (GFP) under the control of the *fli1* promoter (*fli1::GFP* transgene). GFP-fluorescent larvae were selected to give a family of transgenic siblings either wild type or heterozygous for a *psen1* mutation. Analysis of the brain vascular network was performed at 6 and 24 months of age. **b** 3D image analysis pipeline. The telencephalon was manually segmented from the optic tectum using contour lines to generate a masked surface channel containing only GFP signals from the telencephalon. Then, an additional surface was generated over the vessels to remove background fluorescence. A masked surface channel was generated from this surface as input for the filament trace algorithm. **c** Measured values from the surface and filament trace algorithms for the 6 month old zebrafish and **d** the 24 month old female zebrafish for (left to right) the volume of the telencephalon, the density of *fli1::GFP* positive vessels per telencephalon, the surface area of vessels normalised to the volume of the telencephalon, the overall mean diameter of vessels, the overall straightness of the vessels, and the total length of the vessels normalised to the volume of the telencephalon. Data are presented as the mean \pm standard deviation. Colours of the bars represent the two families of fish used in this analysis. *P*-values were determined by Student's *t*-test assuming unequal variance. **e** Representative images of a 200 μ m section of the right hemisphere of the telencephalon from fish of each age and genotype. Scale bars indicate 100 μ m. Vessels appeared morphologically similar in each age and genotype

Methods

Single zebrafish heterozygous for either *psen1* mutation were mated with single fish bearing the *fli1::GFP* [10] transgene. GFP-fluorescent progeny were selected to form families of siblings either wild type or heterozygous for the *psen1* mutant alleles (Fig. 1a). We used $n=4$ brains of each sibling genotype at 6 months (young adult) and 24 months (aged) of age for tissue clearing using the PACT method [11]. Briefly, PACT involves infusing and crosslinking the brain with an acrylamide-based hydrogel. Then, light scattering lipids are passively removed by incubating the brain with a detergent, allowing light to penetrate deep into the tissue [11, 12]. We imaged the telencephalons (thought to be the region loosely equivalent of the prefrontal cortex in humans) using an Olympus FV3000 confocal microscope, and performed 3D image analysis using Imaris v9.1 (Bitplane) (Fig. 1b). For a detailed description of methods, see Additional File 1.

Results and conclusion

No statistically significant differences between sibling genotypes at each age were observed for any of the measured parameters (see Fig. 1). This does not support that the increased basal levels of HRGs observed previously in our zebrafish *psen1* mutants are due to vascular changes. However, subtle changes to vasculature due to *psen1* genotype may be too small to detect using this method and further experimentation using a larger number of biological replicates may increase statistical power to detect changes to these measured parameters. Alternatively, other factors such as altered γ -secretase activity [13] and/or cellular ferrous iron levels [14] may influence HIF1- α activity to affect basal HRG expression.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13041-021-00734-5>.

Additional file 1. Detailed description of sample preparation, imaging and 3D image analysis.

Additional file 2. Quantified values used to produce the graphs in Fig. 1.

Abbreviations

ACH: Amyloid cascade hypothesis; AD: Alzheimer's disease; A β : Amyloid beta; EOFAD: Early-onset familial Alzheimer's disease; HIF1: Hypoxia-inducible factor 1; HIF1- α : Hypoxia-inducible factor 1, alpha subunit; HRG: Hypoxia response gene; PSEN1: Presenilin 1.

Acknowledgements

Confocal imaging data were generated at Adelaide Microscopy (The University of Adelaide).

Authors' contributions

KB performed the experiments, MN and ML conceived the project, CN provided advice and access for 3D image analysis. All authors read and contributed to the final manuscript.

Funding

Grants GNT1061006 and GNT1126422 from the National Health and Medical Research Council of Australia (NHMRC) funded this work, and supported MN and ML. KB is supported by an Australian Government Research Training Program Scholarship. ML is an academic employee of the University of Adelaide. Funding bodies did not play a role in the design of the study, data collection, analysis, interpretation or in writing the manuscript.

Availability of data and materials

The quantified values used to produce the graphs in Fig. 1 can be found in Additional File 2. Raw microscopy images from the current study are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

Work with zebrafish was conducted under the auspices of the University of Adelaide Animal Ethics Committee (permit numbers: S-2017-073 and S-2017-089) and Institutional Biosafety Committee (permit number 15037).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Alzheimer's Disease Genetics Laboratory, School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, SA 5005, Australia. ² Drug Discovery Biology, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC 3058, Australia.

Received: 19 November 2020 Accepted: 13 January 2021

Published online: 25 January 2021

References

- Hardy JA, Higgins GA. Alzheimer's disease: the amyloid cascade hypothesis. *Science*. 1992;256(5054):184–5.
- Morris GP, Clark IA, Vissel B. Questions concerning the role of amyloid- β in the definition, aetiology and diagnosis of Alzheimer's disease. *Acta Neuropathol*. 2018;136(5):663–89.
- de la Torre JC, Mussivan T. Can disturbed brain microcirculation cause Alzheimer's disease? *Neurol Res*. 1993;15(3):146–53.
- Rius-Pérez S, Tormos AM, Pérez S, Taléns-Visconti R. Vascular pathology: cause or effect in Alzheimer disease? *Neurologia (English Edition)*. 2017. <https://doi.org/10.1016/j.nrleng.2015.07.008>.
- Jayne T, Newman M, Verdile G, Sutherland G, Munch G, Musgrave I, et al. Evidence for and against a pathogenic role of reduced gamma-secretase activity in familial Alzheimer's disease. *JAD*. 2016. <https://doi.org/10.3233/JAD-151186>.
- Hin N, Newman M, Kaslin J, Douek AM, Lumsden A, Nik SHM, et al. Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2. *PLoS ONE*. 2020;15(1):e0227258.
- Newman M, Hin N, Pederson S, Lardelli M. Brain transcriptome analysis of a familial Alzheimer's disease-like mutation in the zebrafish presenilin 1 gene implies effects on energy production. *Mol Brain*. 2019. <https://doi.org/10.1186/s13041-019-0467-y>.
- Newman M, Nik HM, Sutherland GT, Hin N, Kim WS, Halliday GM, et al. Accelerated loss of hypoxia response in zebrafish with familial Alzheimer's disease-like mutation of presenilin 1. *Hum Mol Genet*. 2020;29(14):2379–94.
- Gama Sosa MA, Gasperi RD, Rocher AB, Wang AC, Janssen WG, Flores T, et al. Age-related vascular pathology in transgenic mice expressing presenilin 1-associated familial Alzheimer's disease mutations. *Am J Pathol*. 2010;176(1):353–68.
- Lawson ND, Weinstein BM. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Developmental Biology*. 2002;248(2):307–18.
- Yang B, Treweek Jennifer B, Kulkarni Rajan P, Deverman Benjamin E, Chen C-K, Lubeck E, et al. Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell*. 2014;158(4):945–58.
- Chung K, Wallace J, Kim S-Y, Kalyanasundaram S, Andalman AS, Davidson TJ, et al. Structural and molecular interrogation of intact biological systems. *Nature*. 2013;497:332.
- Le Moan N, Houslay DM, Christian F, Houslay MD, Akassoglou K. Oxygen-dependent cleavage of the p75 neurotrophin receptor triggers stabilization of HIF-1 α . *Mol Cell*. 2011;44(3):476–90.
- Yambire KF, Rostosky C, Watanabe T, Pacheu-Grau D, Torres-Odio S, Sanchez-Guerrero A, et al. Impaired lysosomal acidification triggers iron deficiency and inflammation in vivo. *Elife*. 2019. <https://doi.org/10.7554/eLife.51031>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Chapter 4: Pilot behavioural characterisation of zebrafish early-onset familial Alzheimer's disease models

Statement of Authorship

Title of Paper	Pilot behavioural characterisation of zebrafish early-onset familial Alzheimer's disease models
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input checked="" type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	The work presented in this chapter will ultimately form a manuscript which will serve as a more complete characterisation of the working memory of the EOfAD-like zebrafish generated by the Lardelli group.

Principal Author

Name of Principal Author (Candidate)	Karissa Barthelson		
Contribution to the Paper	Re-analysed the raw data of Jiayu Ruan (Masters thesis). Generated the raw data of the sorl1 mutant fish. Analysed the raw data for the sorl1 mutant fish. Wrote the manuscript.		
Overall percentage (%)	100%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	5/3/21

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author			
Contribution to the Paper			
Signature		Date	

Name of Co-Author			
Contribution to the Paper			
Signature		Date	

Please cut and paste additional co-author panels here as required.

Introduction

Impairment of memory is a defining feature of all dementias. Indeed, impairment of working memory is an early pathological change observed in Alzheimer's disease (AD) [1-3]. In humans, working memory is generally assessed on the basis of language. However, for animal models, this is not feasible and testing of working memory is heavily reliant on assessment of spatial working memory. Spatial working memory requires an intact working memory, and construction of a cognitive spatial "map" of an individual's surroundings in working memory and recalling these to complete a task.

Spatial working memory has been assessed in murine models of AD [4-9]. However, it is worthy of note that the majority of these models do not closely reflect the genetic state of AD (i.e. they often express multiple mutations causing early-onset-familial AD (EOfAD) in the same animal under novel promoters). Despite the fact that these animals show pathologies reminiscent of human AD pathologies (e.g. neuritic plaques), they likely do not accurately represent the molecular state of the human disease [10]. Our knock-in genetic models of EOfAD in zebrafish (e.g. in [11-15]) avoid this possibly confounding assumption that phenotypes arising from transgenic/overexpression of EOfAD-mutant genes are simply more extreme versions of heterozygote phenotypes. We believe our approach delivers better insight into the early disease mechanisms which eventually progress to AD.

Ruan (2019) [16], previously assessed whether heterozygosity for the Q96_K97del mutation of presenilin 1 (*psen1*) resulted in impairment of spatial working memory with age using the free movement pattern (FMP) Y-maze task [17]. The FMP Y-maze task is a recently developed method for assessment of spatial working memory in zebrafish (and also mice, flies and, interestingly, humans) [18]. It involves the

assessment of movement patterns of zebrafish freely exploring a Y-maze in a 1 hour time period. A series of frequencies of “tetragrams” (four consecutive arm entries ranging from LLLL to RRRR [17, 19, 20]) are calculated, and differences in frequencies of the alternation tetragrams (LRLR and RLRL, the dominant search strategy used in vertebrates [18]) are indicative of changes to spatial working memory. Ruan (2019) [16] found that heterozygosity for the Q96_K97del mutation of *psen1* impaired spatial working memory at 12 months of age, but not at any other age.

In the original analysis [16], an improper statistical test was performed to quantify whether the *psen1*^{Q96_K97del/+} genotype had a significant effect on alternation frequency. Therefore, in this Chapter, the raw data of Ruan (2019) [16] was re-analysed with a more appropriate statistical test. Similar levels of statistical significance was found. Replication of the impairment of spatial working memory at 12 months of age was attempted, but not reproduced. Additionally, an investigation was performed to explore whether heterozygosity for the EOfAD-like mutations W1818* or V1482Afs of sortilin-related receptor 1 (*sorl1*) results in an impairment of spatial working memory (the generation of these mutant lines of zebrafish are detailed in **Chapters 5** and **6**). However, no statistically significant differences were found at 12 months of age. Since EOfAD mutations in human *SORL1* are associated with a later age of onset in humans compared to EOfAD mutations in *PSEN1* (reviewed in [21]), differences may not be observable until older ages.

Methods

Zebrafish husbandry and animal ethics

Generation of the EOfAD-like mutant lines of zebrafish are described previously [12, 13, 22]. Fish were maintained in a recirculating water system on a 14 hour light/8 hour dark cycle, fed dry food in the morning and live artemia in the afternoon. All zebrafish work was conducted under the auspices of the Animal Ethics Committee and the Institutional Biosafety Committee of the University of Adelaide.

Free movement pattern (FMP) Y-maze task

The FMP Y-maze task was performed using the Zantiks [AD] fully automated behavioural testing environment (Cambridge, UK) on families of heterozygous EOfAD-like mutants and their non-mutant siblings at various ages. For *psen1* mutant fish, Ruan (2019) performed the analysis at three ages (6, 12 and 24 months of age). Here, for the *sor11* mutant fish, the analysis was performed at 12 months of age. Sample sizes can be found in **Figure 1**.

The FMP Y-maze task is described in [17]. Briefly, fish were isolated two at a time for 30 minutes before being placed alone in the Y-mazes for one hour swimmingly freely. Movements of the fish were tracked by the Zantiks system and the raw data was generated as a spreadsheet containing a series of time points which identify the times each fish entered and exited each arm of the mazes. FMP Y-maze tests were always performed between 11am and 4pm to avoid any differences due to circadian rhythm, and data collection took up to 11 days to complete. After data was collected for each fish, fish were genotyped by allele-specific polymerase reactions (PCRs) on genomic DNA isolated from a small fin clip of each fish (this allowed us to be blinded

to genotype until after data collection) as described below. A schematic of the workflow of the experiment can be found in **Figure 1**.

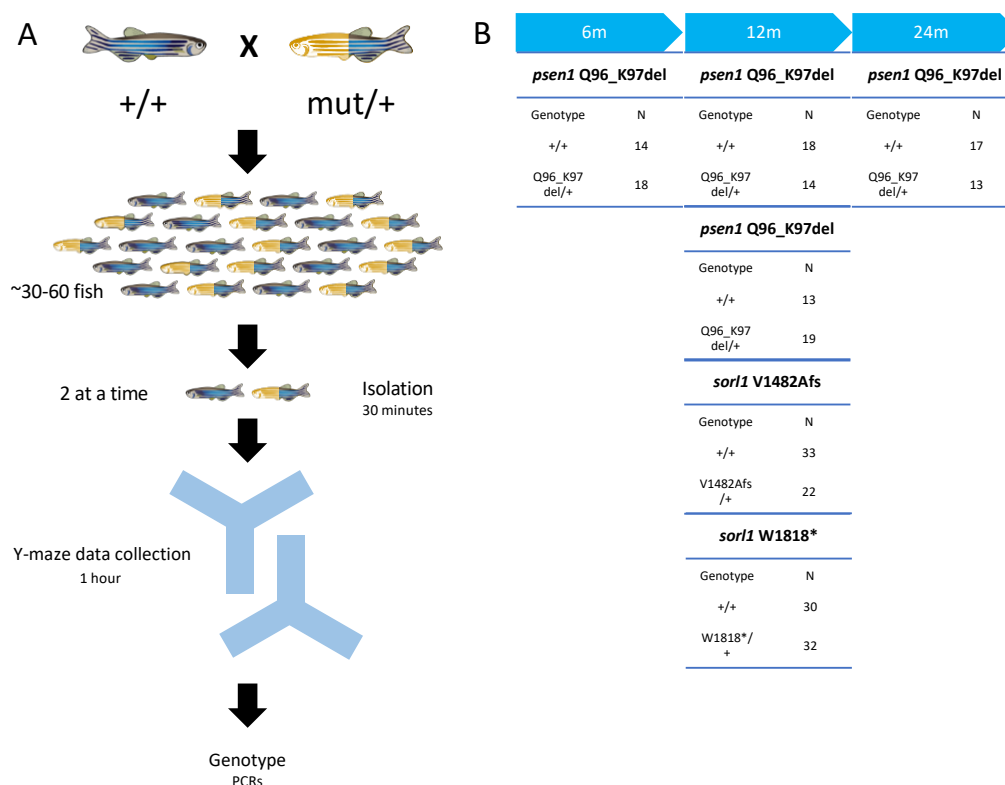


Figure 1: A) Schematic of experimental workflow. We generated families of heterozygous mutants and their non-mutant siblings by pair mating wild type fish (+/+) with fish with genotypes *psen1*^{Q96_K97del/+}, *sor1*^{V1482Afs/+} or *sor1*^{W1818*/+}. The resulting family of fish were allowed to develop to either 6, 12 or 24 months of age (Note that the families of fish analysed had different parents). Each fish were placed in the Y maze alone for 1 hour and their movements were recorded. After data collection, fish were genotyped by PCRs. **B)** Table showing the sample sizes for each family of fish we analysed (including both male and female fish).

PCR genotyping

After data collection, fish were anaesthetised in tricaine solution [23], and a small piece of the tail was cut off using a sterile blade. Genomic DNA (gDNA) was extracted from the tail clip by incubating in a solution containing recombinant

Proteinase K (Roche) diluted to ~2 mg/mL in 1 x tris-EDTA (TE) at 55 °C for 3 hours. Proteinase K was inactivated by incubating at 95 °C for 5 minutes, then the gDNA preps were centrifuged for 3 minutes at maximum speed to pellet cellular debris. The supernatants containing the gDNA were transferred to fresh Eppendorf tubes for subsequent PCRs.

Primers specific to the wild-type and mutant sequences at the *sor11* and *psen1* loci are described in **Table 1**. PCRs were performed using GoTaq DNA polymerase (Promega) following the manufacturer's protocol.

Table 1: PCR primers used for genotyping in this study

Primer name	Sequence (5' – 3')
psen1 Q96_K97del mutation specific F	TGTCAGCTTCTACACAGACGGA
psen1 Q96 wild type specific F	TCTGTCAGCTTCTACACACAGAAGG
psen1 Q96 common R	CCATCCCTAAACTGCTCCTACTC
sor11 V1482Afs mutation specific F	AGATCAGTTTCTGTGTGCCTCC
sor11 V1482 wild type specific F	GTGTGTGAAGCCTCCAGCAT
sor11 V1482 common R	GCGAAGATGCACAAGGGAA
sor11 W1818* mutation specific F	GTGGCGGTGTGATGATGG
sor11 W1818 wild type specific F	TGGCGGTGTGGGCTCAC
sor11 W1818 common R	GCAGGACAAAATAAAAGTGTATGTG

Statistical analyses

The raw Y-maze spreadsheets were batch-processed using an *R* [24] (www.r-project.org) script available from GitHub (https://github.com/thejamesclay/ZANTIKS_YMaze_Analysis_Script, [25]). The output of the batch analysis gives a series of frequencies of 16 tetragrams (four consecutive arm entries, consisting of all possible combinations from LLLL to RRRR) over six time intervals of 10 minutes per fish, and also a series of values of time each fish spent in each arm of the maze over the six 10 minute time intervals. All statistical analyses for this study were performed in *R*.

To determine whether there was a change in the alternation (LRLR and RLRL tetragrams) frequency due to *psen1* or *sorl1* genotype, we fitted the tetragram frequency data from each dataset to generalised linear mixed effect models using the package *glmmTMB* [26] specifying a beta-binomial distribution with logit link function. The beta-binomial variance function accounts for extra binomial dispersion by estimating an overdispersion parameter that is associated with the varying sample sizes among the grouped observations. We specified the fixed effects to be *psen1* or *sorl1* genotype (+/+ or EOfAD-like/+), time intervals of the hour spent in the maze (six 10 minute intervals), an interaction effect of genotype with time interval and the time of day that data recording began. Random intercept effects, which account for variability among the levels of each factor, were specified for day of data collection, fish identifier (to account for the same fish having a data point in each time interval) and an interaction effect between the day and time of data collection. Tests of the fixed effects in the model used Type II Wald χ^2 tests, using the *Anova* function from the package *car* [27]. Planned contrasts were used to determine whether there were differences between group means and a reference, or control level, and adjustments for multiple testing used the Dunnett method [28]. Data was visualised using *ggplot2* [29]. Fixed effects were considered to be significant if they had a p-value < 0.05.

Next, a test was performed to determine whether the genotype of the fish had an effect on the total number of arm entries each fish performed as a measure of any locomotor deficits. The total number of arm entries each fish performed was calculated and fitted to a generalised linear model with a Poisson distribution. However, on inspection of the model, it was found to be over-dispersed. Therefore, the data was fitted to a generalised linear model with a negative binomial distribution and log link function using the function *glm.nb* of the package *MASS* [30]. The effect of genotype was considered to be significant if the coefficient had a p-value < 0.05.

To determine whether the genotype of the fish had a significant effect on the average time fish spent in each zone of the Y-maze, the average time each fish spent in each zone of the Y-maze over the six 10 minute time intervals was calculated. These data were then fitted to a linear model. The response variable was specified as the log of the average time in zone and fixed interaction effect of genotype by zone. The effect of genotype was considered to be significant if the genotype or any of the genotype by zone coefficients had a p-value < 0.05.

Finally, an investigation was performed to determine whether fish displayed behavioural lateralisation. The total number of left turns against the total number of right turns per fish was calculated for the hour it spent in the Y-maze. A fish was considered to show a lateral bias if it performed left or right turns more than 60% of the time. To determine whether having a left or right turn bias has an effect on alternation frequency, the alternation tetragram frequency data was fitted to a generalised linear mixed effect model specifying a beta-binomial distribution with logit link function as described above. However, a bias fixed effect term was added for each fish with levels of “Neither”, “Left” or “Right”. Type II Wald χ^2 tests on the fixed effects were performed to determine their significance as described above.

Results

Zebrafish show spontaneous alternation behaviour throughout their life

Figures 2 and **3** show the total frequencies of fish performing the 16 possible tetragrams during the FMP Y-maze task. A clear increase is observed in frequency of fish performing the alternation tetragrams (LRLR and RLRL) relative to all other tetragrams at all ages and in all genotypes. This increase in frequencies is most

evident in the *sor11* mutant families, likely due to the larger sample sizes in these families ($N_{sor11} \approx 50-60$, $N_{psen1} \approx 30$).

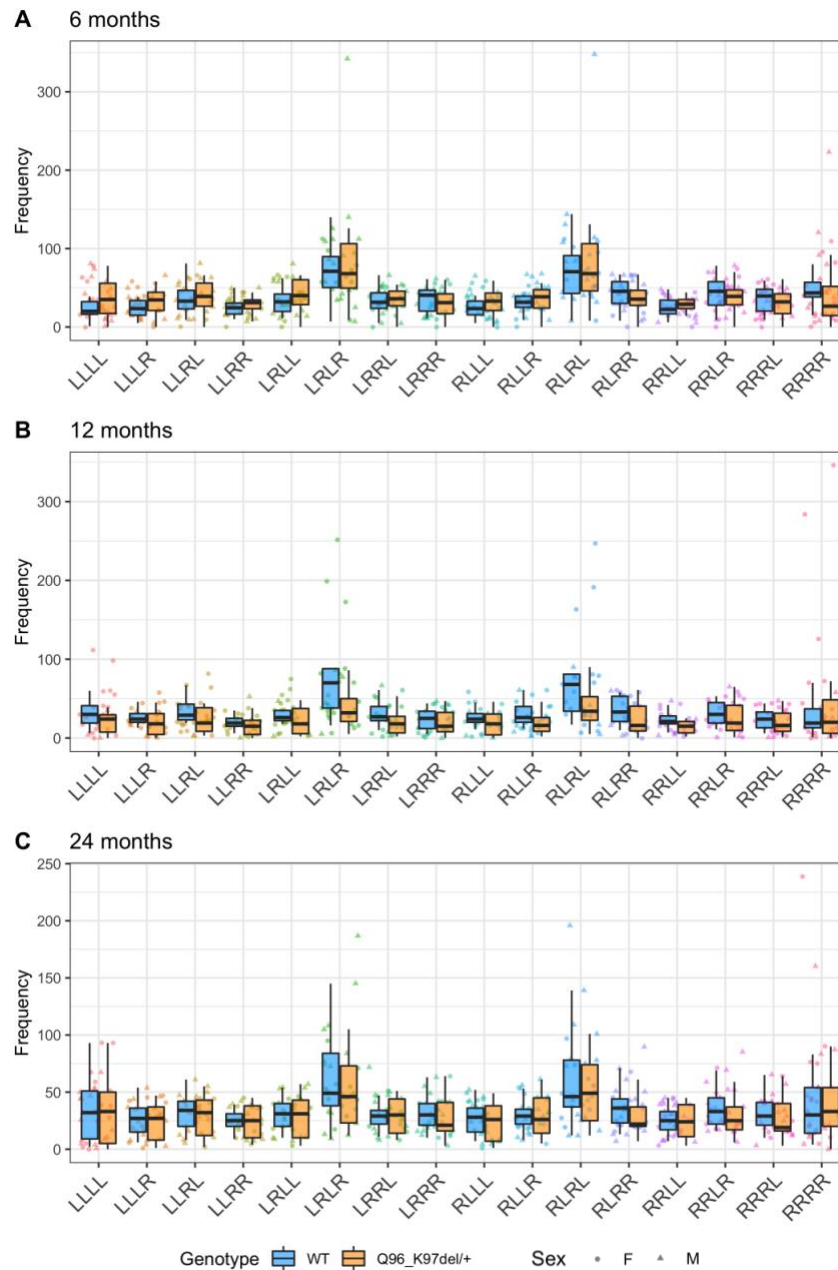


Figure 2: Total tetragram frequencies of zebrafish *psen1* mutants during the FMP Y-maze task. A, B and C show the frequencies of the 16 tetragrams at 6, 12 and 24 months of age respectively. Points are colour coded per tetragram and are shaped for gender of the fish. Boxplots are also indicated showing the summary statistics for each tetragram and are colour coded by genotype.

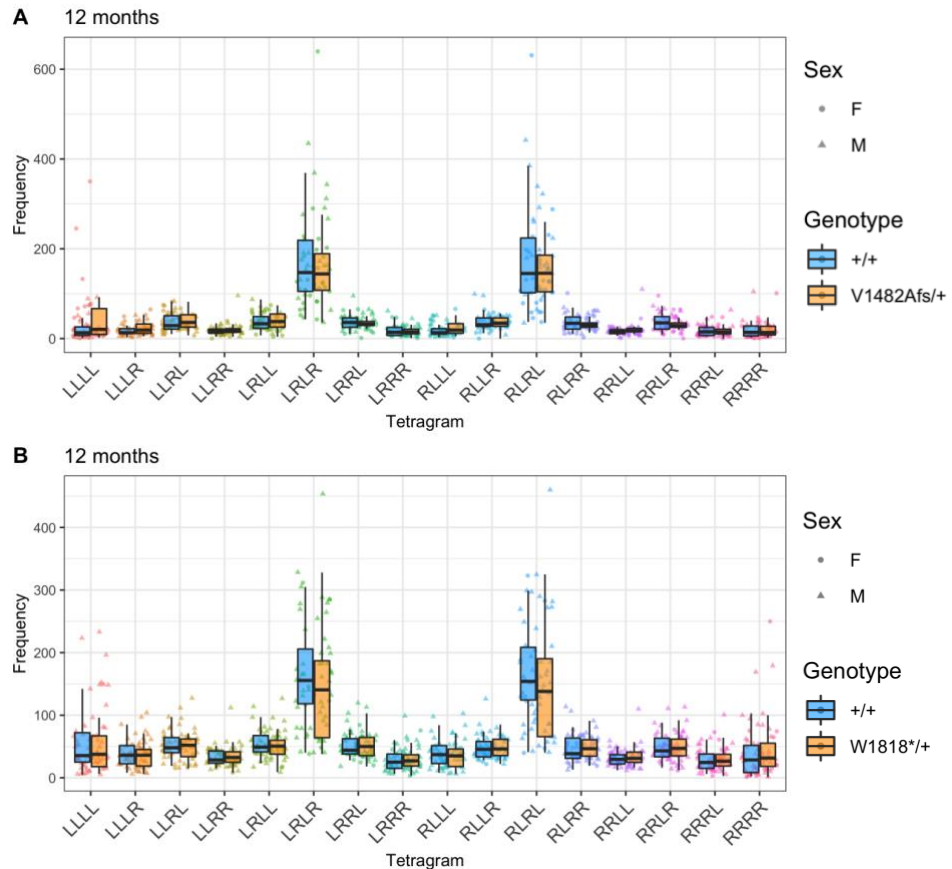


Figure 3: Total tetragram frequencies of 12 month old zebrafish *sorl1* mutants during the FMP Y-maze task. Points are colour coded per tetragram and are shaped for gender of the fish. Boxplots are also indicated showing the summary statistics for each tetragram and are colour coded by genotype.

An EOfAD-like mutation in *psen1* can reduce alternation frequency at 12 months of age.

Does alternation frequency alter due to heterozygosity for the EOfAD-like mutation Q96_K97del in *psen1* at 6, 12 and 24 months of age? To address this question, the movement patterns of the fish were collected and summarised into series of frequencies of tetragrams. **Figure 4** shows the predicted probability of *psen1*^{+/+} and *psen1*^{Q96_K97del/+} fish performing an alternation tetragram during a the FMP Y-maze task at 6, 12 and 24 months of age from generalised linear mixed effect models.

The probability of performing an alternation appears to increase marginally with aging from 6 to 12 months, then decrease from 12 to 24 months for *psen1*^{+/+} fish. At 6 months of age, the *psen1*^{Q96_K97del/+} fish appear to perform alternations slightly more often than their *psen1*^{+/+} siblings, although this difference was not significant ($\chi^2 = 0.27$, $df = 1$, $p = 0.6$). At 12 months of age, the probability of fish performing an alternation is decreased in the 12 month old *psen1*^{Q96_K97del/+} fish relative to their *psen1*^{+/+} siblings ($\chi^2 = 4.1$, $df = 1$, $p = 0.04$). At 24 months of age, the *psen1*^{Q96_K97del/+} fish and their *psen1*^{+/+} siblings appeared to have similar probability of performing alternations ($\chi^2 = 0.12$, $df = 1$, $p = 0.7$).

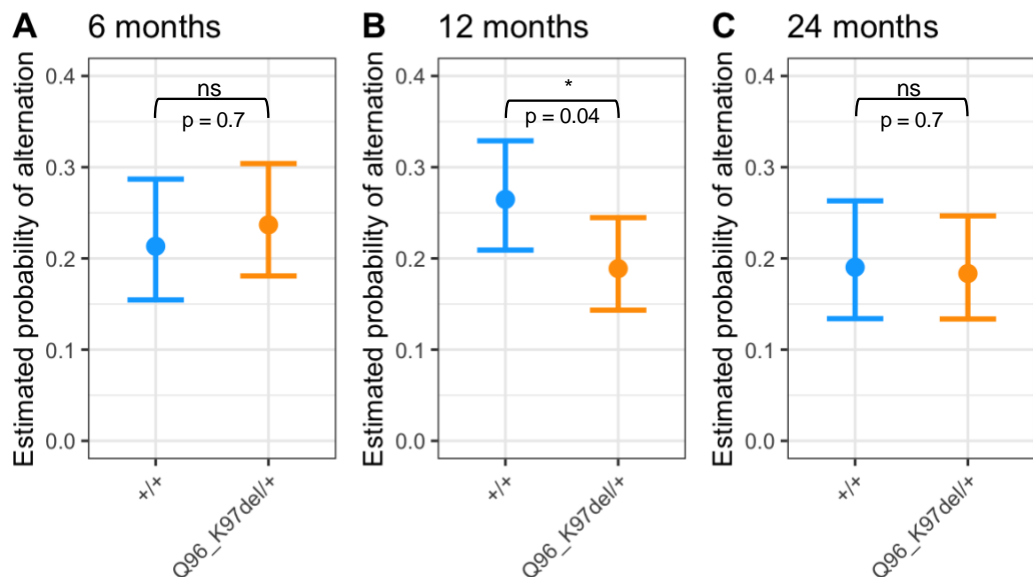


Figure 4: Probability of zebrafish *psen1* mutants performing an alternation tetragram overall in the FMP Y-maze task. Points indicate the estimated mean probability of fish performing an alternation tetragram and the 95% confidence interval from generalised linear mixed effect models. **A** depicts *psen1*^{+/+} zebrafish and their *psen1*^{Q96_K97del/+} siblings at 6 months of age. **B** depicts *psen1*^{+/+} zebrafish and their *psen1*^{Q96_K97del/+} siblings at 12 months of age. **C** depicts *psen1*^{+/+} zebrafish and their *psen1*^{Q96_K97del/+} siblings at 24 months of age. The effect of genotype was considered significant if it had a p value < 0.05 from Type II Wald χ^2 tests on the generalised linear mixed effect models.

Replication of this result was attempted in 12 month old *psen1*^{Q96_K97del/+} fish.

However, no significant difference was observed between the *psen1*^{Q96_K97del/+} fish relative to their *psen1*^{+/+} siblings in the second family ($p = 0.8$, **Figure 5**).

In summary, heterozygosity for the Q96_K97del mutation in *psen1* possibly impairs spatial working memory at 1 year of age, but not at 6 or 24 months of age.

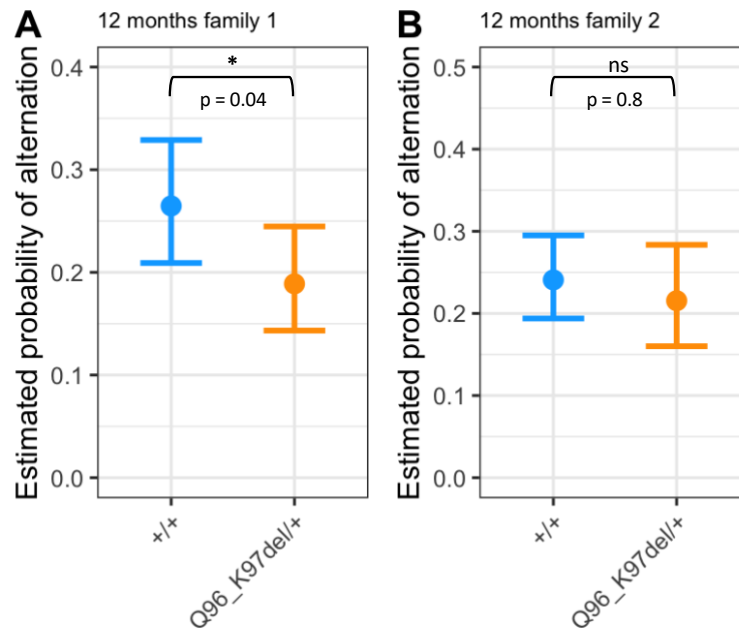


Figure 5: Replication of reduced alternation due to heterozygosity for the Q96_K97del mutation in *psen1* at 12 months of age. **A** and **B** show the estimated mean probability and the 95% confidence interval of 2 separate families of *psen1*^{Q96_K97del/+} fish and their *psen1*^{+/+} siblings performing an alternation tetragram during the FMP Y-maze task at 12 months of age. The two families each contained 32 fish and data was collected 1 month apart. Type II Wald χ^2 tests of fixed effects of generalised linear mixed effect models found that family 1 had a significant effect of *psen1* genotype ($p = 0.04$), but not in family 2 ($p = 0.8$).

EOfAD-like mutations in *sor11* do not appear to affect spatial working memory at 12 months of age.

Tentative statistical evidence was observed for EOfAD-like mutations in *psen1* affecting alternation frequency at 12 months of age. Does heterozygosity for EOfAD-like mutations in another gene, *sor11*, also have an effect at this age?

To address this question, the probability of heterozygous EOfAD-like mutant fish (*sor11*^{V1482Afs/+} or *sor11*^{W1818*/+}) and their *sor11*^{+/+} siblings performing an alternation tetragram in the hour spent in the Y-maze was calculated. No statistical evidence was found for the genotype effect alone significantly affected the probability of performing an alternation tetragram (*V1482Afs*: $\chi^2 = 0.20$, *df* = 1, *p* = 0.65. *W1818**: $\chi^2 = 0.56$, *df* = 1, *p* = 0.45) (**Figure 6**). Consequently, the hour that the fish spent in the maze was divided into six 10 minute time intervals and the data was re-analysed as described in the Methods section. Alternations appeared, overall, to increase with time spent in the maze for both families of fish. Generally, the highest number of alternations occurred during the 20-30 minutes interval. The probability of any fish performing an alternation was, overall, significantly different across the time intervals in the *V1482Afs* family ($\chi^2 = 42$, *df* = 5, *p* = 5.6e-08) and the *W1818** family ($\chi^2 = 35$, *df* = 5, *p* = 1.3e-06). It appears that the *sor11* EOfAD-like mutant fish broadly perform less alternations in the first 30 minutes in the Y-maze. However, I did not find any statistical evidence that the *V1482Afs* mutation or the *W1818** mutation affect the probability of performing an alternation tetragram across the time intervals (*V1482Afs*: $\chi^2 = 6.3$, *df* = 5, *p* = 0.28. *W1818**: $\chi^2 = 8.7$, *df* = 5, *p* = 0.12) (**Figure 7**).

Together, these results indicate that EOfAD-like mutations in *sor11* likely do not affect spatial working memory at 1 year of age.

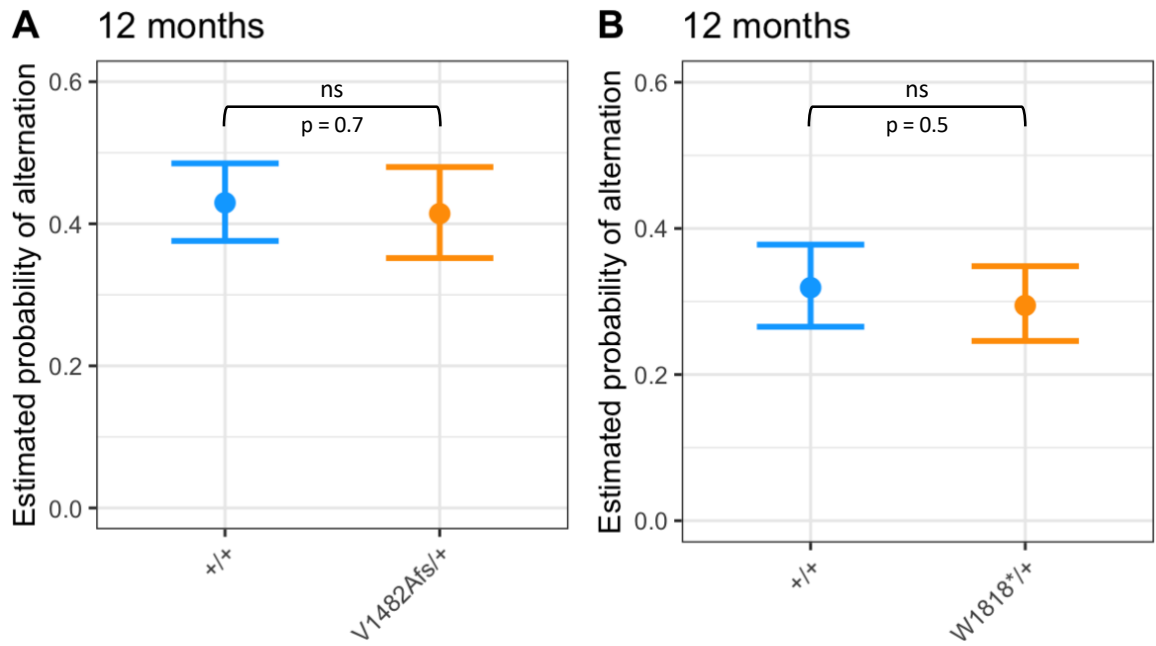


Figure 6: Probability of zebrafish *sorl1* mutants performing an alternation tetragram overall during the FMP Y-maze task at 12 months of age. Points indicate the estimated mean probability of **A)** *sorl1*^{V1482Afs/+} or **B)** *sorl1*^{W1818*/+} zebrafish, and their *sorl1*^{+/+} siblings performing tetragram alternation and the 95% confidence interval from generalised linear mixed effect models at 12 months of age.

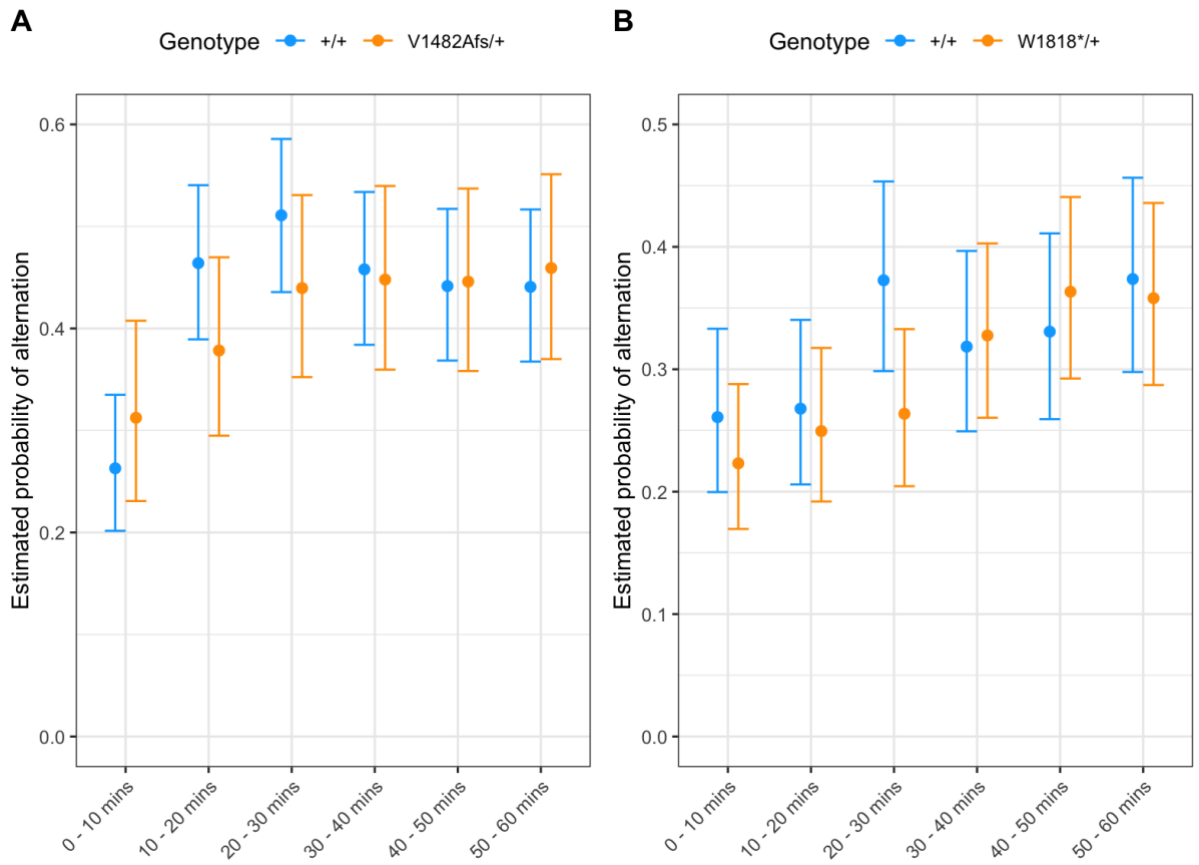


Figure 7: Probability of zebrafish *sorl1* mutants performing an alternation tetragram throughout the FMP Y-maze task. Points indicate the estimated mean probability of fish performing an alternation tetragram and the 95% confidence interval for **A**, the *sorl1*^{V1482Afs/+} mutants and **B**, the *sorl1*^{W1818*/+} mutants relative their respective *sorl1*^{+/+} siblings at 12 months of age.

Changes to alternation frequency are likely not due to locomotor deficits

Alteration to locomotor activity could be a confounding effect with genotype to alternation frequency. If the mutant fish had some form of locomotor defect, they could be performing less alternation tetragrams not due to impairment of spatial working memory. To account for this, I tested whether the genotype of the fish had a significant effect on 1) the total number of turns the fish performed in the hour spent in the maze or 2) the average time spent in each zone of the maze. No statistical evidence was found to support that the mutant fish performed more or less total numbers of turns in the maze ($p > 0.1$), or spent on average more or less time in each arm of the maze ($p > 0.1$). Therefore, it is likely that the mutant fish do not have any confounding locomotor impediments (**Figures 8 and 9**).

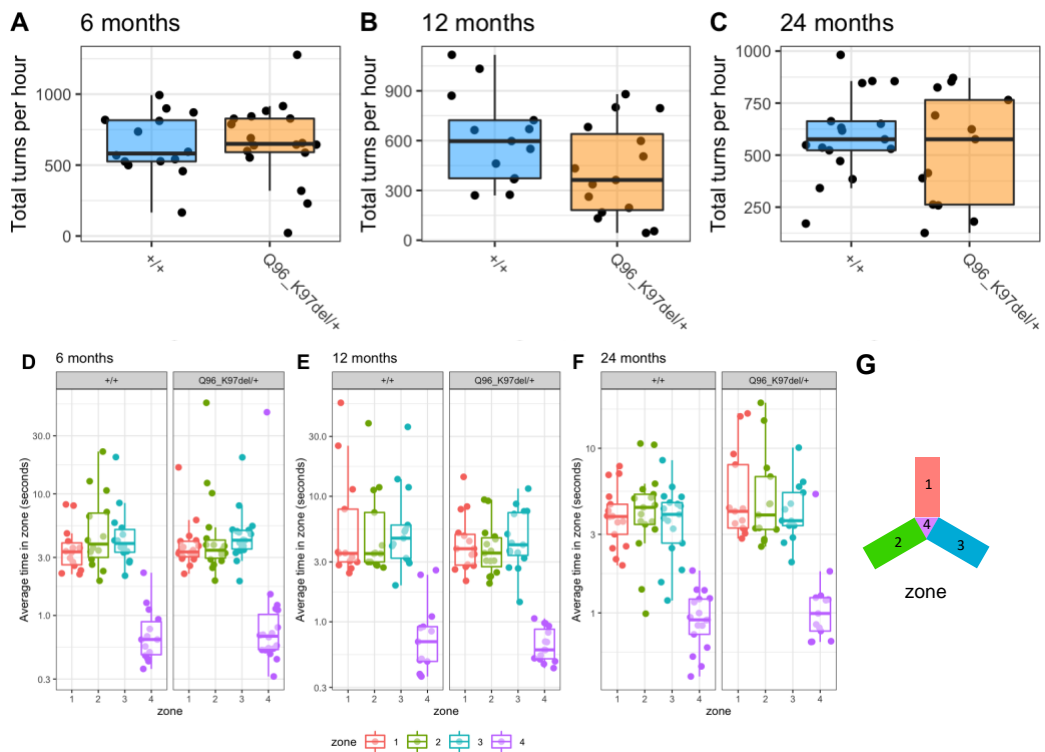


Figure 8: Heterozygosity for the Q96_K97del mutation in *psen1* does not appear to give any locomotor deficits **A, B** and **C** show the total number of turns performed by zebrafish in *psen1* mutant families at 6, 12 and 24 months of age respectively during the FMP Y-maze. We fitted this data to generalised linear models and found no statistical evidence that the Q96_K97del mutation in *psen1* affects the total number of turns performed by fish during the FMP Y-maze task at 6 months ($p = 0.867$), 12 months ($p = 0.107$) or 24 months of age ($p = 0.440$). **D, E** and **F** show the average time each fish spent in each of the zones depicted in **G**, during a 1 hour period in the Y-maze. If fish spent longer on average in each arm of the maze, this could suggest some locomotor deficits. We fitted these datasets to linear models and did not find any statistical evidence that heterozygosity for the Q96_K97del mutation in *psen1* has a significant effect on the average time spent in each arm of the maze at 6 months ($p > 0.634$), 12 months ($p > 0.314$) or 24 months of age ($p > 0.131$). Together, these results indicate that our *psen1* EOfAD-like mutant fish likely do not have any locomotor deficits. Data are presented as boxplots showing summary statistics as well as the raw data points. For the plots in D, E and F, the average time in zone is plotted on the log scale.

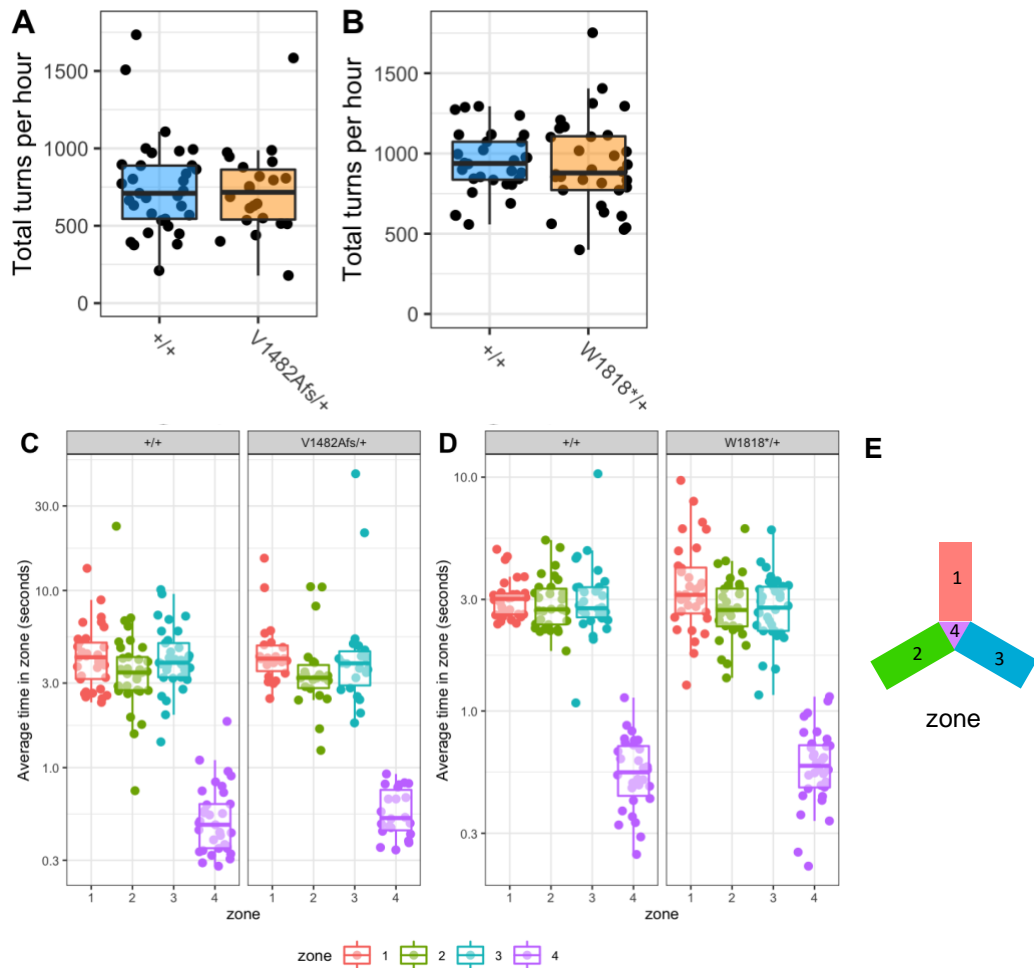


Figure 9: EOofAD-like mutations in *sorl1* do not appear to give any locomotor deficits. **A** and **B** show the total number of turns performed by zebrafish in *sorl1* mutant families during the FMP Y-maze task. We fitted this data to a generalised linear model and found that the V1482Afs mutation ($p = 0.736$) or the W1818* mutation ($p = 0.812$) do not have a significant effect on the total number of turns performed by zebrafish during the FMP Y-maze task. **C** and **D** show the average time each fish spent in each of the zones depicted in **E**. We fitted this data to a linear model and did not find any statistical evidence that the V1482Afs mutation ($p > 0.651$) or the W1818* mutation ($p > 0.127$) affect the average amount of time fish spend in each zone of the maze. Together, these results indicate that our *sorl1* EOofAD-like mutant fish likely do not have any locomotor deficits. Data are presented as boxplots showing summary statistics as well as the raw data points. For plots **C** and **D**, the average time in zone is plotted on the log scale.

A left or right turn bias during the FMP Y-maze task is not commonly observed

Zebrafish (AB strain) have been shown to display some behavioural lateralization in the FMP Y-maze at 3 months of age (i.e., a left or right turn bias) [25]. If individual fish had a preference for left or right turns, this could affect the number of alternations it performs without the use of spatial working memory. I investigated whether our fish display a behavioural laterality and found variable proportions of fish within families show a bias for left or right turns (**Figure 10** and **11**). I observed that fish at 12 months of age in the *psen1* Q96_K97del family who had a turn bias performed significantly less alterations ($p = 4.26e-6$). Similarly, fish at 12 months of age in the *sor11* W1818* family also performed significantly less alternations ($p = 0.00231$). However, including a turn bias fixed effect in the model only marginally affected the significance of the genotype effect, suggesting that overall, containing a L or R turn bias did not significantly effect alternation frequency.

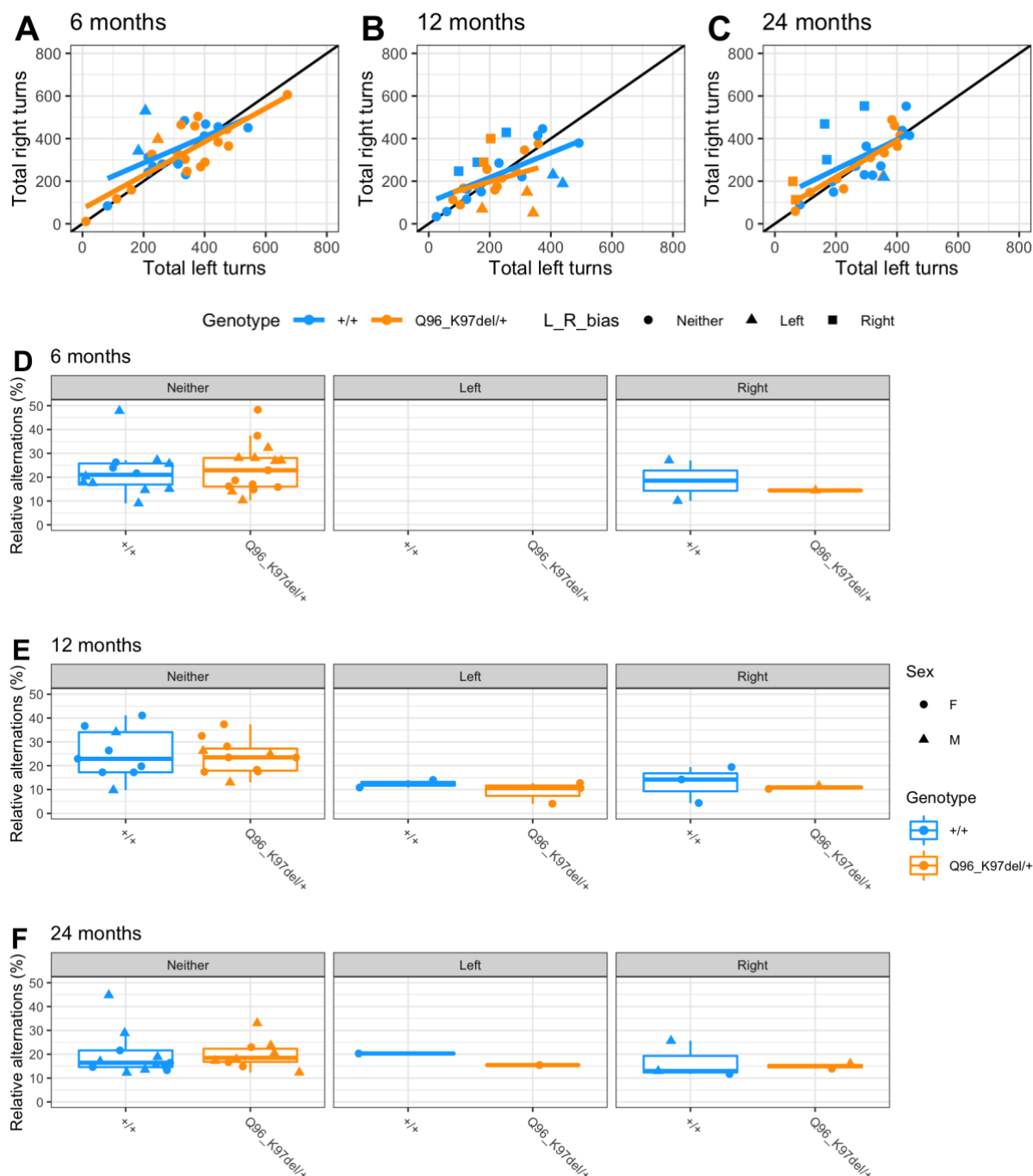


Figure 10: Behavioural laterality of zebrafish $psen1^{Q96_K97del/+}$ mutants in the FMP Y-maze task with aging. A-C show the total left turns against the total right turns performed by $psen1^{Q96_K97del/+}$ zebrafish and their $psen1^{+/+}$ siblings during the FMP- Y maze task at 6 months of age, 12 months of age and 24 months of age. Points are coloured by genotype and shaped by whether they show a bias towards left or right turns more than 60% of the time. Linear-model fitted trendlines are also shown and coloured by genotype. The black line indicates when left turns = right turns. D-F show the relative proportion of alternations (%) performed by $psen1^{Q96_K97del/+}$ fish and their $psen1^{+/+}$ siblings at each age, grouped by whether they had a bias for left turns, right turns or no bias. For the 6 month old family in D, only 3

out of 32 fish showed a bias for right turns and this is insufficient to formally test the effect of a turn bias on alternation frequency. For the 12 month old family in **E**, 10 out of 32 fish showed a bias towards a left or right turns. We added whether fish had left, right or no turn bias as a fixed effect to a generalized linear mixed effect model (with genotype by time interval of the hour in the maze, and also the time of day data was collected as fixed effects, and day of data collection, fish id and an interaction effect of day with time of data collection as random effects) and found that having a left or right turn bias has a significant effect on the number of alternations fish perform ($\chi^2 = 24.7$, $df = 2$, $p = 4.36e-6$). The significance of the genotype effect (which was found to be significant before including the turn bias as a fixed effect) was reduced ($\chi^2 = 3.32$, $df = 1$, $p = 0.068$). We repeated this for the 24 month old family. However, the effect of having a left or right turn bias was not significant ($\chi^2 = 2.7$, $df = 2$, $p = 0.25$).

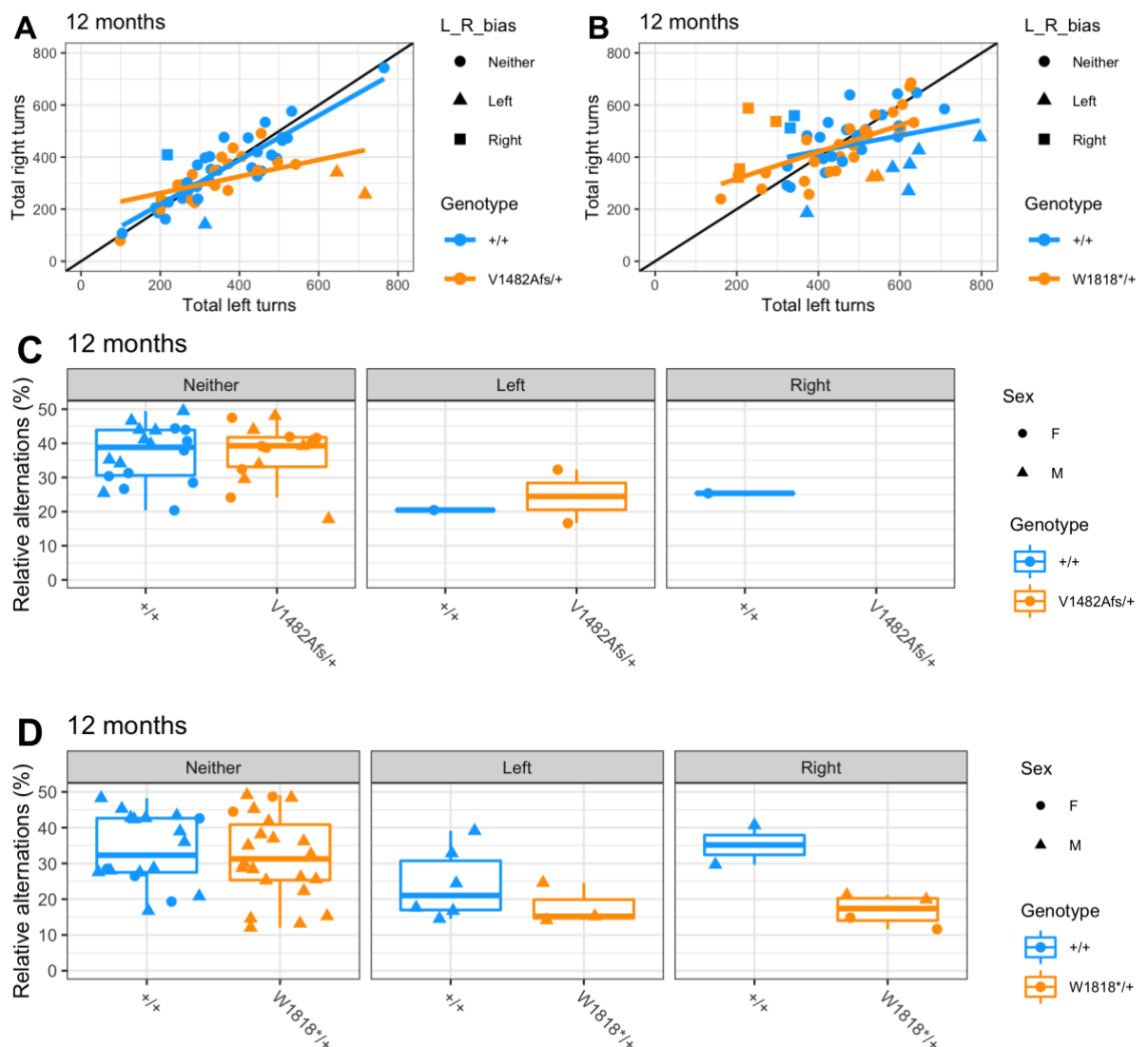


Figure 11: Behavioural laterality of zebrafish *sor11* EOfAD-like mutants in the FMP Y-maze task at 12 months of age. The plots above show the total left turns against the total right turns performed by zebrafish during the FMP Y-maze task in the **A**, V1482Afs family and **B**, the W1818* family. Points are coloured by genotype and shaped by whether they show a bias towards left or right turns more than 60% of the time. Linear-model fitted trendlines are also shown and coloured by genotype. The black line indicates when L = R. **C** and **D** show the relative number of alternations (%) fish performed for each *sor11* genotype and whether they had a bias for left turns, right turns or no bias. Boxplots are indicated showing summary statistics. For the V1482Afs family in **C**, only 4 out of a 55 fish showed a bias for left or right turns and this is insufficient to formally test the effect of a turn bias on alternation frequency. It appears these four fish overall perform a lower number of alternations. For the W1818* family in **D**, 15 out of 62 fish showed a bias towards a left or right bias. We added whether fish had left, right or no turn bias as a fixed effect to a generalized linear mixed effect model (with genotype by time interval of the hour in the maze, and also the time of day data was collected as fixed effects, and day of data collection, fish id and an interaction effect of day with time of data collection as random effects) and found that having a left or right turn bias has a significant effect on the number of alternations fish perform ($\chi^2 = 12.1$, $df = 2$, $p = 0.00231$), but did not change the significance of the genotype by time interval ($\chi^2 = 8.7$, $df = 5$, $p = 0.12$).

Discussion

In this study, the FMP Y-maze was performed to test whether spatial working memory is altered due to heterozygosity for EOfAD-like mutations in *psen1* and *sor11* in zebrafish of various ages. Behavioural tests which measure spontaneous alternation are commonly used to determine whether there are changes to spatial working memory in rodents [6, 19, 31-33]. They have also been shown to be applicable to *Drosophila* [34]. A decrease in spontaneous alternation is thought to indicate impairment in spatial working memory.

Spontaneous alternation behaviour is the tendency of animals to alternate their (non-reinforced) choices. It is thought to reflect the innate curiosity of animals to recall

where they have previously been and to explore novel environments to search for resources such as food or a mate or to escape [35]. Spontaneous alternation behaviour has been observed in both larval [36] and adult [17, 25, 37] zebrafish. However, it is not as extensively characterised in zebrafish as it is in rodents.

We used a tetragram configuration to measure the spontaneous alternation frequency, which involved recording movements of the fish, then calculating the frequencies of each tetragram the fish performed. With a random movement strategy, we would expect that fish would perform each tetragram relatively equally. However, a clear increase in the alternation tetragrams (LRLR and RLRL) is observed (**Figures 2 and 3**). This is in line with the results of Cleal et al. [17] who observed this phenomenon in 3 month old zebrafish, and Gross et al. [19] in rodents.

The tetragram configuration is slightly different from the traditional Y-maze data analysis methods. Spontaneous alternation in a Y-maze can be assessed by tracking *arm entries* of an animal [33, 34, 38]. Alternation in this type of analysis method is referred to if the animal entered an arm of the maze it didn't most recently visit (i.e. alternating it's choices of arm entries). This method based around the spatial location of the arms. Conversely, the tetragram configuration is based around the choice the fish makes when it reaches the center of the maze (a left or right turn). Since zebrafish perform the alternation tetragram more often than all other possible tetragrams, it must be able to recall which way it turned previously to be able to perform an alternation tetragram. Indeed, Cleal et al. [20] showed that the relative frequency of zebrafish performing alternation tetragrams is dependent on a functioning working memory in zebrafish by inhibiting long-term potentiation or the dopaminergic system, which are known processes required for memory formation.

An “accelerated aging” pattern is observed in the spontaneous alternation frequency of *psen1*^{Q96_K97del/+} fish. The young adult *psen1*^{Q96_K97del/+} fish were predicted to have a probability of performing an alternation more similar to that of 12 month old *psen1*^{+/+} fish. Then the *psen1*^{Q96_K97del/+} fish at 12 months of age have a predicted probability of performing an alternation more similar to the aged *psen1*^{+/+} fish (**Figure 4**). A similar pattern of “accelerated aging” is observed in the gene expression response to acute hypoxia in *psen1*^{Q96_K97del/+} fish [39]. At 6 months of age, *psen1*^{Q96_K97del/+} fish express hypoxia response genes when exposed to acute hypoxia to a more similar level to aged *psen1*^{+/+} fish.

Ruan (2019) [16] collected FMP Y-maze data for two families of *psen1*^{Q96_K97del/+} and their *psen1*^{+/+} siblings at 12 months of age. However, the re-analysis in this Chapter found that *psen1*^{Q96_K97del/+} fish only showed a significantly decreased probability of performing an alternation in one of the families. The *psen1*^{+/+} fish in the affected family appeared to have a higher probability of performing an alternation than the *psen1*^{+/+} fish in the family that showed no difference. Potentially, either the two families of fish were not the same “age” in molecular terms, or behaved differently. Both these possibilities could be due to cryptic environmental differences during housing of the fish. Alternatively, since the two families were generated from different parents, fish in one of the families may have carried genetic variants not found in the other and this may have influenced the results. Future work will include another replication of the experiment of *psen1*^{Q96_K97del/+} and their *psen1*^{+/+} siblings at 12 months of age with a larger sample size to confirm whether the *psen1*^{Q96_K97del/+} genotype alters spatial working memory.

Human patients who carry EOfAD mutations in *PSEN1* generally have an age-of-onset of between 40 and 50 years of age [40], while EOfAD mutations in *SORL1*,

although not as broadly characterised due to their rarity, tend to be older than 60 years of age [41-43]. This difference in the age-of-onsets could explain why our zebrafish model of an EOfAD mutation in *PSEN1* appears to possibly show impairment of spatial working memory at 12 months of age, but fish that carry EOfAD-like mutations in *sort1* do not. Future work will include analysis of our *sort1* mutant fish at older ages to determine whether spatial working memory is affected.

Some zebrafish show a behavioural lateralisation while performing the FMP Y-maze task. Our results did not fully support the results of Fontana et al. [25], who observed that approximately 1 in 4 zebrafish showed a right turn bias, 1 in 4 zebrafish show a left turn bias and 1 in 2 zebrafish show no preference for left or right turns. We observed much lower proportions of zebrafish showing a behavioural lateralisation (**Figures 10 and 11**). We suspect that these differing results could be due to strain differences (our zebrafish were generated from an inbred Tübingen strain, whilst Fontana et al. [25] used the AB strain of zebrafish), or due to differences from being raised and handled in different laboratories.

Unfortunately, this experiment contained limited throughput due to only having one set of two Y-mazes, meaning that we can only test two fish at a time. This extends out the data collection period over many days and likely introduces extra variation and reduces our statistical power. Additionally, we analysed families of zebrafish with different parents at the different ages for the *psen1*^{Q96_K97del/+} analysis. Ideally, we would like to have performed a longitudinal study, monitoring the alternation of each fish as it ages. However, this is not feasible in zebrafish as to monitor each individual fish, they would need to be separated as they age. Since zebrafish are, naturally, gregarious, this would cause confounding stresses on the fish and would likely mask the effect of genotype.

Conclusion

This work serves as a pilot study for more complete characterisation of the effects of heterozygosity for EOfAD-like mutations on spatial working memory in our zebrafish models of EOfAD. We presented statistical evidence that an EOfAD-like mutation in the zebrafish orthologue of *PSEN1* impairs spontaneous alternation frequency at 12 months of age, which is indicative of impairment of spatial working memory.

However, further replication of this experiment is required to confirm this result. We did not observe any statistical evidence that EOfAD-like mutations in *sor11* affect spatial working memory at 12 months of age. Future work analysing the effect of heterozygosity for EOfAD-like mutations in *sor11* on spatial working memory in older fish will be performed. A complete characterisation of the effects of EOfAD-like mutations on spatial working memory in zebrafish is part of our overall goal to understand the pathological changes which eventually lead to EOfAD.

References

1. Guariglia CC: **Spatial working memory in Alzheimer's disease: A study using the Corsi block-tapping test.** *Dementia & Neuropsychologia* 2007, **1**:392-395.
2. Baddeley AD, Bressi S, Della Sala S, Logie R, Spinnler H: **THE DECLINE OF WORKING MEMORY IN ALZHEIMER'S DISEASE: A LONGITUDINAL STUDY.** *Brain* 1991, **114**:2521-2542.
3. Guariglia CC, Nitrini R: **Topographical disorientation in Alzheimer's disease.** *Arquivos de Neuro-Psiquiatria* 2009, **67**:967-972.
4. Huang H-J, Liang K-C, Ke H-C, Chang Y-Y, Hsieh-Li HM: **Long-term social isolation exacerbates the impairment of spatial working memory in APP/PS1 transgenic mice.** *Brain Research* 2011, **1371**:150-160.
5. Clark JK, Furgerson M, Crystal JD, Fechheimer M, Furukawa R, Wagner JJ: **Alterations in synaptic plasticity coincide with deficits in spatial working memory in presymptomatic 3xTg-AD mice.** *Neurobiology of Learning and Memory* 2015, **125**:152-162.
6. Manocha GD, Floden AM, Rausch K, Kulas JA, McGregor BA, Rojanathammanee L, Puig KR, Puig KL, Karki S, Nichols MR, et al: **APP Regulates Microglial Phenotype in a Mouse Model of Alzheimer's Disease.** *The Journal of Neuroscience* 2016, **36**:8471.
7. Ognibene E, Middei S, Daniele S, Adriani W, Ghirardi O, Caprioli A, Laviola G: **Aspects of spatial memory and behavioral disinhibition in Tg2576 transgenic mice as a model of Alzheimer's disease.** *Behavioural Brain Research* 2005, **156**:225-232.
8. Moran PM, Higgins LS, Cordell B, Moser PC: **Age-related learning deficits in transgenic mice expressing the 751-amino acid isoform of human**

- beta-amyloid precursor protein.** *Proceedings of the National Academy of Sciences* 1995, **92**:5341.
9. Stover KR, Campbell MA, Van Winssen CM, Brown RE: **Early detection of cognitive deficits in the 3xTg-AD mouse model of Alzheimer's disease.** *Behavioural Brain Research* 2015, **289**:29-38.
 10. Hargis KE, Blalock EM: **Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us?** *Behavioural Brain Research* 2017, **322**:311-328.
 11. Barthelson K, Dong Y, Newman M, Lardelli M: **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.** *bioRxiv* 2021:2021.2001.2026.428321.
 12. Barthelson K, Pederson S, Newman M, Lardelli M: **Transcriptome analysis of a protein-truncating mutation in sortilin-related receptor 1 associated with early-onset familial Alzheimer's disease indicates effects on mitochondrial and ribosome function in young-adult zebrafish brains.** *bioRxiv* 2020:2020.2009.2003.282277.
 13. Barthelson K, Pederson SM, Newman M, Lardelli M: **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* 2020, **13**:142.
 14. Barthelson K, Pederson SM, Newman M, Lardelli M: **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* 2020, **Preprint**:1-15.
 15. Hin N, Newman M, Kaslin J, Douek AM, Lumsden A, Nik SHM, Dong Y, Zhou X-F, Mañucat-Tan NB, Ludington A, et al: **Accelerated brain aging towards**

transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2. *PLOS ONE* 2020, **15**:e0227258.

16. Ruan J: **The impacts of age, hypoxia, and the psen1Q96_K97del / + mutation on familial Alzheimer's disease.** University of Adelaide, School of Biological Science, Faculty of Science; 2019.
17. Cleal M, Parker MO: **Moderate developmental alcohol exposure reduces repetitive alternation in a zebrafish model of fetal alcohol spectrum disorders.** *Neurotoxicology and Teratology* 2018, **70**:1-9.
18. Cleal M, Fontana BD, Ranson DC, McBride SD, Swinny JD, Redhead ES, Parker MO: **The Free-movement pattern Y-maze: A cross-species measure of working memory and executive function.** *Behavior Research Methods* 2020.
19. Gross AN, Engel AKJ, Richter SH, Garner JP, Würbel H: **Cage-induced stereotypies in female ICR CD-1 mice do not correlate with recurrent perseveration.** *Behavioural Brain Research* 2011, **216**:613-620.
20. Cleal M, Fontana BD, Ranson DC, McBride SD, Swinny JD, Redhead ES, Parker MO: **A Test of Memory: The Fish, The Mouse, The Fly And The Human.** *bioRxiv* 2020:2020.2002.2015.950816.
21. Barthelson K, Newman M, Lardelli M: **Sorting Out the Role of the Sortilin-Related Receptor 1 in Alzheimer's Disease.** *J Alzheimers Dis Rep* 2020, **4**:123-140.
22. Newman M, Hin N, Pederson S, Lardelli M: **Brain transcriptome analysis of a familial Alzheimer's disease-like mutation in the zebrafish presenilin 1 gene implies effects on energy production.** *Molecular Brain* 2019, **12**.
23. Westerfield M: *The Zebrafish Book*. Fifth Edition edn. University of Oregon Press; 2007.

24. Team RC: **R: A language and environment for statistical computing.** *R Foundation for Statistical Computing, Vienna, Austria* 2019.
25. Fontana BD, Cleal M, Clay JM, Parker MO: **Zebrafish (Danio rerio) behavioral laterality predicts increased short-term avoidance memory but not stress-reactivity responses.** *Animal Cognition* 2019.
26. Brooks ME, Kristensen K, Benthem KJV, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Mächler M, Bolker BM: **glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling.** *The R Journal* 2017, **9**:378.
27. Weisberg JFaS: *An R Companion to Applied Regression.* Thousand Oaks {CA}: Sage; 2019.
28. Dunnett CW: **A Multiple Comparison Procedure for Comparing Several Treatments with a Control.** 1955, **50**:1096.
29. Wickham H: **ggplot2: Elegant Graphics for Data Analysis.** In *Book ggplot2: Elegant Graphics for Data Analysis* (Editor ed.^eds.). City: Springer-Verlag New York; 2016.
30. Venables WN, Ripley BD: *Modern Applied Statistics with S.* Fourth edn. New York: Springer; 2002.
31. Miedel CJ, Patton JM, Miedel AN, Miedel ES, Levenson JM: **Assessment of Spontaneous Alternation, Novel Object Recognition and Limb Clasping in Transgenic Mouse Models of Amyloid- β and Tau Neuropathology.** *J Vis Exp* 2017:55523.
32. Hernández-Rabaza V, Barcia JA, Llorens-Martín M, Trejo JL, Canales JJ: **Spared place and object-place learning but limited spatial working memory capacity in rats with selective lesions of the dentate gyrus.** *Brain Research Bulletin* 2007, **72**:315-323.

33. Maurice T, Hiramatsu M, Itoh J, Kameyama T, Hasegawa T, Nabeshima T: **Behavioral evidence for a modulating role of σ ligands in memory processes. I. Attenuation of dizocilpine (MK-801)-induced amnesia.** *Brain Research* 1994, **647**:44-56.
34. Lewis SA, Negelspach DC, Kaladchibachi S, Cowen SL, Fernandez F: **Spontaneous alternation: A potential gateway to spatial working memory in *Drosophila*.** *Neurobiology of Learning and Memory* 2017, **142**:230-235.
35. Richman CL, Dember WN, Kim P: **Spontaneous alternation behavior in animals: a review.** *Current Psychological Research & Reviews* 1986, **5**:358-391.
36. Bögli SY, Huang MY-Y: **Spontaneous alternation behavior in larval zebrafish.** *The Journal of Experimental Biology* 2017, **220**:171.
37. Yu L, Tucci V, Kishi S, Zhdanova IV: **Cognitive Aging in Zebrafish.** *PLOS ONE* 2006, **1**:e14.
38. Drew WG, Miller LL, Baugh EL: **Effects of δ 9-THC, LSD-25 and scopolamine on continuous, spontaneous alternation in the Y-maze.** *Psychopharmacologia* 1973, **32**:171-182.
39. Newman M, Nik HM, Sutherland GT, Kim WS, Halliday GM, Jayadev S, Smith C, Kittipassorn T, Peet DJ, Lardelli M: **Accelerated loss of hypoxia response and biased allele expression in zebrafish with Alzheimer's disease-like mutations.** *bioRxiv* 2019:526277.
40. Ryman DC, Acosta-Baena N, Aisen PS, Bird T, Danek A, Fox NC, Goate A, Frommelt P, Ghetti B, Langbaum JBS, et al: **Symptom onset in autosomal dominant Alzheimer disease: A systematic review and meta-analysis.** *Neurology* 2014, **83**:253-260.
41. Pottier C, Hannequin D, Coutant S, Rovelet-Lecrux A, Wallon D, Rousseau S, Legallic S, Paquet C, Bombois S, Pariente J, et al: **High frequency of**

potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. *Mol Psychiatry* 2012, **17**:875-879.


42. Champion D, Charbonnier C, Nicolas G: **SORL1 genetic variants and Alzheimer disease risk: a literature review and meta-analysis of sequencing data.** *Acta Neuropathologica* 2019.
43. 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: **Identification and description of three families with familial Alzheimer disease that segregate variants in the SORL1 gene.** *Acta Neuropathologica Communications* 2017, **5**:43.

Chapter 5: 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

Statement of Authorship

Title of Paper	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
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	This manuscript was published in Molecular Brain on 19 October 2020 Barthelson, K., Pederson, S.M., Newman, M. et al. 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. Mol Brain 13, 142 (2020). https://doi.org/10.1186/s13041-020-00681-7


Principal Author


Name of Principal Author (Candidate)	Karissa Barthelson		
Contribution to the Paper	Generation of the mutant fish, isolation of RNA for RNA-seq, data analysis, generated the figures and drafted of the manuscript		
Overall percentage (%)	95%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	9/1/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Stephen Martin Pederson		
Contribution to the Paper	Supervision of bioinformatic analysis. Editing of manuscript		
Signature		Date	12/01/2021

Name of Co-Author	Morgan Newman		
Contribution to the Paper	Supervision of zebrafish work. Editing of manuscript		
Signature		Date	15/02/2021

Name of Co-Author	Michael Lardelli		
Contribution to the Paper	Supervision of project, editing of manuscript		
Signature		Date	19/02/2021

Please cut and paste additional co-author panels here as required.

RESEARCH

Open Access



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

Karissa Barthelson^{1*} , Stephen Martin Pederson², Morgan Newman¹ and Michael Lardelli^{1*}

Abstract

To prevent or delay the onset of Alzheimer's disease (AD), we must understand its molecular basis. The great majority of AD cases arise sporadically with a late onset after 65 years of age (LOAD). However, rare familial cases of AD can occur due to dominant mutations in a small number of genes that cause an early onset prior to 65 years of age (EOfAD). As EOfAD and LOAD share similar pathologies and disease progression, analysis of EOfAD genetic models may give insight into both subtypes of AD. Sortilin-related receptor 1 (*SORL1*) is genetically associated with both EOfAD and LOAD and provides a unique opportunity to investigate the relationships between both forms of AD. Currently, the role of *SORL1* mutations in AD pathogenesis is unclear. To understand the molecular consequences of *SORL1* mutation, we performed targeted mutagenesis of the orthologous gene in zebrafish. We generated an EOfAD-like mutation, V1482Afs, and a putatively null mutation, to investigate whether EOfAD-like mutations in *sorl1* display haploinsufficiency by acting through loss-of-function mechanisms. We performed mRNA-sequencing on whole brains, comparing wild type fish with their siblings heterozygous for EOfAD-like or putatively loss-of-function mutations in *sorl1*, or transheterozygous for these mutations. Differential gene expression analysis identified a small number of differentially expressed genes due to the *sorl1* genotypes. We also performed enrichment analysis on all detectable genes to obtain a more complete view on changes to gene expression by performing three methods of gene set enrichment analysis, then calculated an overall significance value using the harmonic mean p-value. This identified subtle effects on expression of genes involved in energy production, mRNA translation and mTORC1 signalling in both the EOfAD-like and null mutant brains, implying that these effects are due to *sorl1* haploinsufficiency. Surprisingly, we also observed changes to expression of genes occurring only in the EOfAD-mutation carrier brains, suggesting gain-of-function effects. Transheterozygosity for the EOfAD-like and null mutations (i.e. lacking wild type *sorl1*), caused apparent effects on iron homeostasis and other transcriptome changes distinct from the single-mutation heterozygous fish. Our results provide insight into the possible early brain molecular effects of an EOfAD mutation in human *SORL1*. Differential effects of heterozygosity and complete loss of normal *SORL1* expression are revealed.

*Correspondence: karissa.barthelson@adelaide.edu.au; michael.lardelli@adelaide.edu.au

¹ Alzheimer's Disease Genetics Laboratory, School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, SA 5005, Australia
Full list of author information is available at the end of the article



© The Author(s) 2020. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords: Familial Alzheimer's disease, *SORL1*, RNA-seq, Zebrafish, Mitochondria, Iron homeostasis, Harmonic mean p-value

Background

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the most common form of dementia. AD brains show a broad range of pathologies including deposition of intercellular deposits of insoluble amyloid β (A β) peptides within plaques, intracellular tangles primarily consisting of hyperphosphorylated tau proteins, vascular abnormalities [1, 2], mitochondrial dysfunction [3–5], inflammation [6, 7], lipid dyshomeostasis [8–10], metal ion dyshomeostasis [11, 12] and numerous others.

To prevent or delay the onset of AD, we need to understand the early cellular changes which eventually lead to these AD pathologies. This is difficult to investigate in humans, as pre-symptomatic, living AD brain tissue is inaccessible for detailed molecular analysis. Consequently, animal models can be extremely useful for understanding the stresses driving this disease. The commonly used mouse models of AD overexpress human mutant forms of the EOfAD genes to show histopathological phenotypes reminiscent of the human disease. Troublingly however, the brain transcriptomes of these mouse models show low concordance with human AD, and with each other [13]. Thus, these mouse models are unlikely to mimic, accurately, the genetic state of the human disease.

In rare, familial cases of AD, patients can show symptoms of disease onset before 65 years of age. These cases are most often due to single, autosomal dominant mutations in one of three genes: amyloid β A4 precursor protein (*APP*), presenilin 1 (*PSEN1*) and presenilin 2 (*PSEN2*) (EOfAD) [14]. All EOfAD mutations in the *PSENs* and *APP* follow a “reading-frame preservation rule”, where mutations causing truncation of the open reading frame do not cause EOfAD (reviewed in [15]). Recently, mutations in sortilin-related receptor 1 (*SORL1*) have been found that segregate with early onset AD in dominant inheritance patterns [16, 17], suggesting that *SORL1* may represent a fourth EOfAD-causative gene. Interestingly, both missense and reading frame-truncating mutations in *SORL1* have been observed in early-onset AD families [16, 17]. Since reading frame-truncating mutations in *SORL1* have been shown likely to cause nonsense mediated mRNA decay of the mutant alleles' mRNA [16], it is thought the mutations may act through a haploinsufficient, loss-of-function mechanism. However, this has not been explored at the molecular level in vivo.

Most AD cases arise sporadically and have an age of onset of later than 65 years (LOAD). The etiology of

LOAD is still unclear. However, there are variants at numerous loci associated with increased risk of developing LOAD, such as the ϵ 4 allele of the gene encoding apolipoprotein E (*APOE*) [18]. Interestingly, variation at the *SORL1* locus is also associated with LOAD [19–22] so that understanding the function of this gene may illuminate a mechanistic link between EOfAD and LOAD. However, whether or not *SORL1* should be regarded as an EOfAD-causative locus in the manner of the *PSEN* and *APP* genes is still debated [23].

SORL1 protein is a membrane-bound, multi-domain-containing protein and localises mainly in cells' endolysosomal system and the trans-Golgi network. We have previously found that *SORL1* localises to the mitochondrial associated membranes (MAMs) of the endoplasmic reticulum [24]. *SORL1* belongs to the family of vacuolar protein sorting 10 (VPS10)-containing proteins, or sortilins. These proteins all carry a VPS10 domain with homology to the VPS10P domain found in yeast [25]. *SORL1* also belongs to the low density lipoprotein receptor (LDLR) family of proteins and contains both LDLR class A repeats and LDLR class B repeats (reviewed in [26]).

The functional role of *SORL1* has been investigated mostly in the context of the distribution and processing of APP within cells. *SORL1* guides APP throughout the endolysosomal system and is thought to prevent formation of A β by promoting recycling of APP (reviewed in [26]). Mutations in the protein-coding region of *SORL1* have been shown to reduce the capacity of the *SORL1* protein to bind APP and result in increased levels of A β [27, 28]. A single nucleotide polymorphism (SNP) cluster consisting of 6 SNPs spanning a region between exon 6 to intron 9 of *SORL1* is associated with decreased elevation of *SORL1* expression in response to brain-derived neurotrophic factor (BDNF) in neurons derived from human induced pluripotent stem cells (hiPSCs). This results in aberrant processing of APP [29]). However, the non-APP related functions of *SORL1*, and the effects of mutations in *SORL1* on the molecular state of the central nervous system in vivo, remain largely unexplored.

Here, we describe a study addressing two questions: (1) What are the effects on the young-adult zebrafish whole brain transcriptome due to heterozygosity for a mutation modelling the putatively EOfAD-causative *SORL1* mutation C1481*? (2) Are these effects due to loss or gain of function? To address these questions, we introduced a similar mutation, V1482Afs, into the zebrafish orthologue

of *SORL1* (*sorl1*) [16]. We also generated a putatively loss-of-function mutation, R122Pfs, as a control representing haploinsufficient loss-of-function. We performed RNA sequencing (RNA-seq) on mRNAs derived from entire brains from a family of young-adult sibling zebrafish that were either heterozygous for the V1482Afs mutation (for simplicity, hereafter referred to as EOfAD-like/+), heterozygous for the R122Pfs loss-of-function mutation (hereafter referred to as null/+), transheterozygous for both the EOfAD-like and null mutations (i.e. EOfAD-like/null, a complete loss of wild type *sorl1*) or wild type. We found that, in the heterozygous state, the EOfAD-like mutation causes subtle changes to gene expression and appears to act through both loss-of-function and gain-of-function mechanisms. Differences in the EOfAD-like/+, null/+ and transheterozygous mutant sibling brain transcriptomes highlight the importance of analysing animal models which reflect, as closely as possible, the genetic state of the human disease, and illuminate novel cellular processes previously unknown to be affected by loss of normal *SORL1* expression.

Methods

Zebrafish husbandry and animal ethics

Work with zebrafish was performed under the auspices of the Animal Ethics Committee of the University of Adelaide, permit numbers S-2017-089 and S-2017-073. All zebrafish used in this study were maintained in a recirculating water system on a 14 h light/10 h dark cycle, and fed NRD 5/8 dry food (Inve Aquaculture, Dendermonde, Belgium) in the morning and live *Artemia salina* in the afternoon. In total, 36 fish were used over all the experiments described in this study.

Genome editing constructs

To introduce an EOfAD-like frameshift mutation near the C1481 codon of zebrafish *sorl1*, we used a TALEN pair designed by, and purchased from, Zgenebio Biotech Inc. (Taipei City, Taiwan). The genomic DNA recognition sites (5'–3') were TGAGGTGGCGGTGTG (left TALEN) and CTGAAATACATGCTGG (right TALEN) (Additional file 1). The DNAs encoding the TALEN pair protein sequences were supplied in the pZGB2 vector. These constructs were linearised with *Not* I (NEB, Ipswich, USA) and then mRNA was transcribed in vitro using the mMACHINE T7 in vitro Transcription Kit (Invitrogen, Carlsbad, USA) following the manufacturer's protocol.

To generate a putatively null mutation in *sorl1*, we targeted exon 2 of *sorl1* using the CRISPR-Cpf1 system [30]. A crRNA was designed to recognise the sequence 5' GGGTCTGTGGACCAACGCT 3' in exon 2 of *sorl1*, with a PAM sequence of TTTT (Additional file 1). Both

the crRNA and Cpf1 recombinant protein were synthesised by, and purchased from, IDT Technologies (Iowa, USA).

We injected these constructs into zebrafish embryos at the one-cell stage and, ultimately, isolated fish carrying the mutations of interest (V1482AfsTer12, hereafter referred to as V1482Afs or "EOfAD-like", in exon 32 and R122PfsTer118, hereafter referred to as R122Pfs or "null", in exon 2). For a detailed description of the methods used to isolate these mutant lines of fish, see Additional file 2.

Breeding strategy

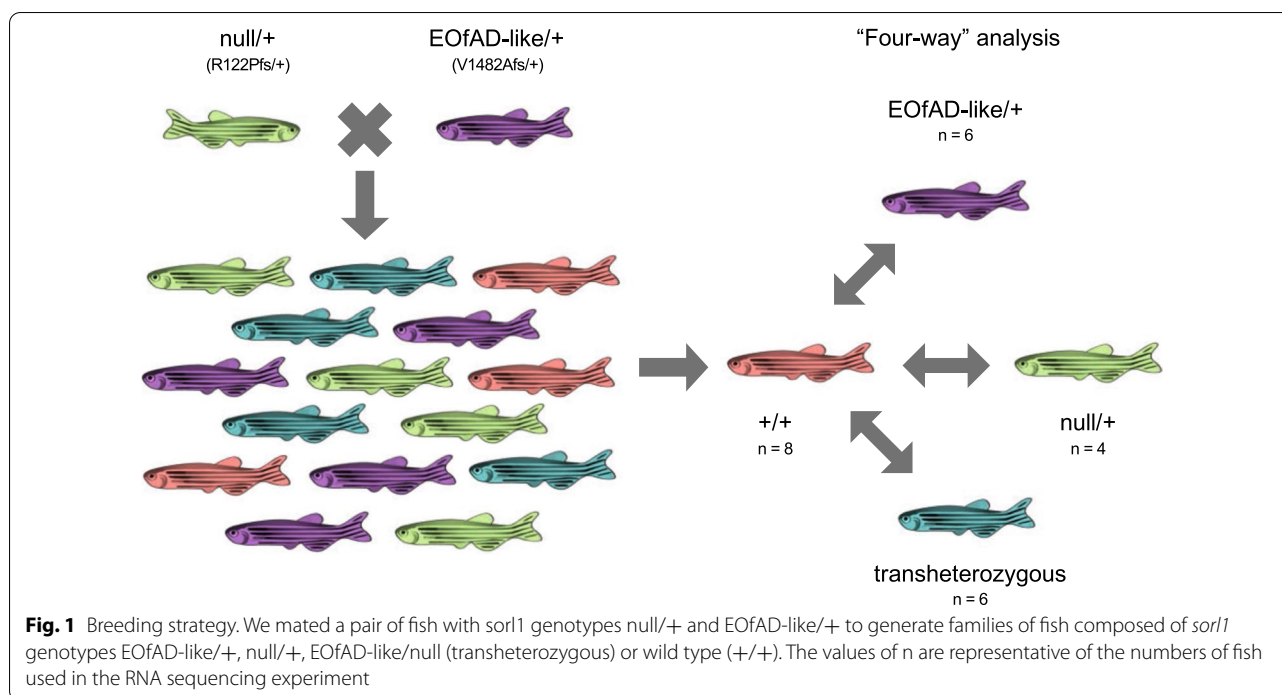
To generate families of fish for analysis, we crossed an EOfAD-like/+ fish to a null/+ fish to generate a family of siblings with one of four *sorl1* genotypes: EOfAD-like/+, null/+, EOfAD-like/null (transheterozygous) or +/+ (Fig. 1). Each family of fish was raised in an 8 L capacity tank, with approximately 40 fish per tank (i.e. approximately 5 fish per litre) to avoid stresses due to overcrowding. The tanks were placed side-by-side in the same recirculated water aquarium system to reduce environmental variation between them.

Allele-specific expression of *sorl1* transcripts

Adult zebrafish (6 months old) were euthanised in a loose ice slurry. Entire heads were cut off and placed in 600 μ L of RNeasyTM Stabilization Solution (Invitrogen, Carlsbad, USA) before incubation at 4 °C overnight. Brains were then removed from the heads using sterile watchmaker's forceps. Total RNA was extracted from the brains using the QIAGEN RNeasy[®] Mini Kit (Qiagen, Venlo, Netherlands) according to the manufacturer's protocol. Recovered RNA concentrations were estimated using a Nanodrop spectrophotometer. RNAs were stored at –80 °C until required. We prepared cDNA using 400 ng of total RNA input (to give a 20 ng/ μ L cDNA solution) and the Superscript III First Strand Synthesis System (Invitrogen, Carlsbad, USA) according to the manufacturer's instructions (random hexamer priming method).

We quantified the absolute copy numbers of *sorl1* transcripts in EOfAD-like/+ mutant brains by allele-specific digital quantitative PCRs (dqPCRs) as described in [31]. The mutant allele-specific forward primers described in Additional file 2 were used with a common reverse primer binding over the *sorl1* exon 33–34 junction (5' GACCTGCTGTTCATCAGTGC 3') to avoid amplification from any genomic DNA carried over during the RNA extraction. We used 50 ng of cDNA as input in each dqPCR reaction.

To compare the *sorl1* transcript expression levels in null/+ mutant brains, we amplified a ~250 bp region spanning the R122 codon in exon 2 from 50 ng of brain-derived cDNA from wild type and null/+ fish using



reverse transcription PCRs (RT-PCRs) and resolved the RT-PCR products on a 3% agarose gel in 1 × tris–acetate-EDTA (TAE) in milliQ water. Since the null mutation deletes 16 bp from the gene, this procedure resolved the two bands representing the two alleles of *sor11* and we could inspect the relative intensity of the bands.

RNA sequencing data generation and analysis

We performed RNA sequencing in the “four-way” analysis described in Fig. 1. Adult zebrafish (6 months old) were euthanised in an ice slurry. Whole heads were removed for preservation in RNAlater solution (Invitrogen, Carlsbad, USA), and a fin biopsy was removed for genotype determination by PCRs (Additional file 2). The whole brains were removed from the preserved heads using sterile watchmaker’s forceps and then total RNA was isolated using the mirVana™ miRNA Isolation Kit (Ambion, Life Technologies, Thermo Fisher Scientific, Waltham, USA) following the manufacturer’s protocol. Genomic DNA was removed from the total RNA by DNase treatment using the DNA-free™ Kit (Ambion, Life Technologies, Thermo Fisher Scientific, Waltham, USA).

500 ng of total RNA was then delivered to the Genomics service at the South Australian Health and Medical Research Institute (SAHMRI, Adelaide, AUS) for poly-A+, stranded library preparation and RNA sequencing using the Illumina Nextseq platform. We used a total of 24 fish at 6 months of age (EOfAD-like/+ : n = 6, null/+ : n = 4, wild type: n = 8 and transheterozygous: n = 6).

We initially aimed to have n=6 fish per genotype, based on a previous power calculation from a brain transcriptome analysis of a mutation in the *psen1* gene of zebrafish. This had indicated that n=6 would provide 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 (data not shown). However, genotype checks of the RNA-seq data revealed that two samples had been misidentified during the genotyping by PCRs and were reclassified as wild type.

Demultiplexed fastq files were provided by SAHMRI as 75 bp single-end reads. All libraries were sequenced to a depth of between 30 and 38 million reads, across four NextSeq lanes which were subsequently merged. The quality of the provided reads was checked using fastQC [32] and ngsReports [33]. We then trimmed adaptors and bases with a PHRED score of less than 20 from the reads using AdapterRemoval (v2.2.0). Reads shorter than 35 bp after trimming were discarded. We then performed a pseudo-alignment to estimate transcript abundances using kallisto v0.43.1 [34] in single-end mode, specifying a forward stranded library, the fragment length as 300 with a standard deviation as 60, and 50 bootstraps. The index file used for the kallisto pseudo-alignment was generated from the zebrafish transcriptome according to the primary assembly of the GRCz11 reference (Ensembl release 96), with the sequences for unspliced transcripts additionally included.

We imported the transcript abundances from kallisto for analysis using R [35] using the catchKallisto function from the package edgeR [36]. We summed the counts of each of the mature transcripts arising from a single gene (i.e. omitting any unspliced transcripts with intronic sequences remaining) to generate gene-level quantifications of transcript abundances. We filtered genes which were undetectable (less than 0.66 counts per million reads (CPM)) in at least 12 of the 24 RNA-seq libraries, leaving library sizes ranging between 13,481,192 and 22,737,049 counts. We normalised for differences in library sizes using the trimmed mean of M-values (TMM) method [37], followed by removal of one factor of unwanted variation using the RUVg method of RUVSeq [38]. The negative control genes for RUVg were specified from an initial differential gene expression analysis using the generalised linear model capabilities of edgeR in an ANOVA-type method. A design matrix was specified with the wild type *sorl1* genotype as the intercept, and the coefficients as each of the other *sorl1* genotypes (EOfAD-like/+, null/+, transheterozygous), the sex of the fish (male or female, with female as the reference level), and the tank in which each fish was raised (tank 1, tank 2 or tank 3, with tank 1 being the reference level). The 5000 least differentially expressed genes (i.e. the genes with the largest p-value) due to all *sorl1* genotypes were then used as the negative control genes. The resulting W_1 offset term, setting $k=1$, from RUVSeq was included in the design matrix for an additional differential gene expression analysis. We considered a gene to be differentially expressed (DE) if the FDR-adjusted p-value was below 0.05 for each specific comparison.

We also checked within the differential expression analysis described above to determine whether there was a bias for GC content or length with the fold change or significance of each gene. No discernible length or GC bias was found (Additional file 3).

For enrichment analysis, the gene sets used were the HALLMARK [39] and KEGG pathway [40] gene sets available from the Molecular Signatures Database (MSigDB, www.gsea-msigdb.org/gsea/msigdb/index.jsp). We downloaded these gene sets as a .gmt file with human Entrez gene identifiers and converted the Entrez identifiers to zebrafish Ensembl identifiers using a mapping file obtained from the Ensembl biomart [41] web interface. We also used the four gene sets of genes with an iron-responsive element (IRE) described in [42] to determine whether there was a possible iron dyshomeostasis signal in our dataset. Finally, the GROSS_HYPOXIA_VIA_HIF1A_DN and GROSS_HYPOXIA_VIA_HIF1A_UP [43] gene sets from MSigDB (C2, CPG subcategory) were used to characterise any changes to expression of genes involved in the cellular response to hypoxia.

Enrichment testing was performed using fry [44] and camera [45] from the limma package [46], and the fast implementation of the GSEA algorithm described in [47] using fgsea [48] for each comparison between the *sorl1* mutants and their wild type siblings.

Since each algorithm of enrichment analysis gave different levels of significance for each gene set, we calculated a consensus p-value by calculating their harmonic mean p-value [49]. The harmonic mean p-value is a method of combining dependent p-values while controlling for the family-wise error rate. We further protected from type I errors by considering gene sets to be significantly altered as a group if the FDR-adjusted harmonic mean p-value was less than 0.05.

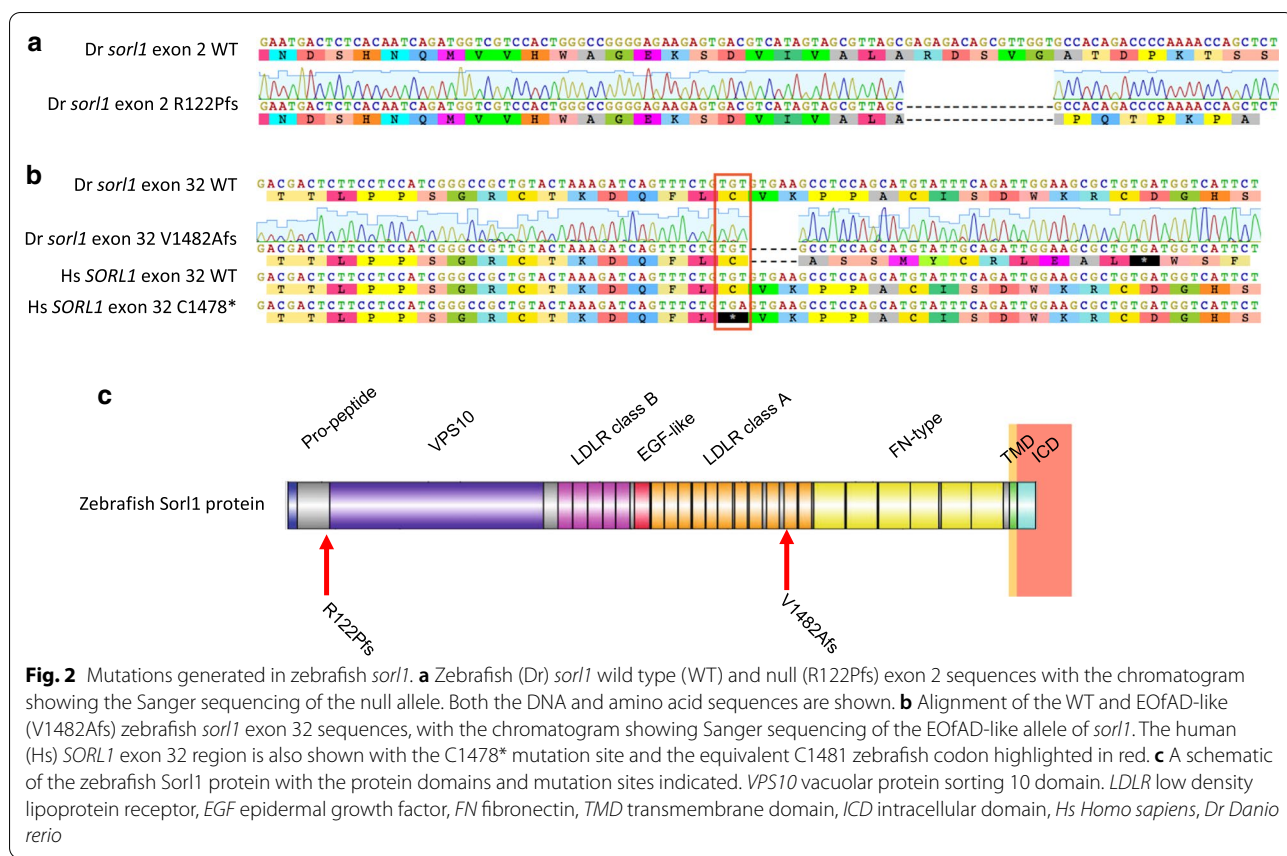
Visualisation for the RNA-seq analysis was performed using ggplot2 [50], pheatmap [51] and upsetR [52].

Results

Creation of in vivo animal models of null and EOfAD-like mutations in *SORL1*

The C1478* mutation in *SORL1* was identified in a French family and appears to segregate with AD with an autosomal dominant inheritance pattern [16]. We aimed to generate a zebrafish model of this mutation by editing the endogenous sequence of zebrafish *sorl1*. The C1478 codon is conserved in zebrafish, C1481 (Fig. 2). Consequently, we used TALENs to generate double stranded breaks at this site in the zebrafish genome (Additional file 1) and then allowed the non-homologous end-joining pathway of DNA repair to generate indel mutations. We identified a family of fish carrying a 5 nucleotide deletion, causing a frameshift in the coding sequence. This results in 11 novel codons followed by a premature termination codon (V1482AfsTer12, or more simply, V1482Afs, hereafter referred to as “EOfAD-like”).

To investigate whether the EOfAD-like mutation acts through loss-of-function, we generated a putatively null mutation in *sorl1*. We targeted exon 2 of *sorl1* using the CRISPR-Cpf1 system, as exon 1 contained three in-frame ATG codons which could, potentially, act as alternative translation initiation codons to allow translation of the majority of the protein. We identified a family of fish carrying a 16 nucleotide deletion, resulting in a frame shift in the coding sequence. This frame shift is predicted to encode 118 novel amino acids followed by a premature termination codon (R122PfsTer118, or more simply, R122Pfs, hereafter referred to as “null”). The protein encoded by the null allele of *sorl1* is predicted to lack most of the functional domains of the wild type Sorl1 protein (Fig. 2).



Protein-truncating mutations in *sorl1* are likely subject to nonsense-mediated mRNA decay

The C1478* mutation in human *SORL1* has been shown likely to cause nonsense mediated mRNA decay (NMD) in lymphoblasts from a human mutation-carrier [16]. Therefore, we investigated whether transcripts of the EOfAD-like allele of *sorl1* are also subject to NMD by allele-specific digital quantitative PCRs (dqPCRs) on cDNA generated from EOfAD-like/+ brains. We observed significantly fewer transcripts of the EOfAD-like allele than the wild type allele in EOfAD-like/+ brains ($p=0.05$). We also observed significantly fewer ($p=0.02$) transcripts of the wild type allele in the EOfAD-like/+ brains than in their wild type siblings (Fig. 3a).

We were also interested to determine whether the null transcript of *sorl1* was subject to NMD. It proved difficult to obtain differential amplification distinguishing the null and wild type alleles of *sorl1* in dqPCRs. Therefore, we used reverse transcription PCRs (RT-PCRs) to amplify a 290 bp region encompassing *sorl1* exon 2 from 50 ng of brain-derived cDNAs from null/+ fish and +/+ fish (Fig. 3b, upper). After electrophoresis through a 3% agarose gel, two RT-PCR products were observed from null/+ fish: one corresponding to transcripts of

the wild type allele (290 bp) and the other corresponding to transcripts of the null allele (274 bp). The signal from the smaller RT-PCR product, from the null allele transcript of *sorl1*, was visibly less intense than that of the larger, wild type transcript RT-PCR product. Additionally, the wild type RT-PCR product from null/+ brains appeared less intense than the wild type RT-PCR product from +/+ brains (Fig. 3b, lower). Although this method of analysis is not strictly quantitative, these results support that the *sorl1* null allele transcript is likely subject to NMD and that there are fewer copies of the wild type transcript in the null/+ brains relative to their +/+ siblings. In summary, the mutant alleles of *sorl1* are likely subject to NMD.

Transcriptome analyses of null/+, EOfAD-like/+ and transheterozygous *sorl1* mutant brains compared to +/+ sibling brains

Do dominant EOfAD mutations of *SORL1* act through loss- or gain-of-function mechanisms or both? To address this question, we sought to make a detailed molecular comparison of the effects of our EOfAD-like mutation and our putatively null mutation. We performed RNA-seq on brain-derived mRNA in a “four-way” analysis (Fig. 1) with the aim of identifying the global changes to

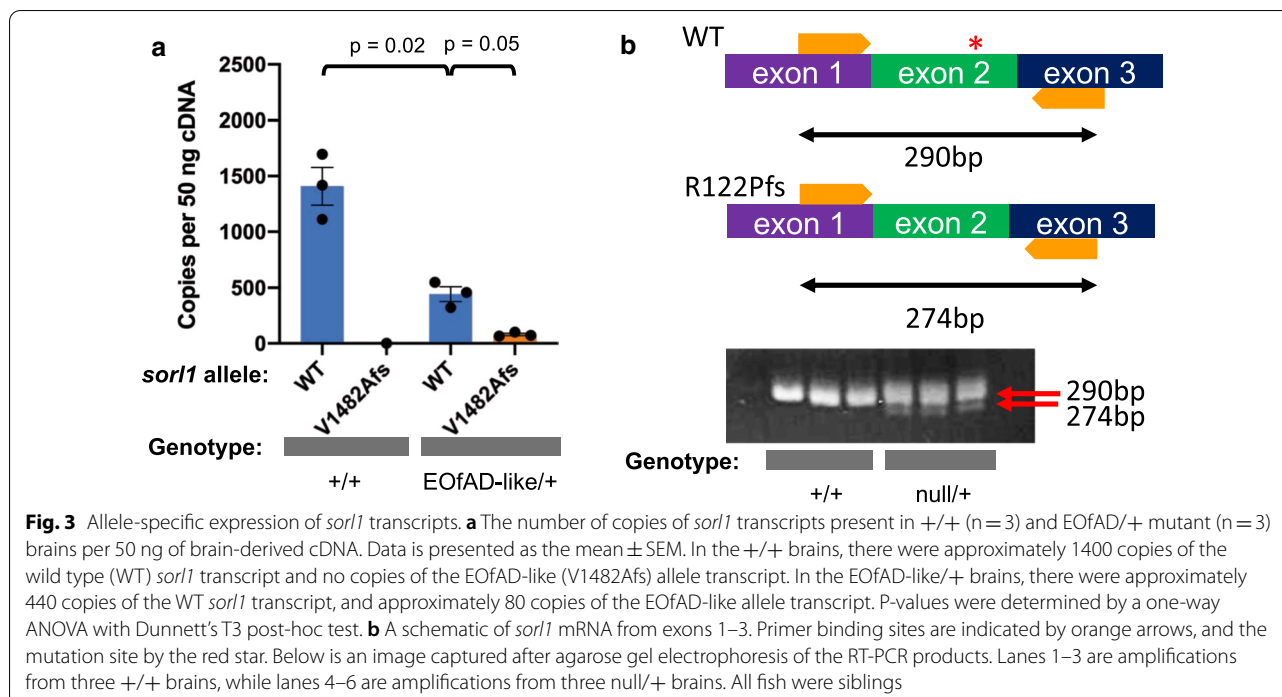


Fig. 3 Allele-specific expression of *sorl1* transcripts. **a** The number of copies of *sorl1* transcripts present in +/+ (n = 3) and EOfAD/+ mutant (n = 3) brains per 50 ng of brain-derived cDNA. Data is presented as the mean \pm SEM. In the +/+ brains, there were approximately 1400 copies of the wild type (WT) *sorl1* transcript and no copies of the EOfAD-like (V1482Afs) allele transcript. In the EOfAD-like/+ brains, there were approximately 440 copies of the WT *sorl1* transcript, and approximately 80 copies of the EOfAD-like allele transcript. P-values were determined by a one-way ANOVA with Dunnett's T3 post-hoc test. **b** A schematic of *sorl1* mRNA from exons 1–3. Primer binding sites are indicated by orange arrows, and the mutation site by the red star. Below is an image captured after agarose gel electrophoresis of the RT-PCR products. Lanes 1–3 are amplifications from three +/+ brains, while lanes 4–6 are amplifications from three null/+ brains. All fish were siblings

gene expression caused by mutations in *sorl1*. We first visualised the relationship between the individual samples by principle component analysis (PCA) (Additional file 4). We observed limited separation of samples across principal component 2 (PC2), which explained only \sim 6% of the total variance in this dataset. This suggests that *sorl1* genotype had only subtle effects on the brain transcriptomes. The PCA also revealed that the first principal component, PC1, i.e. the largest source of variation in this dataset, appeared to be highly correlated with the library size of the samples, even after normalisation using the TMM method. After removal of unwanted variation using RUVSeq, samples mostly separated by genotype across PC1, although some variability was observed within the wild type and transheterozygous mutant samples. Importantly, after RUVSeq, PC1 was no longer correlated with library size (Additional file 4).

Mutations in *sorl1* have subtle effects on gene expression in the brain

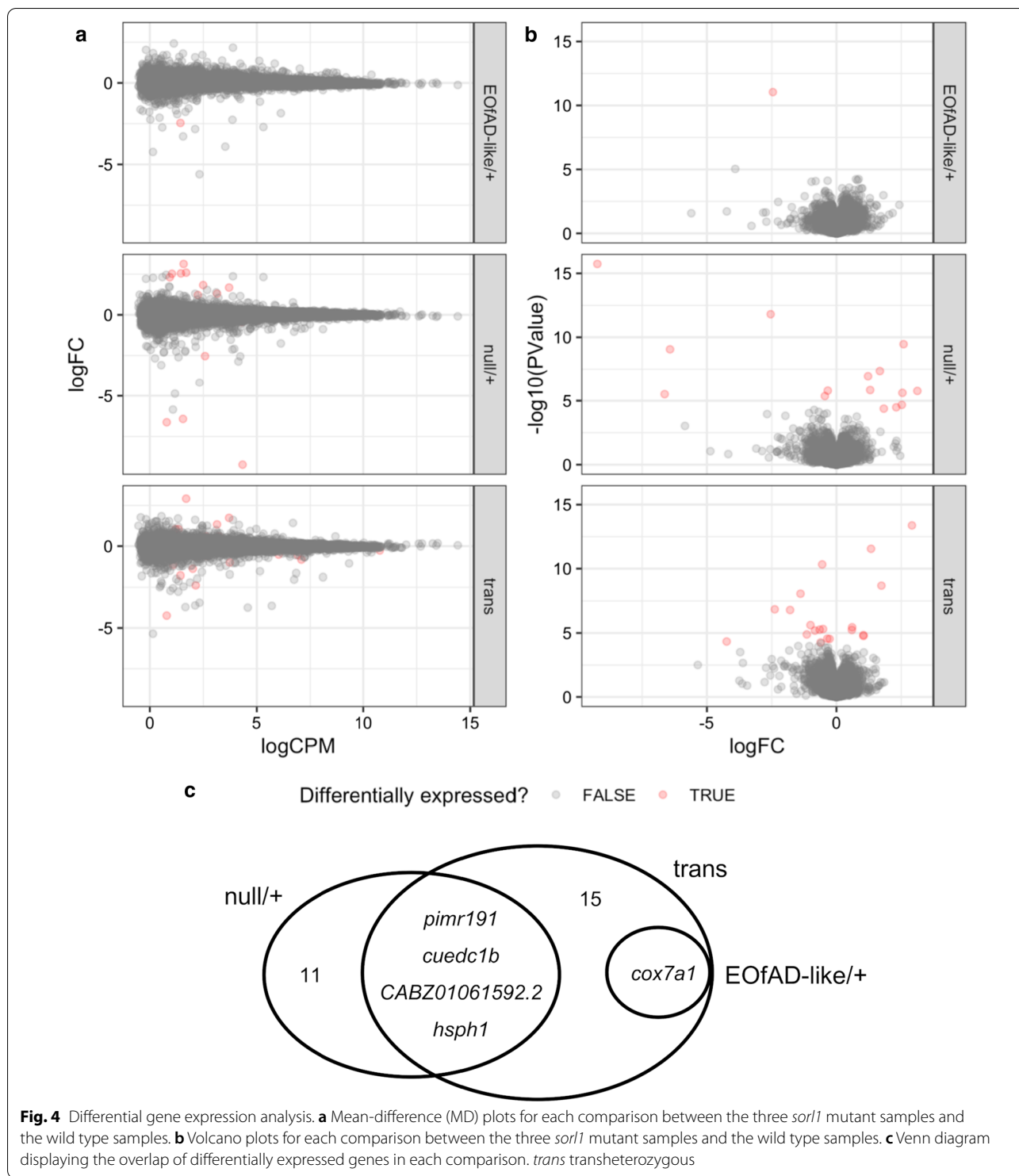
To our knowledge, an in vivo characterisation of the effects of mutations in *sorl1* on the brain transcriptome has not previously been performed. Therefore, we investigated which genes were dysregulated due to heterozygosity for the EOfAD-like mutation, or the null mutation, or complete loss of wild type *sorl1* function (i.e. in the transheterozygous mutant brains).

We could detect differential expression of only one gene, cytochrome c oxidase subunit 7A1 (*cox7a1*) in

EOfAD-like/+ brains relative to wild type brains. In the null/+ brains, we detected 15 differentially expressed (DE) genes relative to wild type, while transheterozygosity for the EOfAD-like and null *sorl1* mutations revealed 20 DE genes relative to wild type (Fig. 4a, b, Additional file 5). Four DE genes were found to be shared between the null/+ and transheterozygous mutant brains: *pim* proto-oncogene, serine/threonine kinase related 191 (*pimr191*), cue domain containing 1b (*cuedc1b*), *CABZ01061592.2* and heat shock 105/110 protein 1 (*hsp1*). Differential expression of *cox7a1* was observed in both the EOfAD-like/+ and transheterozygous mutant samples. No DE genes were shared between the null/+ and EOfAD-like/+ brains relative to their wild type siblings, suggesting that, at the single-gene level, the EOfAD-like mutation does not act through a simple loss-of-function mechanism (Fig. 4c).

Gene set enrichment analysis reveals that an EOfAD-like mutation in *sorl1* appears to have both loss-of-function and gain-of-function properties

Since few DE genes were observed in the *sorl1* mutant brain transcriptomes relative to the wild type brain transcriptomes, we aimed to obtain a more complete view of the changes to gene expression and cellular function by performing enrichment analysis on the entire list of detectable genes in the RNA-seq experiment. If EOfAD-like mutations in *sorl1* caused, exclusively, loss-of-function effects, we would expect similar gene sets to



be altered in the EOFAD-like/+ brains and the null/+ brains.

For enrichment analysis, we used the KEGG and HALLMARK gene sets from MSigDB. The HALLMARK gene sets represent 50 distinct biological processes built

both computationally and manually from the intersections between several gene set collections. These gene sets are useful to generate an overall view of key, distinct biological processes [39]. In contrast, the KEGG gene sets (186 gene sets) give a more precise view of changes

to gene expression in particular cellular processes [40], but pathways are more likely to share many common genes. We used these gene sets in the self-contained gene set testing method fry [44], and the competitive gene set testing method camera [45]. However, we did not find any statistical evidence that any of the KEGG or HALLMARK gene sets were significantly altered in any comparison (for the top 10 most altered pathways, see Additional file 6).

We also used fgsea [48], which is the fast implementation of the self-contained gene set testing method GSEA [47]. Using fgsea, we observed gene sets significantly showing changes as a group after Bonferroni adjustment for multiple testing in each comparison (Additional file 6). However, fgsea does not take into account inter-gene correlations and can be prone to false positives [47]. Therefore, we calculated a consensus p-value derived from the p-values from each of the three methods of enrichment analysis by calculating the harmonic mean p-value [49]. After FDR adjustment of the harmonic mean p-values, we found 11 gene sets to be significantly altered by heterozygosity for the null mutation, 16 gene sets significantly altered by heterozygosity for the EOfAD-like mutation and 11 gene sets significantly altered by transheterozygosity for the null and EOfAD-like mutations. A summary of the significant gene sets is shown in Fig. 5. In contrast, 100 random datasets were generated by permuting the gene labels 100 times and subjected to the enrichment analysis using fry, camera and fgsea, then the harmonic mean p-value calculation. No KEGG, HALLMARK or IRE gene sets were found to be significantly altered after FDR adjustment of the p-values for each individual algorithm, or the combined harmonic mean p-value. (For this analysis, please see the Github repository: https://github.com/karissa-b/sorl1_4way_RNASeq_6m.) This provides further evidence for the statistical significance of the observed alterations in the gene sets shown in Fig. 5.

Out of the 16 gene sets significant in the EOfAD-like/+ brains, 5 were also observed to be significantly altered in the null/+ brains. Intriguingly, three of these gene sets involve mitochondria (the KEGG and HALLMARK gene sets for oxidative phosphorylation, and Parkinson's disease). Genes encoding ribosomal subunits, and genes involved in the mammalian target of rapamycin complex 1 (mTORC1) signalling pathway, also appear

to be affected by the EOfAD-like mutation and the null mutation, indicating that these effects are likely due to decreased *sorl1* function.

Intriguingly, heterozygosity for the EOfAD-like mutation also gives rise to changes in gene expression which do not appear to be affected in the null/+ brains, suggesting gain-of-function action. Interestingly, the KEGG gene set for Alzheimer's disease is one of these significantly altered gene sets. Other gene sets span inflammation, cell adhesion, protein degradation, bile acid metabolism, myc signalling and the tricarboxylic acid (TCA) cycle.

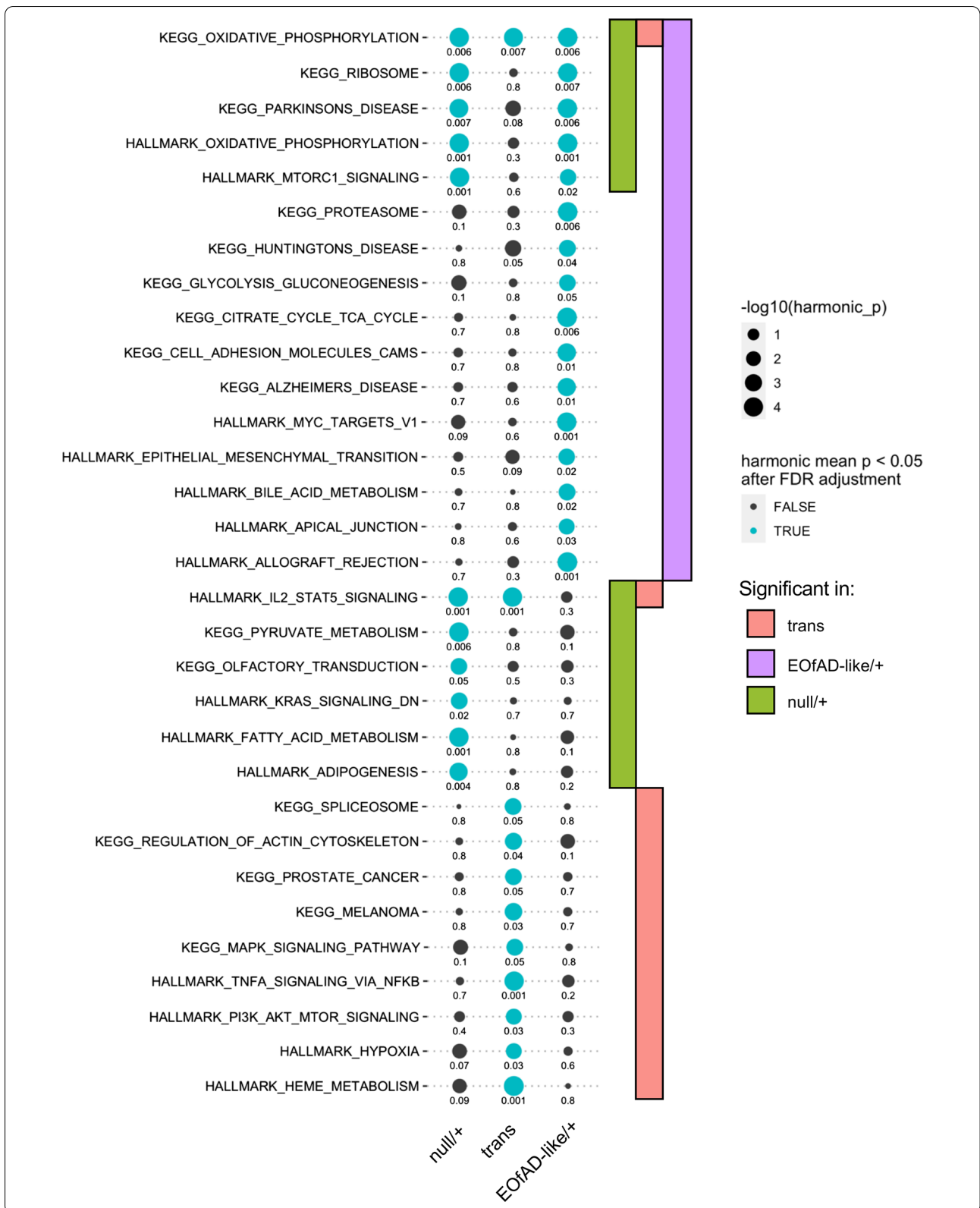
Many HALLMARK and KEGG gene sets consist of shared genes and so are commonly co-identified as significantly altered in enrichment analysis. Therefore, we inspected whether any of the gene sets found to be significantly altered were being driven by changes to expression of the same genes. Genes in the "leading edge" of the fgsea algorithm can be interpreted as the core genes which drive the enrichment of the gene set. We found the leading edge genes within the significant gene sets in each *sorl1* genotype comparison to be mostly independent of one another (Additional file 7). However, we observed some overlap of leading edge genes for the oxidative phosphorylation and neurodegenerative diseases gene sets in the EOfAD-like/+ and null/+ comparisons (i.e. KEGG_ALZHEIMERS_DISEASE, KEGG_HUNTINGTONS_DISEASE and KEGG_PARKINSONS_DISEASE). This indicates that the enrichment of these gene sets is driven mostly by the same gene differential expression signal (genes encoding electron transport chain components) (Additional file 7).

Changes in gene expression due to *sorl1* mutations are not due to broad changes in cell-type distribution

One possible, artifactual explanation for the apparent changes in gene expression we observe could be changes to the proportions of cell-types within the brains of fish with different *sorl1* genotypes. To account for this, we obtained a list of representative marker genes for four broad cell types found in zebrafish brains: neuron, astrocyte and oligodendrocyte markers [53], and microglial markers [54]. We then examined their expression across the brain samples. We did not observe any obvious differences between genotypes, supporting that the changes seen in gene expression are not due to broad changes in cell type proportions within the brains (Additional file 8).

(See figure on next page.)

Fig. 5 Gene set enrichment analysis. Figure 5 depicts the significantly altered KEGG and HALLMARK gene sets in the null/+, EOfAD-like/+ and transheterozygous mutant brains relative their wild type siblings. The sizes of the dots indicate the negative \log_{10} of the harmonic mean p-value (i.e. larger dots indicate greater statistical significance) of three methods of enrichment analysis (fry, camera and fgsea) when combined within each comparison. Gene sets are grouped as to the comparisons in which they are significant. Numbers associated with each dot are FDR adjusted harmonic mean p-values.



The effects of heterozygosity for an EOfAD-like mutation are distinct from those of complete loss of wild type *sor11* function

In research in AD genetics, it is common for cell- and animal-based studies to analyse mutations in their homozygous state. This allows easier identification at contrasting functions between mutant and wild type alleles. However, this ignores that, in humans, EOfAD mutations and LOAD risk variants are rarely homozygous and that interaction between mutant and wild type alleles can generate unique molecular effects. The transheterozygous genotype in our study can be considered somewhat similar to a homozygous loss-of-function, since it lacks the wild type *sor11* allele. We found that genes of the oxidative phosphorylation pathway were the only genes to be significantly affected as a group in common across the EOfAD-like/+, null/+, and transheterozygous mutant brains (Fig. 5). Concerningly, the genes which make up this gene set mostly showed opposite direction of change in the EOfAD-like/+ (and the null/+ brains) compared to the transheterozygous mutant brains (Fig. 6a, Additional file 9). Indeed, we found that the majority of genes within the gene sets significantly altered in the EOfAD-like/+ brains generally showed opposite directions of change in the transheterozygous mutant brains (Fig. 6). This suggests that heterozygosity for this EOfAD-like mutation has very different effects to complete loss of normal *sor11* function, and highlights the importance of analysing genetic models that mimic the heterozygous genetic state of the human disease.

Loss of *sor11* function may cause iron dyshomeostasis

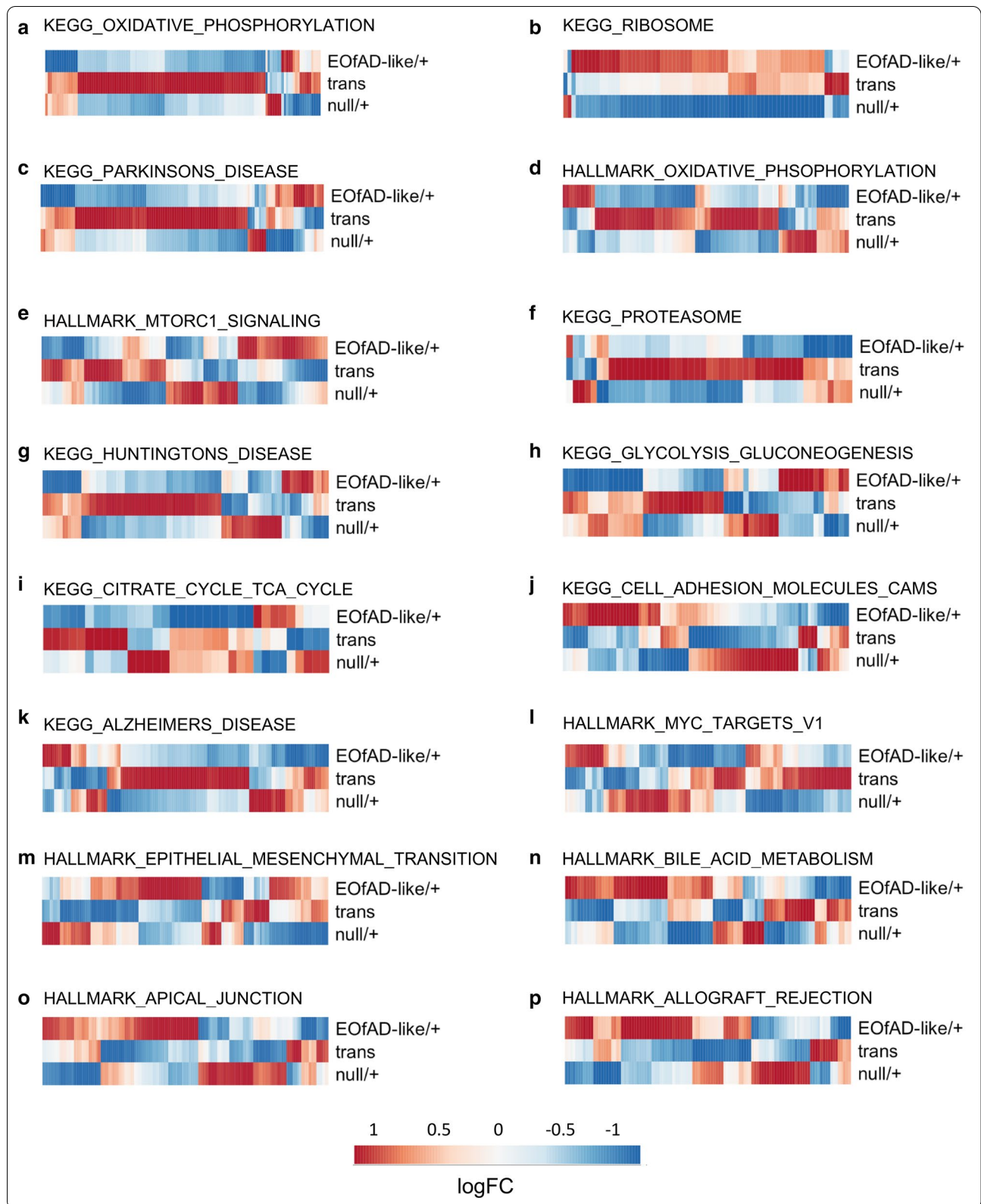
A recent, sophisticated study in [55] demonstrated that impairment of acidification of the endolysosomal system inhibits the ability of cells to reduce ferric iron to more reactive ferrous iron. This results in a pseudo-hypoxic response, since a master regulator protein of the cellular response to hypoxia, HIF1- α , is degraded under normal oxygen conditions by a ferrous iron-dependent mechanism. Ferrous iron deficiency was shown to result in mitochondrial dysfunction and inflammation. We recently developed a method to detect possible iron dyshomeostasis in RNA-seq data, by examining changes to the expression of genes with

iron-responsive elements (IREs) in untranslated regions (UTRs) of their mRNAs [42]. We hypothesised that the changes in gene expression observed in the *sor11* mutant brains in the oxidative phosphorylation pathway may be co-occurring with changes in, or responses to, cellular iron levels. To test this, we obtained the IRE gene sets developed in [42]. These four gene sets consist of genes classified as having a canonical or non-canonical IRE in the UTRs of their mRNAs, and having an IRE motif in the 5' or 3' UTR of the mRNA. We performed an enrichment analysis similar to those performed above for the KEGG and HALLMARK gene sets to determine whether each set of IRE genes showed changes in expression as a group. We found that gene transcripts containing 3' IREs were significantly altered in the transheterozygous mutant brains. Expression of gene transcripts with both canonical and non-canonical 3' IREs generally was decreased. We did not find any statistical evidence that any of the other *sor11* genotypes affected the expression of gene transcripts containing IREs (Fig. 7).

As mentioned earlier, changes to iron homeostasis result in stabilisation of protein HIF1- α . This combines with the protein HIF1- β to form the transcription factor HIF1 that then initiates a pseudo-hypoxic response [55]. The HALLMARK gene set for hypoxia was found to be significantly altered in the transheterozygous mutant brains, but not in the other *sor11* genotypes. We aimed to characterise further whether the *sor11* mutant fish were responding transcriptionally in a hypoxia-like manner by performing an enrichment analysis on the HIF1 target gene sets GROSS_HYPOXIA_VIA_HIF1A_UP and GROSS_HYPOXIA_VIA_HIF1A_DN (hereafter referred to as HIF1 targets). These gene sets were described in [43] and show the genes found to be up- and down-regulated in response to siRNA knockdown of HIF1- α respectively. We found that the downregulated HIF1 targets were significantly altered in the transheterozygous samples, but not in any other mutant samples. However, downregulation of the genes of the GROSS_HYPOXIA_VIA_HIF1A_DN gene set in transheterozygous mutant brains was not consistently observed (Fig. 7d).

(See figure on next page.)

Fig. 6 Heterozygosity for an EOfAD-like mutation in *sor11* often results in changes to gene expression in the opposite direction to more complete loss of wild type *sor11* function. **a–p** LogFC of the genes making up the gene sets in each comparison of *sor11* mutants. Only the gene sets found to be significantly enriched in EOfAD-like/+ brains are shown. The colour legend indicates the magnitude of the logFC in the heatmaps (red is upregulation, blue is downregulation). Remarkable contrast is seen in the directional changes in gene expression between the heterozygous mutants and the transheterozygous mutants for KEGG_PROTEASOME and KEGG/HALLMARK_OXIDATIVE_PHOSPHORYLATION and gene sets in which oxidative phosphorylation genes are an important components (KEGG_HUNTINGTONS_DISEASE and KEGG_ALZHEIMERS_DISEASE, see Additional file 7)



long before cognitive symptoms are predicted to occur. We also wanted to investigate whether the EOfAD-like mutation was deleterious due to haploinsufficiency or neomorphism. Therefore, for comparison, we generated a putatively null mutation in *sor11*, R122Pfs, which is predicted to encode a protein product containing approximately 5% of the coding sequence of wild type Sor11 protein, and lacking any of its functional domains other than the secretory signal sequence. (We assume that expression of this sequence alone has negligible effect on the brain's biology.)

We generated a family of zebrafish composed of sibling fish heterozygous for the EOfAD-like mutation, or heterozygous for the null mutation, as well as fish transheterozygous for the null and EOfAD-like mutations and their wild type siblings (Fig. 1). We initially intended to omit transheterozygous fish from our analysis, as at the time, we believed this genotype was unrepresentative of the human EOfAD genetic state. However, a recent case study presenting a human patient transheterozygous for *sor11* mutations was described in the literature [58]. Additionally, the transheterozygous fish lack wild type *sor11* and, therefore, might be informative as an extreme loss-of-function phenotype. A family of siblings was raised together in a similar environment (in three tanks side-by-side in a re-circulating water system) to reduce genetic and environmental sources of variation in our analysis thereby maximising the possibility of identifying subtle changes in gene expression due to *sor11* genotype.

Mutations in *sor11* have subtle effects on the young-adult brain transcriptome

We identified a small number of DE genes in the brains of each of the different *sor11* mutant fish relative to their wild type siblings (Fig. 4). The transheterozygous mutant fish had the greatest number of genes identified as DE. This was not surprising as these fish lack any wild type *sor11* expression. Only *cox7a1* was found to be significantly differentially expressed in the EOfAD-like/+ brains. This nuclear gene encodes a subunit of the mitochondrial cytochrome c oxidase complex, the terminal component of the respiratory electron transport chain. In humans, this gene has two isoforms, one which is mainly expressed in the heart and skeletal muscle and the other which is mostly expressed in the brain and uterus [59]. Only one of the three isoforms of *cox7a1* known to exist in zebrafish was expressed in the brains of these fish (data not shown).

In comparison to heterozygosity for the EOfAD-like mutation in *sor11*, heterozygosity for an EOfAD-like mutation in *psen1*, the gene most commonly mutated in EOfAD [14, 60], resulted in differential expression of 251 genes in young-adult, zebrafish brains [61]. This suggests

that EOfAD-like mutations in *sor11* are not as disruptive to cell function as EOfAD-like mutations in *psen1*. Campion and colleagues [23], noted that carriers of the C1478* mutation generally had an age of onset closer to the established LOAD threshold of 65 years, which arbitrarily defines the difference between early and late onset AD, and that unaffected, aged (≥ 66 years of age), carriers of this mutation exist. This, together with our results, supports the idea that *SORL1* resembles more closely a LOAD genetic risk locus, rather than an EOfAD-causative locus such as *PSEN1*.

EOfAD-like mutations in *sor11* appear to have both haploinsufficient loss-of-function and neomorphic gain-of-function properties

EOfAD mutations in the *PSENs* and *APP* follow a “reading frame preservation rule”, where mutations preventing translation of C-terminal sequences of the protein do not cause AD (reviewed in [15]). Reading frame-truncating mutations in *SORL1* have been found in EOfAD families so EOfAD mutations in *SORL1* do not follow this rule [16, 17]. Since the protein-truncating mutations have been shown to be subject to NMD [16], it is thought that these mutations have loss-of-function disease phenotypes due to haploinsufficiency. Understanding the mechanism behind the action of EOfAD mutations in *SORL1* is critical to understanding how these mutations increase risk for developing the disease. We observed that heterozygosity for the EOfAD-like mutation resulted in changes to gene expression involved in cellular processes including energy production, mRNA translation and mTORC1 signalling. These changes are also observed in null/+ brains, supporting that they are loss-of-function effects due to haploinsufficiency. Restoring *sor11* expression levels in EOfAD mutation carrier brains could provide a means of ameliorating these apparently pathological changes. We also observed significant changes to gene expression involved in cellular processes in EOfAD-like/+ brains which are not significantly altered in the null/+ brains, suggesting that these effects are due to neomorphic gain-of-function. These cellular processes are likely affected by the residual expression (after NMD) of N-terminal Sor11 protein domains without the C-terminal sequences. These effects span other pathways involved in energy production, such as glycolysis and the TCA cycle, cell adhesion, myc signalling, protein degradation and cell adhesion. However, we cannot rule out that these changes are not occurring in the null/+ brains, as the analysis of the null/+ brains had less statistical power than for the other genotypes ($n=4$ rather than $n=6$). For example, the myc targets and glycolysis gene sets had FDR adjusted p-values of approximately 0.1 in the null/+ brains. With $n=6$ or greater, these gene

sets may have been found to be statistically significantly altered.

Cellular processes potentially dependent on normal *sorl1* function

Our enrichment analysis has identified cellular processes not previously known to require normal *sorl1* function. We observed gene sets linked to energy production to be significantly altered in *sorl1* mutant brains relative to their wild type siblings. Genes involved in the oxidative phosphorylation pathway were observed to have altered expression as a group in all three *sorl1* mutant genotypes (Fig. 5). Expression of genes in this pathway were mostly reduced in the EOfAD-like/+ and the null/+ brains, suggesting that this effect is due to loss-of-function through haploinsufficiency. Intriguingly, expression of genes in the oxidative phosphorylation pathway was mostly increased in the transheterozygous mutant brains relative to their wild type siblings (Additional file 9). This opposite direction of change was also observed in other gene sets involving energy production (i.e. glycolysis in Fig. 6h and the TCA cycle in Fig. 6i). These effects are subtle, as the genes individually were not detected as differentially expressed (other than *cox7a1*) and the signal was only detected at the pathway level. Genes encoding the components of the electron transport chain have been shown previously to be dysregulated in early in AD [62–64], supporting that changes to mitochondrial function are an early pathological change in AD.

One explanation for changes in gene expression in the oxidative phosphorylation pathway could be changes in the function of the mitochondrial associated membranes (MAMs) of the endoplasmic reticulum. MAMs are lipid-raft like regions of the endoplasmic reticulum which form close associations with mitochondria and that regulate mitochondrial activity through calcium ion release (reviewed in [65–67]). It has been shown that the presenilins, APP, BACE1 (β -secretase) and other components of the γ -secretase complex are enriched in MAMs relative to the rest of the endoplasmic reticulum [68, 69]. SORL1 has been shown to form a complex with APP and BACE1 [70], and is also cleaved by γ -secretase [71]. We have shown previously that SORL1 is present in the MAM in the brains of mice [24] (see Additional file 10 for a reproduction of this result). SORL1 may bind APP in the MAM and affect its cleavage by BACE1 and γ -secretase (which can function as a supramolecular complex [72]). Whether mutations in *sorl1* affect normal MAM function is yet to be determined.

We also found evidence that iron homeostasis is affected by complete loss of wild type *sorl1* function. The transcripts of genes possessing a canonical or non-canonical IRE in their 3' UTR were found to be

dysregulated in transheterozygous mutant brains. It has generally been assumed that transcripts with an IRE in their 3' UTR have increased stability under cellular iron deficiency, and decreased stability under iron overload (reviewed in [12]). However, this is not necessarily the case and transcripts with 3' IREs can have increased or decreased stability under cellular iron deficiency [42]. Therefore, whether transheterozygous mutant brains are under cellular iron deficiency or overload is still unclear and further investigation is warranted.

The cellular response to hypoxia is mediated by hypoxia-inducible factor 1, α subunit (HIF1- α) in an iron dependent manner (HIF1- α is less stable under normoxia and in ferrous iron replete conditions). We also observed that genes involved in the cellular response to hypoxia are affected in transheterozygous mutant brains. Since oxygen and iron homeostasis are both involved in the stabilisation of HIF1- α , this further supports a role for *sorl1* in iron homeostasis. Yambire and colleagues, [55], showed that insufficient acidification of the endolysosomal system results in a pseudo-hypoxic response (stabilisation of the HIF1- α not due to a low oxygen environment) and the consequent ferrous iron deficiency led to mitochondrial instability and an increase in markers of inflammation. We also observed that genes involved in inflammation (HALLMARK_TNFA_SIGNALING_VIA_NFKB and HALLMARK_IL2_STAT5_SIGNALING gene sets) are significantly affected as a group in the transheterozygous mutant brains consistent with a role of *sorl1* function in iron homeostasis. Since the effects on iron homeostasis, hypoxic response and inflammation are most evident when both copies of wild type *sorl1* are lost, this implies that these effects may be recessive in nature.

Recently, Jiang and colleagues [73], demonstrated that increased expression of the β -CTF/C99 fragment of APP (produced by β -secretase cleavage, also known as “amyloidogenic processing” of APP) results in decreased acidification of the endolysosomal system in cells. SORL1 physically interacts with APP to modulate its amyloidogenic processing [74–76], and this suggests a molecular mechanism for SORL1's action in iron homeostasis. A more indirect action on iron homeostasis via modulation of APP's C99 fragment would be consistent with the generally later age of AD onset seen for carriers of putative EOfAD mutations in *SORL1* compared to carriers of EOfAD mutations in *APP* [16].

Surprisingly, we did not observe any endolysosomal gene sets to be significantly altered by mutation of *sorl1* mutant genotype despite that one of the primary sites of action of SORL1 protein is thought to be within the endolysosomal system [74]. Loss of *SORL1* by CRISPR-Cas9 mutagenesis in neurons (but not in microglia) derived from human induced pluripotent stem cells

(hiPSCs) resulted in early endosome enlargement [77], a commonly observed pathology in AD brains [78, 79]. To determine whether endolysosomal defects occur in our *sorl1* mutant zebrafish brains requires further investigation, as any effects may be too subtle to detect when analysing RNA-seq data from whole brains, or the effects may not be present until later ages.

Heterozygosity for an EOfAD-like mutation appears to have effects distinct from complete loss of *sorl1* function

Our results support that caution should be used when assuming that homozygous disease mutation model animals have similar, only more severe, phenotypes compared to heterozygous animals. We found little similarity between the transcriptomes of EOfAD-like/+ and transheterozygous mutant brains (that lack wild type *sorl1*, but may retain some mutant protein function). At the pathway level, most processes affected in the transheterozygous mutant brains appeared unaffected in the EOfAD-like/+ mutant brains and vice versa. Only the oxidative phosphorylation pathway was affected in both comparisons. However, as discussed previously, the overall changes in oxidative phosphorylation gene expression appeared to be in opposite directions in the brains of fish from the two genotypes, suggesting different mechanisms of action. This highlights the importance of using a genetic model closely resembling the genetic state of the human disease to make exploratory analyses of the effects of a human mutation in an unbiased manner. Hargis and Blalock [13], showed that some of the commonly used transgenic “mouse models of AD” (which do not closely reflect the genetic state of AD) have very little transcriptome concordance with human LOAD, or even with each other. We recently demonstrated little similarity between brain transcriptome changes in young adults of the popular 5xFAD transgenic mouse model and zebrafish heterozygous for an EOfAD-like, knock-in mutation of *psen1* [42]. The similar apparent effects of EOfAD-like mutations in *psen1* and in *sorl1* on oxidative phosphorylation in the young adult brains of zebrafish supports that knock-in models of EOfAD mutations may provide more consistent modelling of the early molecular pathogenesis of EOfAD.

Transcriptomic analyses of knock-in models of AD mutations are beginning to be appear in the literature. Our analysis of an EOfAD-like mutation in the zebrafish orthologue of *PSEN1* found that heterozygosity for this mutation affects expression of genes involved in acidification of the endolysosomal system, oxidative phosphorylation and iron homeostasis in the brain [42, 61]. Additionally, male knock-in *APOE* mice, which carry a humanised form of *APOE* (with genotype *APOE* $\epsilon 3/\epsilon 4$) associated with LOAD, showed changes to brain

gene expression in the oxidative phosphorylation pathway at 15 months of age [80]. Interestingly, these mice (and *APOE* $\epsilon 4/\epsilon 4$ mice) also show abnormalities in the endolysosomal system, which could mean that iron homeostasis is dysregulated. These results together suggest that iron dyshomeostasis and changes in oxidative phosphorylation may be unifying characteristics of AD, linking the rare, early-onset, familial subtype of AD with the common, late-onset, sporadic subtype.

Candidate genes for regulation by the SORL1 intracellular domain in the brain

Finally, it is worth noting that the genes detected as differentially expressed in transheterozygous mutant brains likely include candidates for regulation by the intracellular domain of Sorl1 protein (Sorl1-ICD). It has been shown that human SORL1 protein is cleaved by both α - and γ -secretases to produce various fragments [71, 81, 82]. The human SORL1-ICD has been demonstrated to enter the nucleus and regulate transcription [81]. However, the genes it regulates are yet to be identified. Both the EOfAD-like and null mutations of zebrafish *sorl1* are predicted to truncate Sorl1 protein so that it lacks the Sorl1-ICD sequences. Therefore, the genes identified as differentially expressed in the transheterozygous mutant brains are possible direct targets for regulation by the Sorl1-ICD and warrant further investigation.

Conclusion

We have made an initial, in vivo characterisation of changes in gene expression in the brain due to mutations in *sorl1*. Our results provide insight into novel cellular processes involving *sorl1*, such as energy production and iron homeostasis. It is possible that these changes are specific to mutations in *sorl1* in zebrafish. However, since this is the first in vivo analysis of a mutation in a *SORL1*-orthologous gene in any model organism, we must await similar studies in other models to support the results' translatability. Nevertheless, the changes observed due to the EOfAD-like mutation in zebrafish *sorl1* may represent early cellular stresses ultimately driving the development of AD in humans. Our results support both loss- and gain-of-function actions for EOfAD-like mutation of *sorl1*. The severity of *sorl1* mutation effects are more consistent with human *SORL1*'s action as a LOAD risk locus, since the effects are subtle compared to an EOfAD-like mutation in *psen1*. Future work may include “four-way” analyses of brain transcriptomes at older ages, where any effects may be more pronounced as well as further characterisation of those cellular processes we have identified as altered by mutations in *sorl1*.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13041-020-00681-7>.

Additional file 1: Genome editing of zebrafish *sorl1*. **a** and **b** show sections of the *sorl1* genomic sequence (ENSG00000137642) with the sense sequence (upper), translation (middle) and anti-sense sequence (lower). **a** *sorl1* exon 2, with the crRNA binding site and PAM sequence for cleavage by Cpf1 indicated by pink arrows, and the R122 site indicated by the orange bar. **b** *sorl1* exon 32, with the C1481 codon indicated by the orange arrow and the TALEN binding sites by blue arrows.

Additional file 2: Generation of mutant zebrafish lines. A detailed description of the isolation of the mutant lines of zebrafish.

Additional file 3: No observed bias for differential expression with GC content or length. Plots showing a ranking metric using the sign of logFC multiplied by $-\log_{10}$ of the p-value against **a** the GC content of the gene and **b** the length of the gene. The blue curve indicates the line of best fit from a generalised additive model (gam), whilst the black dashed line represents the $y=0$ line. Given the gam fit is a nearly horizontal line mostly overlapping $y=0$, a significant bias for GC content or length is likely not present in this dataset. The ranking statistic limits were constrained to -10 and 10 for visualisation purposes, and in **b**, gene length was plotted on the log₁₀ scale.

Additional file 4: Principal component analysis. **a** Plot of principal component 1 (PC1) against PC2 from a principle component analysis (PCA) of the logCPMs from each sample. **b** PC1 against the library size per sample. The linear relationship observed between PC1 and library size suggests that the largest source of variability in this dataset is due to library size. **c** PC1 against the W_1 covariate from RUVseq. A linear relationship between PC1 and W_1 is observed. **d** W_1 against library size. **e** PC1 against PC2 after removal of 1 factor of unwanted variation using RUVseq [38]. **f** PC1 no longer depends on library size after RUVseq transformation.

Additional file 5: Results of differential expression analysis. Each sheet within the spreadsheet gives the entire results of the differential gene expression due to each of the three *sorl1* genotype comparisons with wild type.

Additional file 6: Results of enrichment testing. Results of the individual enrichment testing algorithms (fry, camera and fgsea).

Additional file 7: Results from enrichment analysis are mostly not driven by expression of the same genes. The upset plots show the overlap of the "leading edge" genes from the fgsea algorithm in each of the significant gene sets in **a** EOfAD-like/+, **b** null/+, and **c** transheterozygous mutant brains. Intersections are shown when the leading edge of the gene sets share three or more genes. Overall, the leading edge genes of the significantly altered gene sets are relatively independent of one another. However, genes in the oxidative phosphorylation gene sets and the gene sets for neurodegenerative diseases (Alzheimer's, Parkinson's and Huntington's diseases) all contain genes encoding components of the electron transport chain and are capturing a portion of the same gene expression signal, which is shown in **d** and **e**. Missing positions in **d** and **e** indicate that the gene was not in the leading edge for that gene set.

Additional file 8: Changes to gene expression in young-adult zebrafish brain transcriptomes are likely not due to altered cell type proportions. We obtained representative expression markers of neurons (72 genes), oligodendrocytes (100 genes) and astrocytes (44 genes) from [53] and representative expression markers of microglia (533 genes) from [54]. The logCPM distributions for these marker genes in each of the samples are similar, supporting that, broadly, the distribution of these cell types is consistent between samples. Data are shown as violin plots, displaying the kernel probability density of the logCPMs, overlaid with boxplots, showing summary statistics, and coloured by genotype.

Additional file 9: Pathview visualisation of the KEGG oxidative phosphorylation gene set. The logFC of the genes in the KEGG oxidative phosphorylation gene set are shown in each *sorl1* genotype comparison with wild type. Intensity of the colours indicates the magnitude of the logFC, and

white indicates that a gene was not detected as expressed in the RNA-seq experiment. Plot was adapted from pathview [83].

Additional file 10: SORL1 is localized in the MAM in mouse brains. Western blot analysis of subcellular fractions of mouse brain cellular membranes. Each subcellular fraction probed with antibodies against **a** BACE1, **b** SORL1 and **c** LRP/LR. **d** Shows the identity of fractions western immunoblotting using antibodies recognising proteins MTCO1 in mitochondria (Mito), Na⁺/K⁺-ATPase in the plasma membrane (PM), IP3R3 in mitochondrial-associated membranes (MAM), and KDEL, predominantly in non-MAM endoplasmic reticulum (ER). Some cross contamination was observed as MTCO1 was detected in both the mitochondria and plasma membrane, and Na⁺/K⁺-ATPase in both the mitochondria and plasma membrane. Lim, A. H. L. (2015). Analysis of the subcellular localization of proteins implicated in Alzheimer's Disease. Genetics and Evolution, University of Adelaide. Doctor of Philosophy (PhD): 235. This figure is reproduced from the Ph.D. thesis of Anne Lim for ease of access.

Abbreviations

Aβ: Amyloid beta; AD: Alzheimer's disease; ANOVA: Analysis of variance; APOE: Apolipoprotein E; APP: Amyloid beta 4 precursor protein; BDNF: Brain-derived neurotrophic factor; bp: Base-pair; Cas9: CRISPR associated protein 9; cDNA: Complementary DNA; cox7a1: Cytochrome c oxidase subunit 7A1; cpf1: CRISPR associated protein 12a; cpm: Counts per million; CRISPR: Clustered regularly interspaced short palindromic repeats; cuedc1b: Cue domain containing 1b; DE: Differentially expressed; DNaseI: Deoxyribonuclease I; dqPCR: Digital quantitative polymerase chain reaction; EGF: Epidermal growth factor; EOfAD: Early-onset familial Alzheimer's disease; FC: Fold change; FN: Fibronectin; GEO: Gene Expression Omnibus; GSEA: Gene set enrichment analysis; HIF1-α: Hypoxia-inducible factor 1-alpha; hiPSCs: Human induced pluripotent stem cells; ICD: Intracellular domain; IRE: Iron-responsive element; KEGG: Kyoto Encyclopedia of Genes and Genomes; LDLR: Low density lipoprotein receptor; LOAD: Late-onset Alzheimer's disease; MAM: Mitochondrial-associated membrane; mRNA: Messenger RNA; MSigDB: Molecular signatures database; mTORC: Mammalian target of rapamycin complex; NMD: Nonsense mediated mRNA-decay; nt: Nucleotide; PAM: Protospacer adjacent motif; PC: Principal component; PCA: Principal component analysis; PCR: Polymerase chain reaction; PSEN: Presenilin; RNA-seq: RNA sequencing; RT-PCRs: Reverse transcription polymerase chain reactions; SAHMRI: South Australian Health and Medical Research Institute; siRNA: Small interfering RNA; SORL1: Sortilin-related receptor 1; TALEN: Transcription activator-like effector nucleases; TCA: Tricarboxylic acid; TMD: Transmembrane domain; trans: Transheterozygous; UTR: Untranslated region; VPS10: Vacuolar protein sorting 10; WT: Wild type.

Acknowledgements

The authors would like to acknowledge Seyyed Hani Moussavi Nik for obtaining part of the funding for this work and Nhi Hin for providing the sets of zebrafish genes containing iron-responsive elements.

Authors' contributions

KB performed all experimental and bioinformatic analysis and drafted the manuscript. SMP supervised and provided advice on the bioinformatic analysis and MN and ML supervised all zebrafish work. All authors read and contributed to the final manuscript.

Funding

This work was funded partially by a Alzheimer's Australia Dementia Research Foundation Project Grant titled "Identifying early molecular changes underlying familial Alzheimer's disease" awarded on 1 March 2017 from Alzheimer's Australia Dementia Research Foundation (now called Dementia Australia). KB is supported by an Australian Government Research Training Program Scholarship. ML and MN were both supported by grants GNT1061006 and GNT1126422 from the National Health and Medical Research Council of Australia (NHMRC). ML and SP are employees of the University of Adelaide. Funding bodies did not play a role in the design of the study, data collection, analysis, interpretation or in writing the manuscript.

Availability of data and materials

The raw fastq files and the output from catchKallisto (i.e. the raw transcript abundances) have been deposited in Gene Expression Omnibus database

(GEO) under Accession Number GSE151999 available at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE151999>. All code to reproduce the RNA-seq analysis can be found at https://github.com/karissa-b/sorl1_4way_RNAseq_q_6m.

Ethics approval and consent to participate

All zebrafish work was conducted under the auspices of the University of Adelaide Animal Ethics Committee (permit numbers: S-2017-073 and S-2017-089) and Institutional Biosafety Committee (permit number 15037).

Consent for publication

Not applicable.

Competing interests

The authors have no financial or non-financial competing interests to declare.

Author details

¹ Alzheimer's Disease Genetics Laboratory, School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, SA 5005, Australia. ² Bioinformatics Hub, School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, SA 5005, Australia.

Received: 5 August 2020 Accepted: 6 October 2020

Published online: 19 October 2020

References

- Iturria-Medina Y, Sotero R, Toussaint P, Mateos-Pérez J, Evans A, Initiative AsDN. Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nat Commun*. 2016;7:1–14.
- Buee L, Hof PR, Bouras C, Delacourte A, Perl DP, Morrison JH, et al. Pathological alterations of the cerebral microvasculature in Alzheimer's disease and related dementing disorders. *Acta Neuropathol*. 1994;87(5):469–80.
- Hirai K, Aliev G, Nunomura A, Fujioka H, Russell RL, Atwood CS, et al. Mitochondrial abnormalities in Alzheimer's disease. *J Neurosci*. 2001;21(9):3017–23.
- Swerdlow RH, Khan SM. A "mitochondrial cascade hypothesis" for sporadic Alzheimer's disease. *Med Hypotheses*. 2004;63(1):8–20.
- Flannery PJ, Trushina E. Mitochondrial dysfunction in Alzheimer's disease and progress in mitochondria-targeted therapeutics. *Curr Behav Neurosci Rep*. 2019;6(3):88–102.
- Nordengen K, Kirsebom B-E, Henjum K, Selnes P, Gísladóttir B, Wettergreen M, et al. Glial activation and inflammation along the Alzheimer's disease continuum. *J Neuroinflamm*. 2019;16(1):46.
- Chaney A, Williams SR, Boutin H. In vivo molecular imaging of neuroinflammation in Alzheimer's disease. *J Neurochem*. 2019;149(4):438–51.
- Di Paolo G, Kim T-W. Linking lipids to Alzheimer's disease: cholesterol and beyond. *Nat Rev Neurosci*. 2011;12(5):284–96.
- Arimon M, Takeda S, Post KL, Svirsky S, Hyman BT, Berezovska O. Oxidative stress and lipid peroxidation are upstream of amyloid pathology. *Neurobiol Dis*. 2015;84:109–19.
- Foley P. Lipids in Alzheimer's disease: a century-old story. *Biochim Biophys Acta (BBA) Mol Cell Biol Lipids*. 2010;1801(8):750–3.
- Li Y, Jiao Q, Xu H, Du X, Shi L, Jia F, et al. Biometal dyshomeostasis and toxic metal accumulations in the development of Alzheimer's disease. *Front Mol Neurosci*. 2017;10:339.
- Lumsden AL, Rogers JT, Majd S, Newman M, Sutherland GT, Verdile G, et al. Dysregulation of neuronal iron homeostasis as an alternative unifying effect of mutations causing familial Alzheimer's disease. *Front Neurosci*. 2018;12:533.
- Hargis KE, Blalock EM. Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us? *Behav Brain Res*. 2017;322:311–28.
- Cruts M, Theuns J, Van Broeckhoven C. Locus-specific mutation databases for neurodegenerative brain diseases. *Hum Mutat*. 2012;33(9):1340–4.
- Jayne T, Newman M, Verdile G, Sutherland G, Munch G, Musgrave I, et al. Evidence for and against a pathogenic role of reduced gamma-secretase activity in familial Alzheimer's disease. *J Alzheimer's Dis*. 2016;52:781–99.
- Pottier C, Hannequin D, Coutant S, Rovelet-Lecruc A, Wallon D, Rousseau S, et al. High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. *Mol Psychiatry*. 2012;17(9):875–9.
- Thonberg H, Chiang H-H, Lilius L, Forsell C, Lindström A-K, Johansson C, et al. Identification and description of three families with familial Alzheimer disease that segregate variants in the SORL1 gene. *Acta Neuropathol Commun*. 2017;5(1):43.
- Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. 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*. 1997;278(16):1349–56.
- Lee JH, Barral S, Reitz C. The neuronal sortilin-related receptor gene SORL1 and late-onset Alzheimer's disease. *Curr Neurol Neurosci Rep*. 2008;8(5):384.
- Miyashita A, Koike A, Jun G, Wang L-S, Takahashi S, Matsubara E, et al. SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians. *PLoS ONE*. 2013;8(4):e58618.
- Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat Genet*. 2019;51(3):414–30.
- Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 2013;45:1452.
- Campion D, Charbonnier C, Nicolas G. SORL1 genetic variants and Alzheimer disease risk: a literature review and meta-analysis of sequencing data. *Acta Neuropathol*. 2019.
- Lim AHL. Analysis of the subcellular localization of proteins implicated in Alzheimer's disease. University of Adelaide; 2015.
- Jacobsen L, Madsen P, Moestrup SK, Lund AH, Tommerup N, Nykjaer A, et al. Molecular characterization of a novel human hybrid-type receptor that binds the alpha2-macroglobulin receptor-associated protein. *J Biol Chem*. 1996;271(49):31379–83.
- Barthelson K, Newman M, Lardelli M. Sorting out the role of the sortilin-related receptor 1 in Alzheimer's disease. *J Alzheimer's Dis Rep*. 2020;5:Preprint–18.
- Cuccaro ML, Carney RM, Zhang Y, Bohm C, Kunkle BW, Vardarajan BN, et al. SORL1 mutations in early- and late-onset Alzheimer disease. *Neuro Genet*. 2016;2(6):e116.
- Vardarajan BN, Zhang Y, Lee JH, Cheng R, Bohm C, Ghani M, et al. Coding mutations in SORL1 and Alzheimer disease. *Ann Neurol*. 2015;77(2):215–27.
- Young JE, Boulanger-Weill J, Williams DA, Woodruff G, Buen F, Revilla AC, et al. Elucidating molecular phenotypes caused by the SORL1 Alzheimer's disease genetic risk factor using human induced pluripotent stem cells. *Cell Stem Cell*. 2015;16(4):373–85.
- Zetsche B, Gootenberg Jonathan S, Abudayyeh Omar O, Slaymaker Ian M, Makarova Kira S, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*. 2015;163(3):759–71.
- Jiang H, Newman M, Lardelli M. The zebrafish orthologue of familial Alzheimer's disease gene PRESENILIN 2 is required for normal adult melanotic skin pigmentation. *PLoS ONE*. 2018;13(10):e0206155.
- Andrews S. FastQC: a quality control tool for high throughput sequence data. 2014.
- Ward CM, To TH, Pederson SM. ngsReports: a Bioconductor package for managing FastQC reports and other NGS related log files. *Bioinformatics*. 2020;36(8):2587–8.
- Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34:525.
- R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2019.
- Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2009;26(1):139–40.
- Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25.
- Risso D, Ngai J, Speed TP, Dudoit S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat Biotechnol*. 2014;32(9):896–902.

39. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov Jill P, Tamayo P. The molecular signatures database hallmark gene set collection. *Cell Syst*. 2015;1(6):417–25.
40. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
41. Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, et al. BioMart—biological queries made easy. *BMC Genom*. 2009;10(1):22.
42. Hin N, Newman M, Pederson SM, Lardelli MM. Iron responsive element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease. *bioRxiv*. 2020:2020.05.01.071498.
43. Gross C, Dubois-Pot H, Wasylyk B. The ternary complex factor Net/Elk-3 participates in the transcriptional response to hypoxia and regulates HIF-1 alpha. *Oncogene*. 2008;27(9):1333–411.
44. Wu D, Lim E, Vaillant F, Asselin-Labat M-L, Visvader JE, Smyth GK. ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics*. 2010;26(17):2176–82.
45. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res*. 2012;40(17):e133–e.
46. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47–e.
47. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci*. 2005;102(43):15545.
48. Sergushichev AA. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv*. 2016:060012.
49. Wilson DJ. The harmonic mean p-value for combining dependent tests. *Proc Natl Acad Sci*. 2019;116(4):1195.
50. Wickham H. *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag; 2016.
51. Kolde R. *heatmap: pretty heatmaps*. 1.0.12 ed. 2019.
52. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*. 2017;33(18):2938–40.
53. Cahoy JD, Emery B, Kaushal A, Foo LC, Zamanian JL, Christopherson KS, et al. A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci*. 2008;28(1):264–78.
54. Oosterhof N, Holtman IR, Kuil LE, van der Linde HC, Boddeke EWGM, Eggen BJL, et al. Identification of a conserved and acute neurodegeneration-specific microglial transcriptome in the zebrafish. *Glia*. 2017;65(1):138–49.
55. Yambire KF, Rostovsky C, Watanabe T, Pacheu-Grau D, Torres-Odio S, Sanchez-Guerrero A, et al. Impaired lysosomal acidification triggers iron deficiency and inflammation in vivo. *Elife*. 2019;8:e51031.
56. Frankish A, Vullo A, Zadissa A, Yates A, Thormann A, Parker A, et al. *Ensembl* 2018. *Nucleic Acids Res*. 2017;46(D1):D754–D761761.
57. Lu J, Peatman E, Wang W, Yang Q, Abernathy J, Wang S, et al. Alternative splicing in teleost fish genomes: same-species and cross-species analysis and comparisons. *Mol Genet Genom*. 2010;283(6):531–9.
58. Le Guennec K, Tubeuf H, Hannequin D, Wallon D, Quenez O, Rousseau S, et al. Biallelic loss of function of SORL1 in an early onset Alzheimer's disease patient. *J Alzheimer's Dis*. 2018;62(2):821–31.
59. Arnaudo E, Hirano M, Sedan RS, Milatovich A, Hsieh C-L, Fabrizi GM, et al. Tissue-specific expression and chromosome assignment of genes specifying two isoforms of subunit VIIa of human cytochrome c oxidase. *Gene*. 1992;119(2):299–305.
60. Ryman DC, Acosta-Baena N, Aisen PS, Bird T, Danek A, Fox NC, et al. Symptom onset in autosomal dominant Alzheimer disease: a systematic review and meta-analysis. *Neurology*. 2014;83(3):253–60.
61. Newman M, Hin N, Pederson S, Lardelli M. Brain transcriptome analysis of a familial Alzheimer's disease-like mutation in the zebrafish presenilin 1 gene implies effects on energy production. *Mol Brain*. 2019;12(1):43.
62. Manczak M, Park BS, Jung Y, Reddy PH. Differential expression of oxidative phosphorylation genes in patients with Alzheimer's disease. *NeuroMol Med*. 2004;5(2):147–62.
63. Mastroeni D, Khodour OM, Delvaux E, Nolz J, Olsen G, Berchtold N, et al. Nuclear but not mitochondrial-encoded oxidative phosphorylation genes are altered in aging, mild cognitive impairment, and Alzheimer's disease. *Alzheimer's Dementia*. 2017;13(5):510–9.
64. Lunnon K, Keohane A, Pidsley R, Newhouse S, Riddoch-Contreras J, Thubron EB, et al. Mitochondrial genes are altered in blood early in Alzheimer's disease. *Neurobiol Aging*. 2017;53:36–47.
65. Vance JE. MAM (mitochondria-associated membranes) in mammalian cells: lipids and beyond. *Biochim Biophys Acta (BBA) Mol Cell Biol Lipids*. 2014;1841(4):595–609.
66. Hayashi T, Rizzuto R, Hajnoczky G, Su T-P. MAM: more than just a house-keeper. *Trends Cell Biol*. 2009;19(2):81–8.
67. Schon EA, Area-Gomez E. Mitochondria-associated ER membranes in Alzheimer disease. *Mol Cell Neurosci*. 2013;55:26–36.
68. Area-Gomez E, de Groof AJ, Boldogh I, Bird TD, Gibson GE, Koehler CM, et al. Presenilins are enriched in endoplasmic reticulum membranes associated with mitochondria. *Am J Pathol*. 2009;175(5):1810–6.
69. Newman M, Wilson L, Verdile G, Lim A, Khan I, Moussavi Nik SH, et al. Differential, dominant activation and inhibition of Notch signalling and APP cleavage by truncations of PSEN1 in human disease. *Hum Mol Genet*. 2014;23(3):602–17.
70. Spoelgen R, von Arnim CA, Thomas AV, Peltan ID, Koker M, Deng A, et al. Interaction of the cytosolic domains of sorLA/LR11 with the amyloid precursor protein (APP) and beta-secretase beta-site APP-cleaving enzyme. *J Neurosci*. 2006;26(2):418–28.
71. Nyborg AC, Ladd TB, Zwizinski CW, Lah JJ, Golde TE. Sortilin, SorCS1b, and SorLA Vps10p sorting receptors, are novel γ -secretase substrates. *Mol Neurodegener*. 2006;1(1):3.
72. Liu L, Ding L, Rovere M, Wolfe MS, Selkoe DJ. A cellular complex of BACE1 and γ -secretase sequentially generates A β from its full-length precursor. *J Cell Biol*. 2019;jcb.201806205.
73. Jiang Y, Sato Y, Im E, Berg M, Bordi M, Darji S, et al. Lysosomal dysfunction in down syndrome is APP-dependent and mediated by APP- β CTF (C99). *J Neurosci*. 2019;39(27):5255.
74. Andersen OM, Reiche J, Schmidt V, Gotthardt M, Spoelgen R, Behlke J, et al. Neuronal sorting protein-related receptor sorLA/LR11 regulates processing of the amyloid precursor protein. *Proc Natl Acad Sci USA*. 2005;102(38):13461–6.
75. Andersen OM, Schmidt V, Spoelgen R, Gliemann J, Behlke J, Galatis D, et al. Molecular dissection of the interaction between amyloid precursor protein and its neuronal trafficking receptor SorLA/LR11. *Biochemistry*. 2006;45(8):2618–28.
76. Mehmedbasic A, Christensen SK, Nilsson J, Rüttschi U, Gustafsen C, Poulsen ASA, et al. SorLA complement-type repeat domains protect the amyloid precursor protein against processing. *J Biol Chem*. 2015;290(6):3359–76.
77. Knupp A, Mishra S, Martinez R, Braggin JE, Szabo M, Kinoshita C, et al. Depletion of the AD risk gene SORL1 selectively impairs neuronal endosomal traffic independent of amyloidogenic APP processing. *Cell Rep*. 2020;31(9):107719.
78. Cataldo AM, Petanceska S, Peterhoff CM, Terio NB, Epstein CJ, Villar A, et al. App gene dosage modulates endosomal abnormalities of Alzheimer's disease in a segmental trisomy 16 mouse model of down syndrome. *J Neurosci*. 2003;23(17):6788–92.
79. Cataldo AM, Peterhoff CM, Troncoso JC, Gomez-Isla T, Hyman BT, Nixon RA. Endocytic pathway abnormalities precede amyloid beta deposition in sporadic Alzheimer's disease and Down syndrome: differential effects of APOE genotype and presenilin mutations. *Am J Pathol*. 2000;157(1):277–86.
80. Nuriel T, Peng KY, Ashok A, Dillman AA, Figueroa HY, Apuzzo J, et al. The endosomal-lysosomal pathway is dysregulated by APOE4 expression in vivo. *Front Neurosci*. 2017;11:702.
81. Böhm C, Seibel NM, Henkel B, Steiner H, Haass C, Hampe W. SorLA signaling by regulated intramembrane proteolysis. *J Biol Chem*. 2006;281(21):14547–533.
82. Hampe W, Riedel IB, Lintzel J, Bader CO, Franke I, Schaller HC. Ectodomain shedding, translocation and synthesis of SorLA are stimulated by its ligand head activator. *J Cell Sci*. 2000;113(Pt 24):4475–85.
83. Luo W, Pant G, Bhavnasi YK, Blanchard SG Jr, Brouwer C. Pathview Web: user friendly pathway visualization and data integration. *Nucleic Acids Res*. 2017;45(W1):W501–W508508.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

These files are linked to the published manuscript in *Molecular Brain*.

<https://doi.org/10.1186/s13041-020-00681-7>

Additional file 1:

Genome editing of zebrafish *sorl1*. **a** and **b** show sections of the *sorl1* genomic sequence (ENSG00000137642) with the sense sequence (upper), translation (middle) and anti-sense sequence (lower). **a** *sorl1* exon 2, with the crRNA binding site and PAM sequence for cleavage by Cpf1 indicated by pink arrows, and the R122 site indicated by the orange bar. **b** *sorl1* exon 32, with the C1481 codon indicated by the orange arrow and the TALEN binding sites by blue arrows.

Additional file 2:

Generation of mutant zebrafish lines. A detailed description of the isolation of the mutant lines of zebrafish.

Additional file 3

No observed bias for differential expression with GC content or length. Plots showing a ranking metric using the sign of logFC multiplied by $-\log_{10}$ of the p-value against **a** the GC content of the gene and **b** the length of the gene. The blue curve indicates the line of best fit from a generalised additive model (gam), whilst the black dashed line represents the $y = 0$ line. Given the gam fit is a nearly

horizontal line mostly overlapping $y = 0$, a significant bias for GC content or length is likely not present in this dataset. The ranking statistic limits were constrained to -10 and 10 for visualisation purposes, and in **b**, gene length was plotted on the \log_{10} scale.

[Additional file 4:](#)

Principal component analysis. **a** Plot of principal component 1 (PC1) against PC2 from a principle component analysis (PCA) of the logCPMs from each sample. **b** PC1 against the library size per sample. The linear relationship observed between PC1 and library size suggests that the largest source of variability in this dataset is due to library size. **c** PC1 against the W_1 covariate from RUVseq. A linear relationship between PC1 and W_1 is observed. **d** W_1 against library size. **e** PC1 against PC2 after removal of 1 factor of unwanted variation using RUVSeq [38]. **f** PC1 no longer depends on library size after RUVSeq transformation.

[Additional file 5:](#)

Results of differential expression analysis. Each sheet within the spreadsheet gives the entire results of the differential gene expression due to each of the three *sor11* genotype comparisons with wild type.

[Additional file 6:](#)

Results of enrichment testing. Results of the individual enrichment testing algorithms (fry, camera and fgsea).

[Additional file 7:](#)

Results from enrichment analysis are mostly not driven by expression of the same genes. The upset plots show the overlap of the “leading edge” genes from the fgsea algorithm in each of the significant gene sets in **a** EOfAD-like/+, **b** null/+, and **c** transheterozygous mutant brains. Intersections are shown when the leading edge of the gene sets share three or more genes. Overall, the leading edge genes of the significantly altered gene sets are relatively independent of one another. However, genes in the oxidative phosphorylation gene sets and the gene sets for neurodegenerative diseases (Alzheimer’s, Parkinson’s and Huntington’s diseases) all contain genes encoding components of the electron transport chain and are capturing a portion of the same gene expression signal, which is shown in **d** and **e**. Missing positions in **d** and **e** indicate that the gene was not in the leading edge for that gene set.

[Additional file 8](#)

Changes to gene expression in young-adult zebrafish brain transcriptomes are likely not due to altered cell type proportions. We obtained representative expression markers of neurons (72 genes), oligodendrocytes (100 genes) and astrocytes (44 genes) from [53] and representative expression markers of microglia (533 genes) from [54]. The logCPM distributions for these marker genes in each of the samples are similar, supporting that, broadly, the distribution of these cell types is consistent between samples. Data are shown as violin plots, displaying the kernel probability density of the logCPMs, overlaid with boxplots, showing summary statistics, and coloured by genotype.

[Additional file 9:](#)

Pathview visualisation of the KEGG oxidative phosphorylation gene set. The logFC of the genes in the KEGG oxidative phosphorylation gene set are shown in each *sorl1* genotype comparison with wild type. Intensity of the colours indicates the magnitude of the logFC, and white indicates that a gene was not detected as expressed in the RNA-seq experiment. Plot was adapted from pathview [83].

[Additional file 10:](#)

SORL1 is localized in the MAM in mouse brains. Western blot analysis of subcellular fractions of mouse brain cellular membranes. Each subcellular fraction probed with antibodies against **a** BACE1, **b** SORL1 and **c** LRP/LR. **d** Shows the identity of fractions western immunoblotting using antibodies recognising proteins MTCO1 in mitochondria (Mito), Na⁺/K⁺-ATPase in the plasma membrane (PM), IP3R3 in mitochondrial-associated membranes (MAM), and KDEL, predominantly in non-MAM endoplasmic reticulum (ER). Some cross contamination was observed as MTOC1 was detected in both the mitochondria and plasma membrane, and Na⁺/K⁺-ATPase in both the mitochondria and plasma membrane. Lim, A. H. L. (2015). Analysis of the subcellular localization of proteins implicated in Alzheimer's Disease. Genetics and Evolution, University of Adelaide. Doctor of Philosophy (PhD): 235. This figure is reproduced from the Ph.D. thesis of Anne Lim for ease of access.

Chapter 6: 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

Statement of Authorship

Title of Paper	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
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	This manuscript was published in Journal of Alzheimer's Disease on 26 December 2020. Barthelson, K., Pederson, S.M., Newman, M., and Lardelli, M. (2020). 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 Preprint, 1-15.

Principal Author

Name of Principal Author (Candidate)	Karissa Barthelson		
Contribution to the Paper	Generation of the mutant fish, isolation of RNA for RNA-seq, data analysis, generated the figures and drafted of the manuscript		
Overall percentage (%)	95%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	9/1/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Stephen Martin Pederson		
Contribution to the Paper	Supervision of bioinformatic analysis. Editing of manuscript		
Signature		Date	12/01/2021

Name of Co-Author	Morgan Newman		
Contribution to the Paper	Supervision of zebrafish work. Editing of manuscript		
Signature		Date	15/02/2021

Name of Co-Author	Michael Lardelli		
Contribution to the Paper	Supervision of project, editing of manuscript		
Signature		Date	18/02/2021

Please cut and paste additional co-author panels here as required.

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

Karissa Barthelson^{a,*}, Stephen Martin Pederson^{b,1}, Morgan Newman^a and Michael Lardelli^a
^a*Alzheimer's Disease Genetics Laboratory, School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, SA, Australia*
^b*Bioinformatics Hub, School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, SA, Australia*

Accepted 13 November 2020
Pre-press 26 December 2020

Abstract.

Background: The early cellular stresses leading to Alzheimer's disease (AD) remain poorly understood because we cannot access living, asymptomatic human AD brains for detailed molecular analyses. Sortilin-related receptor 1 (*SORL1*) encodes a multi-domain receptor protein genetically associated with both rare, early-onset familial AD (EOfAD) and common, sporadic, late-onset AD (LOAD). *SORL1* protein has been shown to act in the trafficking of the amyloid β A4 precursor protein (A β PP) that is proteolysed to form one of the pathological hallmarks of AD, amyloid- β (A β) peptide. However, other functions of *SORL1* in AD are less well understood.

Objective: To investigate the effects of heterozygosity for an EOfAD-like mutation in *SORL1* on the brain transcriptome of young-adult mutation carriers using zebrafish as a model organism.

Methods: We performed targeted mutagenesis to generate an EOfAD-like mutation in the zebrafish orthologue of *SORL1* and performed RNA-sequencing on mRNA isolated from the young adult brains of siblings in a family of fish either wild type (non-mutant) or heterozygous for the EOfAD-like mutation.

Results: We identified subtle differences in gene expression indicating changes in mitochondrial and ribosomal function in the mutant fish. These changes appear to be independent of changes in mitochondrial content or the expression of A β PP-related proteins in zebrafish.

Conclusion: These findings provided evidence supporting that EOfAD mutations in *SORL1* affect mitochondrial and ribosomal function and provide the basis for future investigation elucidating the nature of these effects.

Keywords: Alzheimer's disease, mitochondria, ribosome, RNA-seq, *sorl1*, zebrafish

*Correspondence to: Karissa Barthelson, Room 1.24, Molecular Life Sciences Building, North Terrace Campus, The University of Adelaide, SA 5005, Australia. Tel.: +61 83134863; E-mail: karissa.barthelson@adelaide.edu.au.

¹Current affiliation: Dame Roma Mitchell Cancer Research Laboratories, Adelaide Medical School, Faculty of Health and Medical Sciences, University of Adelaide, SA, Australia.

INTRODUCTION

To reduce the prevalence and, therefore, the socio-economic impacts of Alzheimer's disease (AD), we must understand its molecular basis. Analysis of post-mortem, human AD brains can give some insight into the disease mechanism. However, by the end stages of the disease, damage to the brain is considerable and little therapeutic intervention is possible. An understanding of the molecular changes that occur decades before symptoms arise is necessary for development of effective preventative or ameliorative treatments. However, we cannot access living, pre-symptomatic AD patient brain tissue for detailed molecular analysis. Therefore, analysis of animal models of AD is required.

A small number of genes are known to influence strongly the development of AD. Mutations in the *presenilins* (*PSEN1* and *PSEN2*), and the gene encoding amyloid β A4 precursor protein (*A β PP*), are causative for the early-onset (<65 years of age), familial form of AD (EOfAD). EOfAD is an autosomal dominant disorder, and accounts for only a small proportion of all AD cases. The vast majority of AD cases arise sporadically and have an age of onset later than 65 years (LOAD). Variation in at least 20 genes is associated with increased risk of developing LOAD, with the ϵ 4 allele of the *apolipoprotein E* gene (*APOE*) contributing the greatest risk [1, 2]. Since EOfAD and LOAD show similar disease progression and pathology (reviewed in [3]), analysis of EOfAD animal models may give insight into both subtypes of AD. Intriguingly, the gene *sortilin-related receptor 1* (*SORL1*) appears to be associated with both EOfAD and LOAD [1, 2, 4–10] and may provide a mechanistic link between these two AD subtypes. However, *SORL1* is the least studied of the EOfAD genes and its role in AD is unclear.

The majority of research investigating the role of *SORL1* in AD is based around the role of its protein product in the intracellular trafficking of A β PP. It is well accepted that the binding of A β PP to *SORL1* protein determines whether A β PP is directed through a recycling pathway, or is steered through the endolysosomal system to generate amyloid- β (A β , the primary component of the senile neuritic plaques found in AD brains) (reviewed in [11]). *SORL1* also binds A β itself, guides it to the lysosome for degradation [12], and can act as a receptor for *APOE* [13]. These findings have been made mostly using cell lines in which *SORL1* is overexpressed or removed. However, these manipulations do not closely reflect

the pathophysiological state of *SORL1* expression in AD.

We recently published an investigation of the effects on young adult zebrafish brain transcriptome state of heterozygosity for an EOfAD-like mutation and/or a putatively null mutation of the zebrafish orthologue of *SORL1* [14]. Heterozygosity for the EOfAD-like mutation resulted in subtle effects on the young adult brain transcriptome, with only one gene detected as differentially expressed. Analysis at the level of genetic pathways (gene sets) found evidence for changes in cellular processes previously unknown to require *sorl1*, such as energy metabolism, protein translation and degradation. These effects were also observed in the brains of fish heterozygous for the putatively loss-of-function mutation in *sorl1*, suggesting they are due to haploinsufficiency. Transheterozygosity for the EOfAD-like and null mutations gave apparent effects on iron homeostasis and other cellular processes distinct from those detected in the heterozygous fish brains [14].

Here, we aimed to further our understanding of the effects of EOfAD mutations in *SORL1* by generation and analysis of an additional zebrafish *sorl1* mutation model. We performed targeted mutagenesis to generate a line of zebrafish carrying a W1818* mutation in *sorl1*, equivalent to the human *SORL1* EOfAD mutation W1821* [4]. We then compared the brain transcriptomes of young adult *sorl1*^{W1818*} heterozygous and wild type sibling fish to identify, in an unbiased and objective manner, the changes in gene expression caused by this mutation. Consistent with our previous work, heterozygosity for the W1818* mutation results in subtle changes to brain gene expression. Genetic pathways involved in energy production and protein translation were altered. The W1818* mutation of *sorl1* did not appear to affect *Appa/Appb* protein levels, cellular mitochondrial content, or iron homeostasis.

MATERIALS AND METHODS

Zebrafish husbandry and animal ethics

All zebrafish used in this study were maintained in a recirculating water system on a 14 h light/10 h dark cycle, fed dry food in the morning and live brine shrimp in the afternoon. All zebrafish work was conducted under the auspices of the Animal

Ethics Committee (permit numbers S-2017–089 and S-2017–073) and the Institutional Biosafety Committee of the University of Adelaide.

Genome editing

To introduce mutations at the W1818 site in zebrafish *sor11*, we used a TALEN pair designed and purchased from Zgenebio Biotech Inc. (Taipei City, Taiwan). The target genomic DNA sites (5' to 3') were ATGAGGTGGCGGTGTG (left) and GTAGGACTGTCTCCAT (right) (Supplementary Figure 1). The DNAs encoding the TALEN pairs were supplied in the pZGB2 vector, which were linearized with *NotI* (NEB, Ipswich, USA) before transcription of mRNA *in vitro* using the mMESAGE mMACHINE T7 *in vitro* transcription kit following the manufacturer's protocol (Invitrogen, Carlsbad, USA). The mRNAs encoding the TALEN pairs were diluted to a final concentration of 200 ng/ μ L each, and approximately 2–5 nL of the TALEN pair mRNA solution was injected into zebrafish embryos at the one cell stage. We eventually isolated a family of fish carrying an 11 nucleotide deletion in *sor11*, W1818*, following the strategy described in [14]. Briefly, injected embryos were raised until approximately three months of age, then were outcrossed to Tübingen strain fish of approximately the same age to generate F1 families. Since the injected fish are mosaic for any mutations at or near the W1818 codon, we performed a T7 endonuclease I (T7EI) assay (NEB) on ten F1 embryos from a clutch to confirm whether any mutations had occurred in the germline of the injected parent. The PCR primers used to amplify the W1818 region for the T7EI assay have sequences (5' – 3') F: TTAGGACCTCCTGTGTCAGCATTCT and R: ACAAATAAAAAGTGTATGTGC. If the T7EI assay indicated that mutations were indeed in the germline, the remaining F1 embryos of the clutch were raised to adulthood. Sanger sequencing of the W1818 site was performed on genomic DNA extracted from fin biopsies of F1 adult fish to characterize any mutations (sequencing was performed by the Australian Genome Research Facility (AGRF, Adelaide, AUS)). Single heterozygous mutant fish (W1818*/+) from the F1 family were mated with single wild type (+/+) fish to generate F2 families. Heterozygous mutants from F2 families were in-crossed to produce F3 families containing homozygous mutant individuals.

Hypoxia treatment and allele specific digital quantitative PCRs

Adult zebrafish (6 months of age) were subjected to a hypoxic environment by placement in oxygen-depleted water for 3 h (oxygen concentration of 6 ± 0.5 mg/L in normoxia and 1 ± 0.5 mg/L in hypoxia). After treatment, fish were immediately sacrificed in a loose ice slurry and their brains removed. Total RNA was then extracted from whole zebrafish brains using the QIAGEN RNeasy® Mini Kit (Qiagen, Venlo, Netherlands) according to the manufacturer's protocol. Recovered RNA concentrations were estimated using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific Inc, Waltham, MA, USA). cDNA was synthesized using random hexamers in the Superscript III First Strand Synthesis System (Invitrogen) according to the manufacturer's instructions. 50 ng of each resulting cDNA sample was then used in allele-specific digital quantitative PCR (dqPCR) as described in [15]. To detect the two transcripts of *sor11*, we used a W1818* mutation-specific forward primer with sequence 5' GTGGCGGTGTGATGATGG 3', and a wild type-specific primer with sequence 5' TGGCGGTGTGGGCTCAC 3'. A common reverse primer sequence spanning the junction between exons 41 and 42 of *sor11* (to reduce amplification from any genomic DNA carried over during RNA extraction) was 5' GTAGAACACAGCGTACATCTCTGC 3'. Comparisons of mean transcript abundances per 50 ng of brain-derived cDNA were made using a two-way ANOVA with Tukey's *post hoc* test for multiple comparisons.

Western blot analysis

Whole zebrafish brains were placed in 100 μ L of 1X Complete Protease Inhibitor Cocktail (Roche Holding AG, Basel, Switzerland) in RIPA buffer (Sigma-Aldrich Corp. St. Louis, MO, USA), immediately homogenized using a handheld motorized pestle for 1 min on ice, then incubated at 4°C for 2 h with gentle rocking. Cellular debris was sedimented by centrifugation with a relative centrifugal force of 16,100 g for 10 min. 25 μ L of 4X NuPAGE™ LDS Sample Buffer (Thermo Fisher Scientific) was added to the supernatant (75 μ L) containing total protein, then samples were heated at 80°C for 20 min. Total protein concentrations were determined using the EZQ® Protein Quantitation Kit (Molecular Probes, Inc. Eugene, OR, USA) following the manufacturers

Table 1
Antibodies used in this study

Primary antibody	Dilution	Secondary antibody	Dilution
Goat Anti-SORL1 / LR11 (C Terminus (Aviva Systems Biology, San Diego, CA, USA))	1:500	Anti-goat IgG, HRP-conjugated Antibody (Rockland Immunochemicals Inc., Limerick, PA, USA)	1:10,000
Anti-VDAC (Invitrogen)	1:500	Anti-rabbit IgG, HRP-conjugated Antibody (Cell Signalling Technology, Danvers, MA, USA)	1:10,000
Anti-APP (clone 22C11) (Sigma)	1:3000	Anti-mouse IgG, HRP-conjugated Antibody (Rockland Immunochemicals)	1:10,000
Anti- β -Tubulin (E7, deposited to the DSHB by Klymkowsky, M. (DSHB Hybridoma Product E7))	1:200	Anti-mouse IgG, HRP-conjugated Antibody (Rockland Immunochemicals)	1:10,000

protocol. Samples were prepared for SDS-PAGE by adding 2.5 μ L of 10X NuPAGE™ Sample Reducing Agent (Invitrogen) to 75 μ g of total protein. The volume was increased to 25 μ L with 1X LDS buffer in RIPA buffer. Samples were heated at 80°C for 10 min and centrifuged briefly to sediment insoluble material. Supernatants containing the soluble proteins were subjected to SDS-PAGE on NuPAGE™ 4–12% Bis-Tris Protein Gels (Invitrogen) in the Mini Gel Tank and Blot Module Set (Thermo Fisher Scientific), using 1X NuPAGE™ MOPS SDS Running Buffer (Invitrogen). Resolved proteins were then transferred to PVDF membrane in 1X tris glycine SDS, 20% methanol. The PVDF membranes were blocked in 5% Western Blocking Reagent (Roche) then probed with primary and secondary antibodies. Antibody incubations were performed at either 1 hour at room temperature, or overnight at 4°C with gentle rocking. Antibody dilutions can be found in Table 1. Horse-radish peroxidase (HRP) signals were developed using Pierce™ ECL Western Blotting Substrate (Thermo Fisher Scientific) and imaged with a ChemiDoc™ MP Imaging System (Bio-Rad Laboratories, Hercules, CA, USA). The intensities of the signals were measured using the Volume tool in Image Lab 5.1 (Bio-Rad Laboratories). The relative intensities of the signals were normalized to β -tubulin and compared across genotypes by one-way ANOVA tests. Original blot images can be found in Supplementary Figure 7.

RNA-seq analysis

We used a total of 12 fish in the RNA-seq analysis (i.e., $n=6$ per genotype, with mostly equal numbers of males and females), based on a previous power calculation indicating that $n=6$ would provide approximately 70% power to detect the majority of expressed transcripts in a zebrafish brain transcrip-

toe at a fold-change >2 and at a false discovery rate of 0.05 (data not shown). Whole heads were removed and preserved in approximately 600 μ L of RNAlater (Thermo Fisher Scientific). Total RNA was extracted from the brains using the *mirVana* miRNA isolation kit (Thermo Fisher Scientific) following the manufacturer's protocol. Then, any genomic DNA carried over during the total RNA extraction was removed using the DNA-free™ Kit (Ambion, Waltham, USA). RNA was stabilized using RNastable® (Biomatrix, San Diego, CA, USA) following the manufacturer's protocol to minimize RNA degradation. Briefly, RNA was resuspended in DEPC-treated water, applied directly to a RNastable® tube and dried with a vacuum concentrator. Dried RNA was sent to Novogene (Hong King, China) for cDNA library synthesis and RNA-sequencing. The cDNA libraries were generated using NEBNext Ultra RNA Library Prep Kit for Illumina (NEB) and RNA was sequenced using the Illumina Novaseq PE150 platform.

Demultiplexed paired-end fastq files (with adapter sequences removed) were supplied by Novogene. We first inspected the quality of the raw data by *fastQC* and *ngsReports* [16]. We then performed a pseudo alignment using *kallisto* [17] version 0.43.1, in paired end mode specifying 50 bootstraps. The index file for *kallisto* was generated according to the zebrafish transcriptome (primary assembly of GRCz11, Ensembl release 96), with the sequences for all unspliced transcripts additionally included. We imported the transcript counts estimated by *kallisto* for analysis using *R* [18], using the *catchKallisto* function of the package *edgeR* [19]. To obtain gene level counts, we summed the counts of all transcripts arising from a single gene. Normalization factors were calculated using the trimmed mean of M-values (TMM) method [20]. Lowly expressed genes are statistically uninformative as they provide little evidence for differential expression, and we considered genes to be

detectably expressed if they contained at least contained a logCPM of 0.75 in at least 6 of the 12 RNA-seq libraries. After selecting detectable genes, library sizes ranged from 13,050,237 to 50,560,749. Although these library sizes varied considerably, a correlation between library size and the first two principal components was not observed, supporting that variation due to library size does not contribute to the two largest sources of variation in this dataset (Supplementary Figure 6).

We performed an initial differential expression analysis using the exact test function of *edgeR* [19]. A design matrix was specified with an intercept for each brain sample, and *sor11* genotype (W1818*/+) as the common difference. Only one differentially expressed gene was identified in this analysis (data not shown). Therefore, to assist with identification of dysregulated genes due to *sor11* genotype, we removed one factor of unwanted variation using the *RUVg* method of the package *RUVSeq* [21]. For *RUVg*, we set $k=1$, and the negative control genes as the least 5000 differentially expressed genes (i.e., largest p -value) in the initial differential expression test. The W_1 offset term generated was then included in the design matrix for an additional differential expression test using a generalized linear model and likelihood ratio tests with *edgeR* [19, 22]. We considered genes differentially expressed if their FDR adjusted p -value was less than 0.05.

Gene set enrichment testing was performed using three different algorithms: *fry* [23], *camera* [24], and *GSEA* [25, 26]. Since each method gave different levels of significance, we combined the raw p -values from each method by calculating the harmonic mean p -value [27]. We considered a gene set to be significantly altered if the FDR adjusted harmonic mean p -value was less than 0.05. The gene sets we tested in this study were the *KEGG* [28], *HALLMARK* [29], and the zebrafish iron-responsive element [30] gene sets. The *KEGG* and *HALLMARK* gene sets for zebrafish were obtained from *MSigDB* [29] using *msigdbR* [31]. To perform promoter motif enrichment analysis, we used *homer* [32] as described in [33] on the top 100 most statistically significant differentially expressed genes due to *sor11* genotype when including the W_1 covariate in the design matrix.

Data visualization was performed using the packages *ggplot2* [34], *heatmap* [35], and *upsetR* [36].

The raw fastq files and the output from *catch Kallisto* (i.e., the raw transcript abundances) have been deposited in Gene Expression Omnibus database (GEO) under Accession Number GSE156

167. All code to reproduce the RNA-seq analysis can be found at https://github.com/karissa-b/sor11-w1818x_6m.

RESULTS

Generation and characterization of the EOfAD-like mutation W1818 in zebrafish sor11*

In 2012, Pottier et al. [4] identified mutations in *SORL1* segregating with EOfAD in families with autosomal dominant inheritance patterns. We previously analyzed the effects of a coding sequence-truncating mutation from this study (human C1478*) on the young adult brain transcriptome in a zebrafish model (zebrafish V1482Afs) [14]. To further our understanding of the effects of coding sequence-truncating mutations in *SORL1*, we generated an additional zebrafish model of a mutation from the Pottier et al. study, W1821* [4]. The W1821 codon is conserved in zebrafish (W1818). We edited this site in the zebrafish genome using targeted mutagenesis (Supplementary Figure 1) and isolated a line of zebrafish carrying an 11 nucleotide deletion, resulting in a protein coding sequence-level change equivalent to that caused by the human mutation (Fig. 1A).

Protein coding sequence-truncating mutations in *SORL1* have been shown to be subject to nonsense mediated mRNA decay (NMD) in lymphoblasts from human carriers [4]. Additionally, *SORL1* expression is upregulated under hypoxia *in vitro* [37]. Therefore, we asked whether the W1818* transcript of *sor11* is also subject to NMD, and whether its expression is altered *in vivo* by hypoxia treatment. We performed allele-specific, digital quantitative PCRs (dqPCRs) on brain-derived cDNA from fish exposed to normoxia or acute hypoxia at 6 months of age. We found that the W1818* transcript is less abundant in W1818*/+ brains relative to the wild type transcript under both normoxia ($p < 0.0001$) and acute hypoxia ($p = 0.0001$), supporting that the W1818* transcript is subject to NMD. We also observed that the wild type transcript in W1818*/+ brains is less abundant than in +/+ brains under normoxia ($p = 0.004$) and acute hypoxia ($p = 0.02$). We did not observe any significant change of *sor11* transcript levels between normoxia and acute hypoxia (Fig. 1C), despite that these fish were likely showing a transcriptional response to hypoxia, as indicated by upregulation of *pdki* transcript levels (Supplementary Figure 2).

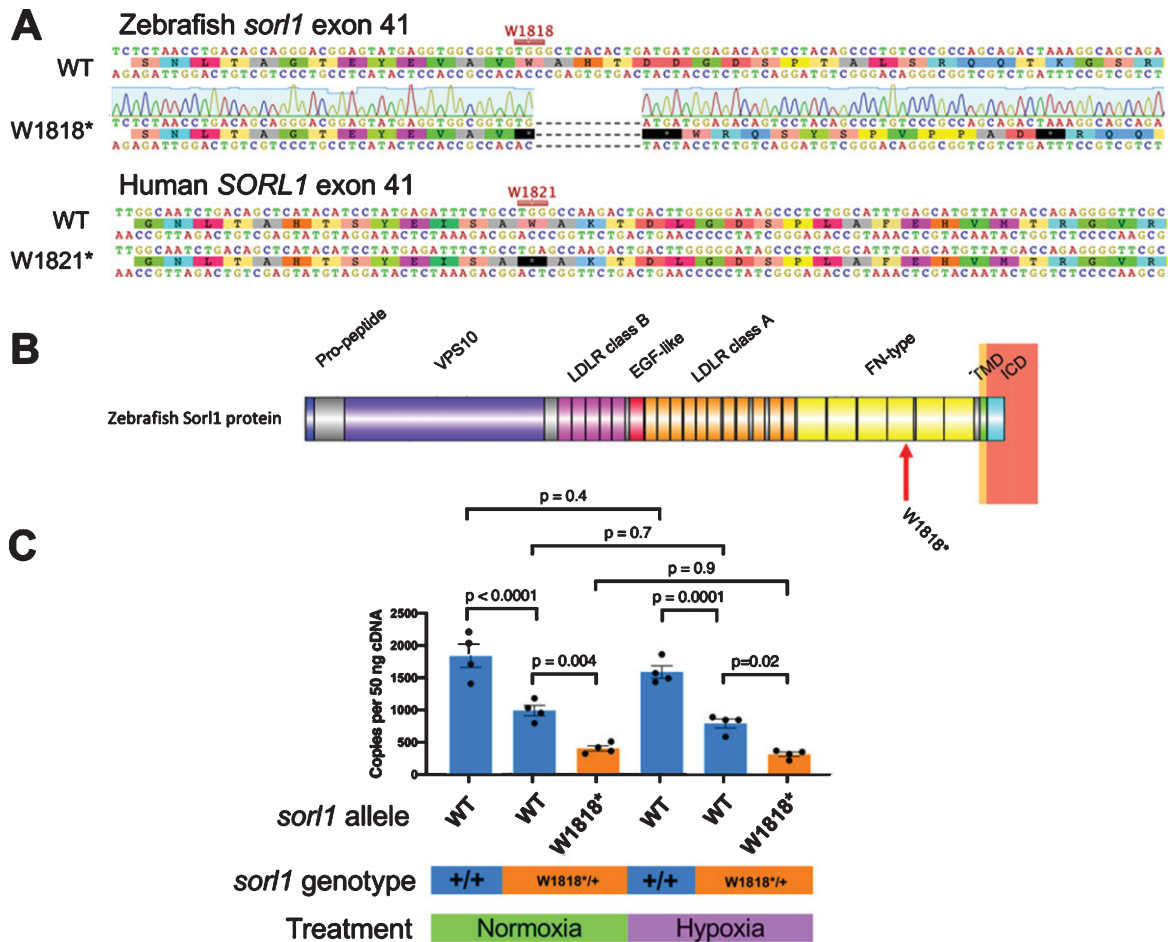


Fig. 1. The W1818* mutation in *sor1l* results in a transcript subject to nonsense mediated mRNA decay. A) Alignment of zebrafish *sor1l* and human *SORL1* wild type and mutant exon 41 sequences. B) Schematic of Sor1l protein with protein domains and the site of the W1818* mutation indicated. VPS10, vacuolar protein sorting 10 domain; LDLR, low density lipoprotein receptor; EGF, epidermal growth factor; FN, fibronectin; TMD, transmembrane domain; ICD, intracellular domain. C) Number of copies of the *sor1l* wild type (WT) and mutant (W1818*) transcripts in 6-month-old wild type (+/) and heterozygous mutant (W1818*/+) sibling brains under normoxia and acute hypoxia. *p*-values were determined by two-way ANOVA with Tukey's *post hoc* test for multiple comparisons.

To determine whether the W1818* mutation alters the expression of Sor1l protein, we performed western immunoblot analysis on zebrafish brain lysates at 6 months of age, using an antibody raised against the C-terminus of human SORL1 (and that cross-reacts with zebrafish Sor1l protein). A ~250 kDa signal is observed in +/+ brains, and is absent in homozygous mutant brains, confirming that the W1818* mutation disrupts translation of Sor1l protein (Supplementary Figure 3).

In summary, *sor1l* does not appear to be upregulated by acute hypoxia *in vivo*, the W1818* results in a transcript which is likely subject to NMD, and the mutation disrupts wild type Sor1l protein translation.

Transcriptome analysis of W1818*/+ and +/+ young adult zebrafish brains

Which genes are dysregulated as a result of heterozygosity for the EOFAD-like, W1818* mutation in *sor1l* in young adult brains? To address this question, we performed mRNA sequencing on the brains from fish either heterozygous for the W1818* mutation and their wild type siblings at 6 months of age ($n = 6$).

The overall similarity between gene expression profiles can be explored using principal component analysis (PCA). Samples with similar gene expression profiles cluster together in a PCA plot. In our RNA-seq dataset, no clear separation of male and

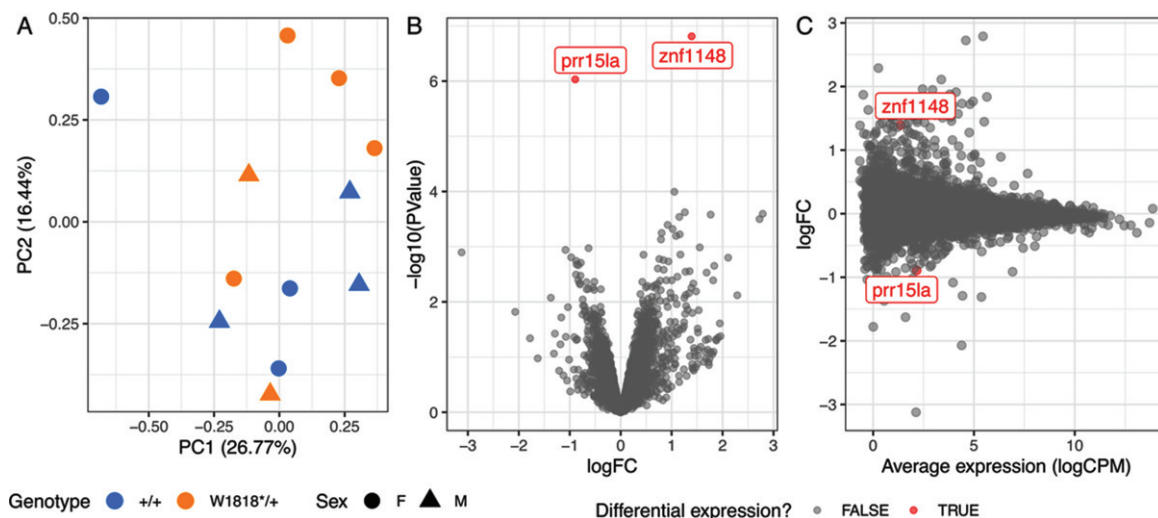


Fig. 2. Heterozygosity for the W1818* mutation in *sorl1* has subtle effects on young-adult brain transcriptomes. A) Plot of principal component 1 (PC1) against PC2 from a principal component analysis. Circles indicate female samples (F) and triangles indicate male (M) samples. B) Volcano plot showing the significance ($-\log_{10}(P\text{-Value})$) and \log_2 fold change (logFC) of genes in W1818*/+ brains relative to +/+ sibling brains. The genes identified as significantly differentially expressed are colored red. C) Mean-difference (MD) plot of gene expression in W1818*/+ brains relative to +/+ sibling brains.

female brain samples was observed, supporting our previous observations that sex does not have a large effect on the zebrafish brain transcriptome [14, 38]. We observed partial separation of W1818*/+ and wild type samples across principal component 2 (PC2), accounting for approximately 16% of the total variation in this dataset (Fig. 2A). Therefore, heterozygosity for the W1818* mutation likely does not have widespread effects on the transcriptome of young adult zebrafish brains.

Differential gene expression analysis supported our conclusions from the PCA. Only two genes were identified as differentially expressed to a statistically significant degree due to the W1818*/+ genotype: *zinc finger protein 1148* (*znf1148*) and *proline rich 15 like a* (*prrr15la*) (Fig. 2B, C and Supplementary Table 1). The functions of these genes are not known. BLAST searches identified *ZNF99* and *PRR15L* as candidate human orthologues of *znf1148* and *prrr15la* respectively. However, broad-scale analysis of conservation of synteny between the genomic regions occupied by the zebrafish and human genes was unable to support these orthologies (Supplementary Figure 4).

To obtain a more complete view of changes to brain gene expression due to heterozygosity for the W1818* mutation in *sorl1* we performed enrichment testing on particular gene sets (below) using the entire list of detectable genes in the RNA-seq experiment.

For insight into the cellular processes represented by the gene expression changes we used the *KEGG* and *HALLMARK* gene sets from the *Molecular Signatures Database* (*MSigDB*, [26, 29]). To test for possible iron dyshomeostasis, we used our recently defined gene sets representing the genes containing iron-responsive elements (IRE) in the untranslated regions (UTRs) of their mRNAs [30]. We applied the self-contained gene set testing methods *fry* [23] and *GSEA* [25, 26], and the competitive gene set testing method *camera* [24], and combined the resulting *p*-values by calculating the harmonic mean *p*-value, a method of combining dependent *p*-values [27]. We further protected against type I errors by performing an FDR adjustment on the resulting harmonic mean *p*-value.

In all, we identified six gene sets significantly altered as a group after FDR adjustment of the harmonic mean *p*-value (Fig. 3). However, the subsets of “leading edge” genes identified by the *GSEA* algorithm (which can be thought of as those genes driving the enrichment of the gene set) for four of these six gene sets (*HALLMARK_* and *KEGG_OXIDATIVE_PHOSPHORYLATION_*, *KEGG_HUNTINGTONS_DISEASE_*, and *KEGG_PARKINSONS_DISEASE_*) all share many genes, showing that the statistical significance of these gene sets is, essentially, being driven by the same signal (genes encoding components of the mitochondrial electron transport chain).

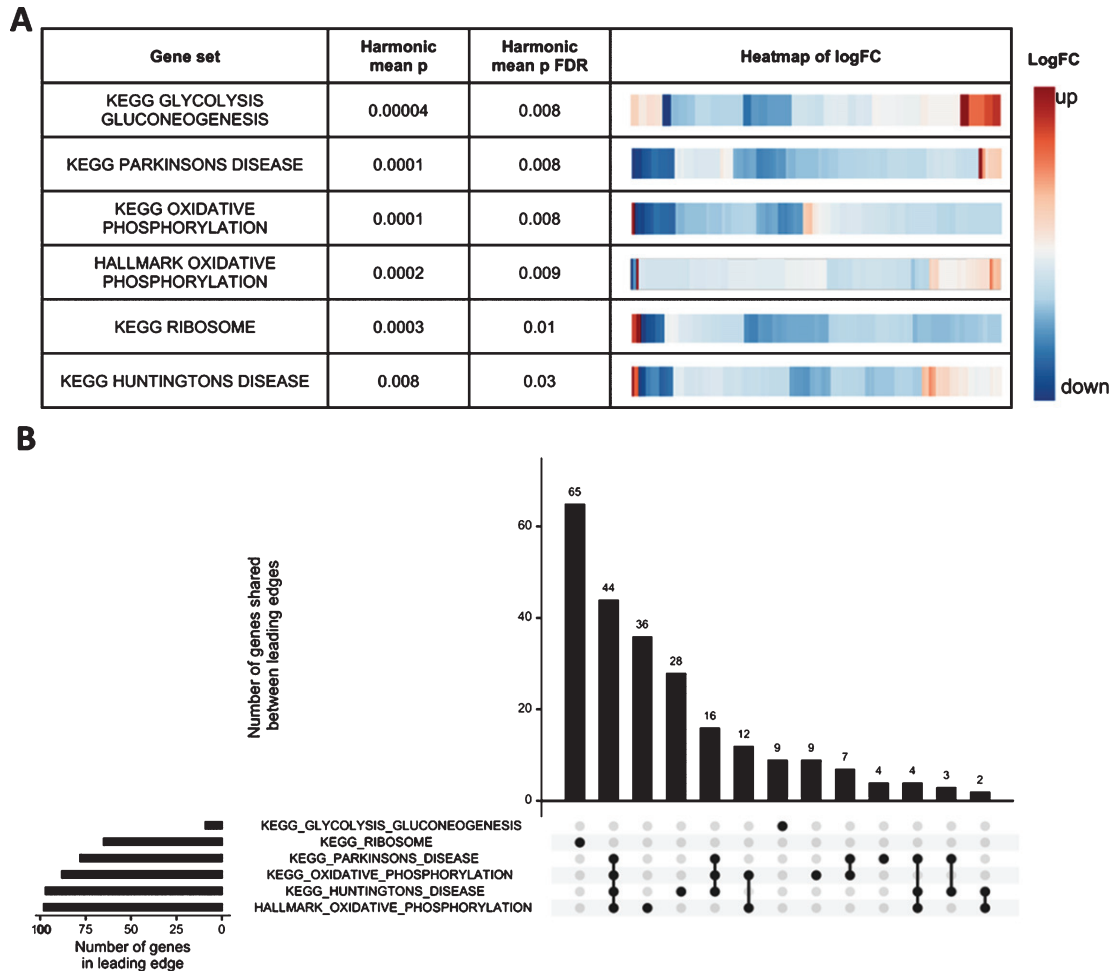


Fig. 3. Mitochondrial and ribosomal function are predicted to be affected by heterozygosity for the W1818* mutation in *sor11*. A) Table indicating the significantly altered KEGG and HALLMARK gene sets (FDR adjusted harmonic mean p -value < 0.05) in W1818*/+ brains relative to +/+ brains. Heatmaps indicate the log₂ fold change (logFC) of all detectable genes within the gene sets, clustered by their Euclidean distance. B) Upset plot of the leading edge genes from the GSEA algorithm for the significant gene sets indicating the number of shared genes driving the enrichment of each gene set. See the online version for color.

No IRE gene sets were found to be significantly altered, consistent with our previous observations for heterozygosity for the V1482Afs EOfAD-like allele of *sor11* [14].

Since our transcriptome analysis was of bulk mRNA isolated from entire zebrafish brains, differences in cell type proportions between W1818*/+ and +/+ brains could cause artefactual apparent changes in gene expression levels. To investigate this possibility, we visualized the expression levels of marker genes present in four broad cell types within zebrafish brains: neurons, astrocytes, oligodendrocytes [39] and microglia [40] (Fig. 4A). Since we did not observe any obvious differences in marker gene

expression levels between genotypes, it is unlikely that changes in cell type proportions cause the differential expression of genes observed in the W1818* heterozygous fish brains.

Differences in transcription factor activity could also be driving the changes to gene expression observed in W1818*/+ brains. To explore this, we performed promoter motif enrichment analysis using *homer* [32] on the 100 most statistically significantly differentially expressed genes due to *sor11* genotype. We identified that the promoter motif for hepatocyte nuclear factor 4 α (Hnf4 α) appears significantly over-represented in the promoters of the top 100 most DE genes (Bonferroni adjusted p -value = 0.04).

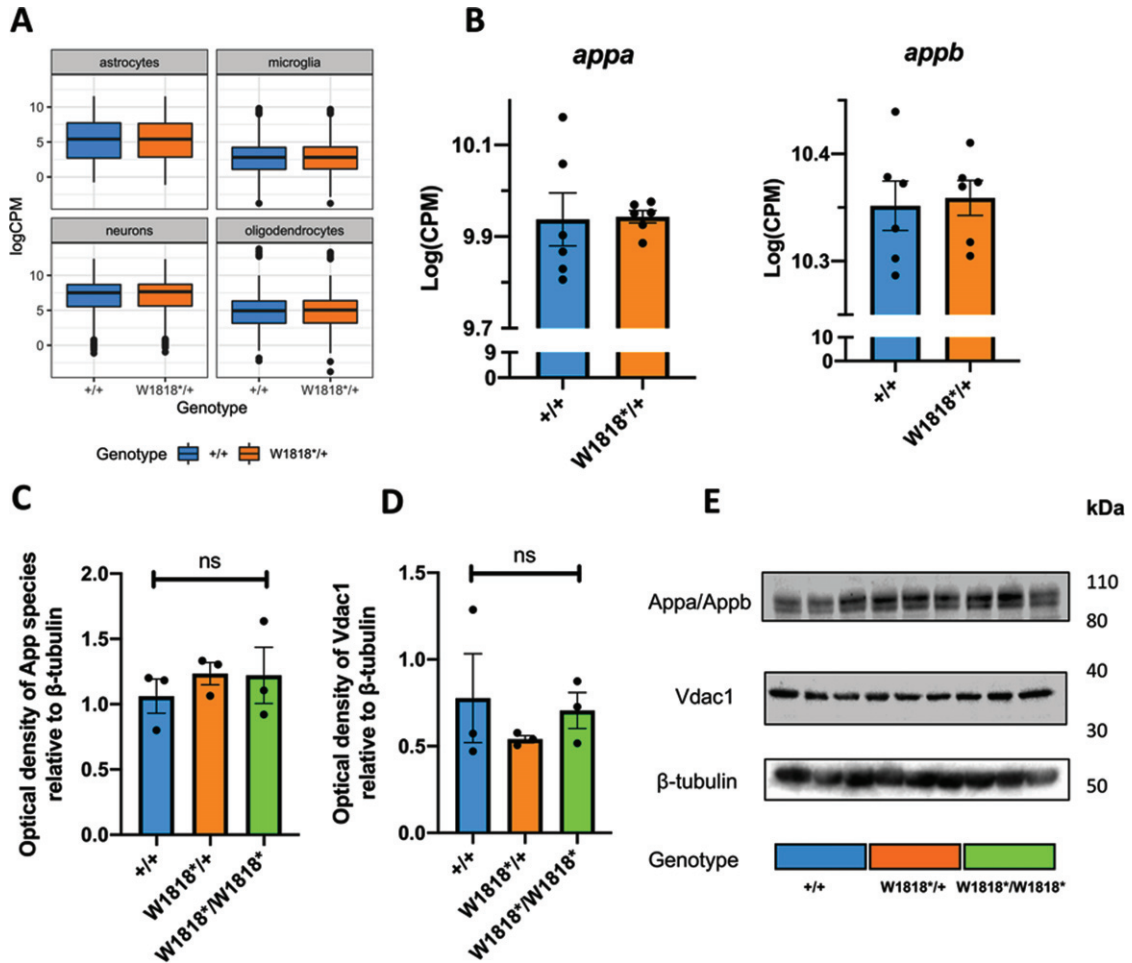


Fig. 4. Brain transcriptome changes in W1818*/+ mutants are unlikely to be due to changes in cell type proportions, altered expression of *appa/b*, or mitochondrial content. A) Distribution of expression (logCPM) of marker genes of astrocytes, microglia, neurons and oligodendrocytes in wild type (+/+) and heterozygous mutant (W1818*/+) brains. B) Expression (logCPM) of A β PP orthologues *appa* and *appb* in wild type and heterozygous mutant brains. C) Quantification of western blot analysis of expression of App proteins (Appa and Appb) in 6-month-old wild type, heterozygous and homozygous mutant zebrafish brains. D) Quantification of western immunoblot analysis of expression of the mitochondrial marker Vdac1 in 6-month-old wild type, heterozygous and homozygous mutant zebrafish brains. E) Representative images of western immunoblots. Significance levels in C) and D) were determined by one-way ANOVA (ns, not significant).

Expression of *hnf4 α* itself was low in these zebrafish brains and was insufficient to be regarded as detectable in this RNA-seq experiment (<0.75 cpm in at least 6 of the 12 RNA-seq libraries). However, examination of the logCPM values for *hnf4 α* before filtering found that one mutant brain displayed far higher expression of *hnf4 α* than the others and is clearly an outlier. A PCA on the expression of all genes predicted to contain a Hnf4 α -binding motif (*HNF4ALPHA_Q6* gene set from the C3 category of *MSigDB*) revealed no obvious separation between W1818*/+ and +/+ samples, suggesting that *sorl1*

genotype does not result in distinct expression patterns for these genes (Supplementary Figure 5).

In summary, heterozygosity for the W1818* mutation appears to have subtle effects on the expression of genes involved in energy production and protein translation in young adult zebrafish brains. The subtle effects of this EOfAD-like mutation are consistent with the typically lesser pathogenicity of such mutations in *SORL1* compared to, for example, EOfAD mutations in *PSEN1* [6, 41] and may reflect the earliest cellular changes that, after decades, lead to overt Alzheimer's disease.

Heterozygosity or homozygosity for W1818 does not appear to affect the abundance of App species, or mitochondrial mass*

The most characterized cellular role of SORL1 protein is the sorting of A β PP throughout the endolysosomal system (reviewed in [11]). Therefore, we sought to investigate whether the W1818* mutation of *sorl1* causes changes in expression of the zebrafish forms of A β PP. In zebrafish, two paralogous genes show orthology with human A β PP: *appa* and *appb*. Inspection of our RNA-seq data did not reveal any differences in the expression levels of *appa* or *appb* between heterozygous mutant and wild type brains (Fig. 4B). We also used an antibody against human A β PP that cross-reacts with both zebrafish Appa and Appb proteins for immunoblot examination of App species in the brains of wild type, heterozygous and homozygous mutant zebrafish at 6 months of age (Fig. 4C). We did not observe any significant differences between genotypes, supporting that the changes to *sorl1* gene function/expression are likely independent of changes to expression of *appa* or *appb*.

Finally, we hypothesized that, in W1818*/+ brains, changes to expression of genes involved in oxidative phosphorylation could cause, or be due to, changes in cellular mitochondrial content. To explore this, we performed western immunoblotting against Vdac1, a protein highly abundant in the outer mitochondrial membrane, in lysates from entire zebrafish brains. We did not observe any significant effect of *sorl1* genotype on the levels of Vdac1, supporting that cellular mitochondrial content is unchanged between genotypes (Fig. 4D).

DISCUSSION

We cannot access tissue from the living brains of *SORL1* mutation carriers for detailed molecular analysis. This restricts our ability to elucidate the early changes which eventually lead to AD and requires that we examine animal models instead. Our particular approach involves generating genetic models of EOfAD mutations in zebrafish which closely mimic the genetic state of human EOfAD (i.e., here we analyzed the effects of a single, heterozygous mutation in the endogenous *sorl1* gene of zebrafish). This approach avoids potentially confounding assumptions such as that homozygous animals simply show more extreme phenotypes than heterozygotes, or that transgenic animals overexpressing mutant genes

cause effects similar to endogenous human mutations (assumptions commonly made in AD genetics research). We also performed our analyses when zebrafish are 6 months old and recently sexually mature. We regard this as the equivalent of early adulthood in humans. At this age changes to gene expression should reflect pathological processes (or the responses to these) occurring long before any cognitive deficits become apparent. Finally, analysis of large families of sibling fish raised in an identical environment (the same tank) reduces both genetic and environmental variation between samples (i.e., noise) and increases our ability to detect subtle changes in gene expression due to heterozygosity for the EOfAD-like mutations.

In this study, we built on our previous analysis [14] investigating the effects of protein coding sequence-truncating mutations in *sorl1* implicated in EOfAD. We generated a zebrafish model of the W1821* mutation in human *SORL1* (zebrafish *sorl1* W1818*). We showed that transcripts of the W1818* allele are likely subject to NMD (consistent with observations of coding sequence-truncating mutations in human *SORL1* [4]), and that W1818* transcripts cannot generate full-length Sorl1 protein.

We did not observe any significant differences between the expression of *sorl1* transcripts in zebrafish brains under normoxia and hypoxia *in vivo*. This contrasts with the observations of Nishii et al. [37], using a hematopoietic stem cell line. These researchers saw increased SORL1 protein levels under hypoxia. However, a recent meta-analysis of RNA-seq datasets investigating transcriptional responses to hypoxia in both humans (128 datasets) and mice (52 datasets) showed that *SORL1* was seldom found to be differentially expressed under hypoxia. When *SORL1* was identified as DE, either up- or downregulation was observed [42]. Therefore, whether *SORL1* is regulated by hypoxia, and in which cell types this may occur, requires further investigation.

Heterozygosity for EOfAD-like mutations in sorl1 causes subtle changes in young adult brains

Transcriptome analysis of young adult, W1818*/+ mutant zebrafish brains relative to their +/+ siblings identified statistical evidence for two DE genes: *znf1148* and *prr15la*. Little or no information on the function of these genes can be found in the scientific literature. However, the candidate orthologue of *prr15la* in humans, *PRR15L*, was seen

as associated with human intelligence in a meta-analysis of genome-wide association studies totaling 78,308 individuals [43]. That only a small number of genes was identified as DE is to be expected from the apparently semi-penetrant effects of EOfAD mutations in *SORL1* (reviewed in [11, 44] and see below) and the young age of the brains examined (long before any overt pathology would be expected). It is also consistent with our previous analysis of another EOfAD-like mutation in *sor11* (zebrafish *sor11*^{V1482Afs} modelling human *SORL1*^{C1478*}), for which heterozygosity gave rise to only one significantly DE (downregulated) gene: cytochrome c oxidase subunit 7A1 (*cox7a1*) [14]. This gene appeared as slightly, but non-significantly, upregulated in W1818*/+ brains (logFC=0.5, p_{FDR} = 1) (Supplementary Table 1) so the differences in the genes identified as DE between the W1818* and V1482Afs heterozygous mutant brains is likely an effect of statistical noise. In contrast, analysis of an EOfAD-like mutation in the zebrafish orthologue of *PSEN1*, the gene most commonly mutated in EOfAD [45], detected 251 DE genes in the brains of young adult heterozygous zebrafish [46]. Thus, it appears that EOfAD mutations in *SORL1* cause less severe effects on cellular state compared to EOfAD mutations in *PSEN1*. This is more consistent with *SORL1*'s action as a LOAD genetic risk locus than as an EOfAD-causative locus. Campion and colleagues [44], noted that the ages of onset of AD in patients with mutations in *SORL1* are proximal to the arbitrary LOAD age threshold of 65 years of age (between 56 and 80 years, compared with *PSEN1* mutation carriers, which have an overall median age of disease onset of approximately 42 years [47]). Also, some *SORL1* mutation carriers with early onset AD carry additional LOAD genetic risk variants in *APOE*, *TREM2*, and *ABCA7*, and/or have aged (>66 years), unaffected siblings who carry the same *SORL1* mutation. Together, these observations, along with our *in vivo* transcriptomic studies, support that *SORL1* is more likely a genetic risk locus for LOAD than an EOfAD locus.

A promotor enrichment analysis of the 100 genes showing the strongest signal (smallest *p*-values) for differential expression due to *sor11* genotype identified the DNA-binding motif for Hnf4 α as significantly enriched. *HNF4 α* , a gene associated with maturity-onset diabetes of the young (MODY, a form of non-insulin-dependent diabetes mellitus [48]), is primarily expressed in the liver (reviewed in [49]) and has been reported to play roles in gluconeogene-

sis [50] and cholesterol homeostasis [51]. Yamanishi and colleagues [52] reported that *HNF4 α* is also expressed in the brains of mice where it appears to drive transcriptome changes in mice housed under stressful conditions. Expression of the *hnf4 α* gene in the brains of our zebrafish was very low so that its transcripts were excluded during the pre-processing preceding our differential gene expression analysis. However, inspection of the logCPM values from before exclusion of genes with low expression identified one W1818*/+ male mutant sample with a relatively higher expression of *hnf4 α* (likely an outlier). Nevertheless, a PCA on the expression of all detectable genes predicted to contain the Hnf4 α DNA-binding motif (250 genes) did not reveal any convincing evidence of global dysregulation of Hnf4 α -regulated genes in W1818*/+ brains.

Changes to mitochondrial and ribosomal function are observed in young-adult, sor11 mutant zebrafish, and in human AD

Enrichment testing of all the detectable genes in the RNA-seq experiment identified subtle changes to expression of genes involved in energy production and protein translation. Reassuringly, we previously identified these processes as altered due to heterozygosity for the EOfAD-like V1482Afs mutation of *sor11*, and for a putatively null mutation in *sor11* in young adult zebrafish brains [14], suggesting that changes to mitochondrial and ribosomal function are an effect-in-common of EOfAD mutations in *SORL1*, and arise through a haploinsufficiency mechanism (i.e., due to decreased *sor11* function). Whether these mitochondrial and ribosomal functions are directly reliant on Sor11 protein expression, or change as part of a homeostatic response to a deficiency of another, unknown, Sor11-dependent function, is unclear. To our knowledge, the role of *SORL1* in the context of mitochondrial and ribosomal function has not been investigated previously. However, it is well accepted that these processes are affected early in AD. Studies exploiting positron emission tomography with 2-[¹⁸F] fluoro-2-deoxy-d-glucose (FDG-PET) (measuring glucose metabolism in the brains of living subjects) have shown that glucose metabolism gradually declines during the conversion from mild cognitive impairment (MCI) to AD, particularly in the parieto-temporal and posterior cingulate cortices [53–57]. In our zebrafish model, genes encoding the components of the mitochondrial electron transport chain are mostly downregulated. Interestingly, the

overall direction of change for genes involved in glycolysis and gluconeogenesis is up, supporting that a metabolic shift occurs from oxidative phosphorylation towards glycolysis as the primary source of energy for the brain.

In W1818*/+ brains, we observed downregulation of genes encoding ribosomal proteins. Ribosomes isolated from postmortem MCI and AD patients show less capacity for protein synthesis [58–60], increased levels of oxidized ribosomal RNA (rRNA), and bound ferrous iron (Fe^{2+}) relative to control ribosomes [61]. Oxidative stress within cells (such as due to increased redox active metals like Fe^{2+} , or free radicals generated during normal mitochondrial respiration) oxidises rRNA and decreases ribosomal activity [62]. Further investigation is needed to determine whether protein synthesis is actually affected in W1818*/+ brains.

We recently developed an enrichment analysis-based method to detect signs of iron dyshomeostasis in RNA-seq data. The method exploits gene sets encompassing genes containing iron-responsive elements (IREs) in the 5' or 3' UTRs of their mRNAs [30]. We did not observe any evidence for changes to expression of these gene sets, suggesting that intracellular iron (Fe^{2+}) levels are not significantly altered in W1818*/+ brains relative to their +/+ siblings. This is consistent with our previous analysis, where only complete loss of wild type *sor11* resulted in significant changes to expression of genes which IREs in the 3' UTRs of their mRNAs [14]. Nevertheless, we cannot exclude that iron homeostasis is affected in the heterozygous mutant brains. Due to the nature of bulk RNA-seq, we may have failed to observe opposite directions of change in gene expression in different cell types.

No evidence that mutation of sor11 significantly changes expression of genes affecting the endo-lysosomal system in young adult zebrafish brains

We did not observe any gene sets involved in endolysosomal system function to be significantly altered in W1818*/+ brains, despite that SORL1 protein is thought to act, primarily, within the endolysosomal system (reviewed in [11]). Knupp and colleagues [63] showed that complete loss of *SORL1* in neurons (and not microglia) derived from human induced pluripotent stem cells (hiPSCs) resulted in early endosome enlargement, a phenomenon previously observed in post mortem EOfAD and LOAD

brains [64–66]. Our previous four-genotype analysis also did not reveal any direct evidence for dysregulation of genes in the endolysosomal system [14] although iron homeostasis appeared disturbed in brains lacking any wild type *sor11* expression, and that might be due to a disturbance of cellular iron importation in which the endolysosomal system plays an important role (e.g., [67]). The results of our two independent zebrafish studies do not support that heterozygosity for mutations in *sor11* affects the endolysosomal system at the level of gene regulation. Nevertheless, changes may be occurring in our mutant fish at the protein level without affecting the transcriptome. Also, as mentioned previously, humans heterozygous for *SORL1* mutations generally show ages of AD onset (if affected) close to the arbitrary late onset threshold of 65 years of age [4, 44]. Therefore, it is possible that any endolysosomal defects in heterozygous W1818* zebrafish brains are not be observable until later ages.

In conclusion, we have provided evidence supporting that mutation of *sor11* affects mitochondrial and ribosomal function. Our bioinformatic analysis provides the basis for future experiments to elucidate the nature of these effects.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dr. Seyyed Hani Moussavi Nik for obtaining part of the funding for this work, and his assistance in the genome editing to generate the W1818* mutant line of zebrafish, Ms. Nhi Hin for providing the zebrafish iron-responsive element (IRE) gene sets and her valuable advice and discussions on the bioinformatic analysis, and Dr. Zijng Zhou for his assistance in performing the western blots. This work was supported with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide.

This work was funded partially by an Alzheimer's Australia Dementia Research Foundation Project Grant titled "Identifying early molecular changes underlying familial Alzheimer's disease" awarded on 1 March 2017 from Alzheimer's Australia Dementia Research Foundation (now named Dementia Australia). KB is supported by an Australian Government Research Training Program Scholarship. ML and MN were both supported by grants GNT1061006 and GNT1126422 from the National Health and Medical Research Council of Australia (NHMRC). ML and SP are employees of the University of Adelaide.

Authors' disclosures available online (<https://www.j-alz.com/manuscript-disclosures/20-1383>).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: <https://dx.doi.org/10.3233/JAD-201383>.

REFERENCES

- [1] Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, Boland A, Vronskaya M, Van Der Lee SJ, Amlie-Wolf A, et al. (2019) Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat Genet* **51**, 414-430.
- [2] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, DeStefano AL, Bis JC, Beecham GW, et al. (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452.
- [3] Masters CL, Bateman R, Blennow K, Rowe CC, Sperling RA, Cummings JL (2015) Alzheimer's disease. *Nat Rev Dis Primers* **1**, 15056.
- [4] Pottier C, Hannequin D, Coutant S, Rovelet-Lecrux A, Wallon D, Rousseau S, Legallic S, Paquet C, Bombois S, Pariente J, Thomas-Anterion C, Michon A, Croisile B, Etcharry-Bouyx F, Berr C, Dartigues JF, Amouyel P, Dauchel H, Boutoleau-Bretonniere C, Thauvin C, Frebourg T, Lambert JC, Campion D, Collaborators PG (2012) High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. *Mol Psychiatry* **17**, 875-879.
- [5] Thonberg H, Chiang H-H, Lilius L, Forsell C, Lindström A-K, Johansson C, Björkström J, Thordardottir S, Slegers K, Van Broeckhoven C, Rönnebeck A, Graff C (2017) Identification and description of three families with familial Alzheimer disease that segregate variants in the SORL1 gene. *Acta Neuropathol Commun* **5**, 43.
- [6] Verheijen J, Van den Bossche T, van der Zee J, Engelborghs S, Sanchez-Valle R, Lladó A, Graff C, Thonberg H, Pastor P, Ortega-Cubero S, Pastor MA, Benussi L, Ghidoni R, Binetti G, Clarimon J, Lleó A, Fortea J, de Mendonça A, Martins M, Grau-Rivera O, Gelpi E, Bettens K, Mateiu L, Dillen L, Cras P, De Deyn PP, Van Broeckhoven C, Slegers K (2016) A comprehensive study of the genetic impact of rare variants in SORL1 in European early-onset Alzheimer's disease. *Acta Neuropathol* **132**, 213-224.
- [7] Lee JH, Cheng R, Schupf N, Manly J, Lantigua R, Stern Y, Rogaeva E, Wakutani Y, Farrer L, St George-Hyslop P, Mayeux R (2007) The association between genetic variants in SORL1 and Alzheimer disease in an urban, multiethnic, community-based cohort. *Arch Neurol* **64**, 501-506.
- [8] Reitz C, Cheng R, Rogaeva E, Lee JH, Tokuhiko S, Zou F, Bettens K, Slegers K, Tan EK, Kimura R, Shibata N, Arai H, Kamboh MI, Prince JA, Maier W, Riemenschneider M, Owen M, Harold D, Hollingworth P, Cellini E, Sorbi S, Nacmias B, Takeda M, Pericak-Vance MA, Haines JL, Younkin S, Williams J, van Broeckhoven C, Farrer LA, St George-Hyslop PH, Mayeux R (2011) Meta-analysis of the association between variants in SORL1 and Alzheimer disease. *Arch Neurol* **68**, 99-106.
- [9] Rogaeva E, Meng Y, Lee JH, Gu Y, Kawarai T, Zou F, Katayama T, Baldwin CT, Cheng R, Hasegawa H, Chen F, Shibata N, Lunetta KL, Pardossi-Piquard R, Bohm C, Wakutani Y, Cupples LA, Cuenco KT, Green RC, Pinessi L, Rainero I, Sorbi S, Bruni A, Duara R, Friedland RP, Inzelberg R, Hampe W, Bujo H, Song Y-Q, Andersen OM, Willnow TE, Graff-Radford N, Petersen RC, Dickson D, Der SD, Fraser PE, Schmitt-Ulms G, Younkin S, Mayeux R, Farrer LA, St George-Hyslop P (2007) The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nat Genet* **39**, 168-177.
- [10] Wen Y, Miyashita A, Kitamura N, Tsukie T, Saito Y, Hattuta H, Murayama S, Kakita A, Takahashi H, Akatsu H, Yamamoto T, Kosaka K, Yamaguchi H, Akazawa K, Ihara Y, Kuwano R (2013) SORL1 is genetically associated with neuropathologically characterized late-onset Alzheimer's disease. *J Alzheimers Dis* **35**, 387-394.
- [11] Barthelson K, Newman M, Lardelli M (2020) Sorting out the role of the sortilin-related receptor 1 in Alzheimer's disease. *J Alzheimers Dis Rep* **4**, 123-140.
- [12] Caglayan S, Takagi-Niidome S, Liao F, Carlo A-S, Schmidt V, Burgert T, Kitago Y, Fuchtbauer E-M, Fuchtbauer A, Holtzman DM, Takagi J, Willnow TE (2014) Lysosomal sorting of amyloid- β by the SORLA receptor is impaired by a familial Alzheimer's disease mutation. *Sci Trans Med* **6**, 223ra220.
- [13] Yajima R, Tokutake T, Koyama A, Kasuga K, Tezuka T, Nishizawa M, Ikeuchi T (2015) ApoE-isoform-dependent cellular uptake of amyloid- β is mediated by lipoprotein receptor LR11/SorLA. *Biochem Biophys Res Commun* **456**, 482-488.
- [14] Barthelson K, Pederson SM, Newman M, Lardelli M (2020) 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. *Mol Brain* **13**, 142.
- [15] Jiang H, Newman M, Lardelli M (2018) The zebrafish orthologue of familial Alzheimer's disease gene PRESENILIN 2 is required for normal adult melanotic skin pigmentation. *PLoS One* **13**, e0206155.
- [16] Ward CM, To TH, Pederson SM (2020) ngsReports: A Bioconductor package for managing FastQC reports and other NGS related log files. *Bioinformatics* **36**, 2587-2588.
- [17] Bray NL, Pimentel H, Melsted P, Pachter L (2016) Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**, 525.
- [18] Team RC (2019) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [19] Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140.
- [20] Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25.
- [21] Risso D, Ngai J, Speed TP, Dudoit S (2014) Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat Biotechnol* **32**, 896-902.
- [22] McCarthy DJ, Chen Y, Smyth GK (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* **40**, 4288-4297.

- [23] Wu D, Lim E, Vaillant F, Asselin-Labat M-L, Visvader JE, Smyth GK (2010) ROAST: Rotation gene set tests for complex microarray experiments. *Bioinformatics* **26**, 2176-2182.
- [24] Wu D, Smyth GK (2012) Camera: A competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res* **40**, e133-e133.
- [25] Sergushichev AA (2016) An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv*, 060012.
- [26] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545.
- [27] Wilson DJ (2019) The harmonic mean *p*-value for combining dependent tests. *Proc Natl Acad Sci U S A* **116**, 1195.
- [28] Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30.
- [29] Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov Jill P, Tamayo P (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425.
- [30] Hin N, Newman M, Pederson SM, Lardelli MM (2020) Iron Responsive Element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease. *bioRxiv*, 2020.2005.2001.071498.
- [31] Dolgalev I (2020) *msigdb: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format*. R package 7.1.1. <https://CRAN.R-project.org/package=msigdb>
- [32] Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589.
- [33] Hin N, Newman M, Kaslin J, Douek AM, Lumsden A, Nik SHM, Dong Y, Zhou X-F, Mañucat-Tan NB, Ludington A, Adelson DL, Pederson S, Lardelli M (2020) Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2. *PLoS One* **15**, e0227258.
- [34] Wickham H (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.
- [35] Kolde R (2019) *pheatmap: Pretty Heatmaps*. R Package. <https://CRAN.R-project.org/package=pheatmap>
- [36] Conway JR, Lex A, Gehlenborg N (2017) UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938-2940.
- [37] Nishii K, Nakaseko C, Jiang M, Shimizu N, Takeuchi M, Schneider WJ, Bujo H (2013) The soluble form of LR11 protein is a regulator of hypoxia-induced, urokinase-type plasminogen activator receptor (uPAR)-mediated adhesion of immature hematological cells. *J Biol Chem* **288**, 11877-11886.
- [38] Newman M, Nik HM, Sutherland GT, Hin N, Kim WS, Halliday GM, Jayadev S, Smith C, Laird AS, Lucas CW, Kittipassorn T, Peet DJ, Lardelli M (2020) Accelerated loss of hypoxia response in zebrafish with familial Alzheimer's disease-like mutation of presenilin 1. *Hum Mol Genet* **29**, 2379-2394.
- [39] Cahoy JD, Emery B, Kaushal A, Foo LC, Zamanian JL, Christopherson KS, Xing Y, Lubischer JL, Krieg PA, Krupenko SA, Thompson WJ, Barres BA (2008) A transcriptome database for astrocytes, neurons, and oligodendrocytes: A new resource for understanding brain development and function. *J Neurosci* **28**, 264-278.
- [40] Oosterhof N, Holtman IR, Kuil LE, van der Linde HC, Boddeke EWGM, Eggen BJJ, van Ham TJ (2017) Identification of a conserved and acute neurodegeneration-specific microglial transcriptome in the zebrafish. *Glia* **65**, 138-149.
- [41] Karch CM, Goate AM (2015) Alzheimer's disease risk genes and mechanisms of disease pathogenesis. *Biol Psychiatry* **77**, 43-51.
- [42] Bono H, Hirota K (2020) Meta-analysis of hypoxic transcriptomes from public databases. *Biomedicine* **8**, 10.
- [43] Sniekers S, Stringer S, Watanabe K, Jansen PR, Coleman JRI, Krapohl E, Taskesen E, Hammerschlag AR, Okbay A, Zabaneh D, Amin N, Breen G, Cesarini D, Chabris CF, Iacono WG, Ikram MA, Johannesson M, Koellinger P, Lee JJ, Magnusson PKE, McGue M, Miller MB, Ollier WER, Payton A, Pendleton N, Plomin R, Rietveld CA, Tiemeier H, van Duijn CM, Posthuma D (2017) Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat Genet* **49**, 1107-1112.
- [44] Champion D, Charbonnier C, Nicolas G (2019) SORL1 genetic variants and Alzheimer disease risk: A literature review and meta-analysis of sequencing data. *Acta Neuropathol* **138**, 173-186.
- [45] Cruts M, Theuns J, Van Broeckhoven C (2012) Locus-specific mutation databases for neurodegenerative brain diseases. *Hum Mutat* **33**, 1340-1344.
- [46] Newman M, Hin N, Pederson S, 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. *Mol Brain* **12**, 43.
- [47] Ryman DC, Acosta-Baena N, Aisen PS, Bird T, Danek A, Fox NC, Goate A, Frommelt P, Ghetti B, Langbaum JBS, Lopera F, Martins R, Masters CL, Mayeux RP, McDade E, Moreno S, Reiman EM, Ringman JM, Salloway S, Schofield PR, Sperling R, Tariot PN, Xiong C, Morris JC, Bateman RJ (2014) Symptom onset in autosomal dominant Alzheimer disease: A systematic review and meta-analysis. *Neurology* **83**, 253-260.
- [48] Hani EH, Suaud L, Boutin P, Chèvre JC, Durand E, Philippi A, Demenais F, Vionnet N, Furuta H, Velho G, Bell GI, Laine B, Froguel P (1998) A missense mutation in hepatocyte nuclear factor-4 alpha, resulting in a reduced transactivation activity, in human late-onset non-insulin-dependent diabetes mellitus. *J Clin Invest* **101**, 521-526.
- [49] Cereghini S (1996) Liver-enriched transcription factors and hepatocyte differentiation. *FASEB J* **10**, 267-282.
- [50] Rhee J, Inoue Y, Yoon JC, Puigserver P, Fan M, Gonzalez FJ, Spiegelman BM (2003) Regulation of hepatic fasting response by PPAR γ coactivator-1 α (PGC-1): Requirement for hepatocyte nuclear factor 4 α in gluconeogenesis. *Proc Natl Acad Sci U S A* **100**, 4012.
- [51] Yin L, Ma H, Ge X, Edwards PA, Zhang Y (2011) Hepatic hepatocyte nuclear factor 4 α is essential for maintaining triglyceride and cholesterol homeostasis. *Arterioscler Thromb Vasc Biol* **31**, 328-336.
- [52] Yamanishi K, Doe N, Sumida M, Watanabe Y, Yoshida M, Yamamoto H, Xu Y, Li W, Yamanishi H, Okamura H, Matsunaga H (2015) Hepatocyte nuclear factor 4 alpha is a key factor related to depression and physiological homeostasis in the mouse brain. *PLoS One* **10**, e0119021.
- [53] Mosconi L (2005) Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease. FDG-PET

- studies in MCI and AD. *Eur J Nucl Med Mol Imaging* **32**, 486-510.
- [54] Mosconi L, Mistur R, Switalski R, Tsui WH, Glodzik L, Li Y, Pirraglia E, De Santi S, Reisberg B, Wisniewski T, de Leon MJ (2009) FDG-PET changes in brain glucose metabolism from normal cognition to pathologically verified Alzheimer's disease. *Eur J Nucl Med Mol Imaging* **36**, 811-822.
- [55] Drzezga A, Lautenschlager N, Siebner H, Riemenschneider M, Willoch F, Minoshima S, Schwaiger M, Kurz A (2003) Cerebral metabolic changes accompanying conversion of mild cognitive impairment into Alzheimer's disease: A PET follow-up study. *Eur J Nucl Med Mol Imaging* **30**, 1104-1113.
- [56] Minoshima S, Giordani B, Berent S, Frey KA, Foster NL, Kuhl DE (1997) Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann Neurol* **42**, 85-94.
- [57] Iturria-Medina Y, Sotero R, Toussaint P, Mateos-Pérez J, Evans A, Alzheimer's Disease Neuroimaging Initiative (2016) Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nat Commun* **7**, 11934.
- [58] Ding Q, Markesbery WR, Chen Q, Li F, Keller JN (2005) Ribosome dysfunction is an early event in Alzheimer's disease. *J Neurosci* **25**, 9171-9175.
- [59] Ding Q, Markesbery WR, Cecarini V, Keller JN (2006) Decreased RNA, and increased RNA oxidation, in ribosomes from early Alzheimer's disease. *Neurochem Res* **31**, 705-710.
- [60] Langstrom NS, Anderson JP, Lindroos HG, Winbland B, Wallace WC (1989) Alzheimer's disease-associated reduction of polysomal mRNA translation. *Mol Brain Res* **5**, 259-269.
- [61] Honda K, Smith MA, Zhu X, Baus D, Merrick WC, Tartakoff AM, Hattier T, Harris PL, Siedlak SL, Fujioka H, Liu Q, Moreira PI, Miller FP, Nunomura A, Shimohama S, Perry G (2005) Ribosomal RNA in Alzheimer disease is oxidized by bound redox-active iron. *J Biol Chem* **280**, 20978-20986.
- [62] Willi J, K pfer P, Ev quoz D, Fernandez G, Katz A, Leumann C, Polacek N (2018) Oxidative stress damages rRNA inside the ribosome and differentially affects the catalytic center. *Nucleic Acids Res* **46**, 1945-1957.
- [63] Knupp A, Mishra S, Martinez R, Braggin JE, Szabo M, Kinoshita C, Hailey DW, Small SA, Jayadev S, Young JE (2020) Depletion of the AD risk gene SORL1 selectively impairs neuronal endosomal traffic independent of amyloidogenic APP processing. *Cell Rep* **31**, 107719.
- [64] Cataldo AM, Petanceska S, Peterhoff CM, Terio NB, Epstein CJ, Villar A, Carlson EJ, Staufenbiel M, Nixon RA (2003) App gene dosage modulates endosomal abnormalities of Alzheimer's disease in a segmental trisomy 16 mouse model of down syndrome. *J Neurosci* **23**, 6788-6792.
- [65] Cataldo AM, Peterhoff CM, Schmidt SD, Terio NB, Duff K, Beard M, Mathews PM, Nixon RA (2004) Presenilin mutations in familial Alzheimer disease and transgenic mouse models accelerate neuronal lysosomal pathology. *J Neuropathol Exp Neurol* **63**, 821-830.
- [66] Cataldo AM, Peterhoff CM, Troncoso JC, Gomez-Isla T, Hyman BT, Nixon RA (2000) Endocytic pathway abnormalities precede amyloid β deposition in sporadic Alzheimer's disease and Down syndrome: Differential effects of APOE genotype and presenilin mutations. *Am J Pathol* **157**, 277-286.
- [67] Yambire KF, Rostovsky C, Watanabe T, Pacheu-Graud D, Torres-Odio S, Sanchez-Guerrero A, Senderovich O, Meyron-Holtz EG, Milosevic I, Frahm J, West AP, Raimundo N (2019) Impaired lysosomal acidification triggers iron deficiency and inflammation *in vivo*. *Elife* **8**, e51031.

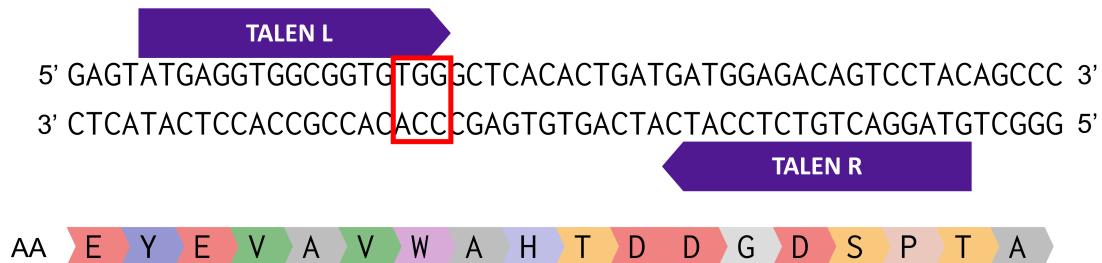
Reprinted from Barthelson, Karissa et al. 'Brain Transcriptome Analysis of a Protein-Truncating Mutation in Sortilin-Related Receptor 1 Associated With Early-Onset Familial Alzheimer's Disease Indicates Early Effects on Mitochondrial and Ribosome Function'. 1 Jan. 2021 : 1105 – 1119 with permission from IOS Press.

The publication is available from IOS Press through <http://dx.doi.org/10.3233/JAD-201383>

Supplementary Material

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

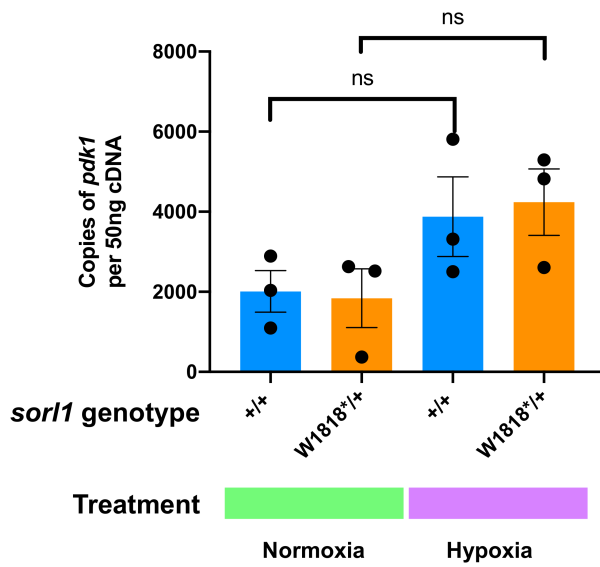
Danio rerio sorl1 exon 41 region



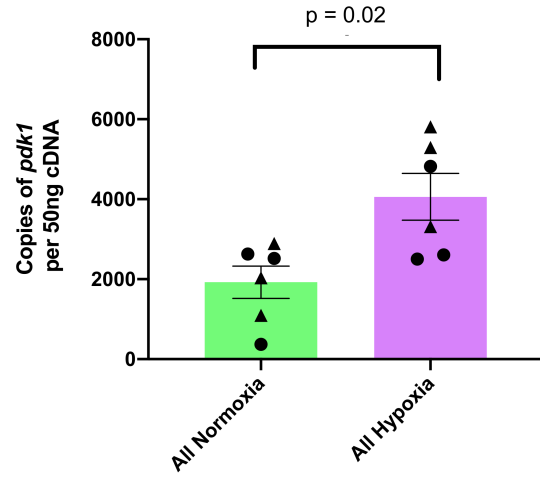
Supplementary Figure 1. Genome editing of *sortl* exon 41.

A 55 nucleotide section of zebrafish *sortl* exon 41 (ENSDARE00000237070). Both the sense, anti-sense and the translated (AA) sequences are indicated. The TALEN recognition sites are indicated by purple arrows, and the W1818 site is highlighted by the red box.

A



B



C *sorl1* copy numbers per 50ng brain-derived cDNA

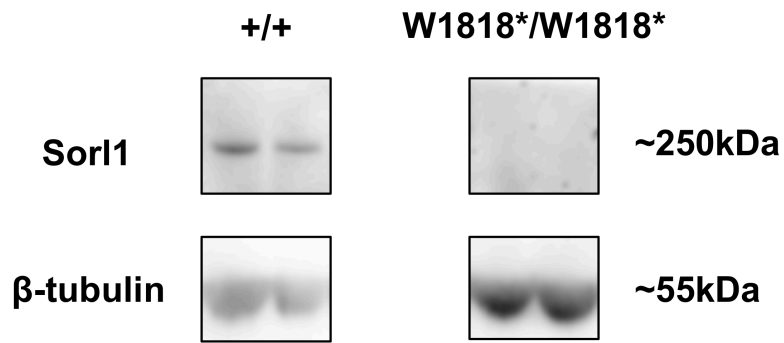
Treatment	Normoxia				Hypoxia			
	+/+		W1818*/+		+/+		W1818*/+	
<i>sorl1</i> allele	WT	W1818*	WT	W1818*	WT	W1818*	WT	W1818*
Number of copies	2209.6	0.07	1193.2	521.3	1581.6		857.8	379.5
	1420.8		965.3	432.5	1502.6		857.6	329.4
	1721.9		807.9	337.2	1873.2		901.8	362.4
			1043.9	375.8			596.4	230.2

D *pdk1* copy numbers per 50ng brain-derived cDNA

Treatment	Normoxia		Hypoxia	
	+/+	W1818*/+	+/+	W1818*/+
Number of copies	2037.5	2629.6	5812.6	5295.6
	2892.7	2519	2505.2	2605.2
	1094.9	370.83	3313.4	4822

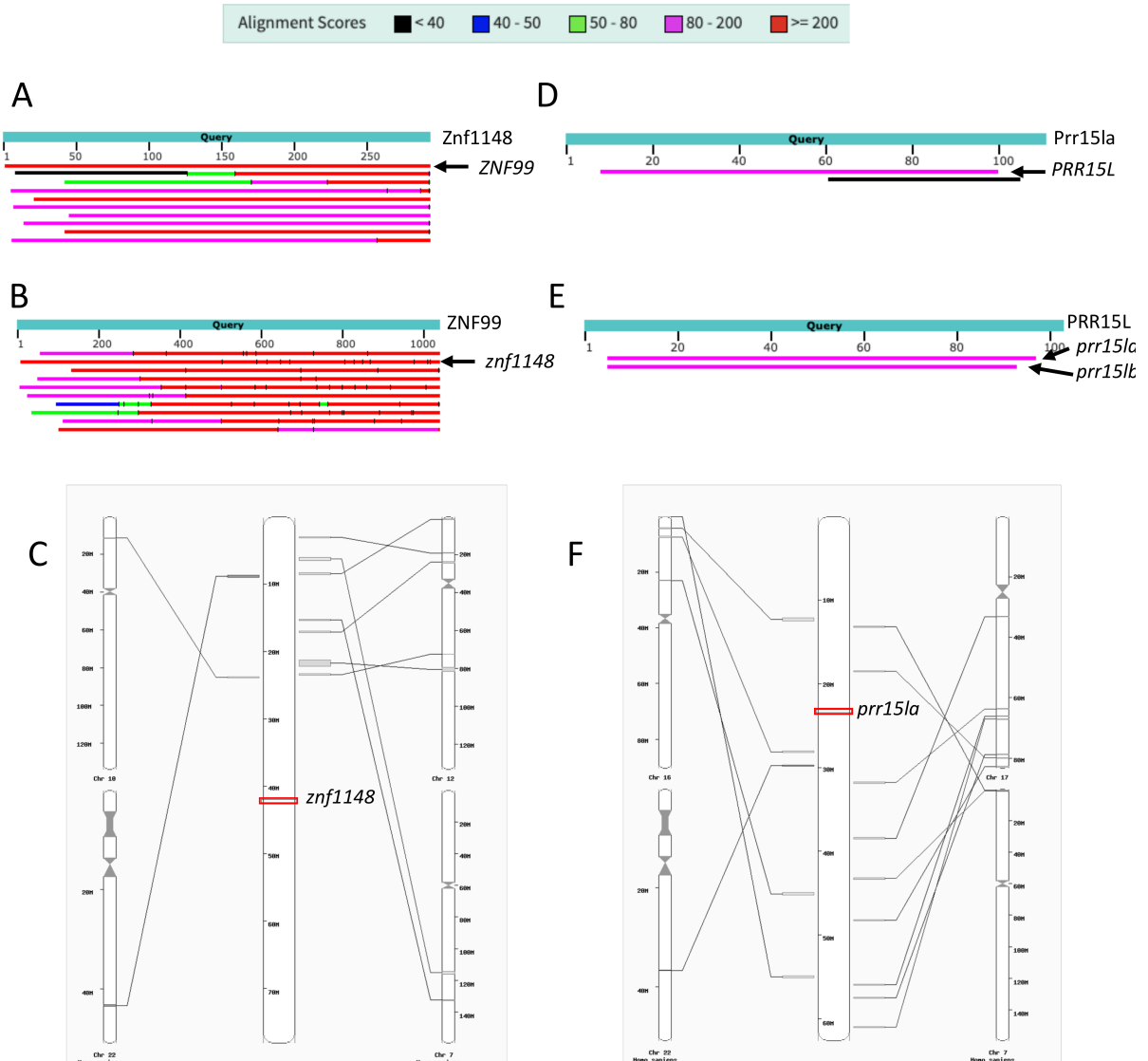
Supplementary Figure 2. Expression of *sor11* and *pdk1* in young adult zebrafish brains exposed to hypoxia.

A) Mean \pm SEM copies of pyruvate dehydrogenase kinase (*pdk1*) per 50 ng of brain derived cDNA of zebrafish either heterozygous for the W1818* mutation in *sor11* (W1818*/+), or their wild type siblings (+/+), and exposed to either normoxia or hypoxia. Expression of *pdk1* appeared to increase after exposure to acute hypoxia in each *sor11* genotype. However, these comparisons did not reach statistical significance (one-way ANOVA with Dunnet's post-hoc test). B) Re-analysis of the data in A, where data points were combined by treatment. A Student's *t*-test identified a significant increase in *pdk1* expression. +/+ samples are indicated by triangles, and W1818*/+ samples are indicated by circles. C) Expression levels of *sor11* transcripts per 50 ng of brain-derived cDNA under normoxia and hypoxia. Numbers represent the copy numbers on the QuantStudio™ 3D Digital PCR 20K chip after processing within QuantStudio™ 3D AnalysisSuite Cloud Software (version 3.0, Applied Biosystems, Thermo Fisher Scientific). Note that only one digital PCR was performed to detect the W1818* allele of *sor11* in +/+ brains under normoxia to confirm that the W1818* mutation specific primers did not amplify from the wild type (WT) sequence of *sor11*. D) Expression levels of *pdk1* transcripts per 50 ng of brain-derived cDNA under normoxia and hypoxia.



Supplementary Figure 3. Expression of Sorl1 protein in young adult zebrafish brains.

Western immunoblot against C-terminal sequences of Sorl1 protein in 6-month-old wild type and homozygous mutant sibling brains. Loss of the signal at approximately 250 kDa supports that the W1818* mutation results in disruption of translation of Sorl1 protein. Immunoblotting against β-tubulin revealed a loading bias, with an apparent increase in total protein levels in the homozygous mutant brains. Entire blots are shown in Supplementary Figure 7.



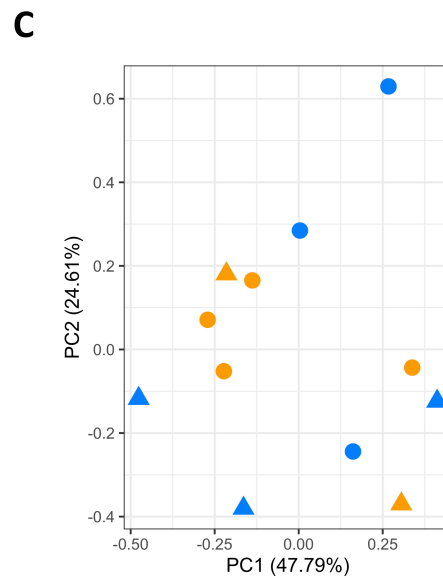
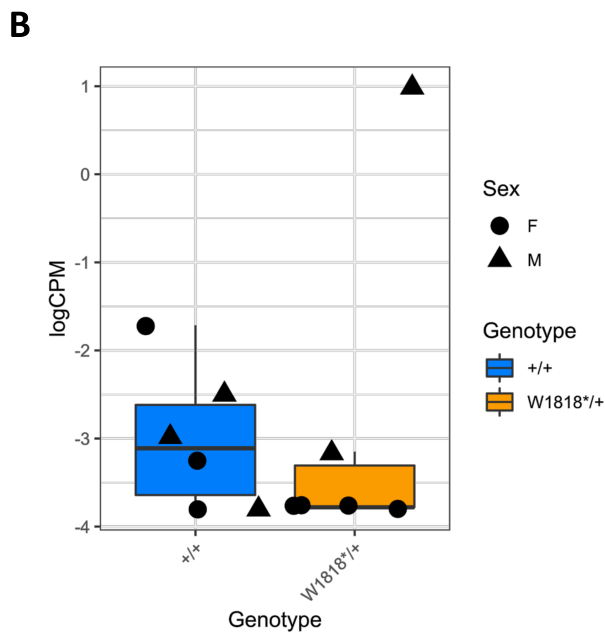
Supplementary Figure 4. BLAST and analysis of synteny of the gene differentially expressed due to heterozygosity for the W1818* mutation of *sor11*.

A) Graphical summary of the top 10 hits of a tblastn search using the protein sequence of Znf1148 (ENSDART00000163798.2) against the human genome GRCh38.p13. The best hit corresponds to *ZNF99* on human chromosome 19 (chr19). B) Graphical summary of the best 10 hits of a tblastn search using the protein sequence of human ZNF99 (ENST00000397104.5) against the zebrafish genome GRCz11. As a reverse viewpoint to A, the second hit refers to *znf1148*. C) Analysis of synteny between zebrafish chromosome 4 (chr4) and the human genome. The location of *znf1148* is shown in red. No conservation of synteny is observed in the human genome with the *znf1148* region. D) Graphical summary of a tblastn search using the protein sequence of zebrafish Prr15la (ENSDART00000134723.2) against the human genome GRCh38.p13. The top hit corresponds to *PRR15L* on human chromosome 17 (chr17). E) Graphical summary of the top 10 hits of a tblastn search using the protein sequence of PRR15L (ENST00000300557.3) against the zebrafish genome GRCz11. The only significant homology is to the paralogues *prr15la* and *prr15lb*. F) Analysis of synteny between zebrafish chromosome 3

(chr3) and the human genome. The location of *prp15la* is shown in red. No conservation of synteny is observed in the human genome with the *prp15la* region. BLAST searches and analysis of synteny were performed using the Ensembl web site (<https://m.ensembl.org/>).

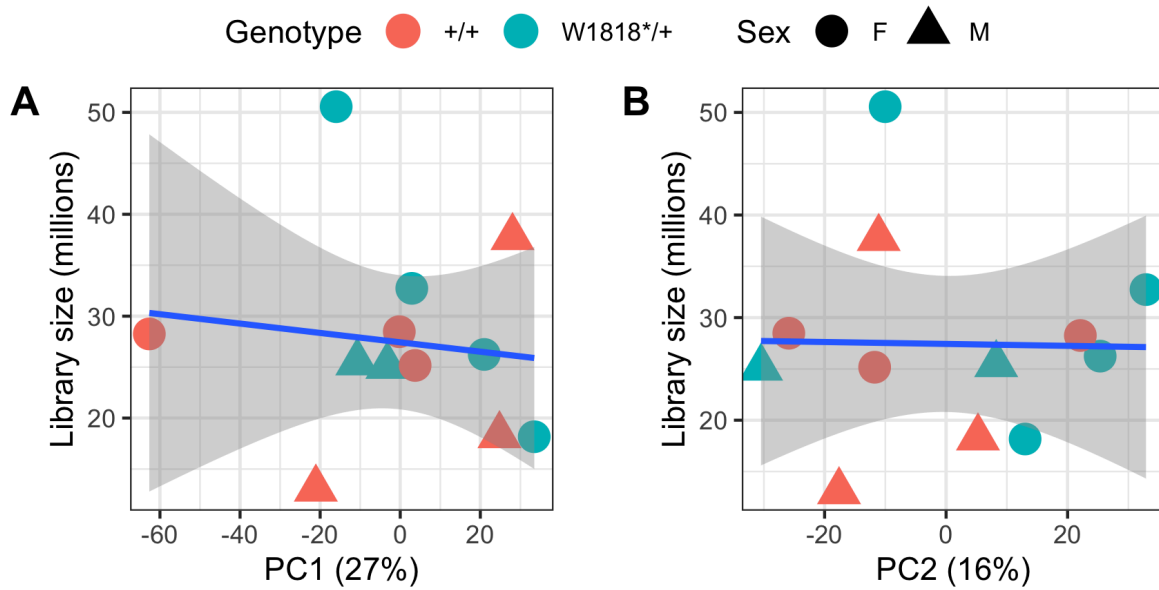
A

Top 5 most significantly enriched promoter motifs in the top 100 most differentially expressed genes due to <i>sor11</i> genotype.							
Motif Name	Sequence	P-value	Bonferroni adjusted p-value	# of Target Sequences with Motif	% of Target Sequences with Motif	# of Background Sequences with Motif	% of Background Sequences with Motif
HNF4a		1.00E-04	0.0428	29	26.36%	1303	12.26%
MyoG		0.01	1	49	44.55%	3331	31.35%
TR4		0.01	1	8	7.27%	252.8	2.38%
ZNF382		0.01	1	5	4.55%	106	1.01%
NF1:FOXA1		0.01	1	10	9.09%	387	3.64%



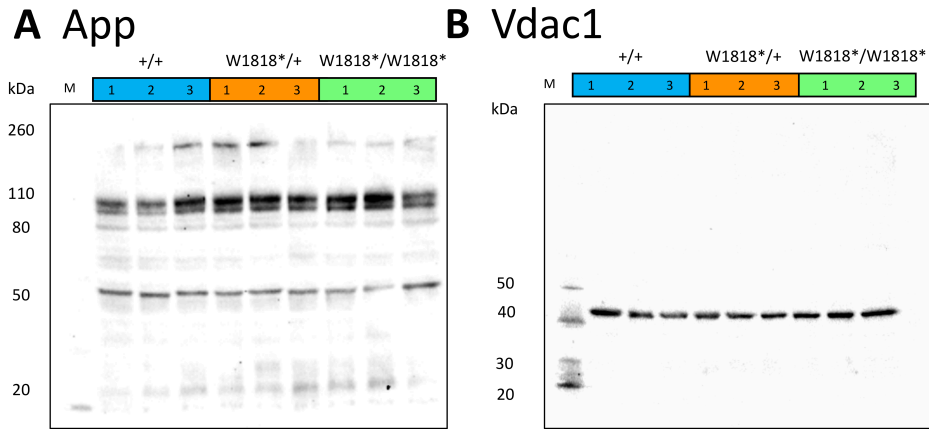
Supplementary Figure 5. Promotor motif analysis using *homer*.

A) The top 5 most significantly enriched motifs in the promoters of the top 100 most statistically significant gene differentially expressed due to *sor11* genotype, relative to all genes detected in the RNA-seq experiment. B) Expression of *hnf4a* before filtering lowly expressed genes in the RNA-seq dataset. C) Principal component (PC) analysis plot of the *RUVseq* adjusted expression values of genes containing a motif for binding of Hnf4α (*HNF4ALPHA_Q6* gene set from *MSigDB*, C3 category, TFT:TFT_Legacy subcategory).

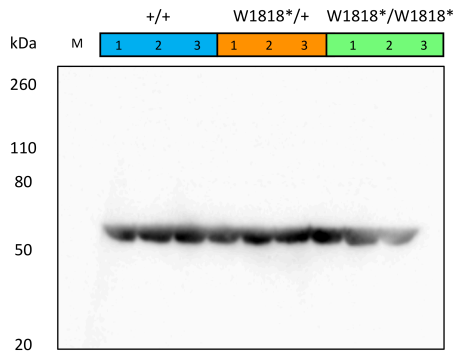


Supplementary Figure 6, PC1 and PC2 do not correlate with library size.

Plot of the A) principal component 1 (PC1) and B) PC2 value against the size of each RNA-seq library. The blue lines indicate a linear model regression lines with standard error shading. No correlation is observed, suggesting that the varying library sizes does not contribute overly to the two largest sources of variation in this dataset.

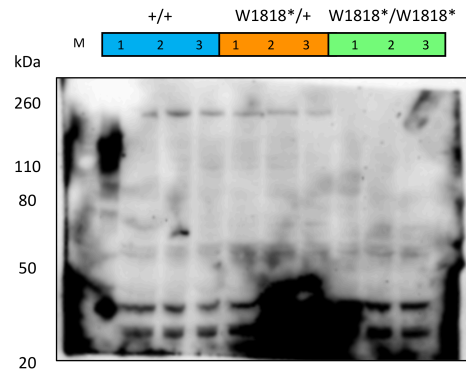


C β -tubulin

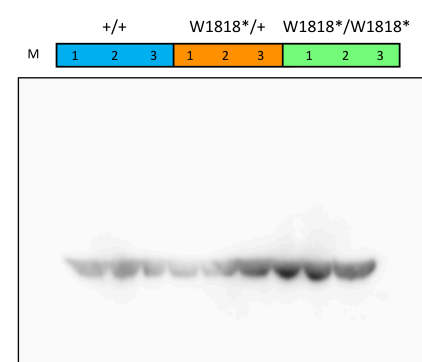


A-C show raw images of the PVDF membrane described in Figure 4 incubated in primary and secondary antibodies and visualized using chemiluminescent substrates as described in the main text.

D Sorl1



E β -tubulin



D and E show raw images of the PVDF membrane described in Supplementary Figure 3 incubated in primary and secondary antibodies and visualized using chemiluminescent substrates as described in the main text.

Supplementary Figure 7. Original images of western blot membranes.

Supplementary Table 1 can be found at

<https://content.iospress.com/articles/journal-of-alzheimers-disease/jad201383>

Chapter 7: Frameshift and frame-preserving mutations in zebrafish *presenilin 2* affect different cellular functions in young adult brains

Statement of Authorship

Title of Paper	Frameshift and frame-preserving mutations in zebrafish presenilin 2 affect different cellular functions in young adult brains		
Publication Status	<input type="checkbox"/> Published	<input checked="" type="checkbox"/> Accepted for Publication	<input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	This manuscript has undergone peer review, and is accepted pending minor revisions in the Journal of Alzheimer's Disease Reports. These revisions will be performed after the thesis has been submitted.		

Principal Author

Name of Principal Author (Candidate)	Karissa Barthelson		
Contribution to the Paper	Performed the secondary bioinformatic analysis, drafted the manuscript and prepared the figures		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature	_____	Date	9/1/21

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Stephen Martin Pederson		
Contribution to the Paper	performed the initial bioinformatic analysis, edited the manuscript and supervised the secondary bioinformatic analysis.		
Signature	_____	Date	12/01/2021

Name of Co-Author	Morgan Newman		
Contribution to the Paper	Supervised zebrafish work. edited the manuscript		
Signature	_____	Date	15/02/2021

Please cut and paste additional co-author panels here as required.

Name of Co-Author	Haowei Jiang	
Contribution to the Paper	generated the mutant zebrafish and edited the manuscript	
Signature	Date	12/01/2021

Name of Co-Author	Michael Landelli	
Contribution to the Paper	Supervision of entire project, edited the manuscript	
Signature	Date	19/02/2021

Frameshift and frame-preserving mutations in zebrafish *presenilin 2* affect different cellular functions in young adult brains

Karissa **Barthelson**^{1*}, Stephen Martin **Pederson**², Morgan **Newman**¹, Haowei **Jiang**³, and Michael **Lardelli**¹

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

2. Bioinformatics Hub, School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, SA 5005, Australia

3. School of Pharmacy, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China

* Corresponding author

Complete correspondence address: Room 1.24, Molecular Life Sciences

Building. North Terrace Campus. The University of Adelaide, SA 5005,

Australia.

Telephone: 83134863 Email: karissa.barthelson@adelaide.edu.au

Present affiliation of Stephen Martin Pederson: Dame Roma Mitchell

Cancer Research Laboratories, Adelaide Medical School, Faculty of Health and

Medical Sciences, Level 8, Adelaide Health & Medical Sciences Building,

University of Adelaide, SA 5005, Australia

Running title: RNA-seq analysis of *psen2* mutations

Abstract

Background: Mutations in *PRESENILIN 2* (*PSEN2*) cause early onset familial Alzheimer's disease (EOfAD) but their mode of action remains elusive. One consistent observation for all *PRESENILIN* gene mutations causing EOfAD is that a transcript is produced with a reading frame terminated by the normal stop codon – the “reading frame preservation rule”. Mutations that do not obey this rule do not cause the disease. The reasons for this are debated.

Methods: A frameshift mutation (*psen2*^{N140fs}) and a reading frame-preserving mutation (*psen2*^{T141_L142delinsMISLISV}) were previously isolated during genome editing directed at the N140 codon of zebrafish *psen2* (equivalent to N141 of human *PSEN2*). We mated a pair of fish heterozygous for each mutation to generate a family of siblings including wild type and heterozygous mutant genotypes. Transcriptomes from young adult (6 months) brains of these genotypes were analysed.

Objective: To predict cellular functions affected by heterozygosity for each mutation using bioinformatic techniques.

Results: The reading frame preserving mutation uniquely caused subtle, but statistically significant, changes to expression of genes involved in oxidative phosphorylation, long term potentiation and the cell cycle. The frameshift mutation uniquely affected genes involved in Notch and MAPK signalling, extracellular matrix receptor interactions and focal adhesion. Both mutations affected ribosomal protein gene expression but in opposite directions.

Conclusion: A frameshift and frame-preserving mutation at the same position in zebrafish *psen2* cause discrete effects. Changes in oxidative phosphorylation, long term potentiation and the cell cycle may promote EOfAD pathogenesis in humans.

Key words

Alzheimer's disease, RNA-seq, zebrafish, PSEN2, mitochondria.

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder which develops silently over decades. The pathological hallmarks of AD include the presence of extracellular senile neuritic plaques consisting, primarily, of amyloid β ($A\beta$) peptides, intracellular neurofibrillary tangles (primarily consisting of hyperphosphorylated tau proteins), and progressive neuronal loss (reviewed in [1]). The majority of therapeutics for AD are aimed at reducing $A\beta$ levels (reviewed in [2]). However, all compounds to date show little or no effect on cognitive symptoms (reviewed in [3, 4]). This likely reflects our ignorance of the pathogenic mechanism underlying AD. Also, once cognitive changes occur, damage to the brain is considerable and may be irreversible. A comprehensive understanding of the early molecular changes/stresses occurring many years before disease onset is required to develop preventative treatments to reduce the prevalence of AD.

The majority of AD cases arise sporadically with an age of onset after 65 years (late onset AD, LOAD). Genome-wide association studies (GWAS) have identified at least 20 genes associated with increased risk for LOAD, with the

most influential being the *APOLIPOPROTEIN E (APOE)* locus [5, 6]. Rare, familial forms of AD also exist. Early onset familial AD (EOfAD) cases have an age of onset before 65 years, and arise due to mutations in the *PRESENILIN* genes (*PSEN1* and *PSEN2*), and the genes *AMYLOID B A4 PRECURSOR PROTEIN (A β PP)* and *SORTILIN-RELATED RECEPTOR 1 (SORL1)* (reviewed in [7-9]).

PSEN2 is the less commonly mutated *PSEN* gene in EOfAD. To date, only 13 pathogenic mutations have been described in *PSEN2* compared to over 185 pathogenic mutations in the homologous gene *PSEN1* [10]. The first *PSEN2* EOfAD mutation characterised was N141I, found in Volga German families and affecting the second transmembrane domain of the *PSEN2* protein [11]. We previously attempted to introduce this mutation into zebrafish by genome editing. While we were not successful at introducing an exact equivalent of the human N141I mutation into zebrafish, we did isolate a reading frame-preserving mutation at this position: *psen2*^{T141_L142delinsMISLISV} [12]. This mutation alters two codons and inserts five additional novel codons and was predicted to largely preserve the transmembrane structure of the resultant protein. We also isolated a frameshift mutation at the same position: *psen2*^{N140fs}. This is a deletion of 7 nucleotides resulting in a premature stop codon at novel codon position 142 [12]. Since all EOfAD mutations in the *PSENs* follow a “reading frame preservation rule” [13], *psen2*^{T141_L142delinsMISLISV} models such a mutation in *PSEN2*. In contrast, the *psen2*^{N140fs} allele is predicted to express a truncated protein, is not EOfAD-like, and so can act as a negative control to investigate the differences between reading frame-preserving and -destructive mutations of *PSEN2*.

The overall goal of the Alzheimer's Disease Genetics Laboratory has been to identify the cellular processes affected in common by EOfAD mutations in different genes in young adult, heterozygous mutant brains (i.e. to establish an early EOfAD brain transcriptomic signature). We exploit zebrafish as a model organism for this work as the EOfAD genes are conserved in zebrafish, and large families containing both EOfAD-like mutant fish and their wild type siblings can be generated from single mating events between single pairs of fish (reviewed in [14]). Analysing the transcriptomes of sibling fish raised in the same environment (aquaria within a single recirculated water system) reduces genetic and environmental sources of variation, (noise) and allows subtle effects of mutations to be identified. After introducing EOfAD-like mutations into the endogenous, zebrafish orthologues of EOfAD genes (i.e. knock-in models), we analyse their effects on brain transcriptomes in heterozygous fish to mimic the genetic state of such mutations in the human disease. We term this experimental strategy Between Sibling Transcriptome (BeST) analysis. We have previously performed BeST analyses for EOfAD-like mutations in *psen1* [15] and *sort11* [16, 17] and for a complex (but possibly EOfAD-like) mutation in *psen2*: *psen2*^{S4Ter} [18].

In the work described in this report, we performed a BeST analysis using a family of sibling fish generated by mating two fish with genotypes *psen2*^{T141_L142delinsMISLISV/+} and *psen2*^{N140fs/+}. The family included fish heterozygous for the mutations *psen2*^{T141_L142delinsMISLISV} (for simplicity, hereafter referred to as "EOfAD-like") and *psen2*^{N140fs} (for simplicity, hereafter referred to as "FS") and their wild type siblings, all raised together in the same tank until 6 months of age. This strategy allowed direct comparisons of the brain

transcriptomes from the two *psen2* mutant genotypes to their wild type siblings while reducing confounding effects from genetic and environmental variation. Despite this, considerable gene expression variation (noise) is still observed between individuals and the effects of the mutations are subtle so that few genes are identified initially as differentially expressed. However, when analysed at the pathway level, heterozygosity for each mutation shows distinct transcriptomic effects. The sets of genes involved in oxidative phosphorylation, long-term potentiation, and the cell cycle are only altered significantly in the EOfAD-like brains, while the FS mutant brains shows apparent changes in Notch and MAPK signalling, focal adhesion, and ECM receptor interaction. Both mutation states affected genes encoding ribosomal subunits, but the effect of each mutation was in opposite directions. Our results are consistent with the concept that only dominant, reading frame-preserving mutations of the *PRESENILIN* genes cause EOfAD and support that effects on oxidative phosphorylation may be a common signature of such mutations.

Methods

Zebrafish lines and ethics statement.

Generation of the *psen2*^{T141_L142delinsMISLISV} and *psen2*^{N140fs} lines of zebrafish used in this study is described in [12]. All experiments involving zebrafish were conducted under the auspices of the Animal Ethics Committee of the University of Adelaide, permit numbers S-2017-089 and S-2017-073.

Breeding strategy

Two fish heterozygous for these mutations were mated to generate a family of sibling fish with *psen2* genotypes EOfAD-like/+, FS/+, EOfAD-like/FS (transheterozygous) or wild type (**Figure 1A**). 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). The heads of these fish were removed and stored in 600 μ L of RNAlater solution, while their tails were removed for genomic DNA extraction and genotyping by polymerase chain reactions (PCRs).

PCR genotyping

Allele-specific genotyping PCRs were performed on genomic DNA extracted from fin biopsies as described in [17]. The allele-specific primer sequences can be found in **Table 1**. Primers were synthesised by Sigma Aldrich (St. Louis, MO, USA).

Table 1: Genotyping primer sequences.	
Primer Name	Sequence (5' to 3')
Wild type specific forward primer	TGAATTCGGTGCTCAACACTC
T141_L142delinsMISLISV specific forward primer	TGAATTCGGTGCTCAACATG
N140fs specific forward primer	TGCTGAATTCGGTGCTCTG

Common reverse primer	TCACCAAGGACCACTGATTCAGC
-----------------------	-------------------------

RNA-seq data generation and analysis

Preparation of RNA for RNA-seq and the subsequent analysis of the data was performed mostly as in [17]. Briefly, brains of $n = 5$ fish of wild type, EOfAD-like/+ and FS/+ genotypes (each group contained 3 females and 2 males) were carefully removed from the heads preserved RNAlater Stabilization Solution (Invitrogen, Carlsbad, CA, USA), followed by total RNA extraction using the mirVana™ miRNA Isolation Kit (Ambion, Austin, TX, USA) and DNase treatment using the DNA-free™ Kit (Ambion). Total RNA was then delivered to the Genomics service at the South Australian Health and Medical Research Institute (SAHMRI, Adelaide, AUS) for polyA+, stranded library preparation and RNA sequencing using the Illumina Nextseq platform.

Processing of the demultiplexed fastq files provided by SAHMRI (75 bp single-end reads) was performed as in [17]. However, the pseudo-alignment using *kallisto* (v0.43.1) [19] was performed using a modified version of the zebrafish transcriptome (Ensembl release 94) which additionally included the two novel mutant *psen2* transcript sequences.

Transcript counts as estimated by *kallisto* were imported into *R* [20] using *tximport* [21], summing the transcript-level counts to gene-level counts. We also imported the transcript-level counts using the *catchKallisto* function from the package *edgeR* [22] to assess simply the allele-specific expression of the different *psen2* transcripts.

We removed genes considered as undetectable (less than one count per million in at least five of the samples) from downstream analysis, leaving library sizes ranging between 17 million and 27 million reads. We normalised for these differences in library sizes using the trimmed mean of M-values (TMM) method [23], followed by removal of one factor of unwanted variation using the *RUVg* method *RUVseq* [24] as described in [17].

Differential gene expression analysis was performed using a generalised linear model (GLM) and likelihood ratio test using *edgeR* [22], specifying a design matrix with an intercept as the wild type genotype, and the coefficients as the FS/+ and EOfAD-like/+ genotypes, and the *W_1* coefficient from *RUVg*. Genes were considered differentially expressed (DE) in each specific comparison if the FDR-adjusted p-value was below 0.05.

Enrichment analysis was performed on the KEGG and GO gene sets from the Molecular Signatures Database (MSigDB) [25], with GO terms excluded if the shortest path back to the root node was < 3 steps. We also tested for possible iron dyshomeostasis in the RNA-seq data by performing an enrichment analysis on the iron-responsive element (IRE) gene sets described in [26]. To test for over-representation of these gene sets within the DE gene lists, we used *goseq* [27], using the average transcript length per gene to estimate sampling bias for DE genes. For enrichment testing on the entire list of detectable genes, we calculated the harmonic mean p-value [28] of raw p-values from *fry* [29], *camera* [30] and *fgsea* [31, 32]. Gene sets were considered significantly altered if the FDR-adjusted harmonic mean p-value was below 0.05. Visualisation of RNA-seq data analysis was performed using *ggplot2* [33], *pheatmap* [34] and *UpSetR* [35].

Reproducibility and data availability

The raw fastq files and the gene-level counts have been deposited in the GEO database with accession number GSE158233. All code to reproduce this analysis can be found at

https://github.com/UofABioinformaticsHub/20181113_MorganLardelli_mRNaseq.

Results

We first inspected the transcript-level counts to ensure that expression of the *psen2* alleles was as expected (**Supplementary File 1**). Consistent with observations in [12], expression of the EOfAD-like transcript was at a similar level to that of the wild type *psen2* transcript in the EOfAD-like/+ brains, while the decreased expression of the FS allele of *psen2* relative to the wild type allele was suggestive of nonsense-mediated decay (NMD). We observed that one of the samples had been incorrectly genotyped during sample preparation and was actually transheterozygous for the EOfAD-like and FS alleles of *psen2*. Therefore, we omitted this sample from subsequent analyses. We also examined the relationship between samples by principal component analysis (PCA) on the gene-level counts. Samples did not cluster by genotype in the PCA plot, supporting that heterozygosity for the *psen2* mutations in this study does not give widespread effects on the brain transcriptome. Additionally, samples did not cluster by sex, consistent with our previous observations that sex does not show extensive effects in zebrafish brain transcriptomes [16, 17, 36]. In the PCA plot, we noticed that one of the wild type brain samples (12_WT_4) separated greatly from the others, appearing to be an outlier

(**Supplementary File 1**). Sample weights, as calculated using the *voomWithQualityWeights* algorithm from the *limma* package [37] on the gene-level counts, revealed that this wild type sample was highly down-weighted. Therefore, we also omitted this sample from subsequent analyses.

Heterozygosity for an EOfAD-like or a FS mutation in *psen2* results in limited differential expression of genes

To identify which genes were dysregulated due to heterozygosity for the EOfAD-like or FS alleles of *psen2*, we performed differential gene expression analysis using a GLM and likelihood ratio test with *edgeR*. This revealed 32 differentially expressed (DE) genes due to the FS mutation, and 5 DE genes due to the EOfAD-like mutation relative to wild type (**Figure 1, Supplementary File 2**). Only two genes were seen to be significantly upregulated in both comparisons relative to wild type: *AL929206.1* and *BX004838.1*. However, these genes are currently not annotated and have no known function. We tested for over-representation within the DE genes using *goseq* (using the average transcript length per gene as the predictor variable) of gene ontology (GO) terms which, as the name suggests, use ontologies to annotate gene function; KEGG gene sets which can give insight into changes in various biological pathways and reactions; and our recently defined IRE gene sets [26] which can give insight to possible iron dyshomeostasis. However, due to the low numbers of DE genes in both comparisons, no significantly over-represented GO terms or gene sets were identified. (For the top 10 most significantly over-represented GO terms and gene sets in the DE genes, see **Supplementary File 3**).

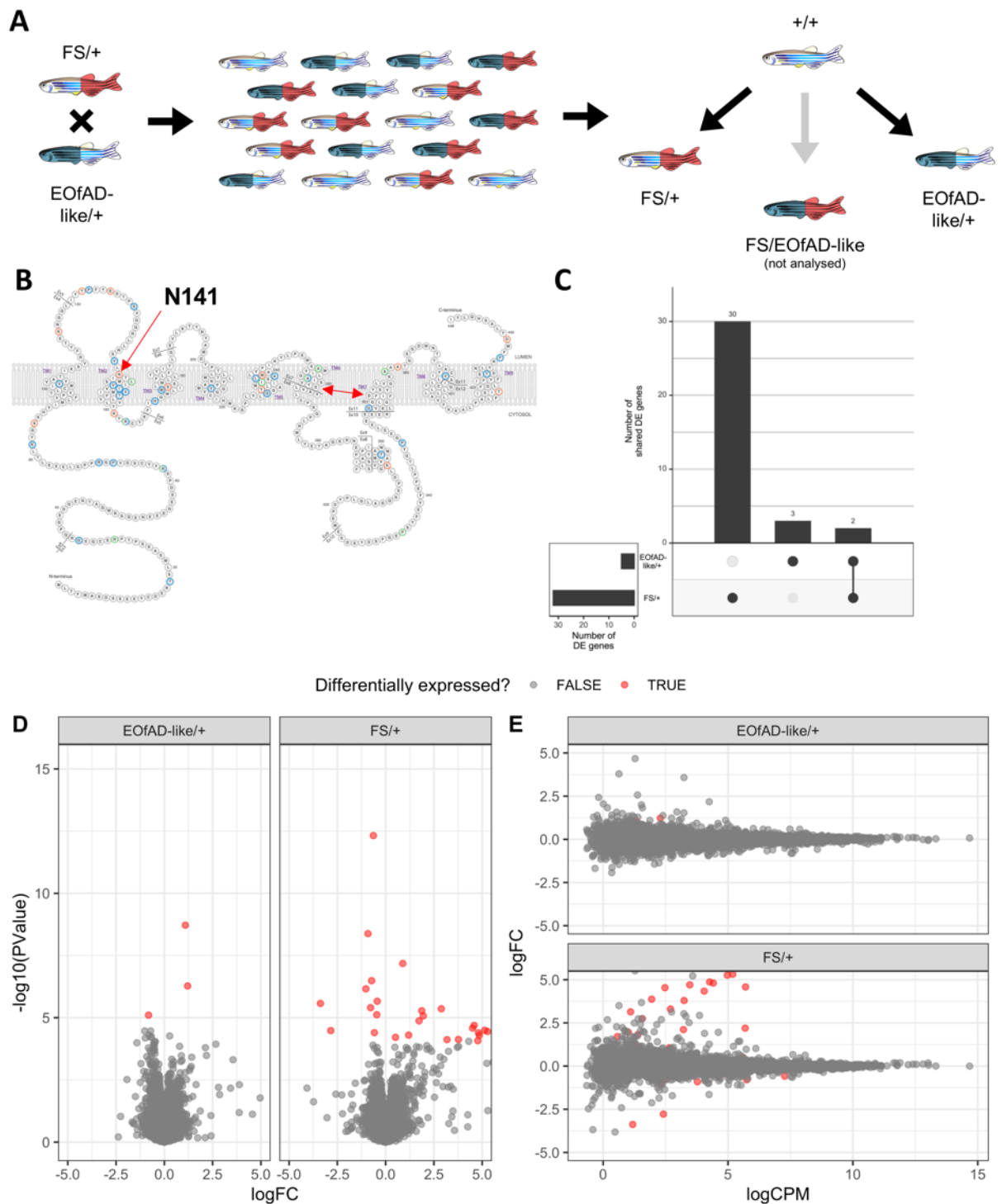


Figure 1: A) Mating strategy and experimental design. **B)** Schematic of the human PSEN2 protein (adapted from <https://www.alzforum.org/mutations/psen-2> with permission from FBRI LLC (Copyright © 1996–2020 FBRI LLC. All Rights Reserved. Version 2.7 – 2020)). Amino acid residues are color-coded to indicate whether

substitutions have been observed to be pathogenic (orange), non-pathogenic (green) or of uncertain pathogenicity (blue). The site of the human EOfAD mutation N141I, in the second transmembrane domain (TMD) is indicated by the red single-headed arrow. The aspartate residues critical for γ -secretase catalysis are indicated by a red double-headed arrow. **C)** Upset plot indicating the number of differentially expressed (DE) genes in each comparison of the *psen2* mutant genotypes to wild type noting that only 2 genes appear to be DE in common between both comparisons. **D)** Mean-difference (MD) plot and **E)** volcano plot of DE genes in each comparison. Note that the logFC axis limits in D) and E) are constrained to -5 and 5 for visualisation purposes.

Heterozygosity for an EOfAD-like or an FS mutation in *psen2* has distinct effects on biological processes in young adult zebrafish brains

We next performed enrichment analysis on the entire list of detectable genes in the RNA-seq experiment to obtain a more complete view on the changes to gene expression due to heterozygosity for the EOfAD-like or FS mutations in *psen2*. Due to the highly overlapping nature of GO terms (many GO terms include many of the same genes), we tested only the KEGG and IRE gene sets. We found statistical evidence for 8 KEGG gene sets in the EOfAD/+ brains and 6 KEGG gene sets in the FS/+ brains to be significantly altered as a group. Notably, no IRE gene sets were found to be significantly altered (**Figure 2A**). Genes in the KEGG_RIBOSOME gene set were significantly altered in both comparisons. However, the genes were mostly downregulated in EOfAD-like/+ brains and upregulated in FS/+ brains (**Figure 2B**). For additional visualisations of our enrichment analysis, see **Supplementary File 4**.

To determine whether the statistical significance of the enrichments of the gene sets was being driven by the expression of the same genes, we inspected the genes from the “leading edge” from *fgsea*. (These can be thought of as the core genes driving the enrichment of each gene set). The leading edge genes in each comparison were mostly independent of one another, except for the KEGG gene sets for oxidative phosphorylation, Parkinson’s disease and Alzheimer’s disease, which share 39 genes (**Figure 2C and D**).

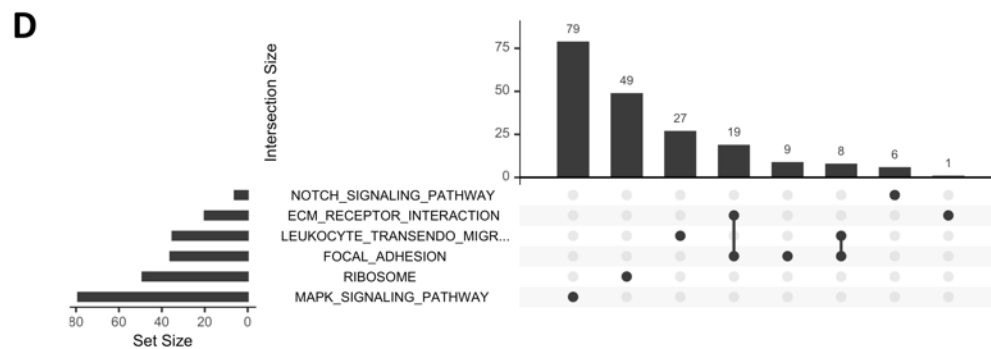
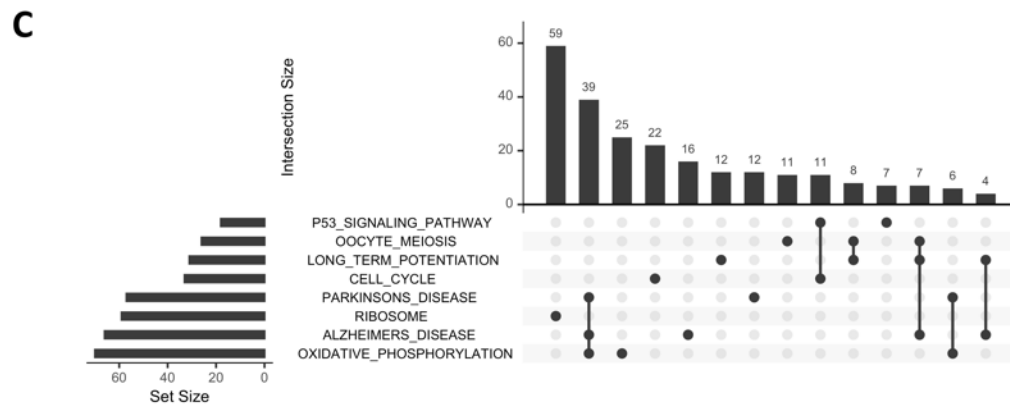
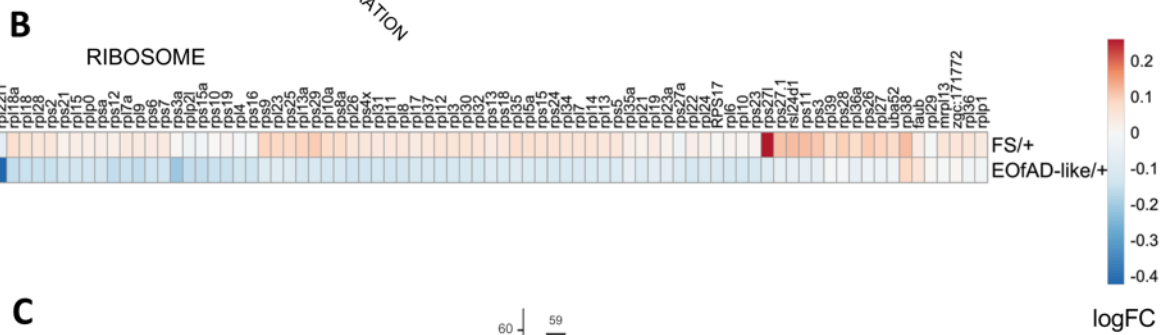
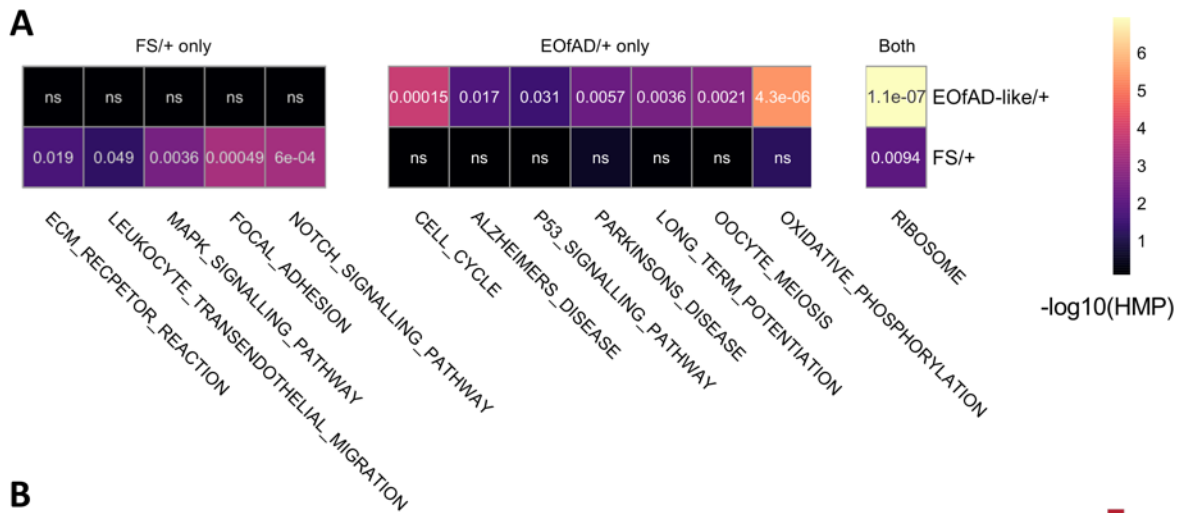


Figure 2: A) Heatmap showing the significant KEGG and IRE gene sets in *psen2* heterozygous mutant brains, clustered based on their Euclidean distance. The columns represent the gene sets which are significant in at least one comparison, and the rows are the two comparisons. The colour of each cell represents the significance value with more significant gene sets appearing lighter. The numbers within each cell are the FDR-adjusted harmonic mean p-values (HMP). ns, not significant. **B)** Heatmap indicating the logFC of genes in the KEGG_RIBOSOME gene set in EOfAD/+ and FS/+ brains. **C)** Upset plot showing the overlap of leading edge genes from the *fgsea* algorithm in gene sets found to be altered in EOfAD-like/+ brains, and **D)** FS/+ brains. See the online version for colour.

Discussion

Here, we aimed to identify similarities and differences in the effects on brain transcriptomes due to heterozygosity for an EOfAD-like mutation compared to a non-EOfAD-like mutation in *psen2*. To this end we used our BeST strategy, involving analysis of sibling fish raised together in an identical environment. BeST analysis reduces confounding effects in direct comparisons of two heterozygous *psen2* mutations relative to wild type, allowing detection of subtle transcriptome state differences.

We observed subtle, but highly statistically significant, changes to the transcriptomes of the *psen2* mutant fish relative to wild type. At the single gene level, relatively few genes were observed to be statistically significantly DE in each mutant genotype, and differentially enriched biological pathways were not detected in these DE gene lists (**Additional File 3**). As a comparison, heterozygosity for an EOfAD-like mutation in zebrafish *psen1*, (orthologous to the human gene most commonly mutated in EOfAD [38]), gives 251 DE genes

in young adult zebrafish brains [15]. However, we were able to identify statistically significant changes in the *psen2* mutant brains by performing enrichment analysis on the entire list of detectable genes in the RNA-seq experiment using an ensemble approach. This involves combining raw p-values by calculation of a harmonic p-value – a strategy thought to be less restrictive than other methods of combining p-values. An additional benefit of this method is that it has been shown specifically to be robust to dependent p-values [28]. The subtlety of the *psen2* mutation effects is consistent with a generally later age of onset, and variable penetrance of EOfAD mutations in *PSEN2* [39-41] relative to EOfAD mutations in *PSEN1* and *A β PP*.

Interestingly, the pathways which were affected by the two *psen2* mutations showed very little overlap. Gene sets which were significantly altered in EOfAD-like/+ brains were not significantly altered in the FS/+ brains and vice versa. Only one biological process appears to be significantly affected by both mutations: ribosome function (the KEGG_RIBOSOME gene set). However, in general, the direction of differential expression of genes in this gene set was observed to be opposite in the two mutants, implying different mechanisms of action of each mutation (**Figure 2B**). The low concordance between the brain transcriptome effects of these two mutations supports that presenilin EOfAD mutations do not act through a simple loss of function but, instead, through a gain of function mechanism connected with encoding an abnormal, but “full-length” protein [13]. Notably, the EOfAD-like mutation of *psen2* is predicted to affect the process of long term potentiation thought to be critical for memory formation. We also observed the dysregulation of genes involved in controlling the cell cycle. Genes driving the enrichment of this gene set (i.e. the leading

edge genes) include the *minichromosome maintenance (mcm)* genes (see **Supplementary File 4**), which we observed to be significantly dysregulated due to heterozygosity for an EOfAD-like mutation in *psen1* in 7 day old zebrafish larvae [42]. Together, these results provide support for the “two-hit hypothesis” of Mark Smith and colleagues [43], which postulates that “cell cycle events” (CCEs) involving inappropriate attempts at cell division by neurons are critical for development of AD. The second “hit” of this hypothesis is oxidative stress, which is an anticipated outcome of disturbance of oxidative phosphorylation. Notably, the EOfAD-like mutation of *psen2* is predicted to affect oxidative phosphorylation, which is in common with all the other EOfAD-like mutations of genes we have examined in zebrafish [16, 18, 26, 44].

One of the most characterised functions of presenilin proteins is their role as the catalytic core of gamma-secretase (γ -secretase) complexes. γ -secretase activity is responsible for cleavage of a wide range of type I transmembrane proteins including A β PP and NOTCH. We observed that genes involved in the Notch signalling pathway were only significantly altered in FS/+ genotype brains, indicating that γ -secretase activity may be altered by the FS mutation. It is difficult to discern from these data whether γ -secretase activity is increased or decreased, as both up- and downregulation of the Notch target genes is observed (**Supplementary File 4**). The FS mutation results in a truncated protein lacking critical aspartate residues needed for γ -secretase activity (**Figure 1B**). Therefore, a simplistic view would be that γ -secretase activity would be reduced. However, a minor, (although statistically non-significant) increase in expression of *psen1* is observed in FS/+ brains (also in EOfAD-like/+ brains) (**Supplementary File 4**) raising the possibility that transcriptional

adaptation (formerly known as “genetic compensation”) might explain an apparent increase in γ -secretase activity [45]. Since presenilin holoproteins are known to form multimeric complexes, we have suggested previously that this may be involved in the regulation of the conversion of the holoprotein into its γ -secretase-active form by endoproteolysis [13]. N-terminal fragments of presenilin proteins are also known to multimerise [46] and the possibility exists that these may disrupt holoprotein multimer formation/stability. In our previous publication identifying the EOfAD-like and FS mutations of *psen2* analysed here, we saw that both mutations reduced adult melanotic surface pigmentation when homozygous [12]. This implied that both mutant alleles encode Psen2 proteins with reduced intrinsic γ -secretase activity, at least in melanosomes. Further research is required to discern whether γ -secretase activity is increased or decreased in *psen2* mutant brains.

In conclusion, our results support that frameshift and reading frame-preserving mutations in the *presenilins* have distinct effects on the brain transcriptome, consistent with the reading frame preservation rule of presenilin EOfAD-causative mutations. The data presented here, along with our growing collection of BeST analyses, indicates that changes to mitochondrial and ribosomal functions are effects-in-common of heterozygosity for EOfAD-like mutations in different genes. These may be early cellular stresses which eventually lead to AD pathology, and warrant investigation for discovery of therapeutic targets.

Declarations

The authors have no conflict of interest to declare.

Acknowledgements

The authors would like thank FBRI LLC for use of the presenilin protein schematic in Figure 1. This work was supported with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide, and by grants GNT1061006 and GNT1126422 from the National Health and Medical Research Council of Australia (NHMRC). KB was supported by an Australian Government Research Training Program Scholarship. HJ was supported by an Adelaide Scholarship International scholarship from the University of Adelaide.

References

- [1] Masters CL, Bateman R, Blennow K, Rowe CC, Sperling RA, Cummings JL (2015) Alzheimer's disease. *Nature Reviews Disease Primers* **1**, 15056.
- [2] Cummings J, Lee G, Ritter A, Sabbagh M, Zhong K (2019) Alzheimer's disease drug development pipeline: 2019. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* **5**, 272-293.
- [3] Elmaleh DR, Farlow MR, Conti PS, Tompkins RG, Kundakovic L, Tanzi RE (2019) Developing Effective Alzheimer's Disease Therapies: Clinical Experience and Future Directions. *Journal of Alzheimer's Disease* **71**, 715-732.
- [4] Modrego P, Lobo A (2019) A good marker does not mean a good target for clinical trials in Alzheimer's disease: the amyloid hypothesis questioned. *Neurodegenerative Disease Management* **9**, 119-121.
- [5] Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, Boland A, Vronskaya M, Van Der Lee SJ, Amlie-Wolf A, Bellenguez C, Frizatti A, Chouraki V, Martin ER, Sleegers K, Badarinarayan N, Jakobsdottir J, Hamilton-Nelson KL, Moreno-Grau S, O'Laso R, Raybould R, Chen Y, Kuzma AB, Hiltunen M, Morgan T, Ahmad S, Vardarajan BN, Epelbaum J, Hoffmann P, Boada M, Beecham GW, Garnier J-G, Harold D, Fitzpatrick AL, Valladares O, Moutet M-L, Gerrish A, Smith AV, Qu L, Bacq D, Denning N,

Jian X, Zhao Y, Del Zompo M, Fox NC, Choi S-H, Mateo I, Hughes JT, Adams HH, Malamon J, Sanchez-Garcia F, Patel Y, Brody JA, Dombroski BA, Naranjo MCD, Daniilidou M, Eiriksdottir G, Mukherjee S, Wallon D, Uphill J, Aspelund T, Cantwell LB, Garzia F, Galimberti D, Hofer E, Butkiewicz M, Fin B, Scarpini E, Sarnowski C, Bush WS, Meslage S, Kornhuber J, White CC, Song Y, Barber RC, Engelborghs S, Sordon S, Voijnovic D, Adams PM, Vandenberghe R, Mayhaus M, Cupples LA, Albert MS, De Deyn PP, Gu W, Himali JJ, Beekly D, Squassina A, Hartmann AM, Orellana A, Blacker D, Rodriguez-Rodriguez E, Lovestone S, Garcia ME, Doody RS, Munoz-Fernandez C, Sussams R, Lin H, Fairchild TJ, Benito YA, Holmes C, Karamujić-Čomić H, Frosch MP, Thonberg H, Maier W, Roshchupkin G, Ghetti B, Giedraitis V, Kawalia A, Li S, Huebinger RM, Kilander L, Moebus S, Hernández I, Kamboh MI, Brundin R, Turton J, Yang Q, Katz MJ, Concari L, Lord J, Beiser AS, Keene CD, Helisalmi S, Kloszewska I, Kukull WA, Koivisto AM, Lynch A, Tarraga L, Larson EB, Haapasalo A, Lawlor B, Mosley TH, Lipton RB, Solfrizzi V, Gill M, Longstreth WT, Montine TJ, Frisardi V, Diez-Fairen M, Rivadeneira F, Petersen RC, Deramecourt V, Alvarez I, Salani F, Ciaramella A, Boerwinkle E, Reiman EM, Fievet N, Rotter JI, Reisch JS, Hanon O, Cupidi C, Andre Uitterlinden AG, Royall DR, Dufouil C, Maletta RG, De Rojas I, Sano M, Brice A, Cecchetti R, George-Hyslop PS, Ritchie K, Tsolaki M, Tsuang DW, Dubois B, Craig D, Wu C-K, Soininen H, Avramidou D, Albin RL, Fratiglioni L, Germanou A, Apostolova LG, Keller L, Koutroumani M, Arnold SE, Panza F, Gkatzima O, Asthana S, Hannequin D, Whitehead P, Atwood CS, Caffarra P, Hampel H, Quintela I, Carracedo Á, Lannfelt L, Rubinsztein DC, Barnes LL, Pasquier F, Frölich L, Barral S, McGuinness B, Beach TG, Johnston JA, Becker JT, Passmore P, Bigio EH, Schott JM, Bird TD, Warren JD, Boeve BF, Lupton MK, Bowen JD, Proitsi P, Boxer A, Powell JF, Burke JR, Kauwe JSK, Burns JM, Mancuso M, Buxbaum JD, Bonuccelli U, Cairns NJ, McQuillin A, Cao C, Livingston G, Carlson CS, Bass NJ, Carlsson CM, Hardy J, Carney RM, Bras J, Carrasquillo MM, Guerreiro R, Allen M, Chui HC, Fisher E, Masullo C, Crocco EA, Decarli C, Bisceglia G, Dick M, Ma L, Duara R, Graff-Radford NR, Evans DA, Hodges A, Faber KM, Scherer M, Fallon KB, Riemenschneider M, Fardo DW, Heun R, Farlow MR, Kölsch H, Ferris S, Leber M, Foroud TM, Heuser I, Galasko DR, Giegling I, Gearing M, Hüll M, Geschwind DH, Gilbert JR, Morris J, Green RC, Mayo K, Growdon JH, Feulner T, Hamilton RL, Harrell LE, Drichel D, Honig LS, Cushion TD, Huentelman MJ, Hollingworth P, Hulette CM, Hyman BT, Marshall R, Jarvik GP, Meggy A, Abner E, Menzies GE, Jin L-W, Leonenko G, Real LM, Jun GR, Baldwin CT, Grozeva D, Karydas A, Russo G, Kaye JA, Kim R, Jessen F, Kowall NW, Vellas B, Kramer JH, Vardy E, Laferla

FM, Jöckel K-H, Lah JJ, Dichgans M, Leverenz JB, Mann D, Levey AI, Pickering-Brown S, Lieberman AP, Klopp N, Lunetta KL, Wichmann HE, Lyketsos CG, Morgan K, Marson DC, Brown K, Martiniuk F, Medway C, Mash DC, Nöthen MM, Masliah E, Hooper NM, McCormick WC, Daniele A, McCurry SM, Bayer A, McDavid AN, Gallacher J, McKee AC, Van Den Bussche H, Mesulam M, Brayne C, Miller BL, Riedel-Heller S, Miller CA, Miller JW, Al-Chalabi A, Morris JC, Shaw CE, Myers AJ, Wiltfang J, O'Bryant S, Olichney JM, Alvarez V, Parisi JE, Singleton AB, Paulson HL, Collinge J, Perry WR, Mead S, Peskind E, Cribbs DH, Rossor M, Pierce A, Ryan NS, Poon WW, Nacmias B, Potter H, Sorbi S, Quinn JF, Sacchinelli E, Raj A, Spalletta G, Raskind M, Caltagirone C, Bossù P, Orfei MD, Reisberg B, Clarke R, Reitz C, Smith AD, Ringman JM, Warden D, Roberson ED, Wilcock G, Rogaeva E, Bruni AC, Rosen HJ, Gallo M, Rosenberg RN, Ben-Shlomo Y, Sager MA, Mecocci P, Saykin AJ, Pastor P, Cuccaro ML, Vance JM, Schneider JA, Schneider LS, Slifer S, Seeley WW, Smith AG, Sonnen JA, Spina S, Stern RA, Swerdlow RH, Tang M, Tanzi RE, Trojanowski JQ, Troncoso JC, Van Deerlin VM, Van Eldik LJ, Vinters HV, Vonsattel JP, Weintraub S, Welsh-Bohmer KA, Wilhelmsen KC, Williamson J, Wingo TS, Woltjer RL, Wright CB, Yu C-E, Yu L, Saba Y, Pilotto A, Bullido MJ, Peters O, Crane PK, Bennett D, Bosco P, Coto E, Boccardi V, De Jager PL, Lleo A, Warner N, Lopez OL, Ingelsson M, Deloukas P, Cruchaga C, Graff C, Gwilliam R, Fornage M, Goate AM, Sanchez-Juan P, Kehoe PG, Amin N, Ertekin-Taner N, Berr C, Debette S, Love S, Launer LJ, Younkin SG, Dartigues J-F, Corcoran C, Ikram MA, Dickson DW, Nicolas G, Champion D, Tschanz J, Schmidt H, Hakonarson H, Clarimon J, Munger R, Schmidt R, Farrer LA, Van Broeckhoven C, C. O'Donovan M, Destefano AL, Jones L, Haines JL, Deleuze J-F, Owen MJ, Gudnason V, Mayeux R, Escott-Price V, Psaty BM, Ramirez A, Wang L-S, Ruiz A, Van Duijn CM, Holmans PA, Seshadri S, Williams J, Amouyel P, Schellenberg GD, Lambert J-C, Pericak-Vance MA (2019) Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nature Genetics* **51**, 414-430.

[6] Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, DeStefano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thornton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin C-F, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau M-T, Choi S-H, Reitz C, Pasquier F, Hollingworth P, Ramirez A, Hanon O, Fitzpatrick AL, Buxbaum JD, Champion D, Crane PK, Baldwin C, Becker T, Gudnason V, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston

JA, Evans D, Lovestone S, Letenneur L, Morón FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fiévet N, Huentelman MJ, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuinness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, St George-Hyslop P, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossù P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Naranjo MCD, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, European Alzheimer's Disease I, Genetic, Environmental Risk in Alzheimer's D, Alzheimer's Disease Genetic C, Cohorts for H, Aging Research in Genomic E, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannfelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley Jr TH, Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RFAG, Passmore P, Montine TJ, Bettens K, Rotter JI, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JSK, Boerwinkle E, Riemenschneider M, Boada M, Hiltunen M, Martin ER, Schmidt R, Rujescu D, Wang L-S, Dartigues J-F, Mayeux R, Tzourio C, Hofman A, Nöthen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM, Van Broeckhoven C, Moskvin V, Seshadri S, Williams J, Schellenberg GD, Amouyel P (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genetics* **45**, 1452.

[7] Dourlen P, Kilinc D, Malmanche N, Chapuis J, Lambert J-C (2019) The new genetic landscape of Alzheimer's disease: from amyloid cascade to genetically driven synaptic failure hypothesis? *Acta Neuropathologica* **138**, 221-236.

[8] Neuner SM, Tcw J, Goate AM (2020) Genetic architecture of Alzheimer's disease. *Neurobiology of Disease* **143**, 104976.

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

[10] Mutations, <https://www.alzforum.org/mutations>, Accessed 1 November.

[11] Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, Yu CE, Jondro PD, Schmidt SD, Wang K, et al. (1995)

Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* **269**, 973-977.

[12] Jiang H, Newman M, Lardelli M (2018) The zebrafish orthologue of familial Alzheimer's disease gene PRESENILIN 2 is required for normal adult melanotic skin pigmentation. *PLOS ONE* **13**, e0206155.

[13] Jayne T, Newman M, Verdile G, Sutherland G, Munch G, Musgrave I, Moussavi Nik SH, 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*.

[14] Newman M, Ebrahimie E, Lardelli M (2014) Using the zebrafish model for Alzheimer's disease research. *Frontiers in Genetics* **5**.

[15] Newman M, Hin N, Pederson S, 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**.

[16] Barthelson K, Pederson S, Newman M, Lardelli M (2020) Transcriptome analysis of a protein-truncating mutation in sortilin-related receptor 1 associated with early-onset familial Alzheimer's disease indicates effects on mitochondrial and ribosome function in young-adult zebrafish brains. *bioRxiv*, 2020.2009.2003.282277.

[17] Barthelson K, Pederson SM, Newman M, Lardelli M (2020) 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.

[18] Jiang H, Pederson SM, Newman M, Dong Y, Barthelson K, 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.

[19] Bray NL, Pimentel H, Melsted P, Pachter L (2016) Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* **34**, 525.

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

[21] Sonesson C, Love M, Robinson M (2015) Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences [version 1; peer review: 2 approved]. *F1000Research* **4**.

- [22] Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140.
- [23] Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* **11**, R25.
- [24] Risso D, Ngai J, Speed TP, Dudoit S (2014) Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature Biotechnology* **32**, 896-902.
- [25] Liberzon A (2014) A description of the Molecular Signatures Database (MSigDB) Web site. *Methods Mol Biol* **1150**, 153-160.
- [26] Hin N, Newman M, Pederson SM, Lardelli MM (2020) Iron Responsive Element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease. *bioRxiv*, 2020.2005.2001.071498.
- [27] Young MD, Wakefield MJ, Smyth GK, Oshlack A (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* **11**, R14.
- [28] Wilson DJ (2019) The harmonic mean p-value for combining dependent tests. *Proceedings of the National Academy of Sciences* **116**, 1195.
- [29] Wu D, Lim E, Vaillant F, Asselin-Labat M-L, Visvader JE, Smyth GK (2010) ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics* **26**, 2176-2182.
- [30] Wu D, Smyth GK (2012) Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic acids research* **40**, e133-e133.
- [31] Sergushichev AA (2016) An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv*, 060012.
- [32] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545.
- [33] Wickham H (2016) Springer-Verlag New York.
- [34] Kolde R (2019).

- [35] Conway JR, Lex A, Gehlenborg N (2017) UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938-2940.
- [36] Newman M, Nik HM, Sutherland GT, Hin N, Kim WS, Halliday GM, Jayadev S, Smith C, Laird AS, Lucas CW, Kittipassorn T, Peet DJ, Lardelli M (2020) Accelerated loss of hypoxia response in zebrafish with familial Alzheimer's disease-like mutation of presenilin 1. *Human Molecular Genetics* **29**, 2379-2394.
- [37] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47-e47.
- [38] Cruts M, Theuns J, Van Broeckhoven C (2012) Locus-specific mutation databases for neurodegenerative brain diseases. *Human Mutation* **33**, 1340-1344.
- [39] Thordardottir S, Rodriguez-Vieitez E, Almkvist O, Ferreira D, Saint-Aubert L, Kinhult-Ståhlbom A, Thonberg H, Schöll M, Westman E, Wall A, Eriksson M, Zetterberg H, Blennow K, Nordberg A, Graff C (2018) Reduced penetrance of the PSEN1 H163Y autosomal dominant Alzheimer mutation: a 22-year follow-up study. *Alzheimer's Research & Therapy* **10**, 45.
- [40] Rossor M, Fox N, Beck J, Campbell T, Collinge J (1996) Incomplete penetrance of familial Alzheimer's disease in a pedigree with a novel presenilin-1 gene mutation. *The Lancet* **347**, 1560.
- [41] Sherrington R, Froelich S, Sorbi S, Campion D, Chi H, Rogaeva EA, Levesque G, Rogaev EI, Lin C, Liang Y, Ikeda M, Mar L, Brice A, Agid Y, Percy ME, Clerget-Darpoux F, Piacentini S, Marcon G, Nacmias B, Amaducci L, Frebourg T, Lannfelt L, Rommens JM, St George-Hyslop PH (1996) Alzheimer's Disease Associated with Mutations in Presenilin 2 is Rare and Variably Penetrant. *Human Molecular Genetics* **5**, 985-988.
- [42] Dong Y, Newman M, Pederson S, Hin N, Lardelli M (2020) Transcriptome analyses of 7-day-old zebrafish larvae possessing a familial Alzheimer's disease-like mutation in psen1 indicate effects on oxidative phosphorylation, mcm functions, and iron homeostasis. *bioRxiv*, 2020.2005.2003.075424.
- [43] Zhu X, Lee H-g, Perry G, Smith MA (2007) Alzheimer disease, the two-hit hypothesis: An update. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1772**, 494-502.

[44] Barthelson K, Pederson SM, Newman M, Lardelli M (2020) 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. *bioRxiv*, 2020.2007.2017.207787.

[45] El-Brolosy MA, Kontarakis Z, Rossi A, Kuenne C, Günther S, Fukuda N, Kikhi K, Boezio GLM, Takacs CM, Lai S-L, Fukuda R, Gerri C, Giraldez AJ, Stainier DYR (2019) Genetic compensation triggered by mutant mRNA degradation. *Nature* **568**, 193-197.

[46] Cervantes S, González-Duarte R, Marfany G (2001) Homodimerization of presenilin N-terminal fragments is affected by mutations linked to Alzheimer's disease. *FEBS Lett* **505**, 81-86.

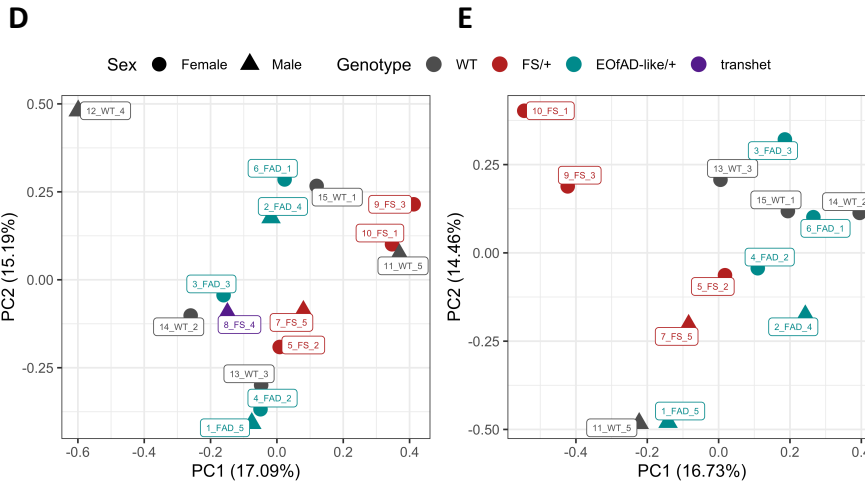
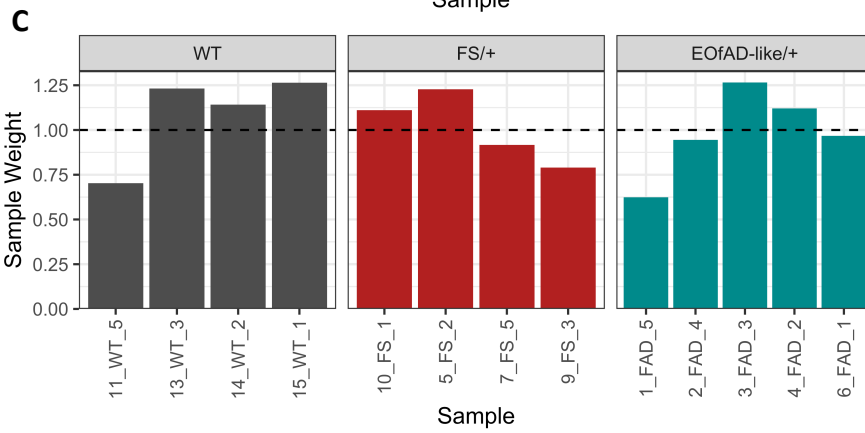
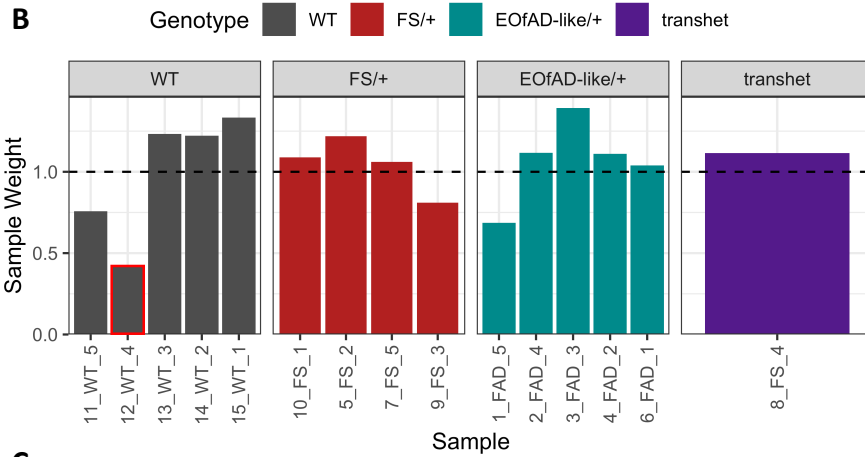
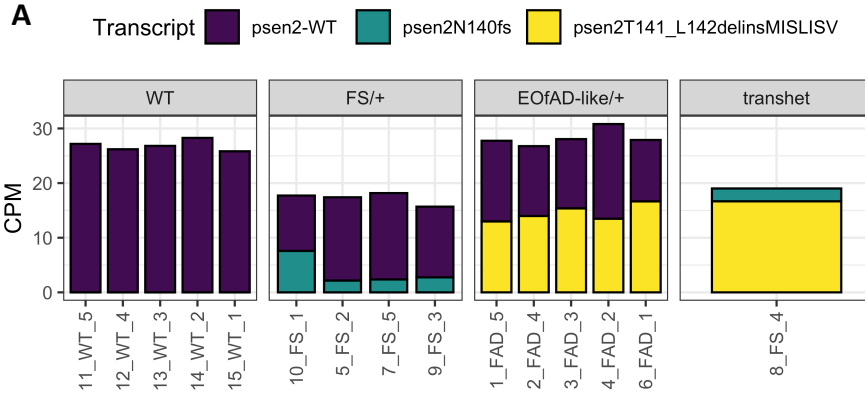
[47] Luo W, Pant G, Bhavnasi YK, Blanchard SG, Jr., Brouwer C (2017) Pathview Web: user friendly pathway visualization and data integration. *Nucleic Acids Research* **45**, W501-W508.

Supplementary material

Supplementary File 1: RNA-seq data quality control

A. Allele-specific expression (in counts per million, CPM) of *psen2* transcripts in young adult zebrafish brains. Reduced expression of the *psen2*^{N140fs} allele is observed in FS/+ brains, consistent with our previous observation that transcripts of this allele are subject to nonsense-mediated decay. Sample 8_FS_4 does not express the wild type allele of *psen2* and is actually a transheterozygous (transhet) sample. Therefore, it was omitted from the analysis. **B.** Sample weights as calculated by the *voomWithQualityWeights* algorithm on all samples sequenced. Sample 12_WT_4 is highly downweighted relative to all other samples and was omitted from subsequent analysis. **C.** Sample weights recalculated after exclusion of samples 8_FS_4 and 12_WT_4. **D.** Principal component 1 (PC1) against PC2 from a principal component analysis (PCA) on all samples of the experiment. Sample 12_WT_4 does not

cluster with the other samples. **E.** PCA plot of the *RUVseq*-normalised counts after exclusion of samples 8_FS_4 and 12_WT_4.



Supplementary File 2: Full results of differential gene expression analysis

This can be found at

<https://www.biorxiv.org/content/10.1101/2020.11.21.392761v1.full>

Supplementary File 3: Over-representation analysis using goseq.

The KEGG, GO and IRE gene sets approaching being over-represented in the DE genes list for the FS mutation. The 5 DE genes due to the EOfAD-like mutations are not found in any of the gene sets. Therefore, the results of the over-representation analysis are not shown.

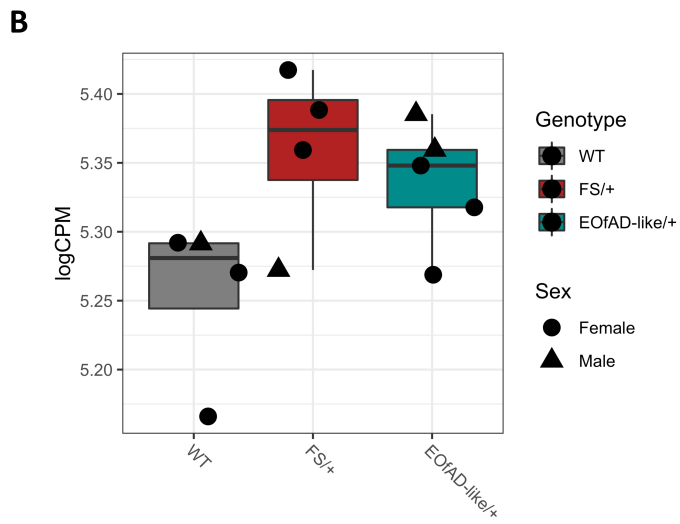
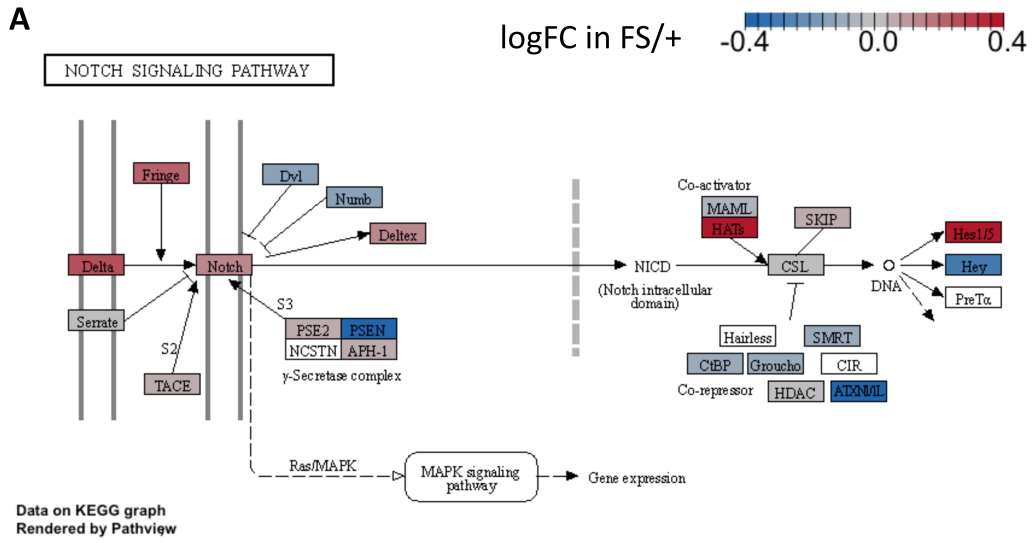
Table 1: The KEGG, GO and IRE gene sets approaching being over-represented in the DE genes list for the FS mutation. The 5 DE genes due to the EOofAD-like mutations are not found in any of the gene sets. Therefore, the results of the over-representation analysis are not shown.

Gene set	p-value	Number of DE genes in gene set	Number of genes in gene set	FDR adjusted p-value
GO_PHOSPHOTRANSFERASE_ACTIVITY_NITROGENOUS_GROUP_AS_ACCEPTOR	1.94E-05	2	5	0.28952
GO_TRNA_3_END_PROCESSING	6.94E-05	2	9	0.517902
GO_NUCLEOSIDE_DIPHOSPHATE_METABOLIC_PROCESS	1.01E-03	3	142	1
KEGG_ALZHEIMERS_DISEASE	1.11E-03	3	146	1
GO_NCRNA_3_END_PROCESSING	1.39E-03	2	39	1
GO_PYRIMIDINE_NUCLEOBASE_TRANSPORT	2.85E-03	1	2	1
GO_UREA_TRANSMEMBRANE_TRANSPORTER_ACTIVITY	2.85E-03	1	2	1
GO_TRNA_BINDING	3.58E-03	2	63	1
GO_GLYCEROL_CHANNEL_ACTIVITY	4.26E-03	1	3	1
GO_UREA_TRANSPORT	4.27E-03	1	3	1

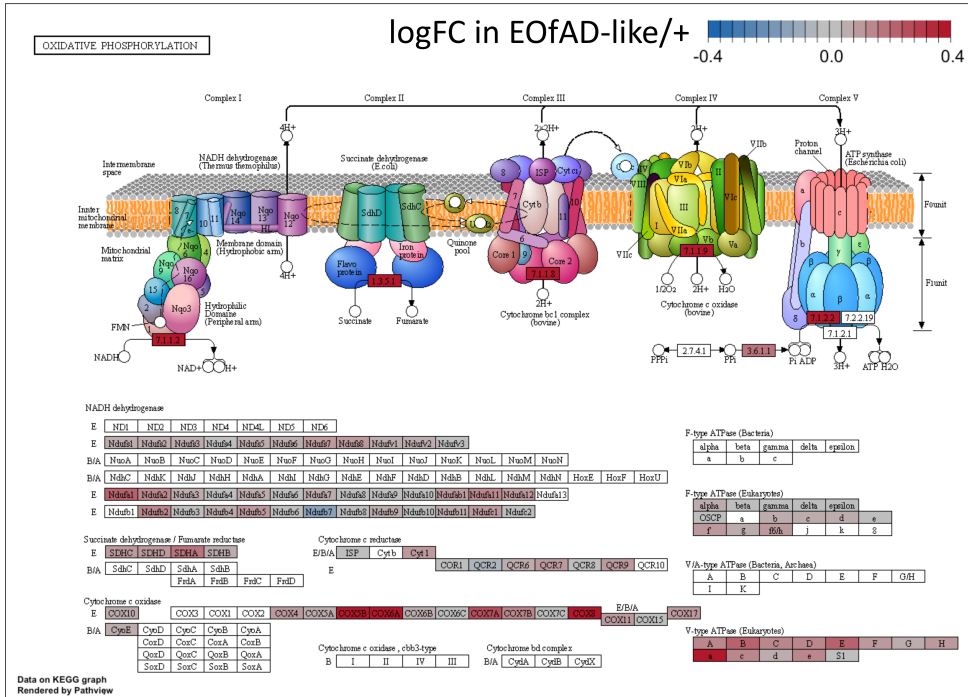
Supplementary File 4: Additional RNA-seq visualisations

A. Pathview [1] visualisation of the logFC of genes in the KEGG_NOTCH_SIGNALLING_PATHWAY gene set in FS/+ brains. **B.** Expression of *psen1* in log counts per million (logCPM). **C.** Pathview [1] visualisation of the logFC of genes in the KEGG_OXIDATIVE_PHOSPHORYLATION gene set in EOfAD-like/+ brains. **D.** Pathview [1] visualisation of the logFC of genes in the KEGG_CELL_CYCLE gene set in EOfAD-like/+ brains. **E.** Heatmap of the logFC of genes in the KEGG_LONG_TERM_POTENTIATION gene set in EOfAD-like/+ brains.

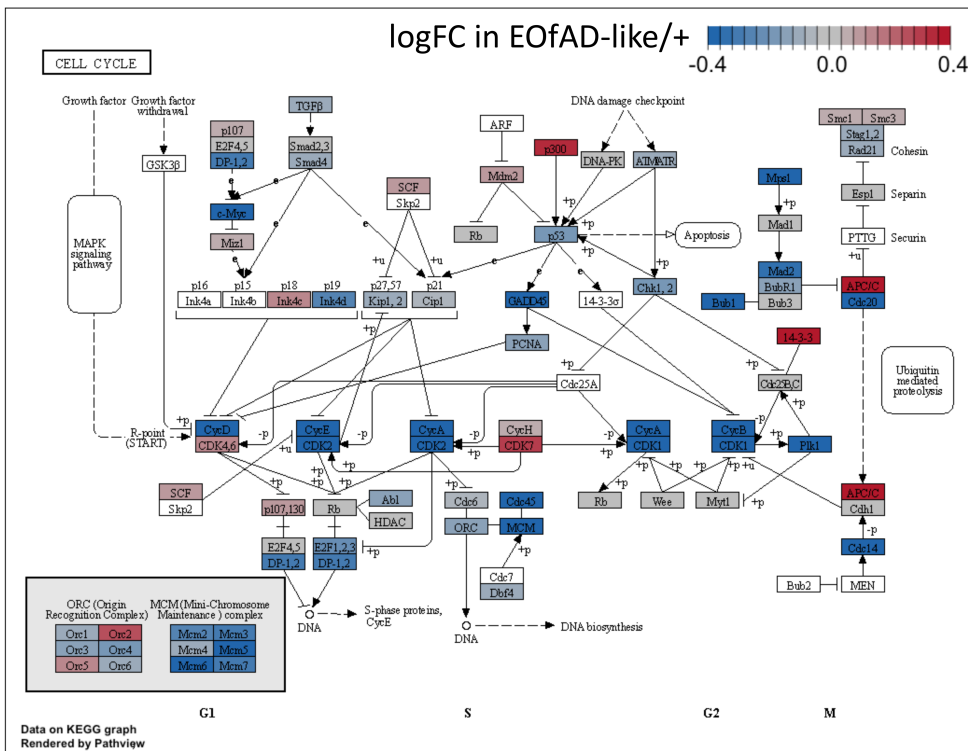
[1] Luo W, Pant G, Bhavnasi YK, Blanchard SG, Jr., Brouwer C (2017) Pathview Web: user friendly pathway visualization and data integration. *Nucleic Acids Research* **45**, W501-W508.



C

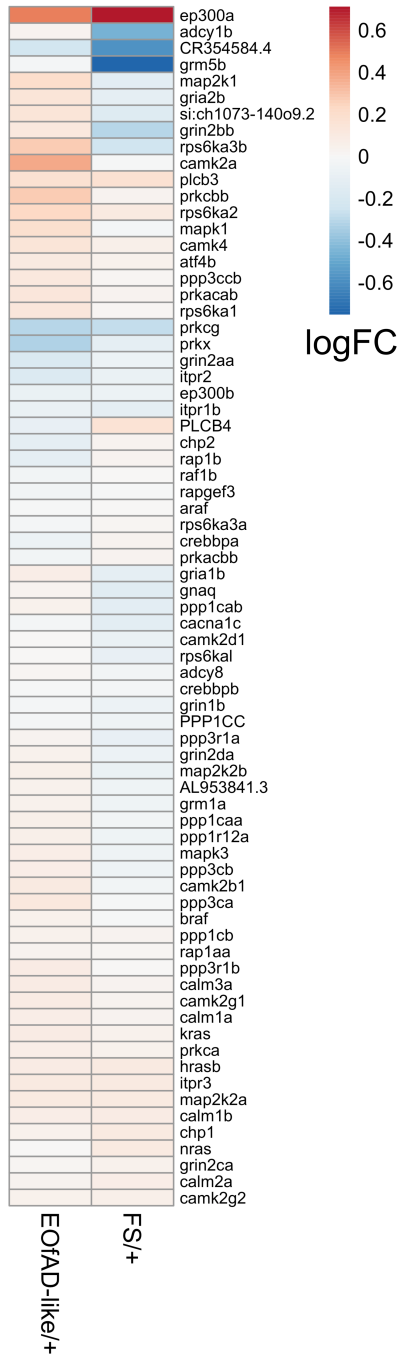


D



E

KEGG_LONG_TERM_POTENTIATION



Chapter 8: *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

Statement of Authorship

Title of Paper	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
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input checked="" type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Submitted to the Journal of Alzheimer's Disease

Principal Author

Name of Principal Author (Candidate)	Karissa Barthelson		
Contribution to the Paper	Generated one of the mutant lines of zebrafish, prepared the RNA samples for sequencing. Analysed the RNA-seq data. Drafted the manuscript		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	11/2/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Yang Dong		
Contribution to the Paper	Generated one of the mutant lines of zebrafish		
Signature		Date	12/2/2021

Name of Co-Author	Morgan Newman		
Contribution to the Paper	Supervision of zebrafish work. Editing manuscript		
Signature		Date	15/02/2021

Name of Co-Author	Michael Lardelli		
Contribution to the Paper	Supervision of project, editing of manuscript		
Signature		Date	18/02/2021

Please cut and paste additional co-author panels here as required.

***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**

Karissa **Barthelson**^{a*}, Yang **Dong**^a, Morgan **Newman**^a and Michael **Lardelli**^a

^a Alzheimer's Disease Genetics Laboratory, School of Biological Sciences, University of Adelaide, North Terrace, Adelaide, SA 5005, Australia

* Corresponding author

Complete correspondence address: Room 1.24, Molecular Life Sciences Building. North Terrace Campus. The University of Adelaide, SA 5005, Australia.

Telephone: 83134863

Email: karissa.barthelson@adelaide.edu.au

Running title: RNA-seq analysis of psen1 mutations

Abstract

Background: The most common cause of early-onset familial Alzheimer's disease (EOfAD) is mutations in *PRESENILIN 1* (*PSEN1*) allowing production of mRNAs encoding full-length, but mutant, proteins. In contrast, a single known frameshift mutation in *PSEN1* causes familial acne inversa (fAI) without EOfAD. The molecular consequences of heterozygosity for these mutation types, and how they cause completely different diseases, remains largely unexplored.

Objective: To analyse brain transcriptomes of young adult zebrafish to identify similarities and differences in the effects of heterozygosity for *psen1* mutations causing EOfAD or fAI.

Methods: RNA sequencing was performed on mRNA isolated from the brains of a single family of 6-month-old zebrafish siblings either wild type or possessing a single, heterozygous EOfAD-like or fAI-like mutation in their endogenous *psen1* gene.

Results: Both mutations downregulate genes encoding ribosomal subunits, and upregulate genes involved in inflammation. Genes involved in energy metabolism appeared significantly affected only by the EOfAD-like mutation, while genes involved in Notch, Wnt and neurotrophin signalling pathways appeared significantly affected only by the fAI-like mutation. However, investigation of direct transcriptional targets of Notch signalling revealed possible increases in γ -secretase activity due to heterozygosity for either *psen1* mutation. Transcriptional adaptation due to the fAI-like frameshift mutation was evident.

Conclusion: We observed both similar and contrasting effects on brain transcriptomes of the heterozygous EOfAD-like and fAl-like mutations. The contrasting effects may illuminate how these mutation types cause distinct diseases.

Key words

Presenilin 1, RNA-seq, zebrafish, gamma-secretase, Alzheimer's disease, acne inversa

Introduction

Cases of Alzheimer's disease (AD) can be classified by age of onset and mode of inheritance. Dominant mutations in a small number of genes cause AD with an age of onset younger than 65 years (early onset familial AD, EOfAD). On a population basis, around 60% of the mutations causing EOfAD occur in one gene, *PRESENILIN 1 (PSEN1)* [1-4].

PSEN1 encodes a multi-pass transmembrane protein resident in the endoplasmic reticulum, plasma membrane, endolysosomal pathway and other membranes [5, 6]. It has nine recognised transmembrane domains [7]. A tenth transmembrane domain may exist when PSEN1 protein is in its holoprotein state [7], before it undergoes autocatalytic endoproteolysis to form the active catalytic core of γ -secretase [8], an enzyme complex consisting of PSEN1 (or PSEN2), and the proteins NCSTN, PSENEN, and APH1A (or APH1B) [9, 10].

As a locus for genetic disease, *PSEN1* is truly remarkable both for the number of mutations found there, and the variety of diseases these mutations cause.

Mutations have been found associated with Pick's disease [11], dilated cardiomyopathy [12] and acne inversa [13]. However, over 300 mutations of *PSEN1* are known to cause EOfAD (www.alzforum.org/mutations/psen-1). In total, these mutations affect 161 codons of the gene. Remarkably, the mutations are widely distributed in the *PSEN1* coding sequence, but are particularly common in the transmembrane domains. Only three regions of the *PSEN1* protein are mostly devoid of EOfAD mutations: upstream of the first transmembrane domain; a large part of the "cytosolic loop domain" (cytosolic loop 3); and the last two thirds of the 9th transmembrane domain together with the luminal C-terminus (see **Figure 1**).

The most common outcome of mutation of a protein sequence is either no effect or a detrimental effect on the protein's evolved activity. Only rarely are mutations selectively advantageous so that they enhance organismal survival and reproduction. The very large number of EOfAD-causative mutations in *PSEN1* and their wide distribution in the protein coding sequence is consistent with a loss-of-function. However, this cannot be a simple loss of γ -secretase activity, as EOfAD-causative mutations have never been found in the genes encoding the other components of γ -secretase complexes (other than less frequent mutations in the *PSEN1* homologous gene, *PSEN2*, reviewed in [14]). Also, an *in vitro* analysis of 138 EOfAD mutations of *PSEN1* published in 2017 by Sun et al. [15] found that approximately 10% of these mutations actually increased γ -secretase activity.

Currently, the most commonly discussed hypothetical mechanism addressing how EOfAD mutations of *PSEN1* cause disease is that these act through "qualitative

changes” to γ -secretase cleavage of the AMYLOID β A4 PRECURSOR PROTEIN (A β PP) to alter the length distribution of the AMYLOID β (A β) peptides derived from it [16]. However, the comprehensive study of Sun et al. revealed no consistency in the effects of *PSEN1* EOfAD mutations on A β length distribution. The single consistent characteristic of all EOfAD mutations in both *PSEN1* and *PSEN2* is that these permit production of transcripts with coding sequences containing in-frame mutations, but terminated by the wild type stop codons (i.e. they still permit production of a full length protein). This phenomenon was first noted by De Strooper in 2007 [17] and described in detail by us in 2016 (the “reading frame preservation rule” [18]). The universality of this rule, and that it reflects a critical feature of the EOfAD-causative mechanism of *PSEN1* mutations, is shown by the fact that the P242LfsX11 frameshift mutation of *PSEN1* (hereafter referred to as P242fs) causes a completely different inherited disease, familial acne inversa (fAI, also known as hidradenitis suppurativa), *without* EOfAD [13]. (Recently, a frameshift mutation in *PSEN1*, H21PfsX2, was identified in an early-onset AD patient. However, whether the mutation is causative of EOfAD mutation is still uncertain [19]. Questionable additional claims of EOfAD-causative frameshift mutations in PSEN genes have been made and are reviewed in [18].) Critically, fAI can also be caused by mutations in *NCSTN* and *PSENEN* [13], strongly supporting that this disease is due to changes in γ -secretase activity.

Understanding the role of *PSENs* and their mutations is complicated by the partial functional redundancy shared by *PSEN1* and *PSEN2* and their complex molecular biology. For example, the *PSEN1* holoprotein has been shown to have γ -secretase-independent activities required for normal lysosomal acidification [20],

can form multimers [21-24], and may interact with the HIF1 α protein [25-27] that is critical both for responses to hypoxia and for iron homeostasis (reviewed in [28]). Additionally, within γ -secretase complexes, the PSENs act to cleave at least 149 different substrates [29]. To simplify analysis, most previous investigation of PSEN activity has involved inactivation (knock-out) of *PSEN1* and/or *PSEN2* in cells or animals, and expression of only single forms of mutant PSEN (i.e. without simultaneous expression of wild type forms). Forced expression of *PSEN* genes is also usually at non-physiological levels which has unexpected regulatory feedback effects [30]. In humans, investigating *PSEN*'s role in AD at the molecular level is restricted to post-mortem brain tissues. However, these show substantially different patterns of gene expression compared to the brains of people with mild cognitive impairment (MCI) or age-matched healthy controls [31]. Since AD is thought to take decades to develop [32], we must understand the pathological effects of EOfAD mutations in young adult brains if we wish to identify preventative treatments. For this reason, we must model EOfAD mutations in animals.

The overwhelming majority of animal modelling of AD has utilised transgenic models favoured for their apparent, partial reproduction of A β histopathology and easily discernible cognitive disabilities. However, the relationship between A β histopathology and cognitive change in these models is questionable [33]. Additionally, the most detailed form of molecular phenotyping currently available, transcriptome analysis, shows little consistency between the disturbed brain gene transcription of various transgenic models and limited concordance between them and human sporadic AD brain transcriptomes [34]. "Knock-in" mouse models of single EOfAD mutations (e.g. [35, 36]) make the fewest assumptions regarding the

pathological mechanism(s) of AD and most closely replicate the human EOfAD genetic state (i.e. incorporating a single, dominant, endogenous mutation in the heterozygous state). However, the brain transcriptomes of these mice have never been analysed, and interest in them waned due to their lack of A β histopathology and mild cognitive effects.

Analysis of mouse brain transcriptomes is complicated by strong effects on gene expression of sex [37, 38] and, potentially, litter-of-origin (i.e. due to environmental and genotypic variation) (K. Barthelson, unpublished results). In contrast, zebrafish brain transcriptomes show only subtle influences of sex, and very large numbers of siblings can be generated from single mating event, alleviating potential litter-of-origin issues [39-43]. In 2014, our laboratory began a program of creating knock-in models of EOfAD-like (and non-EOfAD-like) mutations in the zebrafish genes orthologous to *PSEN1*, *PSEN2*, and *SORL1*. In 2019 we began publishing the results of transcriptome analyses of the young adult brains of these fish [39-46] as we attempt to establish what effect(s) all the EOfAD mutations have in common (and differentiate them from the non-EOfAD-like mutations).

Our previous analyses of an EOfAD-like mutation in the zebrafish *psen1* gene, Q96_K97del, revealed very significant effects on the expression of genes involved in mitochondrial function, lysosomal acidification, and iron homeostasis [45, 46]. Although Q96_K97del follows the reading-frame preservation rule of EOfAD mutations in *PSEN1*, it is not an exact equivalent of any human EOfAD-causative mutation. Consequently, in this study we aimed to generate an additional, exactly equivalent, model of a human *PSEN1* EOfAD mutation. For technical reasons, the T428 codon of zebrafish *psen1* (equivalent to the T440 codon of human *PSEN1*)

was predicted to be readily targetable using CRISPR-Cas9 technology, and we subsequently deleted this codon in the zebrafish gene. This generated a zebrafish model of the human EOfAD mutation *PSEN1*^{T440del} [47]. This mutation was identified in a Japanese man classified as displaying a mixed dementia phenotype (variant AD with spastic paraparesis, Parkinson's disease and dementia with Lewy bodies). Also, to understand how reading-frame preserving and frameshift mutations can cause completely different diseases we generated a frameshift mutation in zebrafish *psen1*, *psen1*^{W233fs}, very similar to the fAI-causative P242fs mutation of human *PSEN1*. We then performed an RNA-seq analysis with high read depth and large sample numbers to compare the brain transcriptomes of fish from a single family of young adult siblings heterozygous for either mutation or wild type. We observed subtle, and mostly distinct, effects of the two mutations. In particular, changes in the fAI-like brain transcriptomes implied significant effects on Notch, Wnt, neurotrophin, and Toll-like receptor signalling, while changes in the EOfAD-like brain transcriptomes implied effects on oxidative phosphorylation similar to those previously seen for EOfAD-like mutations in *psen1* [45], *psen2* [40, 43], and *sor11* [39, 41].

Materials and methods

Zebrafish husbandry and animal ethics

All zebrafish (Tübingen strain) used in this study were maintained in a recirculating water system on a 14 hour light/10 hour dark cycle, fed dry food in the morning and live brine shrimp in the afternoon. All zebrafish work was conducted under the

auspices of the Animal Ethics Committee (permit numbers S-2017-089 and S-2017-073) and the Institutional Biosafety Committee of the University of Adelaide.

CRISPR-Cas9 genome editing

To mutate zebrafish *psen1*, we used the Alt-R® CRISPR-Cas9 system (Integrated DNA Technologies, Coralville, IA, USA). To generate the T428del mutation (EOfAD-like) in exon 11 of *psen1*, we used a custom-designed crRNA recognising the sequence 5' CTCCCCATCTCCATAACCTT 3' and a PAM of CGG. For the W233fs mutation (fAI-like), the crRNA was designed to recognise the sequence 5' GATGAGCCATGCGGTCCACT 3' in exon 6 of *psen1*, with a PAM sequence of CGG. We aimed to generate exact equivalents of the human P242fs mutation causing fAI, and the T440del mutation causing EOfAD by homology directed repair (HDR). For the P242fs mutation, we used a plasmid DNA template as described in [48] (synthesised by Biomatik, Kitchener, Ontario, Canada). For the T440del mutation, we used an antisense, asymmetric single-stranded oligonucleotide with phosphorothioate modifications (synthesised by Merck, Kenilworth, NJ, USA) as described in [49] (HDR template DNA sequences are given in **Supplementary File 1**).

Each crRNA was annealed with an equal amount of Alt-R® CRISPR-Cas9 tracrRNA (IDT) in nuclease free duplex buffer (IDT) by heating at 95°C for 5 minutes, then allowed to cool to room temperature, giving sgRNA solutions of 33 µM (assuming complete heteroduplex formation of the RNA molecules). Then, 1 µL of the sgRNA solution was incubated with 1 µL of Alt-R® S.p.Cas9 Nuclease 3NLS (IDT) at 64 µM at 37°C for 10 minutes to form ribonucleoprotein (RNP)

complexes. The final concentration for the linear ssODN for the T428del mutation was 1 μ M, and the final concentration of the plasmid DNA for the W233fs mutation was 25 ng/ μ L. Approximately 2-5 nL of RNP complexes in solution with the respective template DNAs were injected into Tübingen strain zebrafish embryos at the one cell stage. The procedures followed for testing of the mutagenesis efficiencies of CRISPR-Cas9 systems using allele-specific polymerase chain reactions and T7 endonuclease I assays, and the breeding strategy to isolate the mutations of interest, are described in [41, 50].

RNA-seq raw data generation and processing

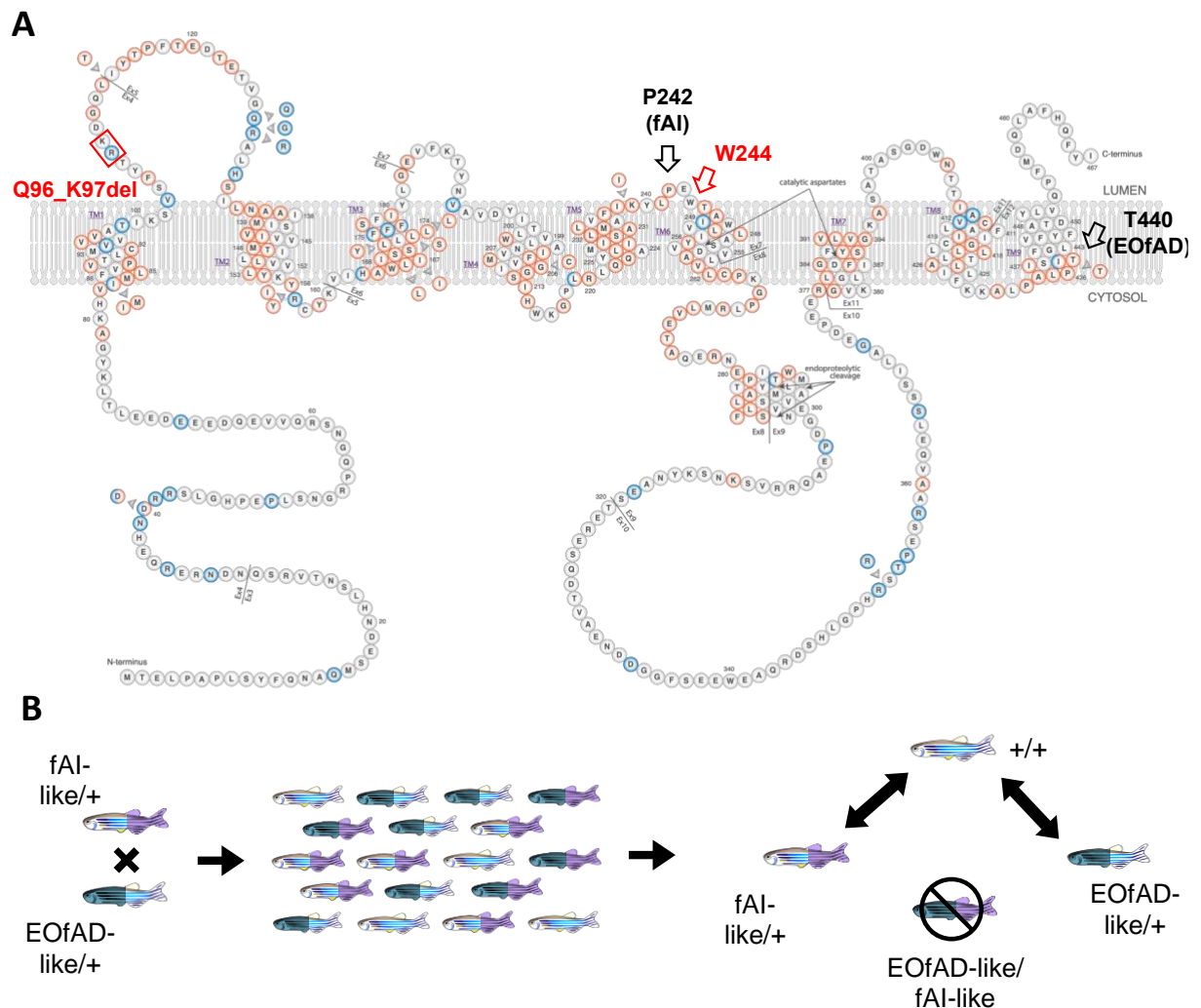


Figure 1: Experimental design. **A** Schematic of the human PSEN1 protein adapted from <https://www.alzforum.org/mutations/psen-1> with permission from FBRI LLC (Copyright © 1996–2020 FBRI LLC. All Rights Reserved. Version 3.3 – 2020). Amino acid residues are colour-coded as to whether they are pathogenic for Alzheimer’s disease (red) or their pathogenicity is unclear (blue). The human mutation sites (P242 (fAI) and T440 (EOfAD)) are indicated by black arrows. The site of the zebrafish W233fs-equivalent codon (W244) is shown by the red arrow. Note that the human T440 codon is equivalent to the zebrafish T428 codon. The residues equivalent to those deleted in the Q96_K97del mutation of zebrafish *psen1* analysed previously are indicated by a red box. **B** A fish heterozygous for the W233fs mutation (fAI-like/+) was mated with a fish heterozygous for the T428del mutation (EOfAD-like/+). The resulting family of fish contain genotypes fAI-like/+, EOfAD-like/+, EOfAD-like/fAI-like and their wild type siblings. The pairwise comparisons performed in the RNA-seq experiment are depicted. Since the EOfAD-like/fAI-like genotype is not representative of any human disease, it was not analysed. See online version for colour.

We performed RNA-seq on a family of zebrafish as described in **Figure 1**. Total RNA (with genomic DNA depleted by DNaseI treatment) was isolated from the brains of $n = 4$ fish per genotype and sex as described in [41]. Then, 500 ng of total RNA ($RIN_e > 9$) was delivered to the South Australian Genomics Centre (SAGC, Adelaide, Australia) for polyA+ library construction (with unique molecular identifiers (UMIs)) and RNA-sequencing using the Illumina Novaseq S1 2x100 SE platform.

The raw fastq files from SAGC were provided as 100 bp paired end reads as well as an index file containing the UMIs for each read (over two Novaseq lanes which were subsequently merged). The merged raw data was processed using a developed Nextflow [51] RNA-seq workflow (see <https://github.com/sagc->

[bioinformatics/sagc-rnaseq-nf](#)). Briefly, UMIs were added to headers of each read using *fastp* (v0.20.1). Alignment of the reads to the zebrafish genome (GRCz11, Ensembl release 98) was performed using *STAR* (v2.5.3a). Then, reads which contained the same UMI (i.e. PCR duplicates) were deduplicated using the *dedup* function of *umi_tools* (version 1.0.1). Finally, the gene-level counts matrix was generated using *featureCounts* from the *Subread* (version 2.0.1) package.

Differential gene expression

Statistical analysis of the RNA-seq data was performed using *R* [52]. Since lowly expressed genes are considered uninformative for differential expression analysis, we omitted genes with less than 0.1 counts per million (CPM) (following the 10/minimum library size in millions rule described in [53]). Library sizes after omitting the lowly expressed genes ranged between 61 and 110 million reads. These were normalised using the trimmed mean of M-values (TMM) method [54]. To test for differential expression of genes due to heterozygosity for the T428del or W233fs mutation, we used a generalised linear model and likelihood ratio tests using *edgeR* [55, 56]. A design matrix was specified with the wild type genotype as the intercept, and the T428del/+ and W233fs/+ genotypes as the coefficients. We considered a gene to be differentially expressed (DE) due to each *psen1* mutant genotype if the FDR adjusted p-value was less than 0.05.

Enrichment analysis

We tested for over-representation of gene ontology (GO) terms within the DE gene lists using *goseq* [57], using the average transcript length per gene to calculate the

probability weighting function (PWF). We considered a GO term to be significantly over-represented within the DE gene lists relative to all detectable genes in the RNA-seq experiment if the FDR-adjusted p-value generated by *goseq* was less than 0.05.

We also performed enrichment analysis on the entire list of detectable genes by calculating the harmonic mean p-value from the raw p-values calculated from *fry* [58], *camera* [59] and *fgsea* [60, 61] as described in [41]. To test for changes to gene expression in a broad range of biological processes, we used the KEGG [62] gene sets obtained from MSigDB [63] using the *msigdb* package [64]. We also used *msigdb* to obtain gene sets which contain genes that show changed expression in response to changes in the Notch signalling pathway (*NGUYEN_NOTCH1_TARGETS_UP* and *NGUYEN_NOTCH1_TARGETS_DN*, *NOTCH_DN.V1_UP*, *NOTCH_DN.V1_DN* and *RYAN_MANTLE_CELL_LYMPHOMA_NOTCH_DIRECT_UP*). We also tested for evidence of possible iron dyshomeostasis using gene sets containing genes encoding transcripts which contain iron-responsive elements in their untranslated regions (described in [45]).

Comparison of the T428del and Q96_K97del mutations in *psen1*

Isolation of the zebrafish Q96_K97del mutation in zebrafish *psen1* and analysis of its effects on zebrafish brain transcriptomes have been described previously [45, 46]. That dataset is comprised of brain RNA-seq data for fish heterozygous for the Q96_K97del mutation and their wild type siblings, at 6 months old (young adult) and 24 months old (aged), and under normoxia or hypoxia treatment (n = 4 fish

per genotype, age and treatment). In the analysis presented here, we performed enrichment analysis using the methods described above on the entire dataset, but presented the results for the pairwise comparison between 6-month-old Q96_K97del/+ fish and wild type fish under normoxia.

To obtain a broader comparison on the effects of the Q96_K97del and T428del mutations, we performed adaptive, elastic-net sparse PCA (AES-PCA) [65] as implemented in the *pathwayPCA* package [66]. For this analysis, we utilised the HALLMARK [67] gene sets from MSigDB to generate the pathway collection. The pathway principal components (PCs) were calculated only on the gene expression data from samples heterozygous for an EOfAD-like mutation (Q96_K97del or T428del) and their wild type siblings under normoxia at 6 months of age. Then, the categorical effect of genotype was tested for association with the pathway PCs using a permutation-based regression model as described in [66].

Data availability

The paired end fastq files and the output of *featureCounts* have been deposited in GEO under accession number GSE164466. Code to reproduce this analysis can be found at https://github.com/karissa-b/psen1_EOfAD_fAI_6m_RNA-seq.

Results

Generation of an EOfAD-like and a fAI-like mutation in zebrafish *psen1*

An unsolved puzzle regarding the dominant EOfAD mutations of human *PSEN1* (and *PSEN2*) is why these are consistently found to permit production of

transcripts in which the reading frame is preserved, while heterozygosity for mutations causing frameshifts (or deleting the genes) does not cause EOfAD. To investigate this quandary in an *in vivo* model, we initially aimed to generate mutations in zebrafish *psen1* which would be exact equivalents of the T440del and P242fs mutations using homology directed repair (HDR). While screening for the desired mutations, we identified the mutations W233fs and T428del, both likely generated by the non-homologous end joining (NHEJ) pathway of DNA repair. T428del is a 3 nucleotide deletion which, nevertheless, produces a protein-level change exactly equivalent to that observed for the human T440del mutation. Hereafter, for simplicity, we refer to the T428del mutation as “EOfAD-like”. W233fs is an indel mutation causing a frameshift in the second codon downstream of the zebrafish *psen1* proline codon equivalent to human *PSEN1* codon P242. This change is still within the short third luminal loop of the Psen1 protein (see **Figure 1**). Assuming no effect on splicing, the frameshift mutation results in a premature stop codon 36 codons downstream of W233 (**Figure S1 in Supplementary File 2**). Hereafter, we refer to this mutation as “fAI-like”. The alignment of the wild type and mutant sequences in humans and zebrafish is shown in **Figure 2**.

Heterozygosity or homozygosity for neither the EOfAD-like nor the fAI-like mutation produces any obvious morphological defects. However, this is unsurprising considering that rare examples of humans homozygous for EOfAD mutations are known [68, 69] and loss of *PSEN1* γ -secretase activity is, apparently, compatible with viability in zebrafish [70] and rats [71] (although not in mice [72]).

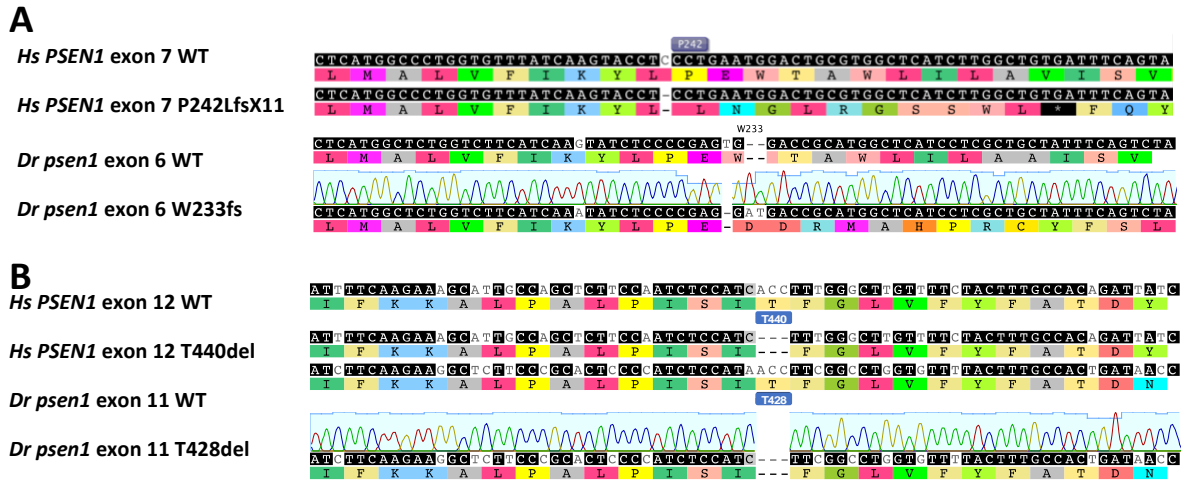


Figure 2: A Alignment of a region of wild type human (*Hs*) *PSEN1* exon 7, and the same region containing the human P242fs (P242LfsX11) mutation, and the equivalent zebrafish (*Dr*) *psen1* exon 6 wild type and W233fs sequences. **B** Alignment of a region of wild type human (*Hs*) *PSEN1* exon 12, and the same region containing the human T440del mutation, and the equivalent zebrafish (*Dr*) *psen1* exon 11 wild type and T428del sequences.

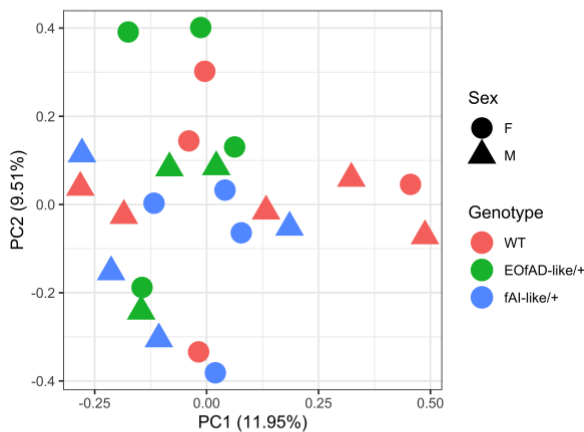
Transcriptome analysis

To investigate global changes to the brain transcriptome due to heterozygosity for the EOfAD-like or fAl-like mutations in *psen1*, we performed mRNA-seq on a family of fish as described in **Figure 1**. The family of sibling fish were raised together in a single tank, thereby reducing sources genetic and environmental variation between individuals and allowing subtle changes to the transcriptome to be detected with minimal confounding effects.

To begin our exploration of the similarities and differences between the brain transcriptomes of the mutant and wild type fish, we first performed principal

component analysis (PCA) on the gene level, log transformed counts per million (logCPM) of the zebrafish RNA-seq samples. A plot of principal component 1 (PC1) against PC2 did not show distinct clustering of samples by genotype or sex, indicating that these variables do not result in stark changes to the brain transcriptome. This is consistent with our previous observations of EOofAD-like mutations in other genes [39-41]. However, some separation of the EOofAD-like and fAI-like samples is observed across PC2, indicating distinct, but subtle, differences between these transcriptomes. Notably, the majority of the variation in this dataset is not captured until PC6 (**Figure 3**).

A



B

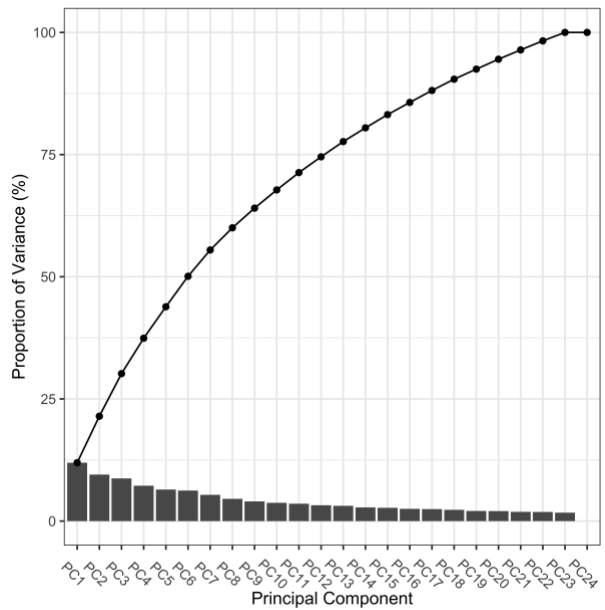


Figure 3: Principal component (PC) analysis of the gene expression values for the RNA-seq experiment. **A** PC1 plotted against PC2 for each sample. Each point represents a sample and is coloured according to *psen1* genotype. Female (F) samples appear as circles and male (M) samples appear as triangles. **B** Scree plot indicating the variance explained by each principal component. The points joined by lines indicate the cumulative variance explained by each PC. See online version for colour.

Heterozygosity for the EOfAD-like or fAI-like mutations of *psen1* causes only subtle effects on gene expression

Which genes are dysregulated due to heterozygosity for the EOfAD-like or the fAI-like mutations? To address this question, we performed differential gene expression analysis using a generalised linear model and likelihood ratio tests with *edgeR*. We observed statistical evidence for 13 genes as significantly differentially expressed (DE) due to heterozygosity for the EOfAD-like mutation, and 5 genes due to the fAI-like mutation (**Figure 4, Supplementary Table 1**). Notably, *psen1* was the most significantly DE gene due to heterozygosity for the fAI-like mutation (logFC = -0.8, FDR = 1.33e-78), consistent with the observation that frame-shift mutations commonly induce nonsense-mediated mRNA decay when they result in premature stop codons (reviewed in [73]). Total levels of *psen1* transcripts were unchanged in EOfAD-like/+ brains (logFC = -0.0065, FDR = 1, **Figure S7 in Supplementary File 2**). No DE genes were found to be shared between the comparisons of either form of heterozygous mutant to wild type, or were found to be significantly overrepresented by any gene ontology (GO) terms by *goseq* (for the top 5 most significantly over-represented GO terms in each comparison, see

Tables S1 and S2 in Supplementary File 2). This is not unexpected due to the relatively low number of significantly DE genes detected.

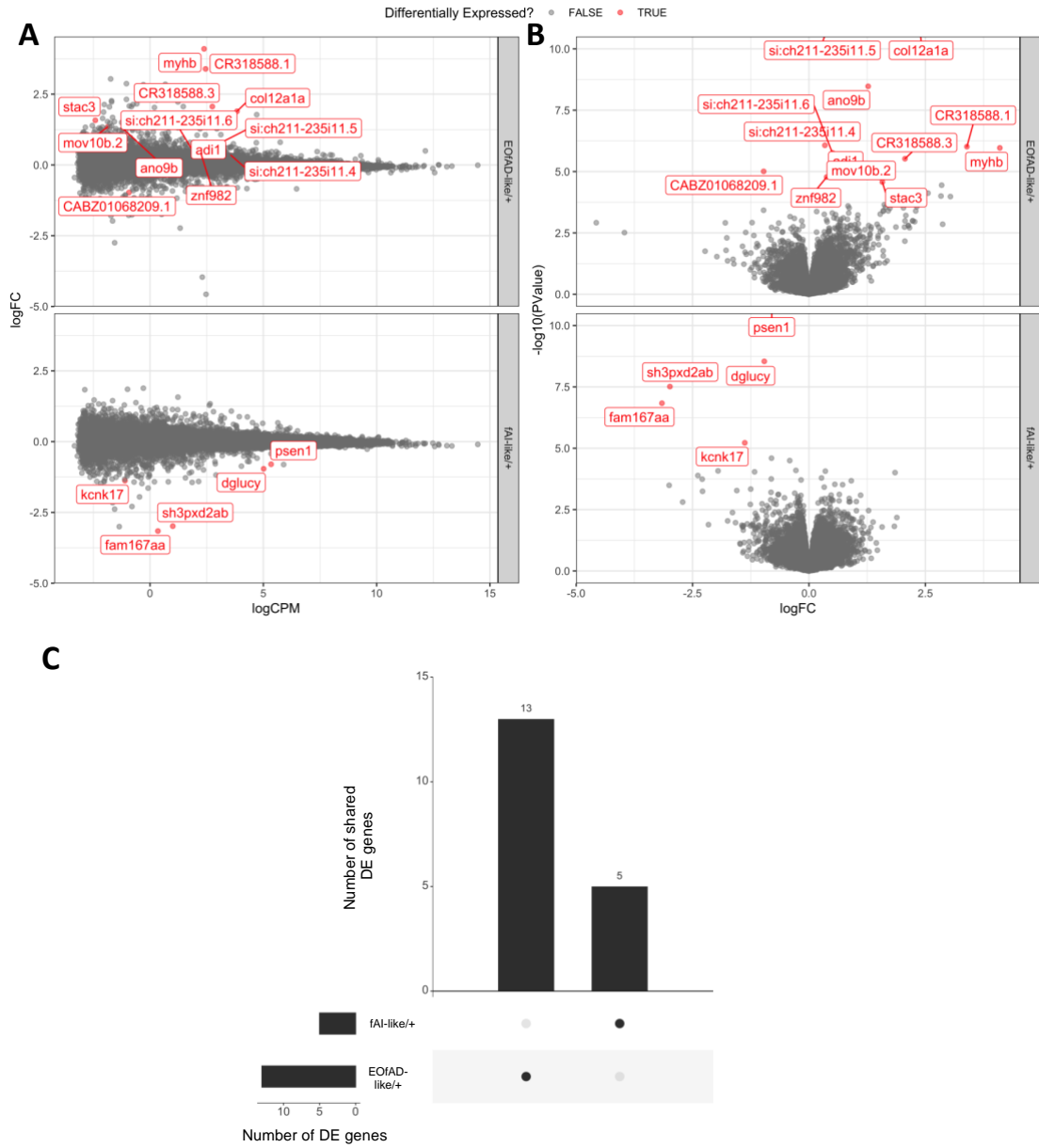


Figure 4: Differential expression analysis. **A** Mean-difference (MD) plots and **B** volcano plots of changes to gene expression in EOfAD-like/+ and fAI-like/+ mutant zebrafish brains. Note that the limits of the y-axis in **B** are restrained to between 0 and 10 for visualisation purposes. **C** Upset plot indicating the low number of genes which are significantly differentially expressed (DE) in either comparison. See online version for colour.

Significant differences in gene expression between the EOfAD-like and fAI-like mutants can be detected at the pathway level.

Since very few DE genes were detected in each comparison of heterozygous *psen1* mutant fish to their wild type siblings, we performed enrichment analysis on all detectable genes in the RNA-seq experiment. Our method, inspired by the *EGSEA* framework [74], involves calculation of the harmonic mean p-value [75] from the raw p-values of three different rank-based gene set testing methods: *fry* [58], *camera* [59] and *GSEA* [60, 61]. Unlike *EGSEA*, we use the harmonic mean p-value to combine the raw p-values, as the harmonic mean p-value has been specifically shown to be robust for combining dependent p-values [75]. We performed enrichment testing using the KEGG gene sets (describing 186 biological pathways and processes) to obtain information on changes to activities for these pathways. We also tested for evidence of iron dyshomeostasis using our recently defined sets of genes containing iron-responsive elements (IREs) in the untranslated regions of their mRNAs [45]. We observed statistical evidence for 7 KEGG gene sets as significantly altered by heterozygosity for the EOfAD-like mutation and 11 KEGG gene sets as significantly altered by heterozygosity for the fAI-like mutation (**Figure 5**, full results of the raw p-values from each algorithm as well as the harmonic mean p-value can be found in **Supplementary Table 2**).

Gene sets significantly altered in the brains of both forms of heterozygous mutant included the KEGG gene sets for cytokine receptor interactions, Jak/Stat signalling, and encoding the components of the ribosomal subunits. Inspection of the leading edge genes (which can be interpreted as the core genes driving the enrichment of a gene set) showed that similar genes were driving the enrichment of the gene sets for cytokine receptor interactions and Jak/Stat signalling. Similar genes were also driving the enrichment of the *KEGG_RIBOSOME* gene set in both heterozygous mutants. However, the magnitude of the logFC was greater in the fAI-like/+ samples, suggesting a stronger effect (**Figure S4-S6** in **Supplementary File 2**). Gene sets which were only altered significantly by heterozygosity for the EOfAD-like mutation were involved in energy metabolism (*KEGG_PARKINSONS_DISEASE*, *KEGG_OXIDATIVE_PHOSPHORYLATION*, and *KEGG_CITRATE_CYCLE_TCA_CYCLE*). Notably, the KEGG gene sets for Parkinson's disease, Huntington's disease, and for oxidative phosphorylation, share 55 leading-edge genes, implying that their enrichment is driven by, essentially, the same gene expression signal (**Figure 5**). Conversely, the 10 KEGG gene sets found to be altered significantly by heterozygosity for the fAI-like mutation appear to be driven mostly by distinct gene expression signals. No IRE gene sets were observed statistically to be altered in the brains of either mutant, suggesting that iron homeostasis is unaffected (at least at 6 months of age). The changes to expression of genes within the KEGG gene sets are likely not due to broad changes in cell-type proportions in the zebrafish brain samples, since the expression of marker genes of neurons, astrocytes, oligodendrocytes and microglia was similar in all samples (**Figure S13** in **Supplementary File 2**).

Figure 5: **A** KEGG gene sets with FDR-adjusted harmonic mean p-values of < 0.05 in *psen1* EOfAD-like/+ and fAI-like/+ mutant brains. The colour of the cells indicates the level of significance (brighter colour indicates greater statistical significance, while dark grey indicates the FDR-adjusted harmonic mean p-value > 0.05). The number within each cell is the FDR-adjusted harmonic mean p-value. **B** Upset plot indicating the overlap of leading edge genes from the fgsea algorithm which drive the enrichment of gene sets significantly altered in EOfAD-like/+ and **C** fAI-like/+ brains. See online version for colour.

The EOfAD-like and fAI-like mutations alter expression of Notch signalling genes

Notch signalling plays a critical role in many cell differentiation events and is dependent on PSEN's γ -secretase activity. Disturbance of Notch signalling due to decreased γ -secretase activity has been suggested to contribute to the changes in skin histology of fAI, as Notch signalling is required for normal epidermal maintenance ([76-78] and reviewed by [79]). However, fAI has not been reported as associated with EOfAD, despite that the T440del mutant form of *PSEN1* appears to have little intrinsic γ -secretase activity [15]. The expression of genes involved in the KEGG gene set for the Notch signalling pathway was observed to be highly significantly altered in the brains of fAI-like/+ mutants, but not of EOfAD/+ mutants, implying that γ -secretase activity might only be affected significantly by the frameshift, fAI-like mutation (**Figure 5**). However, inspection of the logFC of genes in the *KEGG_NOTCH_SIGNALING_PATHWAY* gene set revealed similar patterns of changes to gene expression in both mutants (**Figure 6**). Upregulation of the genes encoding the Notch and Delta receptors is observed in both mutants compared to wild type. In fAI-like/+ brains, we observe

downregulation of the downstream transcriptional targets of the Notch intracellular domain (NICD), implying decreased Notch signalling (and, likely, reduced γ -secretase activity). Genes encoding repressors of Notch signalling are observed to be upregulated (i.e. *dyl* and *numb*), reinforcing this interpretation.

Since the KEGG gene set for Notch signalling only contains two genes that are direct transcriptional targets of the NICD, we investigated further whether Notch signalling is perturbed in both mutants by analysis of gene sets from MSigDB containing information on genes responsive to Notch signalling in different cell lines: *NGUYEN_NOTCH1_TARGETS_UP*; *NGUYEN_NOTCH1_TARGETS_DN*; *NOTCH_DN.V1_UP*; *NOTCH_DN.V1_DN*; and *RYAN_MANTLE_CELL_LYMPHOMA_NOTCH_DIRECT_UP*. The *NGUYEN_NOTCH1_TARGETS_UP* and *_DOWN* gene sets consist of genes which have been observed as up- or downregulated respectively in response to a constitutively active Notch receptor in keratinocytes [80]. The *NOTCH_DN.V1_UP* and *_DN* gene sets contain genes which are up- and down-regulated respectively in response to treatment with the γ -secretase inhibitor DAPT in a T-cell acute lymphoblastic leukemia (T-ALL) cell line [81]. The *RYAN_MANTLE_CELL_LYMPHOMA_NOTCH_DIRECT_UP* gene set contains genes showing both increased expression upon rapid activation of Notch signalling by washout of the γ -secretase inhibitor compound E, and evidence for a NICD binding site in the promotor by ChIP-seq, in mantle cell lymphoma cell lines [82]. (Note that there is no equivalent “*RYAN_MANTLE_CELL_LYMPHOMA*” gene set representing genes downregulated in response to Notch signalling.) Of these 5 gene sets, statistical support was found only for changes to the expression of genes in the *RYAN_MANTLE_CELL_LYMPHOMA_NOTCH_DIRECT_UP* gene

set, and this was found for both the EOfAD-like ($p=0.006$) and the fAl-like ($p=0.008$) mutants. The leading edge genes were mostly observed to be upregulated, which supports increases in Notch signalling (implying increased γ -secretase activity). Transcriptional adaptation (previously known as “genetic compensation”) might contribute to the apparent increase in Notch signalling in the frameshift, fAl-like/+ mutant brains via upregulated expression of the *psen1*-homologous gene, *psen2* [83, 84]. Although no statistically significant differences in expression were observed for *psen2* in the differential expression test using *edgeR* (see **Supplementary Table 1**), a trend towards upregulation in the fAl-like/+ mutants was observed following a simple Student’s t-test ($p=0.074$, **Figure 6D**). El-Brolosy et al. [83] showed that the wild type allele of a mutated gene can also be upregulated by transcriptional adaptation (where the mutation causes nonsense-mediated decay, NMD, of mutant transcripts). Inspection of the number of reads aligning to the W233 mutation site across samples indicates that the expression of the wild type *psen1* allele in fAl-like/+ brains appears to be greater than 50% of the expression of the wild type *psen1* allele in wild type brains ($p = 0.006$), providing further evidence for transcriptional adaptation due to the fAl-like mutation (**Figure S8 in Supplementary File 2**).

Together, these results suggest that Notch signalling and, by implication, γ -secretase activity, may be enhanced in *psen1* mutant brains. However, future biochemical assays should be performed to confirm this prediction.

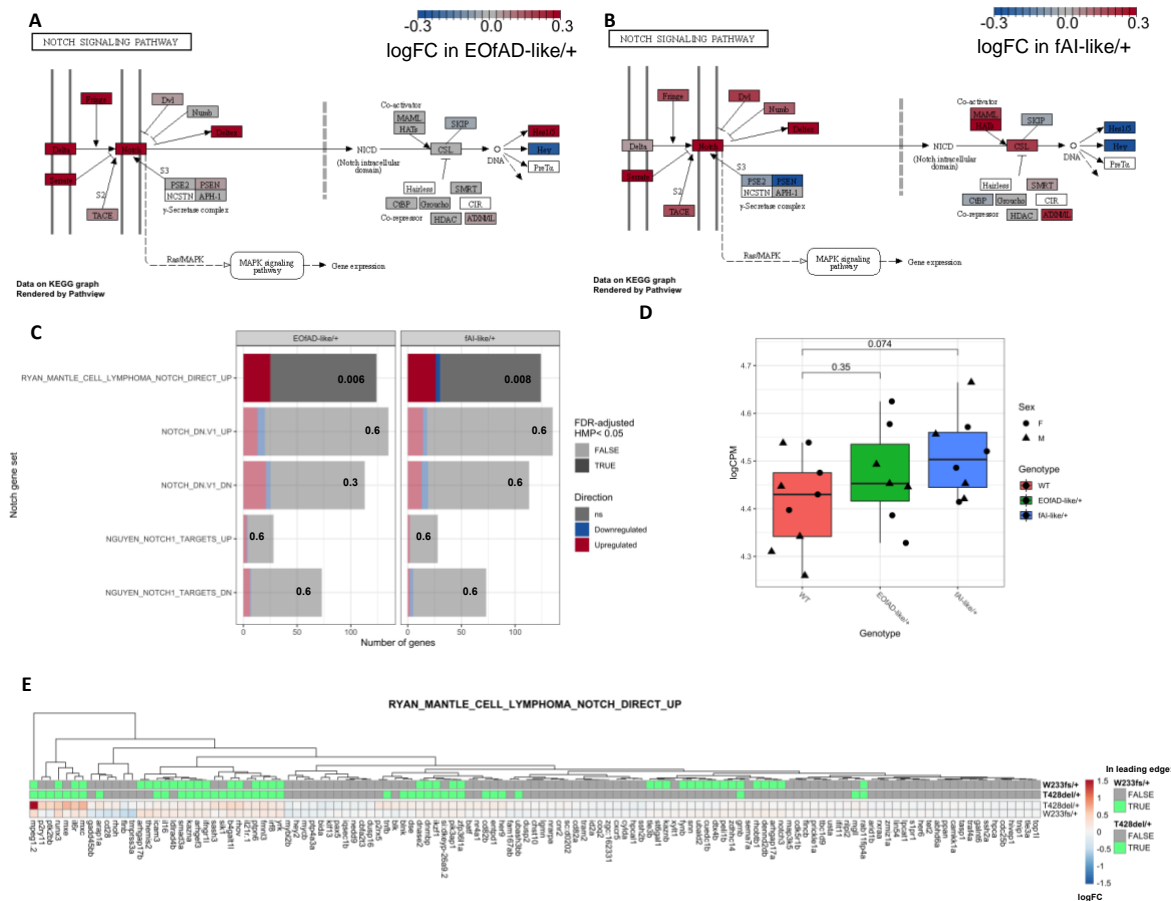


Figure 6: A Pathview [85] visualisation of the changes to gene expression in the *KEGG_NOTCH_SIGNALING_PATHWAY* gene set in EOfAD-like/+ mutants and **B** fAl-like/+ mutants. **C** The proportion of genes with increased expression (red, $z > \sqrt{2}$) and decreased expression (blue, $z < -\sqrt{2}$) in MSigDB gene sets for Notch signalling in EOfAD-like/+ and fAl-like/+ mutant brains. Gene sets which contained a FDR-adjusted harmonic mean p-value (HMP) < 0.05 appear less transparent. The FDR adjust p-values are also listed on the bars. **D** The expression of *psen2* is trending towards upregulation, particularly in fAl-like/+ mutants. Here, p-values were determined by Student's unpaired t-tests. **E** Heatmap indicating the logFC values for genes in the *RYAN_MANTLE_CELL_LYMPHOMA_NOTCH_DIRECT_UP* gene set. Genes are clustered based on their Euclidean distance, and are labelled with green if they appear in the leading edge of the *fgsea* algorithm for each comparison of a *psen1* heterozygous mutant with wild type. See online version for colour.

The EOfAD-like mutation T428del has a milder phenotype than the previously studied Q96_K97del EOfAD-like mutation of *psen1*

The T428del mutation of *psen1* is the first identified zebrafish mutation exactly equivalent, (at the protein sequence level), to a characterised human EOfAD mutation. Therefore, we sought to assess the consistency of its effects with those of a previously studied EOfAD-like mutation, Q96_K97del, and to identify cellular processes affected in common by the two mutations. The Q96_K97del mutation deletes two codons in the sequence encoding the first luminal loop of the Psen1 protein (see **Figure 1**). Comparison of transcriptomes from the 6-month-old brains of Q96_K97del/+ and wild type siblings previously predicted changes to expression of genes involved in energy metabolism, iron homeostasis and lysosomal acidification [45, 46]. To compare which cellular processes are affected by heterozygosity for the Q96_K97del mutation or the T428del mutation, we first performed enrichment analysis on the RNA-seq data previously generated by our analysis of zebrafish heterozygous for the Q96_K97del mutation relative to their wild type siblings. Here, we observed that heterozygosity for the Q96_K97del mutation results in significant alterations in 7 KEGG gene sets (at 6 months of age during normoxia, Figure 7A). We also found statistical evidence for altered expression of genes possessing IREs in their 3' UTRs (see *IRE3_ALL* in **Figure 7A**), consistent with our previous finding using a different method of enrichment analysis [45]. Gene sets affected in common between the two EOfAD-like mutations in *psen1* are involved in energy metabolism and protein translation (**Figure 7A**). The expression of genes involved in protein degradation, and of

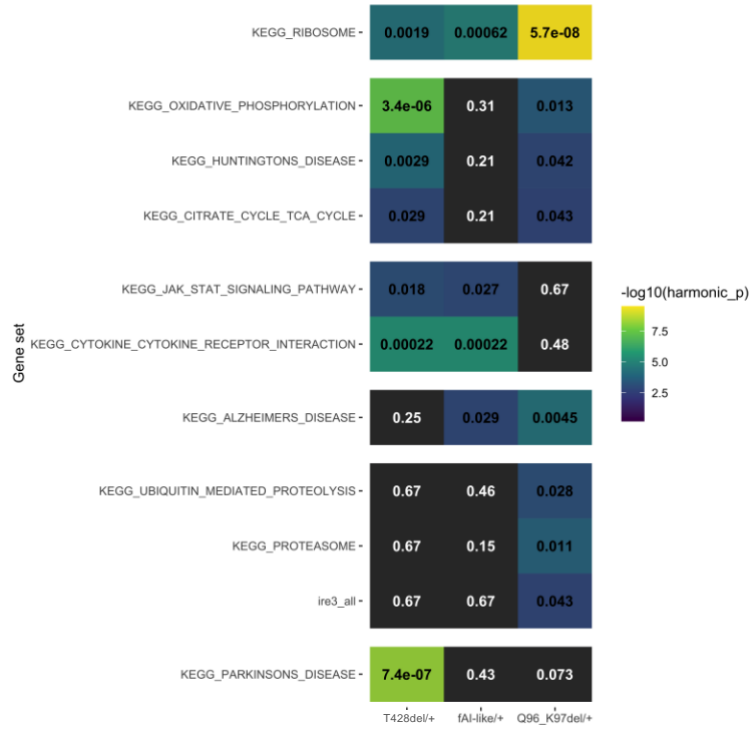
genes containing IREs in the 3' UTRs of their transcripts, appeared significantly altered only by the Q96_K97del mutation (**Figure 7A**).

We also compared the effects of the two EOfAD-like mutations using adaptive, elastic-net sparse PCA (AES-PCA) as implemented in the *pathwayPCA* package [66]. AES-PCA allows reduction of data dimensionality and for the overall activity of predefined gene sets to be observed in a sample-specific manner [65]. To obtain a global view of the changes to gene expression between the two *psen1* EOfAD-like mutations over the two datasets, we utilised the HALLMARK gene sets that encompass 50 distinct biological processes (rather than the 186 KEGG gene sets that share many genes).

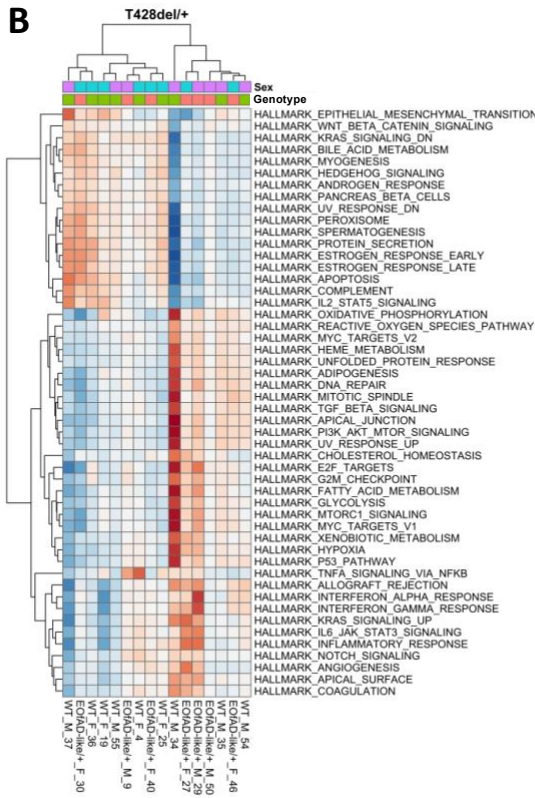
The latent variables estimated by AES-PCA for the HALLMARK gene sets (i.e. the first principal components) in each dataset did not show any significant association with *psen1* genotype, suggesting that changes to gene expression (measured over entire brains) are too subtle to be detected as statistically significant using this method. However, clustering of the calculated PC1 values by AES-PCA for each HALLMARK gene set in each sample and dataset revealed that samples in the Q96_K97del dataset clustered mostly according to genotype (one wild type sample did not follow the trend), supporting that heterozygosity for the Q96_K97del mutation does result in marked effects on gene expression for the HALLMARK gene sets. Conversely, clustering of PC1 values in the T428del dataset resulted in two distinct clusters of samples. However, samples did not group by genotype over the two clusters to the same extent as seen for the Q96_K97del dataset. Intriguingly, the Q96_K97del dataset had less sample numbers per genotype (n = 4), and did not have as great sequencing depth as the

current RNA-seq experiment. Therefore, this supports that heterozygosity for the Q96_K97del mutation has more consistent (more severe) effects on young adult brain transcriptomes than heterozygosity for the T428del mutation (**Figure 7 B,C**).

A



B



C

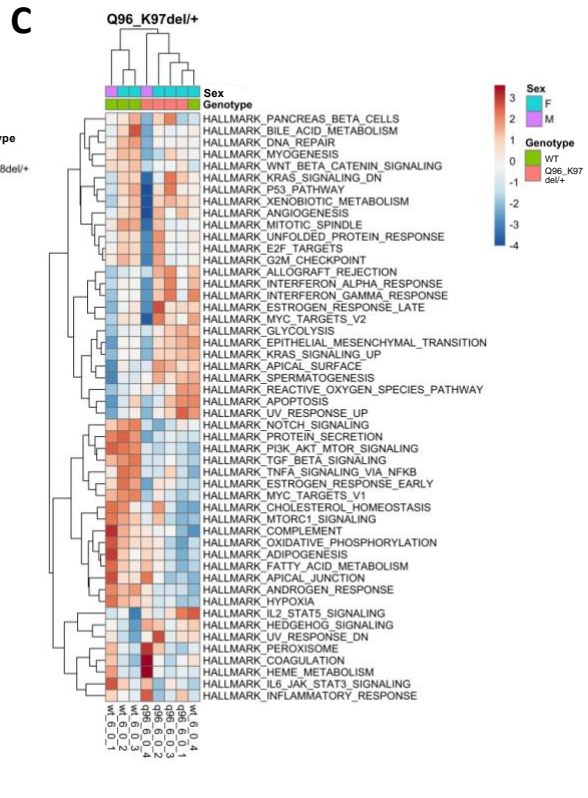


Figure 7: A Comparison of KEGG and IRE gene sets significantly altered by the EOfAD-like mutations T428del and Q96_K97del in 6-month-old zebrafish brains. Each cell is coloured according to statistical significance, and the FDR-adjusted harmonic mean p-value is shown. Gene sets not significantly altered (FDR adjusted harmonic mean p-value > 0.05) in a comparison between a *psen1* mutant zebrafish with their respective wild type siblings appear grey. **B** Principal component 1 (PC1) values for the HALLMARK gene sets as calculated by AES-PCA, clustered based on their Euclidean distance in T428del/+ samples relative to their wild type siblings. **C** PC1 values for the HALLMARK gene sets as calculated by AES-PCA clustered based on their Euclidean distance in Q96_K97del/+ samples relative to their wild type siblings at 6 months of age under normal oxygen conditions. See online version for colour.

Discussion

In this study, we exploited transcriptome analysis of whole brains of young adult zebrafish siblings, to detect differences in molecular state between the brains of fish heterozygous for an EOfAD-like mutation or an fAl-like mutation of *psen1* compared to wild type *in vivo*. The subtlety of the effects observed is consistent with that EOfAD is, despite its designation as “early-onset”, a disease affecting people overwhelmingly at ages older than 30 years [86]. The person reported to carry the T440del mutation of *PSEN1* showed cognitive decline at 41 years [47]. (Overall, EOfAD mutations in *PSEN1* show a median survival to onset of 45 years [86]). In contrast, at 6 months of age, zebrafish are only recently sexually mature. Nevertheless, since AD is thought to take decades to develop [32], it is these subtle, early changes that we must target therapeutically if we wish to arrest the pathological processes driving the progression to AD. As seen in all our previous analyses of EOfAD-like mutations [39-41, 43, 45], changes in expression of genes

involved in oxidative phosphorylation were identified as significant. However, this was not the case for the fAI-like, frameshift mutation. Therefore, oxidative phosphorylation changes appear to be an early signature cellular stress of EOofAD. Changes to mitochondrial function have been observed in heterozygous *PSEN1* mutant astrocytes [87], homozygous *PSEN1* mutant neurons [88], and in neurons differentiated from human induced pluripotent stem cells (hiPSCs) from sporadic AD patients [89], supporting our findings. However, such changes are not always observed [90], possibly due to issues of experimental reproducibility between laboratories when working with hiPSCs [91].

The EOofAD-like mutation also caused very statistically significant changes in the *KEGG_PARKINSONS_DISEASE* gene set (that shares many genes with *KEGG_OXIDATIVE_PHOSPHORYLATION*) and the person carrying the *PSEN1*^{T440del} mutation modelled by zebrafish *psen1*^{T428del} initially showed symptoms of early onset parkinsonism at 34 years of age before those of cognitive decline at 41 years [47].

In contrast to this EOofAD-like mutation, the fAI-like mutation apparently caused very statistically significant changes in Notch signalling and changes in other signal transduction pathways such as those involving Wnt and neurotrophins, as might be expected from changes in γ -secretase activity. Also notable was enrichment for the *KEGG_TOLL_LIKE_RECEPTOR_SIGNALLING_PATHWAY* gene set since acne inversa is a chronic inflammatory skin disorder and, in humans, increased expression of Toll-like receptor 2 has been noted in acne inversa lesions [92].

Both the EOfAD-like and fAl-like mutations caused very statistically significant changes in the gene sets *KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION* and *KEGG_RIBOSOME*. The former gene set reflects that both mutations appear to affect inflammation that is a characteristic of the pathologies of both EOfAD [93] and fAl (reviewed in [94]). Like oxidative phosphorylation, we have also observed effects on ribosomal protein genes sets for every EOfAD-like mutation we have studied [39-41, 43]. This may be due protein synthesis consuming a large proportion of cells' energy budgets [95] and requiring amino acid precursors that can be sourced from lysosomes. Recently, Bordi et al [96] noted that mTOR is highly activated in fibroblasts from people with Down syndrome (DS, trisomy 21). DS individuals commonly develop EOfAD due to overexpression of the A β PP gene (that is resident on human chromosome 21). The consequent increased expression of A β PP's β -CTF/C99 fragment (generated by β -secretase cleavage of A β PP without γ -secretase cleavage) affects endolysosomal pathway acidification [97] in a similar manner to EOfAD mutations of *PSEN1* [98]. The mTOR protein is localised at lysosomes in the mTORC1 and mTORC2 protein complexes (reviewed in [99, 100]) and monitors the energy and nutrient status of cells (reviewed in [101]). It is important for regulating ribosomal activity, partly by regulating transcription of ribosome components (reviewed in [102, 103]). Therefore, one explanation for the consistent enrichment for transcripts of the *KEGG_RIBOSOME* gene set we see in EOfAD mutant brains may be mTOR activation due to effects on lysosomal acidification and/or the energy status of cells. We did not observe any significant changes to the expression of genes involved in the mTOR signalling pathway in the analyses described in this paper (the FDR-adjusted harmonic mean p-value for the

KEGG_MTOR_SIGNALING_PATHWAY was 0.7 for each comparison of the *psen1* mutant fish to their wild type siblings, see **Supplementary Table 2**).

However, these changes could be undetectably subtle in young adult brains and/or occurring at the protein level and therefore not observable in bulk RNA-seq data.

(Statistically significant enrichment for genes in the

HALLMARK_PI3K_AKT_MTOR_SIGNALING gene set was seen previously for the normoxic, 6-month-old brains of fish heterozygous for the more severe EOfAD-like mutation Q96_K97del when compared to wild type siblings [45].)

While the brain transcriptome alterations caused by the EOfAD-like and fAI-like mutations are subtle (as illustrated by the lack of tight clustering of samples in the principal component analysis in **Figure 3**, and the low number of significantly differentially expressed genes), we are reassured in their overall veracity by their similarity to the results of a parallel analysis of sibling brain transcriptomes from 6-month-old zebrafish heterozygous for either a frameshift or a frame-preserving mutation in the zebrafish *psen2* gene relative to wild type [40]. In that similarly structured (but less statistically powered) experiment, only the frameshift mutation significantly affected the *KEGG_NOTCH_SIGNALLING* gene set while only the frame-preserving, EOfAD-like mutation significantly affected the *KEGG_OXIDATIVE_PHOSPHORYLATION* gene set. Both *psen2* genotypes affected the *KEGG_RIBOSOME* gene set, but in overall opposite directions (the frameshift mutation largely upregulated these genes while the frame-preserving mutation did the opposite).

Transcriptome analysis can reveal a great deal of data on differences in gene transcript levels between different genotypes or treatments. However, interpreting

changes in cellular state from this information is not straight forward. Are any changes seen direct molecular effects of a mutation/treatment (e.g. the direct, downstream effects of a change in γ -secretase activity) or homeostatic responses as cells/tissues adjust their internal states to promote survival? For example, in the *KEGG_NOTCH_SIGNALING_PATHWAY* gene set shown in **Figure 6B**, more pathway components are upregulated than are downregulated. However, the direct transcriptional targets of Notch signalling (*her4.2* and *hey1*) are downregulated, as might be expected from reduced expression of wild type, catalytically-competent Psen1 protein. The upregulation of other components of the pathway may represent homeostatic responses attempting to restore normal levels of Notch signalling.

Only two Notch downstream transcriptional target genes are described in the *KEGG_NOTCH_SIGNALING_PATHWAY* gene set. Therefore, in an effort to assess more generally the effects of the EOfAD-like and fAl-like mutations on γ -secretase activity, we also analysed additional sets of genes previously identified (in various systems) as direct transcriptional targets of γ -secretase-dependent signalling. One of these sets, encompassing genes identified as Notch signalling targets by both γ -secretase inhibitor responses and binding of the Notch intracellular domain to chromatin, revealed apparent upregulation of Notch signalling in both the EOfAD-like and fAl-like heterozygous mutant brains relative to wild type siblings. In both EOfAD-like/+ and fAl-like/+ mutant brains, the most highly ranked genes in terms of differential expression tended to be upregulated (although some highly ranked genes were downregulated in fAl-like/+ brains). The idea that putative low levels of a form of Psen1 protein truncated in the third

luminal loop domain could increase Notch signalling is not unexpected, as we previously observed an implied upregulation of Notch signalling in zebrafish embryos with forced expression of the fAl-causative P242Lfs allele of human *PSEN1* [104]. However, a widespread assumption within AD research is that EOfAD-like mutations of *PSEN1* decrease γ -secretase cleavage of A β PP [105], possibly through a dominant negative mechanism [21]. This assumption conflicts with the observation of Sun et al. [15] that approximately 10% of the 138 EOfAD mutations of human *PSEN1* they studied actually increased γ -secretase cleavage of A β PP's β -CTF/C99 fragment (in experiments examining the activities of the mutant proteins in isolation from wild type protein). Zhou et al. [106] also observed increased γ -secretase activity (cleavage of A β PP's β -CTF/C99 fragment) due to an EOfAD mutation of *PSEN1* (S365A, this replicated Sun et al.'s finding for this mutation).

It is important to note that mutations of *PSEN1* need not cause similar effects on Notch and A β PP cleavage [104, 107, 108]. The transmembrane domains of the Notch receptor and the APP's C99 fragment have different conformations [109]. Therefore, changes in the conformation of PSEN1 within γ -secretase due to a mutation may differentially affect Notch and C99 cleavage. Indeed, in our previously mentioned study of forced expression of the human *PSEN1*^{P242Lfs} allele in zebrafish embryos, increased apparent Notch signalling was observed without change in A β PP processing [104]. Conversely, Zhang et al. [108] showed that transgenic expression of an EOfAD mutation S169del in *PSEN1* under the control of the *Thy1* brain specific promotor altered the processing of A β PP *in vivo* without affecting Notch signalling. Notably, both of these studies did not use the *PSEN1*

gene's own promoter to express mutant forms of this gene, and so the effects seen may be distorted by gene/protein over-expression.

Unfortunately, the direct transcriptional targets of the intracellular domain of A β PP (AICD) have not been characterised to the same extent as those of NICD (reviewed in [110]). This constrains transcriptome analysis for detection of differential effects on γ -secretase cleavage of A β PP caused by the EOfAD-like and fAI-like mutations. Future work should include further investigation of how these mutations effect γ -secretase cleavage of A β PP *in vivo*.

In conclusion, we have performed the first direct comparison of an EOfAD-like and a fAI-like mutation of *presenilin 1* in an *in vivo* model. Both forms of mutation cause apparent changes in inflammation, downregulate expression of genes encoding the components of the ribosome subunits, and potentially affect γ -secretase activity as supported by altered expression of Notch signalling pathway transcriptional target genes. We see that changes to mitochondrial function are a specific, common characteristic of EOfAD-like mutations while the fAI-like mutation specifically affects important signal transduction pathways. These differential effects on brain transcriptomes give insight into how reading-frame preserving mutations in *PSEN1* cause EOfAD while frameshift mutations do not.

Conflict of Interest/Disclosure Statement

The authors have no financial or non-financial competing interests to declare.

Acknowledgements

The authors would like to thank Dr. Nhi Hin for providing the Q96_K97del gene expression values and the zebrafish IRE gene sets. We also would like to thank Dr. Jimmy Breen for his assistance with using the Nextflow pipeline. The authors thank Dr Giuseppe Verdile for critical reading of the manuscript.

This work was supported with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide and by grants GNT1061006 and GNT1126422 from the National Health and Medical Research Council of Australia (NHMRC). KB was supported by an Australian Government Research Training Program Scholarship and by funds from the Carthew Family Charity Trust. YD was supported by an Adelaide Graduate Research Scholarship from the University of Adelaide. MN was supported by funds from the grants listed above. ML is an academic employee of the University of Adelaide.

References

- [1] Janssen JC, Beck JA, Campbell TA, Dickinson A, Fox NC, Harvey RJ, Houlden H, Rossor MN, Collinge J (2003) Early onset familial Alzheimer's disease: Mutation frequency in 31 families. *Neurology* **60**, 235-239.
- [2] Żekanowski C, Styczyńska M, Peptońska B, Gabryelewicz T, Religa D, Ilkowski J, Kijanowska-Haładyna B, Kotapka-Minc S, Mikkelsen S, Pfeiffer A, Barczak A, Łuczywek E, Wasiak B, Chodakowska-Żebrowska M, Gustaw K, Łączkowski J, Sobów T, Kuźnicki J, Barcikowska M (2003) Mutations in presenilin 1, presenilin 2 and amyloid precursor protein genes in patients with early-onset Alzheimer's disease in Poland. *Experimental Neurology* **184**, 991-996.
- [3] Cruts M, van Duijn CM, Backhovens H, Van den Broeck M, Wehnert A, Serneels S, Sherrington R, Hutton M, Hardy J, St George-Hyslop PH (1998) Estimation of the genetic contribution of presenilin-1 and-2 mutations in a population-based study of presenile Alzheimer disease. *Human molecular genetics* **7**, 43-51.
- [4] Kamimura K, Tanahashi H, Yamanaka H, Takahashi K, Asada T, Tabira T (1998) Familial Alzheimer's disease genes in Japanese. *Journal of the neurological sciences* **160**, 76-81.
- [5] Area-Gomez E, de Groof AJ, Boldogh I, Bird TD, Gibson GE, Koehler CM, Yu WH, Duff KE, Yaffe MP, Pon LA, Schon EA (2009) Presenilins are enriched in endoplasmic reticulum membranes associated with mitochondria. *American Journal of Pathology* **175**, 1810-1816.
- [6] Schon EA, Area-Gomez E (2010) Is Alzheimer's disease a disorder of mitochondria-associated membranes? *J Alzheimers Dis* **20 Suppl 2**, S281-292.
- [7] Laudon H, Hansson EM, Melén K, Bergman A, Farmery MR, Winblad B, Lendahl U, von Heijne G, Näslund J (2005) A nine-transmembrane domain topology for presenilin 1. *J Biol Chem* **280**, 35352-35360.
- [8] Fukumori A, Fluhrer R, Steiner H, Haass C (2010) Three-amino acid spacing of presenilin endoproteolysis suggests a general stepwise cleavage of gamma-secretase-mediated intramembrane proteolysis. *J Neurosci* **30**, 7853-7862.
- [9] Sun L, Zhao L, Yang G, Yan C, Zhou R, Zhou X, Xie T, Zhao Y, Wu S, Li X, Shi Y (2015) Structural basis of human γ -secretase assembly. *Proceedings of the National Academy of Sciences* **112**, 6003.

- [10] Wolfe MS (2020) Unraveling the complexity of γ -secretase. *Semin Cell Dev Biol* **105**, 3-11.
- [11] Dermaut B, Kumar-Singh S, Engelborghs S, Theuns J, Rademakers R, Saerens J, Pickut BA, Peeters K, van den Broeck M, Vennekens K, Claes S, Cruts M, Cras P, Martin JJ, Van Broeckhoven C, De Deyn PP (2004) A novel presenilin 1 mutation associated with Pick's disease but not beta-amyloid plaques. *Ann Neurol* **55**, 617-626.
- [12] Li D, Parks SB, Kushner JD, Nauman D, Burgess D, Ludwigsen S, Partain J, Nixon RR, Allen CN, Irwin RP, Jakobs PM, Litt M, Hershberger RE (2006) Mutations of presenilin genes in dilated cardiomyopathy and heart failure. *Am J Hum Genet* **79**, 1030-1039.
- [13] Wang B, Yang W, Wen W, Sun J, Su B, Liu B, Ma D, Lv D, Wen Y, Qu T, Chen M, Sun M, Shen Y, Zhang X (2010) γ -Secretase Gene Mutations in Familial Acne Inversa. **330**, 1065-1065.
- [14] Jiang H, Jayadev S, Lardelli M, Newman M (2018) A Review of the Familial Alzheimer's Disease Locus PRESENILIN 2 and Its Relationship to PRESENILIN 1. *Journal of Alzheimer's Disease* **66**, 1323-1339.
- [15] Sun L, Zhou R, Yang G, Shi Y (2017) Analysis of 138 pathogenic mutations in presenilin-1 on the in vitro production of A β 42 and A β 40 peptides by γ -secretase. *Proceedings of the National Academy of Sciences* **114**, E476-E485.
- [16] Szaruga M, Veugelen S, Benurwar M, Lismont S, Sepulveda-Falla D, Lleo A, Ryan NS, Lashley T, Fox NC, Murayama S, Gijsen H, De Strooper B, Chávez-Gutiérrez L (2015) Qualitative changes in human γ -secretase underlie familial Alzheimer's disease. *J Exp Med* **212**, 2003-2013.
- [17] De Strooper B (2007) Loss-of-function presenilin mutations in Alzheimer disease. Talking Point on the role of presenilin mutations in Alzheimer disease. *EMBO Rep* **8**, 141-146.
- [18] Jayne T, Newman M, Verdile G, Sutherland G, Munch G, Musgrave I, Moussavi Nik SH, 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*.
- [19] Wong TH, Seelaar H, Melhem S, Rozemuller AJM, van Swieten JC (2020) Genetic screening in early-onset Alzheimer's disease identified three novel presenilin mutations. *Neurobiology of aging* **86**, 201.e209-201.e214.
- [20] Lee JH, McBrayer MK, Wolfe DM, Haslett LJ, Kumar A, Sato Y, Lie PP, Mohan P, Coffey EE, Kompella U, Mitchell CH, Lloyd-Evans E, Nixon RA (2015) Presenilin 1 Maintains

Lysosomal Ca(2+) Homeostasis via TRPML1 by Regulating vATPase-Mediated Lysosome Acidification. *Cell Report* **12**, 1430-1444.

[21] Heilig EA, Gutti U, Tai T, Shen J, Kelleher RJ, 3rd (2013) Trans-dominant negative effects of pathogenic PSEN1 mutations on γ -secretase activity and A β production. *J Neurosci* **33**, 11606-11617.

[22] Schroeter EH, Ilagan MXG, Brunkan AL, Hecimovic S, Li Y-m, Xu M, Lewis HD, Saxena MT, De Strooper B, Coonrod A, Tomita T, Iwatsubo T, Moore CL, Goate A, Wolfe MS, Shearman M, Kopan R (2003) A presenilin dimer at the core of the γ -secretase enzyme: Insights from parallel analysis of Notch 1 and APP proteolysis. *Proceedings of the National Academy of Sciences* **100**, 13075.

[23] Brautigam H, Moreno CL, Steele JW, Bogush A, Dickstein DL, Kwok JBJ, Schofield PR, Thinakaran G, Mathews PM, Hof PR, Gandy S, Ehrlich ME (2015) Physiologically generated presenilin 1 lacking exon 8 fails to rescue brain PS1^{-/-} phenotype and forms complexes with wildtype PS1 and nicastrin. *Scientific Reports* **5**, 17042.

[24] Li X, Dang S, Yan C, Gong X, Wang J, Shi Y (2013) Structure of a presenilin family intramembrane aspartate protease. *Nature* **493**, 56-61.

[25] De Gasperi R, Sosa MAG, Dracheva S, Elder GA (2010) Presenilin-1 regulates induction of hypoxia inducible factor-1 α : altered activation by a mutation associated with familial Alzheimer's disease. *Molecular Neurodegeneration* **5**, 38-38.

[26] Villa JC, Chiu D, Brandes AH, Escorcía FE, Villa CH, Maguire WF, Hu CJ, de Stanchina E, Simon MC, Sisodia SS, Scheinberg DA, Li YM (2014) Nontranscriptional role of Hif-1 α in activation of gamma-secretase and notch signaling in breast cancer. *Cell Replication* **8**, 1077-1092.

[27] Newman M, Nik HM, Sutherland GT, Hin N, Kim WS, Halliday GM, Jayadev S, Smith C, Laird AS, Lucas CW, Kittipassorn T, Peet DJ, Lardelli M (2020) Accelerated loss of hypoxia response in zebrafish with familial Alzheimer's disease-like mutation of presenilin 1. *Human Molecular Genetics* **29**, 2379-2394.

[28] Lane DJR, Merlot AM, Huang MLH, Bae DH, Jansson PJ, Sahni S, Kalinowski DS, Richardson DR (2015) Cellular iron uptake, trafficking and metabolism: Key molecules and mechanisms and their roles in disease. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1853**, 1130-1144.

[29] Güner G, Lichtenthaler SF (2020) The substrate repertoire of γ -secretase/presenilin. *Seminars in Cell & Developmental Biology* **105**, 27-42.

- [30] Thinakaran G, Harris CL, Ratovitski T, Davenport F, Slunt HH, Price DL, Borchelt DR, Sisodia SS (1997) Evidence that levels of presenilins (PS1 and PS2) are coordinately regulated by competition for limiting cellular factors. *J Biol Chem* **272**, 28415-28422.
- [31] Berchtold NC, Sabbagh MN, Beach TG, Kim RC, Cribbs DH, Cotman CW (2014) Brain gene expression patterns differentiate mild cognitive impairment from normal aged and Alzheimer's disease. *Neurobiology of Aging* **35**, 1961-1972.
- [32] Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Salvado O, Szoek C, Macaulay SL, Martins R, Maruff P, Ames D, Rowe CC, Masters CL (2013) Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *The Lancet Neurology* **12**, 357-367.
- [33] Foley AM, Ammar ZM, Lee RH, Mitchell CS (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**, 787-795.
- [34] Hargis KE, Blalock EM (2017) Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us? *Behavioural Brain Research* **322**, 311-328.
- [35] Guo Q, Fu W, Sopher BL, Miller MW, Ware CB, Martin GM, Mattson MP (1999) Increased vulnerability of hippocampal neurons to excitotoxic necrosis in presenilin-1 mutant knock-in mice. *Nat Med* **5**, 101-106.
- [36] Kawasumi M, Chiba T, Yamada M, Miyamae-Kaneko M, Matsuoka M, Nakahara J, Tomita T, Iwatsubo T, Kato S, Aiso S, Nishimoto I, Kouyama K (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**, 2826-2838.
- [37] Xu J, Burgoyne PS, Arnold AP (2002) Sex differences in sex chromosome gene expression in mouse brain. *Human Molecular Genetics* **11**, 1409-1419.
- [38] Bundy JL, Vied C, Nowakowski RS (2017) Sex differences in the molecular signature of the developing mouse hippocampus. *BMC Genomics* **18**, 237.
- [39] Barthelson K, Pederson S, Newman M, Lardelli M (2020) Transcriptome analysis of a protein-truncating mutation in sortilin-related receptor 1 associated with early-onset familial Alzheimer's disease indicates effects on mitochondrial and ribosome function in young-adult zebrafish brains. *bioRxiv*, 2020.2009.2003.282277.

- [40] Barthelson K, Pederson SM, Newman M, Jiang H, Lardelli M (2020) Frameshift and frame-preserving mutations in zebrafish presenilin 2 affect different cellular functions in young adult brains. *bioRxiv*, 2020.2011.2021.392761.
- [41] Barthelson K, Pederson SM, Newman M, Lardelli M (2020) 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.
- [42] Hin N, Newman M, Kaslin J, Douek AM, Lumsden A, Nik SHM, Dong Y, Zhou X-F, Mañucat-Tan NB, Ludington A, Adelson DL, Pederson S, Lardelli M (2020) Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2. *PLOS ONE* **15**, e0227258.
- [43] Jiang H, Pederson SM, Newman M, Dong Y, Barthelson K, 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.
- [44] Dong Y, Newman M, Pederson S, Hin N, Lardelli M (2020) Transcriptome analyses of 7-day-old zebrafish larvae possessing a familial Alzheimer's disease-like mutation in psen1 indicate effects on oxidative phosphorylation, mcm functions, and iron homeostasis. *bioRxiv*, 2020.2005.2003.075424.
- [45] Hin N, Newman M, Pederson SM, Lardelli MM (2020) Iron Responsive Element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease. *bioRxiv*, 2020.2005.2001.071498.
- [46] Newman M, Hin N, Pederson S, 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**.
- [47] Ishikawa A, Piao Y-S, Miyashita A, Kuwano R, Onodera O, Ohtake H, Suzuki M, Nishizawa M, Takahashi H (2005) A mutant PSEN1 causes dementia with lewy bodies and variant Alzheimer's disease. *Annals of Neurology* **57**, 429-434.
- [48] Zhang Y, Zhang Z, Ge W (2018) An efficient platform for generating somatic point mutations with germline transmission in the zebrafish by CRISPR/Cas9-mediated gene editing. *J Biol Chem* **293**, 6611-6622.
- [49] Prykhozhiy SV, Fuller C, Steele SL, Veinotte CJ, Razaghi B, Robitaille JM, McMaster CR, Shlien A, Malkin D, Berman JN (2018) Optimized knock-in of point mutations in zebrafish using CRISPR/Cas9. *Nucleic Acids Research* **46**, e102-e102.

- [50] Jiang H, Newman M, Lardelli M (2018) The zebrafish orthologue of familial Alzheimer's disease gene PRESENILIN 2 is required for normal adult melanotic skin pigmentation. *PLOS ONE* **13**, e0206155.
- [51] Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C (2017) Nextflow enables reproducible computational workflows. *Nature Biotechnology* **35**, 316-319.
- [52] Team RC (2019) R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.
- [53] Chen Y, Lun A, Smyth G (2016) From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline [version 2; peer review: 5 approved]. *F1000Research* **5**.
- [54] Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* **11**, R25.
- [55] McCarthy DJ, Chen Y, Smyth GK (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**, 4288-4297.
- [56] Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140.
- [57] Young MD, Wakefield MJ, Smyth GK, Oshlack A (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* **11**, R14.
- [58] Wu D, Lim E, Vaillant F, Asselin-Labat M-L, Visvader JE, Smyth GK (2010) ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics* **26**, 2176-2182.
- [59] Wu D, Smyth GK (2012) Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic acids research* **40**, e133-e133.
- [60] Sergushichev AA (2016) An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv*, 060012.
- [61] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545.

- [62] Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28**, 27-30.
- [63] Liberzon A (2014) A description of the Molecular Signatures Database (MSigDB) Web site. *Methods Mol Biol* **1150**, 153-160.
- [64] Dolgalev I (2020), p. R package.
- [65] Chen X (2011) Adaptive elastic-net sparse principal component analysis for pathway association testing. *Statistical applications in genetics and molecular biology* **10**, 48.
- [66] Odom GJ, Ban Y, Liu L, Sun X, Pico AR, Zhang B, Wang L, Chen X (2019) pathwayPCA: an R package for integrative pathway analysis with modern PCA methodology and gene selection. *bioRxiv*, 615435.
- [67] Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov Jill P, Tamayo P (2015) The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Systems* **1**, 417-425.
- [68] Kosik KS, Muñoz C, Lopez L, Arcila ML, García G, Madrigal L, Moreno S, Ríos Romenets S, Lopez H, Gutierrez M, Langbaum JB, Cho W, Suliman S, Tariot PN, Ho C, Reiman EM, Lopera F (2015) Homozygosity of the autosomal dominant Alzheimer disease presenilin 1 E280A mutation. *Neurology* **84**, 206-208.
- [69] Parker J, Mozaffar T, Messmore A, Deignan JL, Kimonis VE, Ringman JM (2019) Homozygosity for the A431E mutation in PSEN1 presenting with a relatively aggressive phenotype. *Neurosci Lett* **699**, 195-198.
- [70] Sundvik M, Chen Y-C, Panula P (2013) Presenilin1 Regulates Histamine Neuron Development and Behavior in Zebrafish, *Danio rerio*. *The Journal of Neuroscience* **33**, 1589.
- [71] Tambini MD, D'Adamio L (2020) Knock-in rats with homozygous PSEN1(L435F) Alzheimer mutation are viable and show selective γ -secretase activity loss causing low A β 40/42 and high A β 43. *J Biol Chem* **295**, 7442-7451.
- [72] Xia D, Watanabe H, Wu B, Lee SH, Li Y, Tsvetkov E, Bolshakov VY, Shen J, Kelleher RJ, 3rd (2015) Presenilin-1 knockin mice reveal loss-of-function mechanism for familial Alzheimer's disease. *Neuron* **85**, 967-981.
- [73] Nickless A, Bailis JM, You Z (2017) Control of gene expression through the nonsense-mediated RNA decay pathway. *Cell & Bioscience* **7**, 26.

- [74] Alhamdoosh M, Law CW, Tian L, Sheridan JM, Ng M, Ritchie ME (2017) Easy and efficient ensemble gene set testing with EGSEA. *F1000Research* **6**, 2010-2010.
- [75] Wilson DJ (2019) The harmonic mean p-value for combining dependent tests. *Proceedings of the National Academy of Sciences* **116**, 1195.
- [76] Pan Y, Lin MH, Tian X, Cheng HT, Gridley T, Shen J, Kopan R (2004) gamma-secretase functions through Notch signaling to maintain skin appendages but is not required for their patterning or initial morphogenesis. *Dev Cell* **7**, 731-743.
- [77] Blanpain C, Lowry WE, Pasolli HA, Fuchs E (2006) Canonical notch signaling functions as a commitment switch in the epidermal lineage. *Genes Dev* **20**, 3022-3035.
- [78] Kopan R, Ilagan MX (2009) The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell* **137**, 216-233.
- [79] Nowell C, Radtke F (2013) Cutaneous Notch signaling in health and disease. *Cold Spring Harbor perspectives in medicine* **3**, a017772-a017772.
- [80] Nguyen B-C, Lefort K, Mandinova A, Antonini D, Devgan V, Della Gatta G, Koster MI, Zhang Z, Wang J, di Vignano AT, Kitajewski J, Chiorino G, Roop DR, Missero C, Dotto GP (2006) Cross-regulation between Notch and p63 in keratinocyte commitment to differentiation. *Genes & Development* **20**, 1028-1042.
- [81] Dohda T, Maljukova A, Liu L, Heyman M, Grandér D, Brodin D, Sangfelt O, Lendahl U (2007) Notch signaling induces SKP2 expression and promotes reduction of p27Kip1 in T-cell acute lymphoblastic leukemia cell lines. *Exp Cell Res* **313**, 3141-3152.
- [82] Ryan RJH, Petrovic J, Rausch DM, Zhou Y, Lareau CA, Kluk MJ, Christie AL, Lee WY, Tarjan DR, Guo B, Donohue LKH, Gillespie SM, Nardi V, Hochberg EP, Blacklow SC, Weinstock DM, Faryabi RB, Bernstein BE, Aster JC, Pear WS (2017) A B Cell Regulome Links Notch to Downstream Oncogenic Pathways in Small B Cell Lymphomas. *Cell Rep* **21**, 784-797.
- [83] El-Brolosy MA, Kontarakis Z, Rossi A, Kuenne C, Günther S, Fukuda N, Kikhi K, Boezio GLM, Takacs CM, Lai S-L, Fukuda R, Gerri C, Giraldez AJ, Stainier DYR (2019) Genetic compensation triggered by mutant mRNA degradation. *Nature* **568**, 193-197.
- [84] Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, Stainier DYR (2015) Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* **524**, 230-233.

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

[86] Ryman DC, Acosta-Baena N, Aisen PS, Bird T, Danek A, Fox NC, Goate A, Frommelt P, Ghetti B, Langbaum JBS, Lopera F, Martins R, Masters CL, Mayeux RP, McDade E, Moreno S, Reiman EM, Ringman JM, Salloway S, Schofield PR, Sperling R, Tariot PN, Xiong C, Morris JC, Bateman RJ (2014) Symptom onset in autosomal dominant Alzheimer disease: A systematic review and meta-analysis. *Neurology* **83**, 253-260.

[87] Oksanen M, Petersen AJ, Naumenko N, Puttonen K, Lehtonen Š, Gubert Olivé M, Shakirzyanova A, Leskelä S, Sarajärvi T, Viitanen M, Rinne JO, Hiltunen M, Haapasalo A, Giniatullin R, Tavi P, Zhang S-C, Kanninen KM, Hämäläinen RH, Koistinaho J (2017) PSEN1 Mutant iPSC-Derived Model Reveals Severe Astrocyte Pathology in Alzheimer's Disease. *Stem Cell Reports* **9**, 1885-1897.

[88] Martín-Maestro P, Sproul A, Martinez H, Paquet D, Gerges M, Noggle S, Starkov AA (2019) Autophagy Induction by Bexarotene Promotes Mitophagy in Presenilin 1 Familial Alzheimer's Disease iPSC-Derived Neural Stem Cells. *Molecular Neurobiology* **56**, 8220-8236.

[89] Birnbaum JH, Wanner D, Gietl AF, Saake A, Kündig TM, Hock C, Nitsch RM, 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.

[90] Kwart D, Gregg A, Scheckel C, Murphy EA, Paquet D, Duffield M, Fak J, Olsen O, Darnell RB, Tessier-Lavigne M (2019) A Large Panel of Isogenic APP and PSEN1 Mutant Human iPSC Neurons Reveals Shared Endosomal Abnormalities Mediated by APP β -CTFs, Not A β . *Neuron* **104**, 256-270.e255.

[91] Volpato V, Smith J, Sandor C, Ried JS, Baud A, Handel A, Newey SE, Wessely F, Attar M, Whiteley E, Chintawar S, Verheyen A, Barta T, Lako M, Armstrong L, Muschet C, Artati A, Cusulin C, Christensen K, Patsch C, Sharma E, Nicod J, Brownjohn P, Stubbs V, Heywood WE, Gissen P, De Filippis R, Janssen K, Reinhardt P, Adamski J, Royaux I, Peeters PJ, Terstappen GC, Graf M, Livesey FJ, Akerman CJ, Mills K, Bowden R, Nicholson G, Webber C, Cader MZ, Lakics V (2018) Reproducibility of Molecular Phenotypes after Long-Term Differentiation to Human iPSC-Derived Neurons: A Multi-Site Omics Study. *Stem Cell Reports* **11**, 897-911.

[92] Hunger RE, Surovy AM, Hassan AS, Braathen LR, Yawalkar N (2008) Toll-like receptor 2 is highly expressed in lesions of acne inversa and colocalizes with C-type lectin receptor. *British Journal of Dermatology* **158**, 691-697.

- [93] Otani K, Shichita T (2020) Cerebral sterile inflammation in neurodegenerative diseases. *Inflammation and Regeneration* **40**, 28.
- [94] Zouboulis CC, Benhadou F, Byrd AS, Chandran NS, Giamarellos-Bourboulis EJ, Fabbrocini G, Frew JW, Fujita H, González-López MA, Guillem P, Gulliver WPF, Hamzavi I, Hayran Y, Hórvath B, Hùe S, Hunger RE, Ingram JR, Jemec GBE, Ju Q, Kimball AB, Kirby JS, Konstantinou MP, Lowes MA, MacLeod AS, Martorell A, Marzano AV, Matusiak Ł, Nassif A, Nikiphorou E, Nikolakis G, Nogueira da Costa A, Okun MM, Orenstein LAV, Pascual JC, Paus R, Perin B, Prens EP, Röhn TA, Szegedi A, Szepietowski JC, Tzellos T, Wang B, van der Zee HH (2020) What causes hidradenitis suppurativa ?—15 years after. *Experimental Dermatology* **29**, 1154-1170.
- [95] Research IoMUCoMN (1999) The Energy Costs of Protein Metabolism: Lean and Mean on Uncle Sam's Team In *The Role of Protein and Amino Acids in Sustaining and Enhancing Performance* National Academies Press (US), Washington (DC).
- [96] Bordi M, Darji S, Sato Y, Mellén M, Berg MJ, Kumar A, Jiang Y, Nixon RA (2019) mTOR hyperactivation in Down Syndrome underlies deficits in autophagy induction, autophagosome formation, and mitophagy. *Cell Death & Disease* **10**, 563.
- [97] Jiang Y, Sato Y, Im E, Berg M, Bordi M, Darji S, Kumar A, Mohan PS, Bandyopadhyay U, Diaz A, Cuervo AM, Nixon RA (2019) Lysosomal Dysfunction in Down Syndrome Is APP-Dependent and Mediated by APP- β CTF (C99). *The Journal of Neuroscience* **39**, 5255.
- [98] Lee JH, Yu WH, Kumar A, Lee S, Mohan PS, Peterhoff CM, Wolfe DM, Martinez-Vicente M, Massey AC, Sovak G, Uchiyama Y, Westaway D, Cuervo AM, Nixon RA (2010) Lysosomal proteolysis and autophagy require presenilin 1 and are disrupted by Alzheimer-related PS1 mutations. *Cell* **141**, 1146-1158.
- [99] Linke M, Fritsch SD, Sukhbaatar N, Hengstschläger M, Weichhart T (2017) mTORC1 and mTORC2 as regulators of cell metabolism in immunity. *FEBS letters* **591**, 3089-3103.
- [100] Inpanathan S, Botelho RJ (2019) The Lysosome Signaling Platform: Adapting With the Times. *Frontiers in Cell and Developmental Biology* **7**, 113.
- [101] Bond P (2016) Regulation of mTORC1 by growth factors, energy status, amino acids and mechanical stimuli at a glance. *J Int Soc Sports Nutr* **13**, 8.
- [102] Loewith R, Hall MN (2011) Target of rapamycin (TOR) in nutrient signaling and growth control. *Genetics* **189**, 1177-1201.
- [103] Lempiäinen H, Shore D (2009) Growth control and ribosome biogenesis. *Current Opinion in Cell Biology* **21**, 855-863.

- [104] Newman M, Wilson L, Verdile G, Lim A, Khan I, Moussavi Nik SH, Pursglove S, Chapman G, Martins RN, Lardelli M (2014) Differential, dominant activation and inhibition of Notch signalling and APP cleavage by truncations of PSEN1 in human disease. *Human Molecular Genetics* **23**, 602-617.
- [105] Kelleher RJ, Shen J (2017) Presenilin-1 mutations and Alzheimer's disease. *Proceedings of the National Academy of Sciences* **114**, 629.
- [106] Zhou R, Yang G, Shi Y (2017) Dominant negative effect of the loss-of-function γ -secretase mutants on the wild-type enzyme through heterooligomerization. *Proc Natl Acad Sci U S A* **114**, 12731-12736.
- [107] Chávez-Gutiérrez L, Bammens L, Benilova I, Vandersteen A, Benurwar M, Borgers M, Lismont S, Zhou L, Van Cleynenbreugel S, Esselmann H, Wiltfang J, Serneels L, Karran E, Gijzen H, Schymkowitz J, Rousseau F, Broersen K, De Strooper B (2012) The mechanism of γ -Secretase dysfunction in familial Alzheimer disease. *The EMBO Journal* **31**, 2261-2274.
- [108] Zhang S, Cai F, Wu Y, Bozorgmehr T, Wang Z, Zhang S, Huang D, Guo J, Shen L, Rankin C, Tang B, Song W (2020) A presenilin-1 mutation causes Alzheimer disease without affecting Notch signaling. *Molecular Psychiatry* **25**, 603-613.
- [109] Deatherage CL, Lu Z, Kroncke BM, Ma S, Smith JA, Voehler MW, McFeeters RL, Sanders CR (2017) Structural and biochemical differences between the Notch and the amyloid precursor protein transmembrane domains. *Science Advances* **3**, e1602794.
- [110] Bukhari H, Glotzbach A, Kolbe K, Leonhardt G, Loose C, Müller T (2017) Small things matter: Implications of APP intracellular domain AICD nuclear signaling in the progression and pathogenesis of Alzheimer's disease. *Progress in Neurobiology* **156**, 189-213.

Supplemental File 1: Template DNA sequences for attempted HDR

Generation of the P242fs mutation in zebrafish *psen1*.

The following sequence was supplied in the pBluescript II SK(+)-Amp vector. The sequences for the sgRNA are indicated in red, PAM sites are highlighted in blue, and the site of the mutation site is highlighted in yellow.

Template plasmid insert sequence 5' – 3':

```
GATGAGCCATGCGGTCCACTCGGTATACATTATTAATAAAGCACACAATACTG
TCTGTTTCTTTTGCATAGACCTGAATAATGAATACACTTTTACAGTGCTGTAAAAT
ACAAGCCTACACACAACCTTTTTTTTTTTGCTATTTTGCTGAAAATAAAAATTTT
TAGGGCAAGTTTGACATTTGCATGGAATTGCTCAGTATTAACGTGATTTCTGC
CTCTTAACGATGTTTCTGTTCCCTCCAGGGAAGTGTTCAAGACGTATAACGTGG
CGATGGATTACTTCACGCTGGCGTTGATCATCTGGAACCTTCGGTGTGGTGGGA
ATGATCTGCATCCACTGGAAGGGGCCGCTGCGGCTCCAGCAGGCCTATCTGAT
CATGATCAGCGCTCTCATGGCTCTGGTCTTCATCAAGTATCTCCTGAATGGACT
GCGTGGCTCATCCTGGCTGTGATTTAGTCTACGGTCAGTCAGTCTGCAAACA
GAGCAACACATTCACTTCTGTGTACTGATAGGTTTATTGTTCTGTAATGCTGCTT
TGGAACAATAATTAGCCTGAAATGAATTGAATAATTTCAAAAACACAGCCCCTTA
AATGTCTTTCCTGCTCCTGTGTAGATCTTCTGGCAGTGTTGTGTCCGAAAGGCC
CTCTGCGAATCCTTGTGGAAACAGCTCAAGAGAGGAATGAGGCCATTTTCCCA
GCGCTCATCTACTCCTGTAAGAAAAACCCCTCAGCACTTCCATCTTCTCAGACT
GTGAATGTCCATCTGTTTAATTGTGTTATCGTCACGTTTCTATGAAATTTCTCCTT
CGATTTGTCAGCTACGATGGTGTGGCTCTTCAATATGGCGGACCGAGTGGACC
GCATGGCTCATC
```

Generation of the T428del mutation in zebrafish *psen1*.

A linear, single-stranded oligonucleotide (ssODN) with phosphorothioated (*) ends.
The PAM is indicated in blue.

ssODN sequence 5' – 3':

```
C*G*ACCTCTCTATATGTAGAACTGATGGACGGCCAGCTGGTCCATGAACGGCC  
GCACGAGGTTATCAGTGGCAAAGTAAAACACCAGGCCGAATATGGAGATGGGG  
AGTGCGGGAAGAG*C*C
```

Supplemental File 2

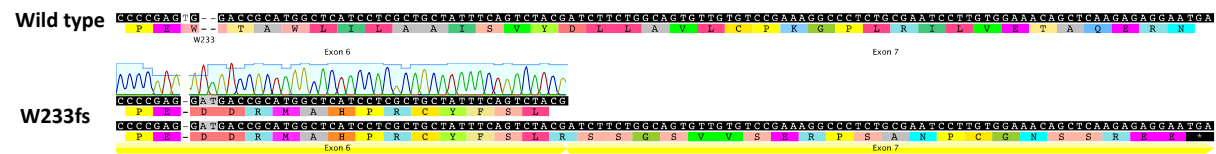


Figure S1: Alignment of the a region of *psen1* cDNA (ENSDART00000149864.3) including a subset of exons 6 and 7. The top sequence is the WT sequence, which is followed by a subset of the Sanger sequencing (from genomic DNA) of exon 6 from W233fs homozygous mutants. The bottom sequence is the predicted coding sequences of exon 6 and exon 7 assuming splicing is unaffected by the W233fs mutation. 36 novel amino acids are generated from this mutation followed by a premature termination codon.

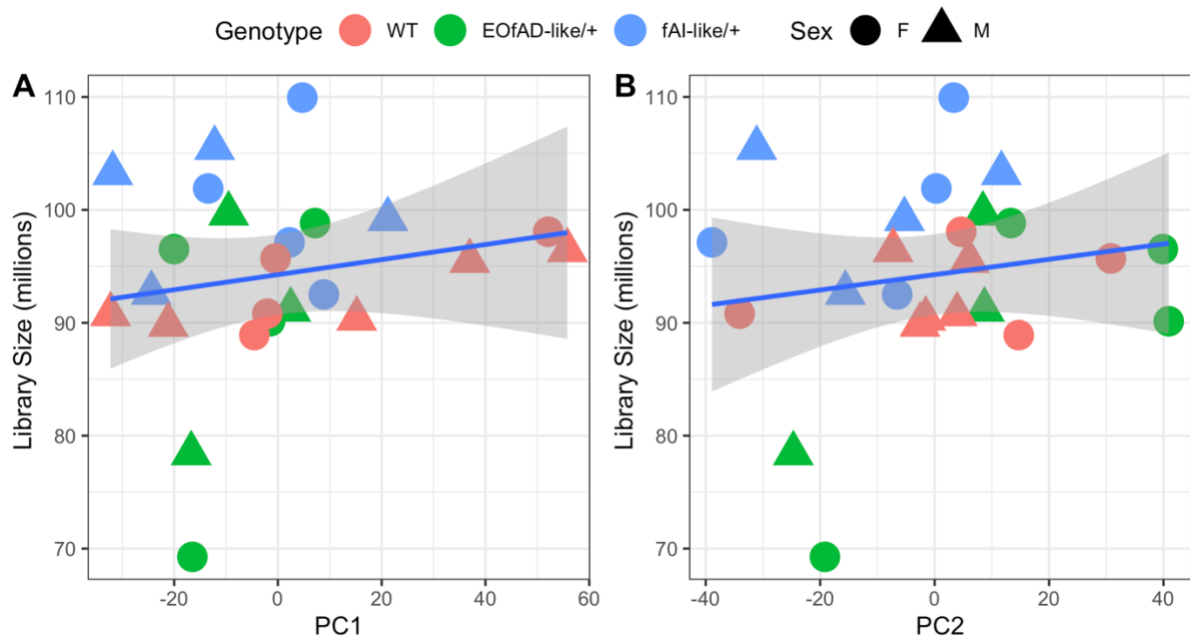


Figure S2: RNA-seq library size does correlate with **A** principal component 1 (PC1) or **B** PC2 after weighted trimmed mean of M-values (TMM) normalisation. Blue linear regression lines are shown with standard error shading.

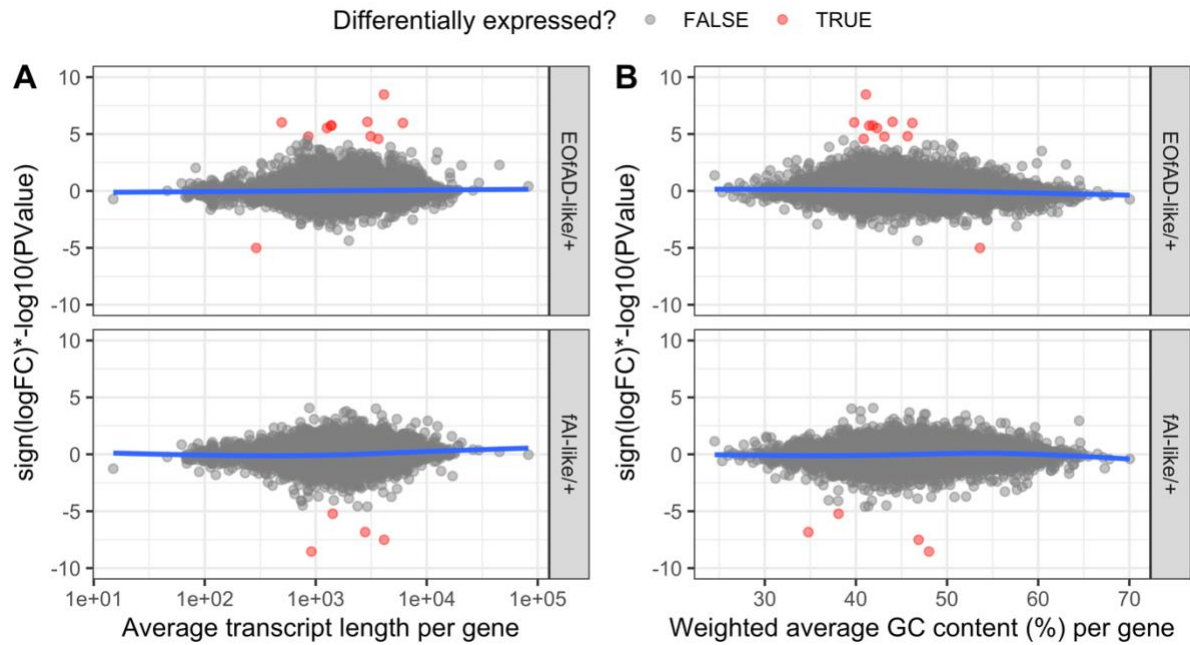


Figure S3: No observed bias for differential expression of genes due to **A** average transcript length or **B** weighted (per transcript length) GC content per gene. A ranking statistic was calculated for each gene. Blue lines indicating generalised additive model fits lie mostly around 0. The y axis in each graph is constrained to between -10 and 10 for visualisation purposes.

Table S1: Top 5 GO terms over-represented in genes significantly differentially expressed (DE) in EOofAD-like/+ relative to wild type brains.

GO term	p.val	Number of DE genes in category	Number of genes in category	FDR
GO SKELETAL MUSCLE CONTRACTION	3.5e-05	2	32	0.35
GO MULTICELLULAR ORGANISMAL MOVEMENT	6.8e-05	2	44	0.35
GO STRIATED MUSCLE CONTRACTION	7.8e-04	2	147	1
GO CALCIUM ACTIVATED PHOSPHOLIPID SCRAMBLING	1.2e-03	1	4	1
GO L METHIONINE SALVAGE FROM METHYLTHIOADENOSINE	1.6e-03	1	5	1

Table S2: Top 5 GO terms over-represented in genes significantly differentially expressed (DE) in fAI-like/+ relative to wild type brains

GO term	p.val	Number of DE genes in category	Number of genes in category	FDR
GO GAMMA SECRETASE COMPLEX	0.0009	1	4	1
GO ASPARTIC ENDOPEPTIDASE ACTIVITY INTRAMEMBRANE CLEAVING	0.001	1	5	1
GO NOTCH RECEPTOR PROCESSING LIGAND DEPENDENT	0.001	1	5	1
GO CATION CHANNEL ACTIVITY	0.001	2	276	1
GO REGULATION OF L GLUTAMATE IMPORT ACROSS PLASMA MEMBRANE	0.001	1	6	1

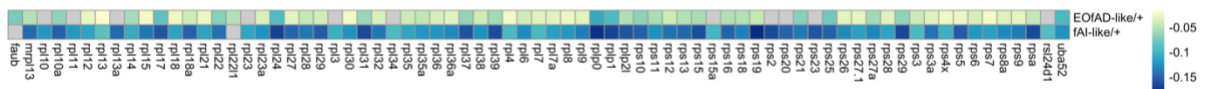


Figure S6: Similar genes are driving the enrichment of the *KEGG_RIBOSOME* gene set in the same direction. Colour of the heatmap indicates the logFC of the leading edge genes. Grey indicates that the gene was not found in the leading edge in the fgsea algorithm.

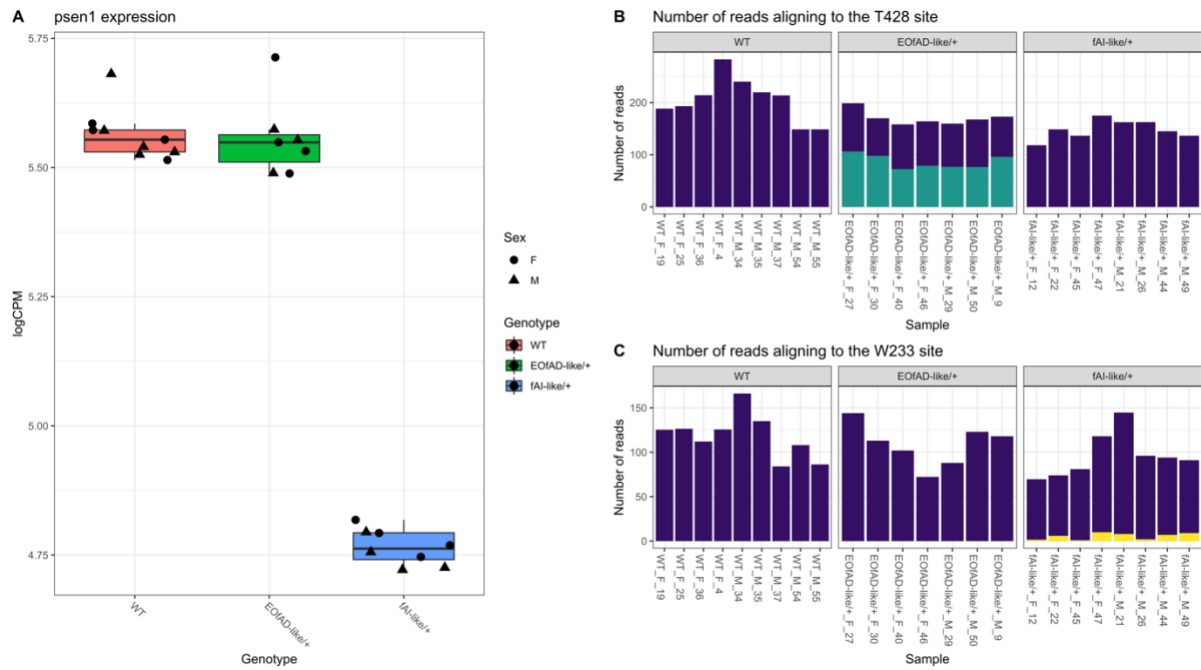


Figure S7: A Overall expression of *psen1* in 6 month old zebrafish brains in log2 counts per million (logCPM). **B** The number of reads aligning to the GRCz11 genome at the T428 site and **C** the W233 site in *psen1*. Bars appear purple if aligning to the wild type allele, blue if the aligned reads contained the T428del mutation, and yellow if the aligned reads contained the W233fs mutation. For each sample, the number of reads was calculated as the average number of reads over the three nucleotides deleted for the T428 site, and the two nucleotides affected at the W233 site.

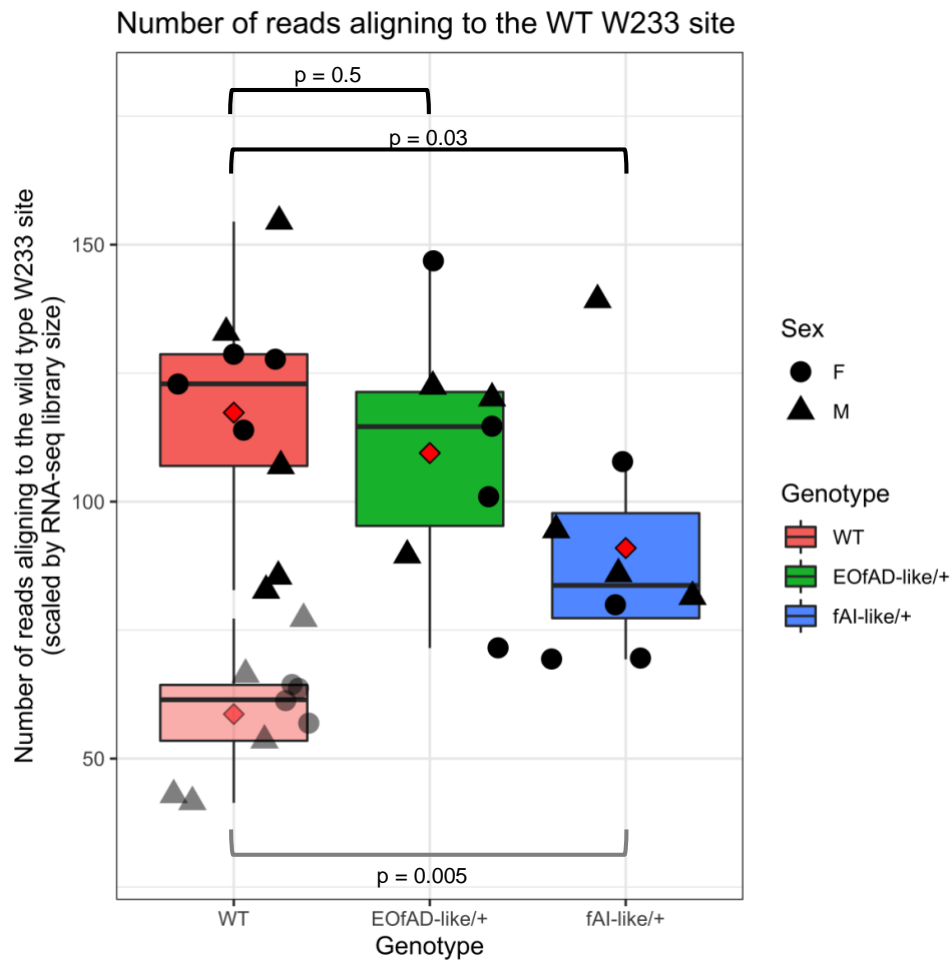


Figure S8: For each sample, the number of reads was calculated as the average number of reads over two nucleotides affected at the W233 site in *psen1*, then scaled for RNA-seq library size (TMM method). Pairwise comparisons between genotypes for these measures (indicated by the red diamonds) was performed using Welch's two sample t-tests. To investigate whether the expression of the wild type *psen1* allele in fAl-like/+ brains was greater than 50% of the expression of the wild type allele in wild type brains, the scaled reads for wild type brains were divided by two (shown in the graph by the more transparent points), then these were subject to Welch's two sample t-test.

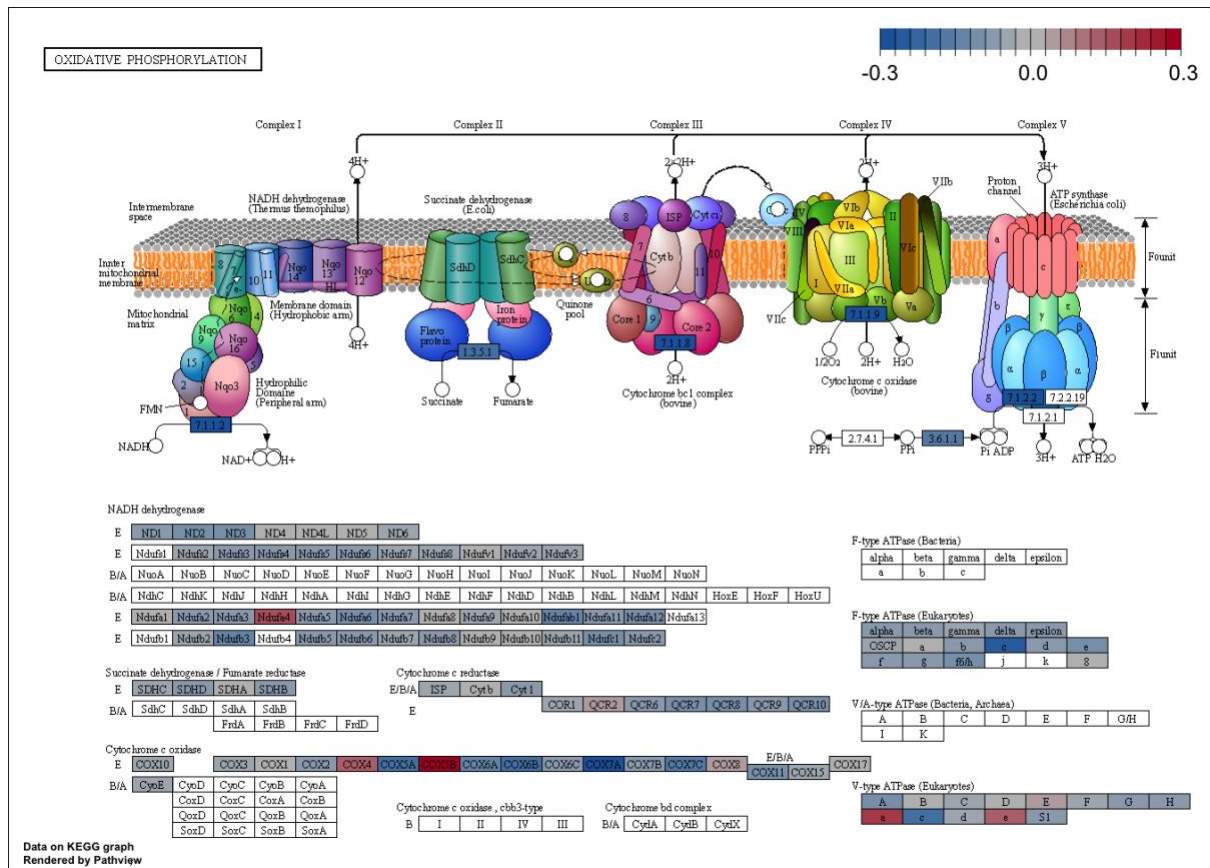


Figure S9: Pathview visualisation of the *KEGG_OXIDATIVE_PHOSPHORYLATION* gene set. Colour of the cells indicates the logFC in EOofAD-like/+ brains.

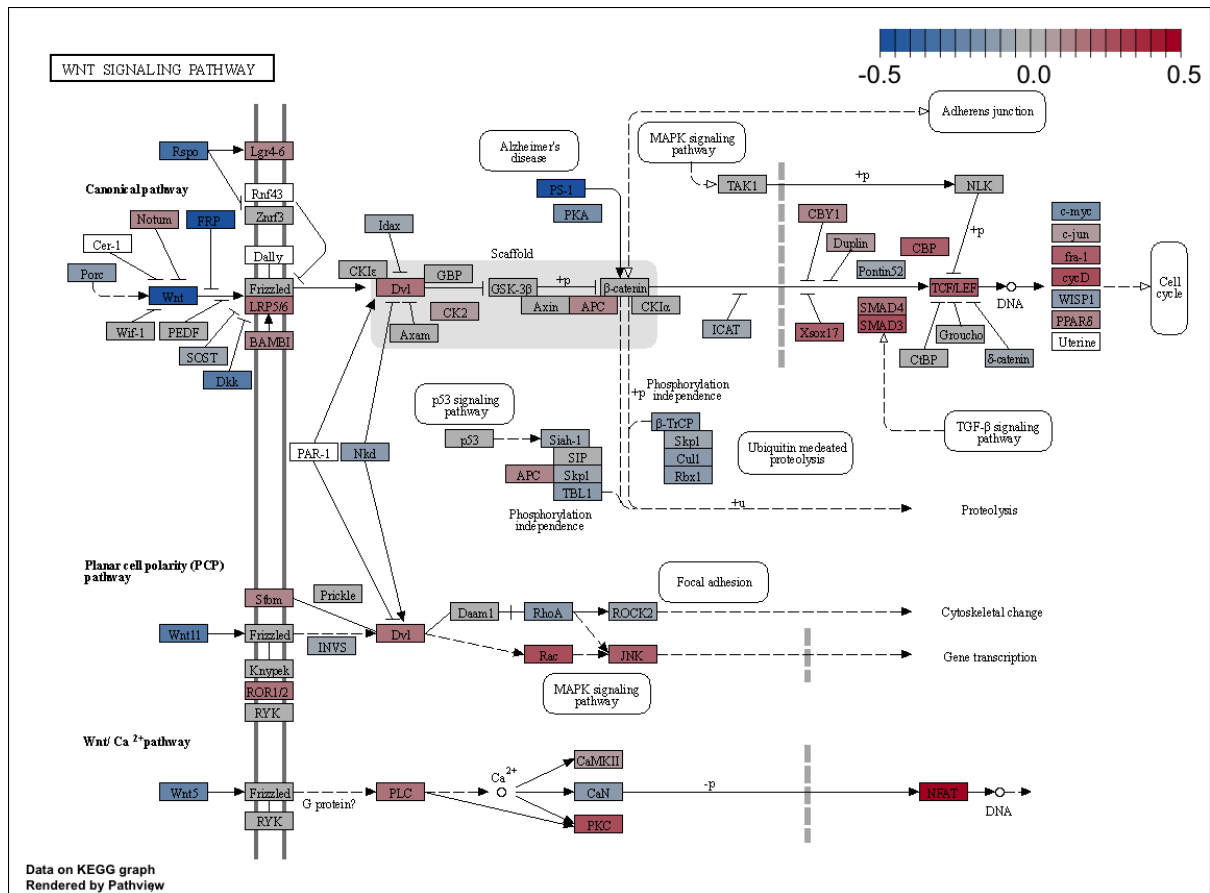


Figure S10: Pathview visualisation of the *KEGG_WNT_SIGNALING_PATHWAY* gene set. Colour of the cells indicates the logFC in fAI-like/+ brains.

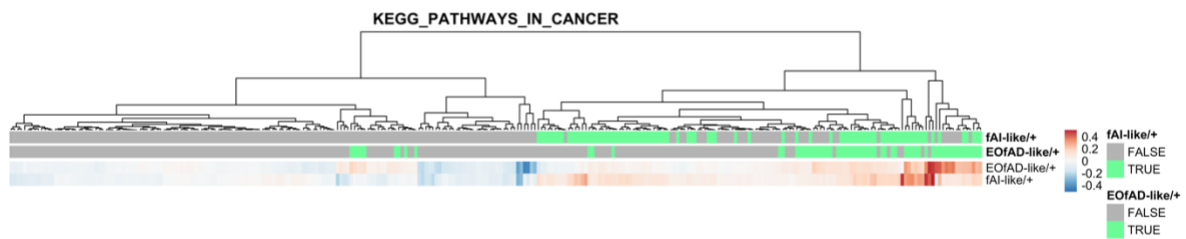


Figure S11: Heatmap indicating the logFC of genes in the *KEGG_PATHWAYS_IN_CANCER* gene set in EOfAD-like/+ and fAI-like/+ brains. Cells in the heatmap are labelled in green if found in the leading edge from the fgsea algorithm and grey if they were not.

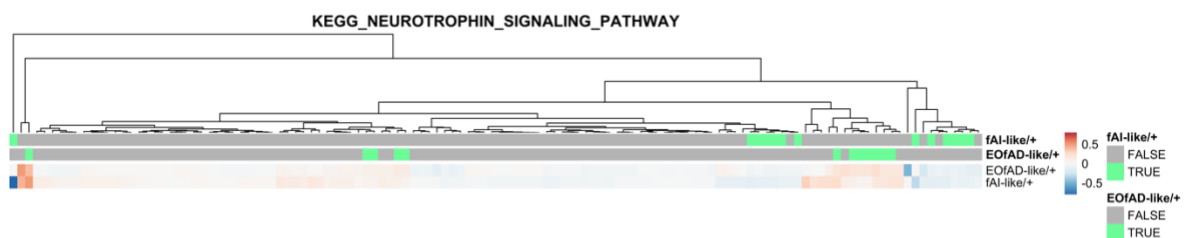


Figure S12: Heatmap indicating the logFC of genes in the *KEGG_NEUROTROPHIN_SIGNALING_PATHWAY* gene set in EOfAD-like/+ and fAI-like/+ brains. Cells in the heatmap are labelled in green if found in the leading edge from the fgsea algorithm and grey if they were not.

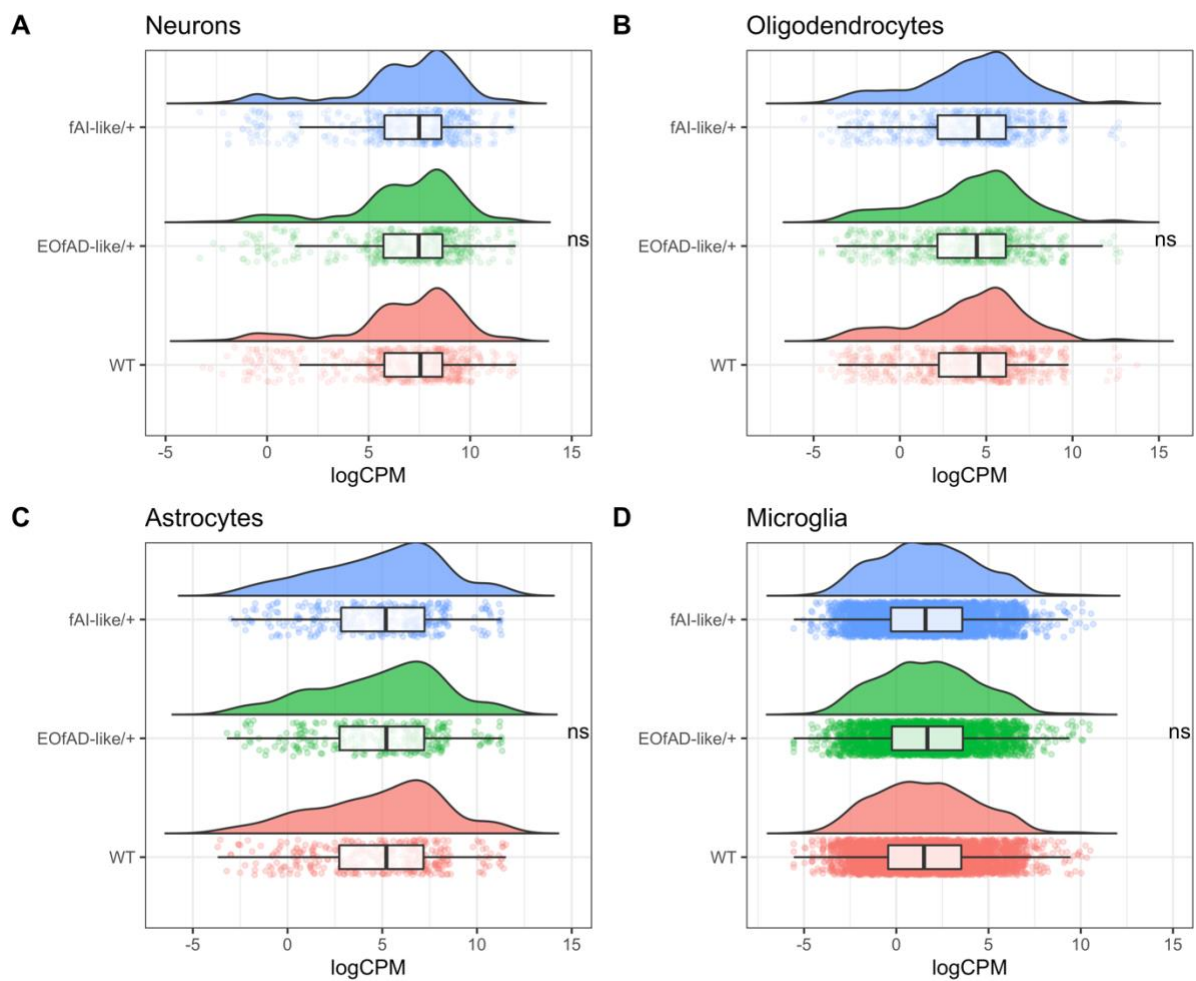


Figure S13: Cell type proportions appear to be consistent across samples. The raincloud plots below summarise the expression (logCPM) of marker genes for four broad cell types in zebrafish brains: neurons, oligodendrocytes, astrocytes and microglia. The marker genes for **A** neurons, **B** oligodendrocytes and **C** astrocytes were obtained from [1], while the **D** microglia markers were obtained from [2]. No significant differences (ns) of expression of these markers was found across genotypes using *fry* with a directional hypothesis, supporting that cell type proportions are consistent.

Supplemental Tables can be found at

<https://www.biorxiv.org/content/10.1101/2021.01.26.428321v2>

References:

1. Cahoy, J.D., et al., *A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function.* J Neurosci, 2008. **28**(1): p. 264-78.
2. Oosterhof, N., et al., *Identification of a conserved and acute neurodegeneration-specific microglial transcriptome in the zebrafish.* Glia, 2017. **65**(1): p. 138-149.

Chapter 9: Comparative analysis of Alzheimer's disease knock-in model brain transcriptomes implies changes to energy metabolism as a causative pathogenic stress

This chapter is a hybrid Results and Discussion chapter. The results of **Chapters 5-8** are summarised, then compared to other knock-in models of Alzheimer's disease.

Statement of Authorship

Title of Paper	A comprehensive, multi-species brain transcriptome analysis of knock-in models of early- and late-onset Alzheimer's disease reveal changes to mitochondrial and ribosome function as fundamental early stresses in disease pathogenesis.		
Publication Status	<input type="checkbox"/> Published	<input type="checkbox"/> Accepted for Publication	<input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
	<input checked="" type="checkbox"/> Submitted for Publication		
Publication Details	Submitted to Neuron (Cell Press)		

Principal Author

Name of Principal Author (Candidate)	Karissa Barthelson		
Contribution to the Paper	Performed the analysis, drafted the manuscript.		
Overall percentage (%)	90%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	11/2/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Morgan Newman		
Contribution to the Paper	Conception of the project, drafting of manuscript		
Signature		Date	15/02/2021

Name of Co-Author	Michael Lardelli		
Contribution to the Paper	Conception of the project, drafting of manuscript		
Signature		Date	19/02/2021

Please cut and paste additional co-author panels here as required.

Comparative analysis of Alzheimer's disease knock-in model brain transcriptomes implies changes to energy metabolism as a causative pathogenic stress

Authors: Karissa Barthelson^{1*}, Morgan Newman¹, Michael Lardelli¹

Affiliations:

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

Author list footnotes: *Lead/corresponding author

Corresponding author e-mail address: karissa.barthelson@adelaide.edu.au

Summary

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 $\epsilon 4$. 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 both zebrafish and knock-in, single EOfAD mutation models for understanding the causes of this disease.

Keywords

Alzheimer's disease, RNA-seq, microarray, APP, APOE, PSEN1, PSEN2, SORL1, mitochondria, ribosome

Introduction

Alzheimer's disease (AD) is a complex and highly heterogeneous neurodegenerative disease, defined by the presence of intracellular neurofibrillary tangles (NFTs), and extracellular plaques consisting of a small peptide, amyloid β ($A\beta$) (Jack et al., 2018). AD was discovered over 100 years ago (Alzheimer, 1906). However, an effective therapeutic to halt, or even to slow disease progression, remains to be developed.

AD has a strong genetic basis (reviewed in (Sims et al., 2020)). In some rare cases, early-onset familial forms of AD (EOAD, 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 (Kunkle et al., 2019; Lambert et al., 2013). The most potent variant is the $\epsilon 4$ 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 access to young, pre-symptomatic brains is limited. Imaging studies have implicated structural and functional changes to the brain long before diagnosis of overt AD (Iturria-Medina

et al., 2016; Quiroz et al., 2015). However, 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., 2021; Barthelson et al., 2020b; Barthelson et al., 2020c; Barthelson et al., 2020e; 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”). Since EOfAD and LOAD present as similar diseases (reviewed in (Blennow et al., 2006; Masters et al., 2015)), despite their different genetic underpinnings, understanding the molecular effects of heterozygosity for EOfAD mutations may also give insight into changes occurring in 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-G-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 ϵ 4 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

We first collated our findings from our zebrafish models of EOfAD-like mutations in *psen1* (Barthelson et al., 2021; Hin et al., 2020a; Hin et al., 2020b; Newman et al., 2019), *psen2* (Barthelson et al., 2020c) and *sorl1* (Barthelson et al., 2020b; Barthelson et al., 2020e). An advantage of using zebrafish for RNA-seq analyses is minimisation of genetic and environmental noise through breeding strategies such as that shown in **Figure 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 **Supplemental File 1**). The detailed analyses can be found in the

publications cited above. However, the outcomes are summarised below and in **Figure 1**.

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., 2021; Barthelson et al., 2020b; Barthelson et al., 2020c; Barthelson et al., 2020e) (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 **Figure 1B**. 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, 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, K97fs (Hin et al., 2020a) models the PS2V isoform of human *PSEN2* (Sato et al., 1999), that shows increased expression in LOAD brains (see (Moussavi Nik et al., 2015) for an explanation)

and so 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* (**Figure 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 *sor11* mutants (i.e. with complete loss of wild type *sor11*), as shown by significant changes to the expression of genes possessing IRE(s) in the 3' untranslated regions of their encoded mRNAs.

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 A β 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. (2017) 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. We re-analysed this microarray dataset to ask:

- 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?

We did not observe alteration of any similar gene sets between the *App*^{NL-G-F} mice and our EOfAD-like zebrafish. 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-G-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 (**Figure 2**).

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* (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 **Figure 1**, we identified changes in the gene set *KEGG_OXIDATIVE_PHOSPHORYLATION* in young

adult zebrafish heterozygous for EOfAD-like mutations in *sor11*, as well as in zebrafish modelling overexpression of the PS2V isoform that is upregulated in LOAD. Therefore, we asked whether changes in oxidative phosphorylation might be associated with the most common genetic risk factor for LOAD, the $\epsilon 4$ allele of the gene *APOE* (Corder et al., 1993; Genin et al., 2011; 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 $\epsilon 4$ 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* promoter: the predominant human allele $\epsilon 3$, the rare AD-protective $\epsilon 2$ allele, and the AD-risk allele, $\epsilon 4$. Zhao et al. (2020) performed a comprehensive brain transcriptome profiling experiment across aging in both male and female mice to assess the effect of homozygosity for the $\epsilon 2$ or $\epsilon 4$ alleles relative to the risk-neutral $\epsilon 3$ allele. 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). For details of this re-analysis see **Supplementary File 3**. We found statistical evidence for significant changes in expression of oxidative phosphorylation and ribosome gene sets in mice homozygous for the humanised $\epsilon 4$ *APOE* allele, consistent with our zebrafish models of EOfAD. These effects were not observed for the AD-protective $\epsilon 2$ allele (**Figure 3**). Note that the effects of *APOE* genotype were highly dependent on the litter-of-origin of the samples and that changes to cell type proportions were observed in the male APOE4 mice. 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

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 Ca²⁺ 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)).

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-EOfAD-like 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., 2020), suggesting changes to mitochondrial function are a very early cellular stress in AD pathogenesis. All of the mutations studied have also resulted in changes in the levels of transcripts required for ribosome formation and we suggest this may reflect changes in mTOR activity (see **Supplementary File 4**). The reason for the different effects of EOfAD-like and non-EOfAD like mutations in these genes is uncertain. However, for the *PRESENILINs*, 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). The EOfAD genes, *PSEN1*, *PSEN2*, *APP*, and *SORL1*, encode proteins which are all expressed within the endolysosomal pathway of cells (Andersen et al., 2005; Kawai et al., 1992; Pasternak et al., 2003; Sannerud et al., 2016) and also 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 Ca²⁺ signalling (Duchen, 1992)), oxidative protein folding (reviewed in (Simmen et al., 2010)), and the initiation of autophagy (Hamasaki et al., 2013). Interestingly, like mutations in *PSEN1* (Lee et al., 2010) and the C99 fragment of APP (Jiang et al., 2019), the ε4 allele of APOE has also been shown to affect lysosomal pH (Prasad and Rao, 2018) and the MAM (Tambini et al., 2016). From the work in

this paper, we have now seen that the “semi-dominant” $\epsilon 4$ 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, early effects on oxidative phosphorylation (and so energy production) appear to be a common, early disturbance associated with both early- and late- onset forms of Alzheimer’s disease. Our analysis supports that the two diseases are part of a single pathological continuum, with age of onset likely influenced by the severity of the energy effect involved. (After all, the 65 year age threshold defined as differentiating “early-onset” from “late-onset” is arbitrary.) 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).

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 ($A\beta$ deposition and neurofibrillary tangles of tau protein (Jack et al., 2018)). By expressing multiple mutant forms of EOfAD genes in transgenic mice, $A\beta$ 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 $A\beta$ levels (Foley et al., 2015) 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 amyloid β 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 likely due to changes in the relative proportions of different cell types in the model.

Changes to cell type proportions are not observed in our zebrafish models (Barthelson et al., 2021; Barthelson et al., 2020c; Barthelson et al., 2020d, e; Hin et al., 2020a; Hin et al., 2020b), revealing another advantage of exploiting zebrafish in transcriptome analysis of bulk RNA from brain tissue (their highly regenerative brains (Kroehne et al., 2011) are likely more resistant to changes in cell type proportions).

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 ϵ 4 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.

Acknowledgements

The authors would like to thank Dr. Nhi Hin for providing the Q96_K97del zebrafish dataset and original analysis. The results for the young APOE-TR mice were based on data obtained from the AD Knowledge Portal (<https://adknowledgeportal.synapse.org/>). Support for these data was provided by the NIH RF1 AG051504 and P01 AG030128. We thank Drs. Patrick Sullivan and Nobuyo Maeda for generating human APOE targeted replacement mice and providing access through Taconic.

Author Contributions

K.B. performed the analysis. M.L. and M.N. conceived the project. All authors contributed to drafting and editing the manuscript.

Declaration of Interests

The authors declare no competing interests.

Tables

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

Gene	Mutation(s)	Total number of zebrafish analysed	Genders analysed	Comment	Reference
<i>psen1</i>	W233fs (fAI-like) and T428del (EOfAD-like)	24	Males and Females	High read depth (between 61 and 110 million reads per sample)	(Barthelson et al., 2021)
<i>psen1</i>	Q96_K97del (EOfAD-like)	32	Males and females	Aged zebrafish and the effect of acute hypoxia treatment was also included	(Hin et al., 2020b; Newman et al., 2019)
<i>psen1</i>	K97fs	12	Females only	Aged zebrafish were also included	(Hin et al., 2020a)
<i>psen2</i>	T141_L142delinsMISLISV (EOfAD-like) and N140fs (not EOfAD-like)	15	Males and females	-	(Barthelson et al., 2020c)
<i>sor11</i>	W1818*	12	Males and Females	-	(Barthelson et al., 2020d)
<i>sor11</i>	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.	(Barthelson et al., 2020e)

Figures

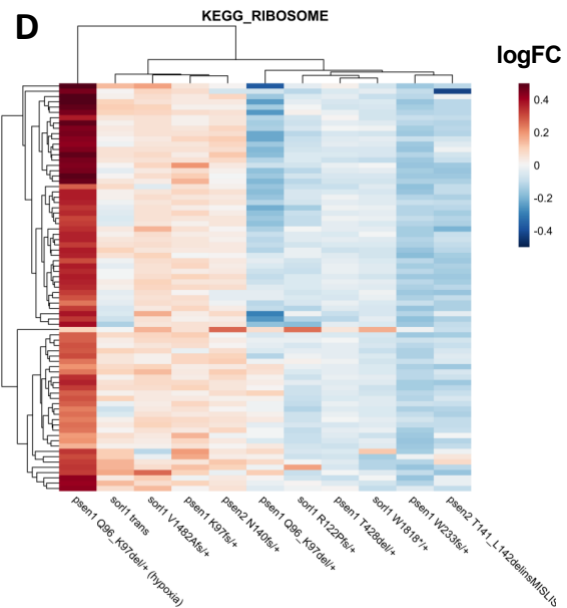
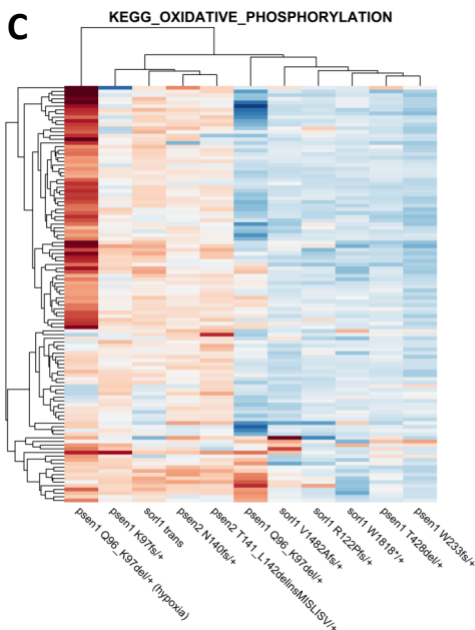
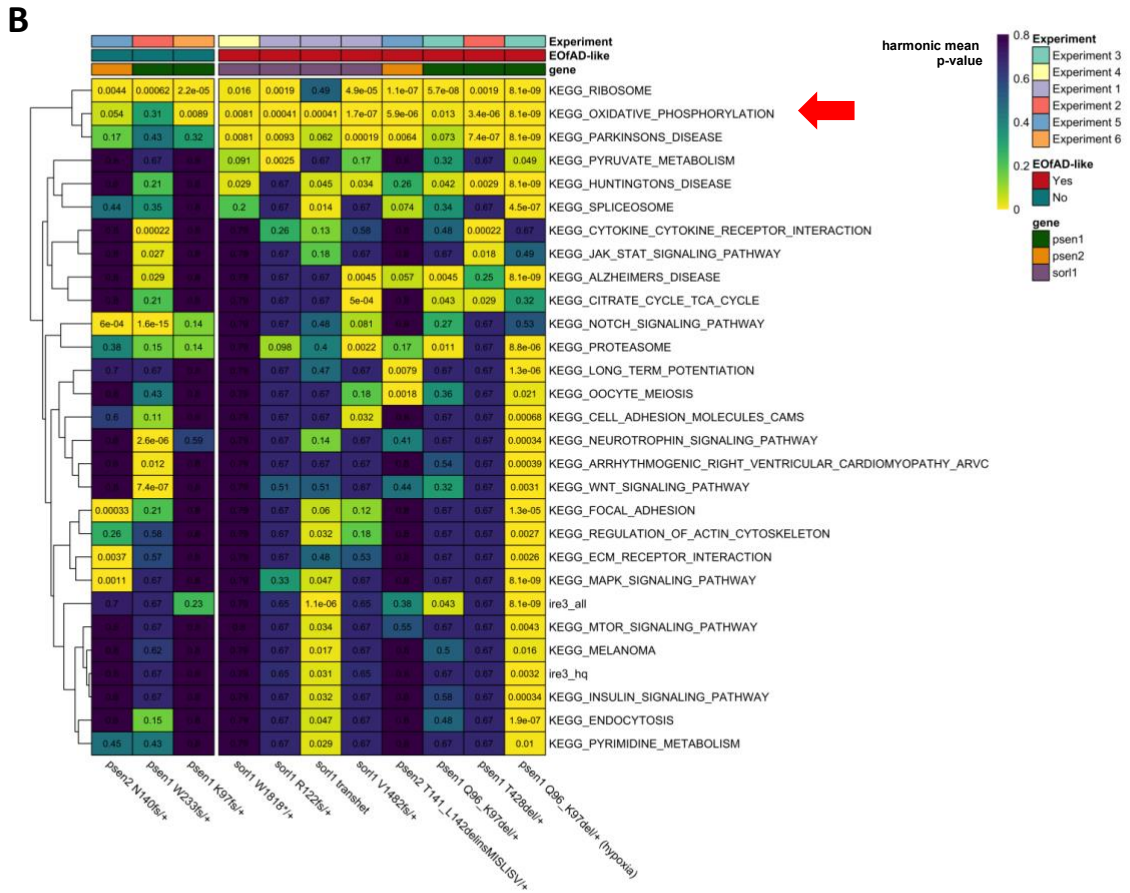
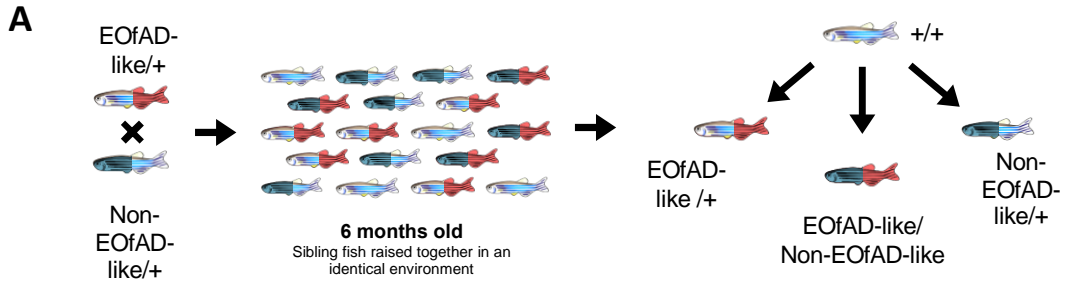


Figure 1: A. Schematic of a RNA-seq experiment using zebrafish. A single mating of a single pair of fish heterozygous for either an EOfAD-like or a non-EOfAD-like mutation results in a family heterozygous mutant, transheterozygous mutant, and wild type siblings. Comparisons made between genotypes in 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 and columns are clustered based on their Euclidean distance. Only genes considered detectable in all RNA-seq experiments are depicted. See **Supplemental File 1** for more information on individual study designs.

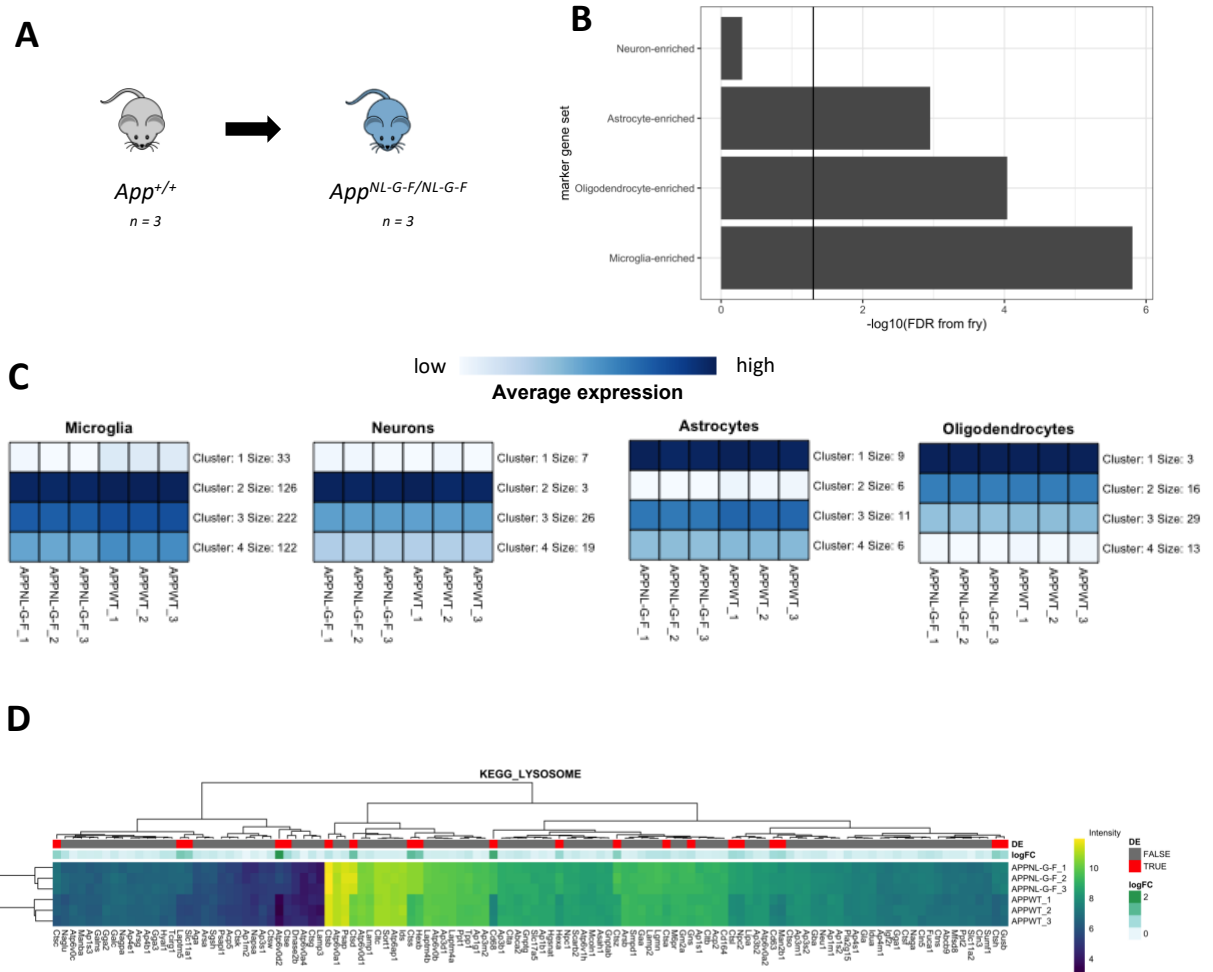


Figure 2: A. Visual representation of comparison of homozygous *App*^{NL-G-F} mice to wild type mice. **B.** Bar chart showing the FDR-adjusted p-value (directional hypothesis) from *fry* on marker genes of neurons, oligodendrocytes, astrocytes and microglia in *App*^{NL-G-F} relative to wild type. **C.** Heatmap indicating the expression 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. See **Supplemental File 2** for more details.

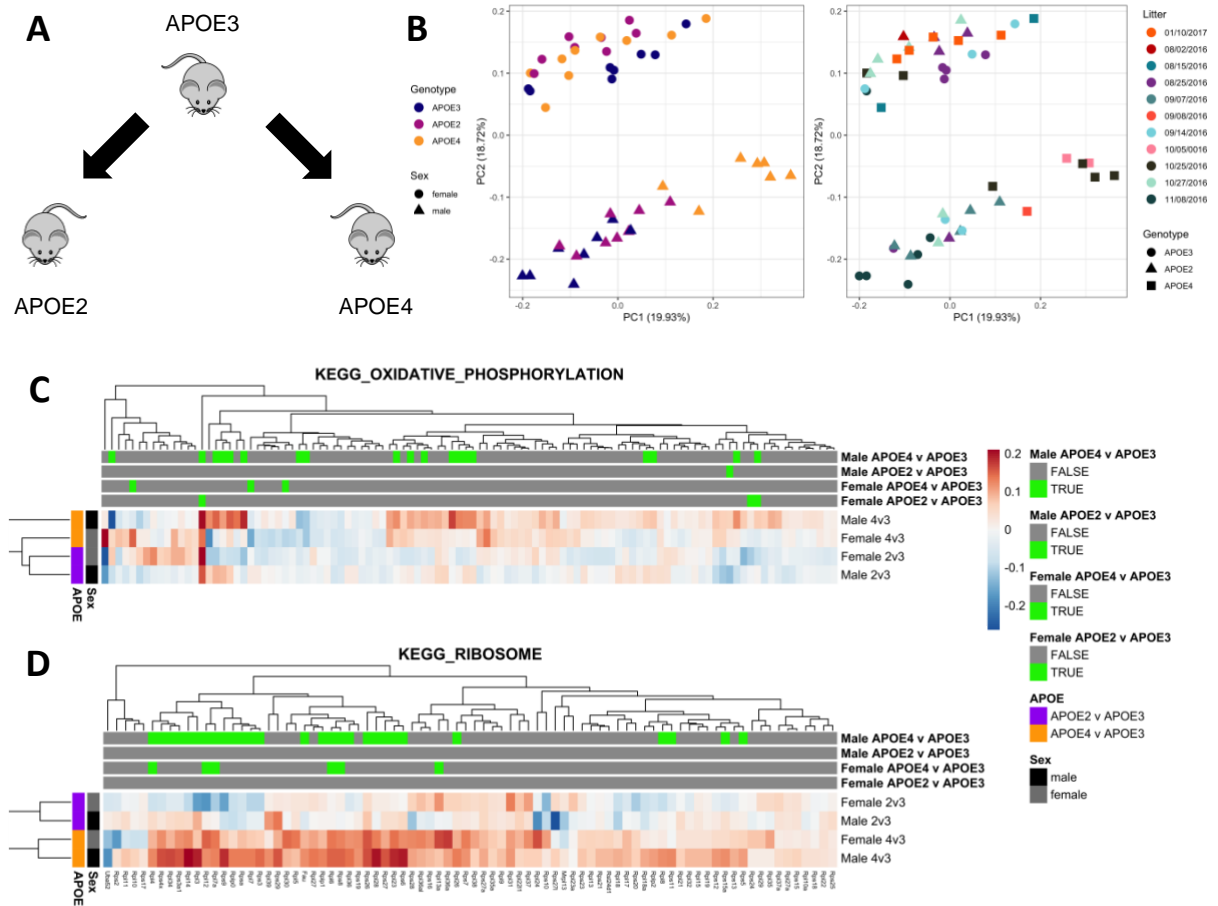


Figure 3: A. Visual representation of comparison of APOE4 or APOE2 mice to APOE3. This was performed for both male and female mice separately. **B.** Principal component analysis (PCA) of three month old APOE-TR mice. Principal component 1 (PC1) is plotted against PC2. The numbers between parentheses 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. Genes are labelled in green above whether they were classified as differentially expressed (FDR < 0.05) in the differential gene expression analysis in a particular comparison. See **Supplemental File 3** for more details.

STAR Methods

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Karissa Barthelson (karissa.barthelson@adelaide.edu.au).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability Statement

The complete list of datasets (with links), software, and relevant algorithms used in this study is provided in the Key Resources Table. The code used to perform the analysis in this study can be found at <https://github.com/karissa-b/AD-signature>

METHOD DETAILS

Zebrafish analysis

The harmonic mean p-values and differential gene expression analysis outputs were obtained from each individual analysis (see the Key Resources Table and (Barthelson et al., 2021; Barthelson et al., 2020b; Barthelson et al., 2020c; Barthelson et al., 2020e)). 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)). 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 and the results of the differential gene expression analysis described in (Hin et al., 2020b) from the first author of the cited paper.

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

The detailed analysis can be found in **Supplemental File 2**, while the R code to reproduce the analysis can be found <https://github.com/karissa-b/AD-signature>. Briefly, the raw .CEL files (obtained from GEO) were imported into analysis with *R*, and pre-processed 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-AD 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 as described for the zebrafish analysis.

APOE-TR RNA-seq re-analysis

The detailed analysis can be found in **Supplemental File 3**. 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* (v1.10) (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.

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 gene length and %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 length alone. We also performed ranked-list based enrichment analysis as described for the zebrafish analysis.

Code to reproduce the analysis for the APOE-TR dataset can be found at <https://github.com/karissa-b/AD-signature>.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis of *App*^{NL-G-F} microarray data

The detailed analysis can be found in **Supplemental File 2**. The sample numbers for the mouse *App*^{NL-G-F} microarray re-analysis consisted of $n = 3$ male mouse cortex transcriptomes per genotype (Castillo et al., 2017). The raw .CEL files were obtained from GEO, then imported into R and pre-processed using the *rma* method (Irizarry et al., 2003), which performs background correction, quantile normalisation and summarisation (median-polish). Differential gene expression analysis was conducted using *limma* v3.44.3 (Ritchie et al., 2015). The generated p-values were corrected for multiple testing using the false discovery rate (FDR). We considered a gene (probeset) to be differentially expressed if the FDR-adjusted p-value was less than 0.05 (no fold change cut-off).

Enrichment of the *KEGG* and *IRE* gene sets within the DE gene lists was performed using the *kegga*. We also performed ranked-list based enrichment analysis by calculation of the harmonic mean p-value (Wilson, 2019) of the raw p-values calculated from *fry* (Wu et al., 2010), *camera* (Wu and Smyth, 2012) and *fgsea* (Sergushichev, 2016). For enrichment analysis, we considered a gene set to be significantly altered if the FDR-adjusted p-value from *goseq*, or the FDR-adjusted harmonic mean p-value was less than 0.05.

Statistical analysis of APOE-TR RNA-seq data

The detailed analysis can be found in **Supplemental File 3**. Sample numbers mostly consisted of $n = 8$ mice per genotype, age and sex. However, upon inspection of the expression of genes of the Y-chromosome, we noted that three mice at 3 months of age had been incorrectly identified. The raw gene counts were calculated from *featureCounts*, normalised using the trimmed mean of M-values (TMM) method (Robinson and Oshlack, 2010), and used for calculation of gene expression values as logCPM (log₂-counts-per-million). Differential gene expression analysis was performed using a generalised linear model and likelihood ratio tests as implemented in *edgeR* (Robinson et al., 2009). The generated p-values were corrected for multiple testing using the false discovery rate (FDR). We considered a gene to be differentially expressed if the FDR-adjusted p-value was less than 0.05 (no fold change cut-off). To test for over-representation of the *KEGG* and *IRE* gene sets within the DE genes, we used *goseq* (Young et al., 2010), specifying the average transcript length to calculate the probability weighting function. We considered a gene set to be significantly over-represented if the FDR-adjusted p-value generated by *goseq* was < 0.05 . The ranked-list enrichment analysis was performed as described for the microarray data.

References

- Alzheimer, A. (1906). Über einen eigenartigen schweren erkrankungsprozeß der hirnrinde. *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, 13461-13466.

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., and Schon, E.A. (2009). Presenilins are enriched in endoplasmic reticulum membranes associated with mitochondria. *American Journal of Pathology* 175, 1810-1816.

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. (2021). 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. *bioRxiv*, 2021.2001.2026.428321.

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., Newman, M., and Lardelli, M. (2020b). Transcriptome analysis of a protein-truncating mutation in sortilin-related receptor 1 associated with early-onset familial Alzheimer's disease indicates effects on mitochondrial and ribosome function in young-adult zebrafish brains. *bioRxiv*, 2020.2009.2003.282277.

Barthelson, K., Pederson, S.M., Newman, M., Jiang, H., and Lardelli, M. (2020c). Frameshift and frame-preserving mutations in zebrafish presenilin 2 affect different cellular functions in young adult brains. *bioRxiv*, 2020.2011.2021.392761.

Barthelson, K., Pederson, S.M., Newman, M., and Lardelli, M. (2020d). 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 Preprint*, 1-15.

Barthelson, K., Pederson, S.M., Newman, M., and Lardelli, M. (2020e). 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.

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

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

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

- Carvalho, B.S., and Irizarry, R.A. (2010). A framework for oligonucleotide microarray preprocessing. *Bioinformatics* 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.
- 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, 921-923.
- 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* 29, 15-21.
- Dong, Y., Newman, M., Pederson, S., Hin, N., and Lardelli, M. (2020). Transcriptome analyses of 7-day-old zebrafish larvae possessing a familial Alzheimer's disease-like mutation in psen1 indicate effects on oxidative phosphorylation, mcm functions, and iron homeostasis. *bioRxiv*, 2020.2005.2003.075424.
- Duchen, M.R. (1992). Ca(2+)-dependent changes in the mitochondrial energetics in single dissociated mouse sensory neurons. *Biochem J* 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.

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, 1349-1356.

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, 2006-2008.

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. *Journal of Alzheimer's disease : JAD* 44, 787-795.

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, 903-907.

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, 101-106.

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*, 2020.2005.2001.071498.

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.

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.

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- β CTF (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, 947-958.

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. *The European journal of neuroscience* 19, 2826-2838.

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 A β , 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, 1146-1158.

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* 25, 2078-2079.

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

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

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.

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.

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.

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. *Human molecular genetics* 24, 3662-3678.

Myers, A., and McGonigle, P. (2019). Overview of Transgenic Mouse Models for Alzheimer's Disease. *Current protocols in neuroscience* 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., Kaye, 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, 409-421.

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

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, 26687-26694.

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.

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* 26, 139-140.
- Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* 11, R25.
- 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.
- 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 A β 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, 2498-2505.
- Saunders, A.M., Schmechel, K., Breitner, J.C., Benson, M.D., Brown, W.T., Goldfarb, L., Goldgaber, D., Manwaring, M.G., Szymanski, M.H., McCown, N., and *et al.* (1993). Apolipoprotein E epsilon 4 allele distributions in late-onset Alzheimer's disease and in other amyloid-forming diseases. *Lancet (London, England)* 342, 710-711.
- 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. *Human molecular genetics* 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., and Mesirov, J.P. (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.

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 reports* 17, 27-36.

Team, R.C. (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 (Barcelona, Spain)* 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., Slegers, 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.

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.

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* 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.e726.

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

Supplementary File 1 can be found at

<https://www.biorxiv.org/content/10.1101/2021.02.16.431539v2.supplementary-material>

Supplementary File 2: *App*^{NL-G-F} microarray re-analysis

Castillo et al. (2017) previously performed transcriptome analysis on the brains of two mouse models of AD: a knock-in model, *App*^{NL-G-F} (Saito et al., 2014), and the transgenic model 3xTg-AD-H (Oddo et al., 2003). The *App*^{NL-G-F} strain of mice carries a total of six mutations in the murine *App* gene: three mutations that humanise the mouse A β sequence, plus the Swedish (KM670/671NL), Iberian (I716F) and Arctic (E693G) EOfAD mutations (Saito et al., 2014). The 3xTg-AD-H model of AD (hereafter referred to as 3xTg) carries the M146V EOfAD mutation within the endogenous mouse *Psen1* gene, and expresses two transgenes: human *APP* with the Swedish EOfAD mutation, and human *MAPT* with the P301L mutation. Transcription of the transgenes is driven by the mouse *Thy1.2* promoter (Oddo et al., 2003).

The cortices of three male mice of each of these mutant strains (both strains of mice were homozygous for their respective mutations/transgenes), as well as wild type *APP*^{+/+} and non-transgenic (non-Tg) controls, were subjected to microarray analysis at 12 months of age using the Affymetrix Mouse Gene 2.0ST Array. All mice used in the study were maintained as inbred lines. 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.

We first obtained the raw microarray data from the GEO database (accession number GSE92926). Initially, we attempted to replicate the results of Castillo et al. (2017) using the Affymetrix Transcriptome Analysis Console software. However, we were unable to find sufficient information to replicate their results. Therefore, we analysed the microarray dataset in a reproducible manner by importing the .CEL files for all twelve mice for analysis with *R* (Team, 2019) using the *oligo* package (Carvalho and Irizarry, 2010). The analysis of this dataset is based on the proposed workflow found in (Klaus and Reisenauer, 2018) and code to reproduce the analysis can be found at <https://github.com/karissa-b/AD-signature/>.

Pre-processing of raw data

Raw intensities were pre-processed using the robust multichip average (rma) method (Irizarry et al., 2003). We next excluded probesets which contained a median log₂ intensity value of < 3.5 (lowly expressed) and also any probesets

assigned to multiple genes as recommended by Klaus and Reisenauer (2018).

The distribution of the intensity values before and after the pre-processing can be found in **Figure 1**.

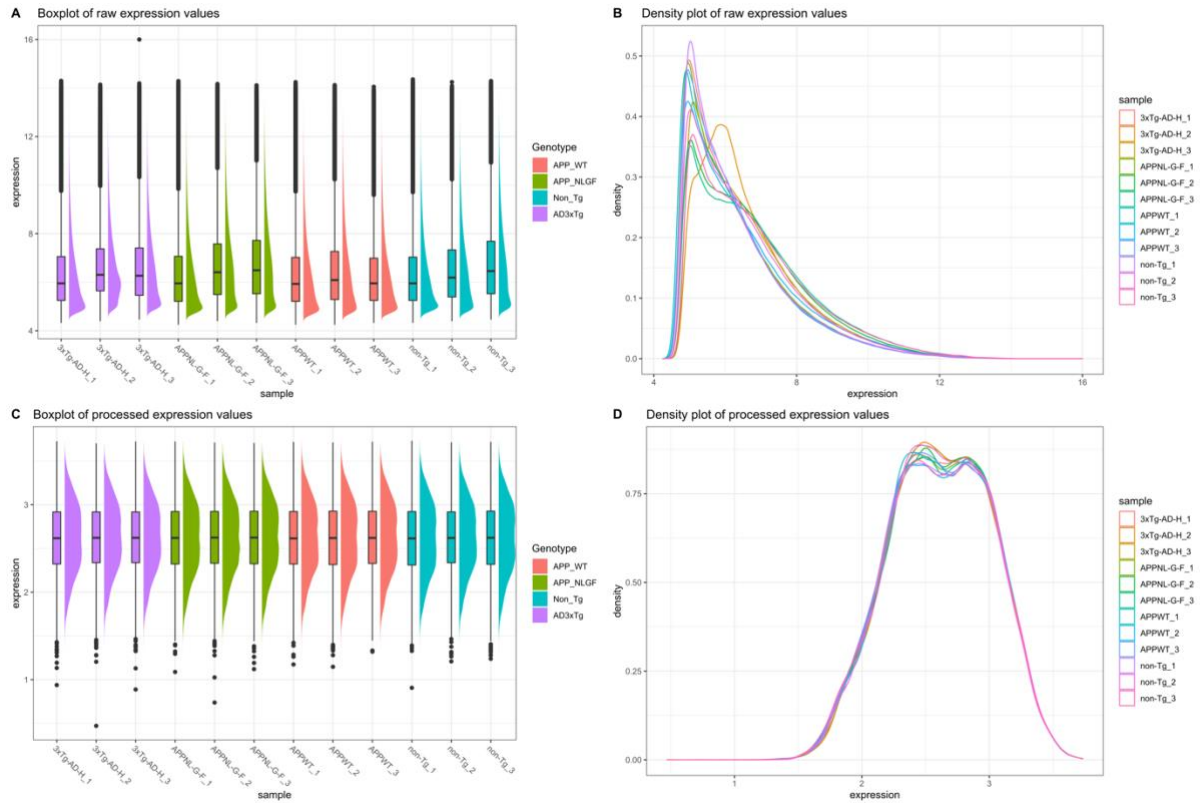


Figure 1: **A)** Boxplot and **B)** density plot of the raw intensity data. **C)** Boxplot and **D)** density plot of the intensity data after rma normalisation and filtering for lowly expressed and multi-mapping probes.

Principal component analysis

We performed principal component analysis (PCA) to explore the overall similarity between samples. **Figure 2** below shows the plot of principal component 1 (PC1) against PC2 for each microarray sample after pre-processing. Samples separated across PC1 by genotype, suggesting that the homozygous genotypes in this study result in distinct transcriptome states. Notably, the *App*^{NL-G-F/NL-G-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. This suggests that the disturbance to the cortex transcriptome in *App*^{NL-G-F/NL-G-F} mice is greater than that in 3xTg mice.

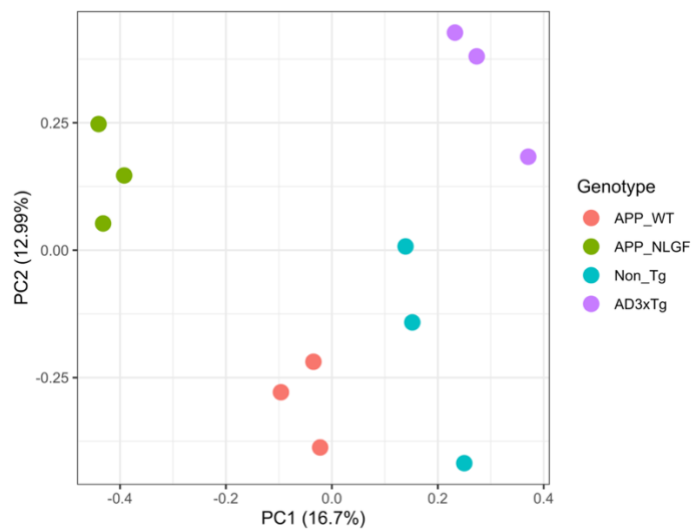


Figure 2: Principal component (PC) analysis on rma-processed intensity values.

Differential Gene Expression Analysis

To determine which probesets (i.e. genes) are differentially expressed (DE) in each of the pairwise comparisons of mutant mice with their respective non-mutant/non-Tg controls, we performed differential gene expression analysis with *limma* (Ritchie et al., 2015). A design matrix was generated to specify sample genotypes, while a contrasts matrix was generated to specify the pairwise contrasts. We considered a gene to be DE if the FDR-adjusted p-value was less than 0.05 (no fold change filter). We identified 158 genes to be differentially expressed (DE) in *App*^{NL-G-F/NL-G-F} mice relative to *App*^{+/+} mice, and 126 genes to

be DE in 3xTg mice relative to non-Tg wild type controls (**Figure 3A, B**). Three downregulated genes: *Fos*, *Gadd45b*, and *Nr4a1* and one upregulated gene, *I133*, were found to be DE in both forms of mutant mice relative to their wild type/non-transgenic counterparts (**Figure 3C, D**).

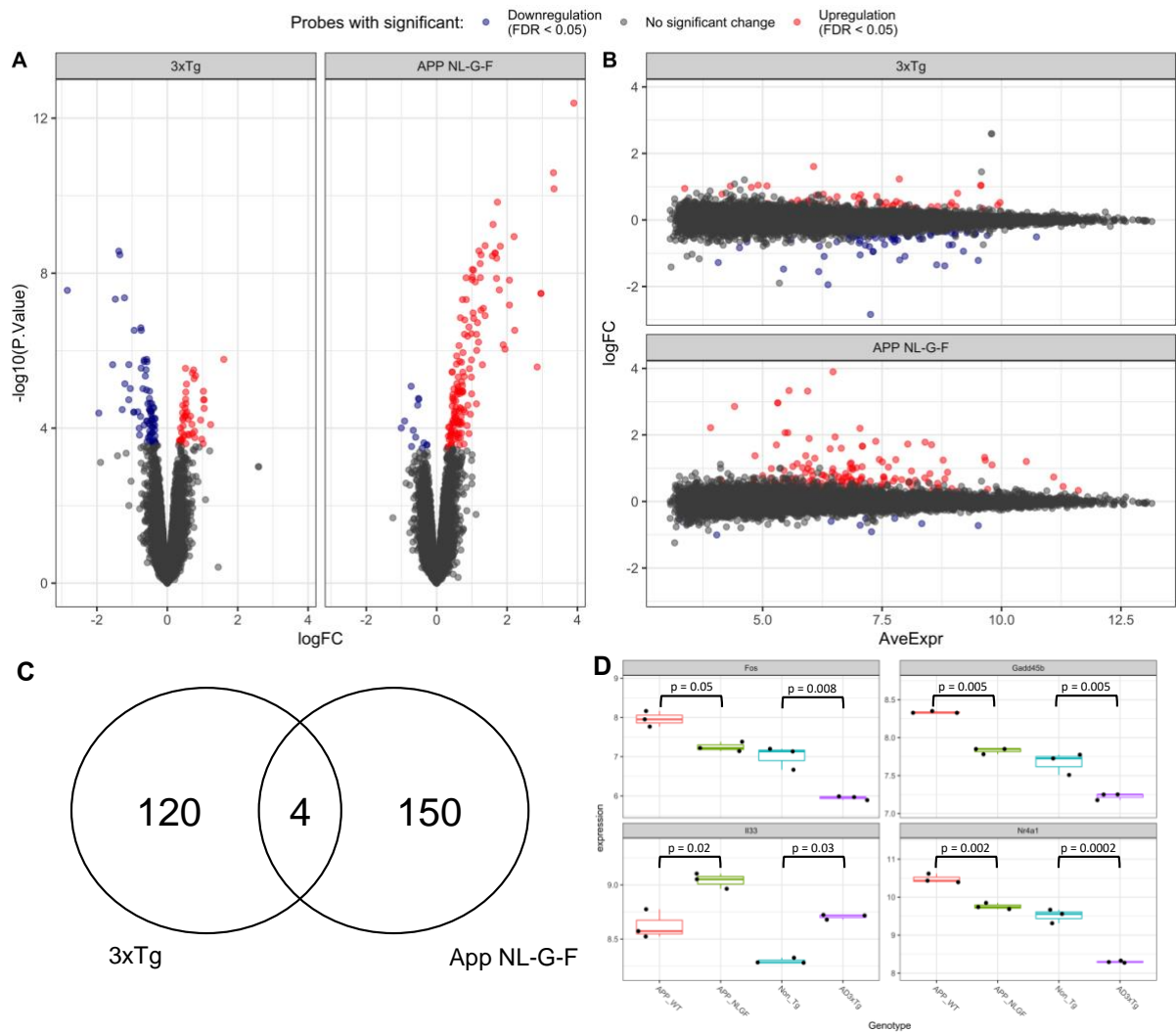
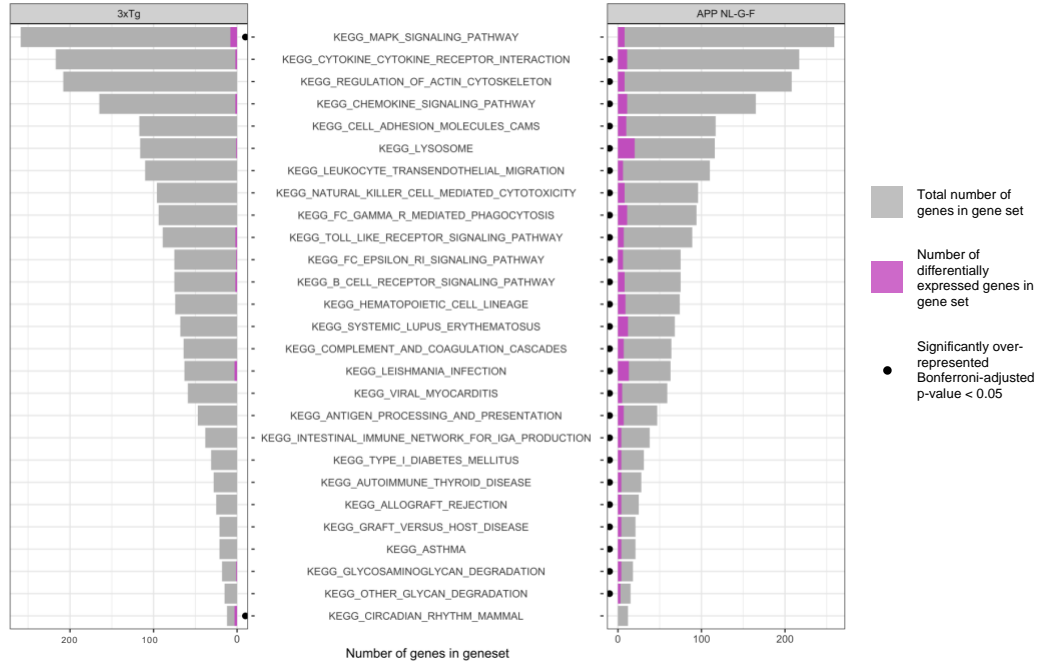


Figure 3: Differential gene expression analysis. **A)** Volcano plot and **B)** MD plot of the changes to gene expression in *App*^{NL-G-F/NL-G-F} and 3xTg mice relative to controls. **C)** Venn diagram showing four genes are identified as differentially expressed (DE) in both comparisons. **D)** Boxplot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

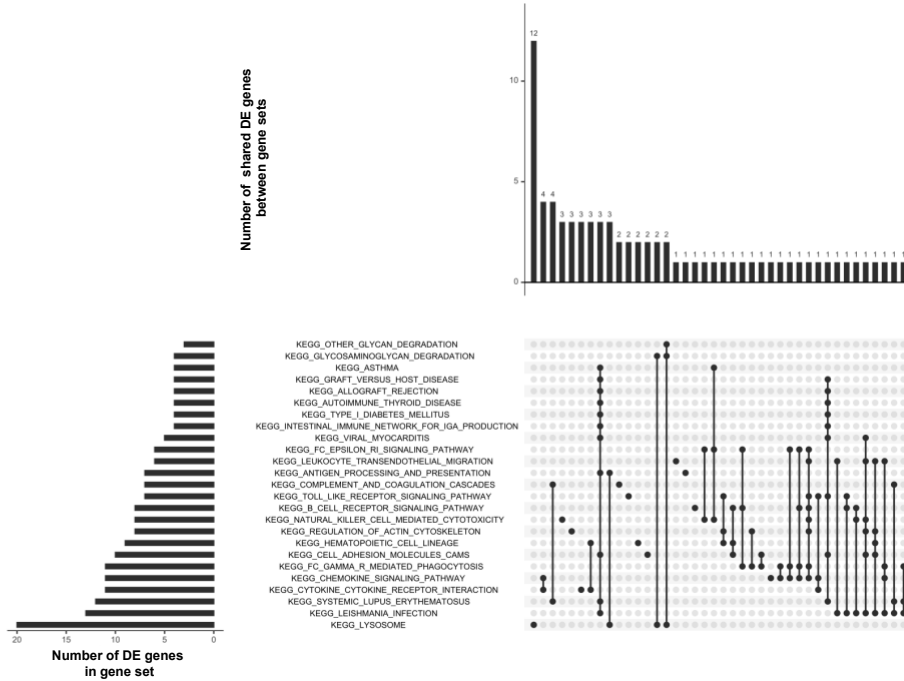
Over-representation analysis

We next tested whether any *KEGG* (Kanehisa and Goto, 2000) or *IRE* (Hin et al., 2020) gene sets were significantly over-represented within the DE genes in each comparison using *kegga* (Young et al., 2010). The *KEGG* gene sets were obtained using the package *msigdbR* (Dolgalev, 2020), and the *IRE* (Hin et al., 2020) gene sets, obtained from <https://github.com/nhihin/ire>. We restricted the gene sets to include only genes that had been tested for differential expression in the *limma* analysis. After correcting for multiple testing using the Bonferroni method, we found statistical evidence (Bonferroni-adjusted p-value < 0.05) for two gene sets as significantly over-represented among the DE genes in 3xTg mice, and 27 gene sets to be significantly overrepresented among the DE genes in *App^{NL-G-F/NL-G-F}* mice (**Figure 4A**). No *IRE* gene sets were found to be significantly enriched among either of the DE gene lists. Notably, the enriched DE genes within each gene set in the *App^{NL-G-F/NL-G-F}* mice appear to be shared across many of the gene sets (**Figure 4B**). Many of these gene sets contain shared DE genes are involved in inflammatory processes, which is not unexpected as gliosis has previously been observed to occur in these mutant mice (Castillo et al., 2017; Saito et al., 2014). Only the *KEGG_LYSOSOME* gene set appears to be enriched in mostly independent DE genes (which are upregulated) (**Figure 4C**). In summary, enrichment testing within the DE genes suggests that, at 12 months of age, the *App^{NL-G-F/NL-G-F}* mice have a more severe phenotype/transcriptome disturbance than the 3xTg mice (due to the larger number of gene sets significantly altered), and this phenotype mostly consists of an inflammatory response, and changes to the lysosome.

A



B



C

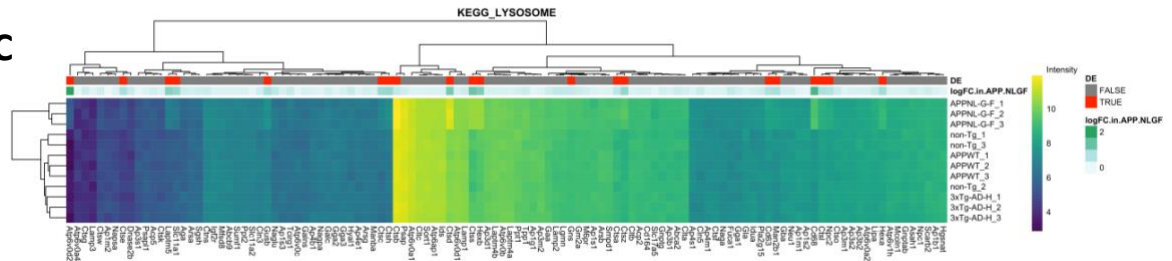


Figure 4: Over-representation analysis. **A)** Pyramid bar plot indicating the number of genes in the significantly enriched KEGG and IRE gene sets in *App*^{NL-G-F/NL-G-F} and 3xTg mice. Only gene sets with a Bonferroni adjusted p-value from *kegga* are shown (and are indicated by a black dot). The total numbers of genes in these gene sets are shown by grey bars, while the numbers of significantly differentially expressed 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-G-F/NL-G-F} mice. **C)** Heatmap of the *KEGG_LYSOSOME* gene set indicating the intensity expression values. Both rows and columns are clustered according to their Euclidean distance. Genes are labelled red if they were found to be differentially expressed by *limma* in *App*^{NL-G-F/NL-G-F} mice, and the magnitude of logFC is shown in green.

Ranked list enrichment analysis

We next obtained a more complete view on the changes to gene expression observed in *App*^{NL-G-F/NL-G-F} and 3xTg mice relative to their respective controls by performing three methods of rank-based, gene set enrichment analysis: *fry* (Wu et al., 2010), *camera* (Wu and Smyth, 2012) and *GSEA* (Sergushichev, 2016; Subramanian et al., 2005). We then calculated the harmonic mean p-value (Wilson, 2019) from the raw p-values from each method as discussed in (Barthelson et al., 2020). In this analysis, we considered a *KEGG* or *IRE* gene set to be significantly altered if the Bonferroni-adjusted harmonic mean p-value was less 0.05. We observed similar gene sets to be significantly altered using this ranked-list approach to the over-representation analysis described above. Additionally, the significance of the gene sets is likely being driven by similar genes, as supported by the observed number of the DE genes across the gene sets (and to a lesser extent, the leading edge genes from the *GSEA* algorithm) (**Figure 5**). A summary of whether a gene set is observed to be significantly

altered in either over-representation analysis or from our ranked-list approach can be found in **Figure 6**.

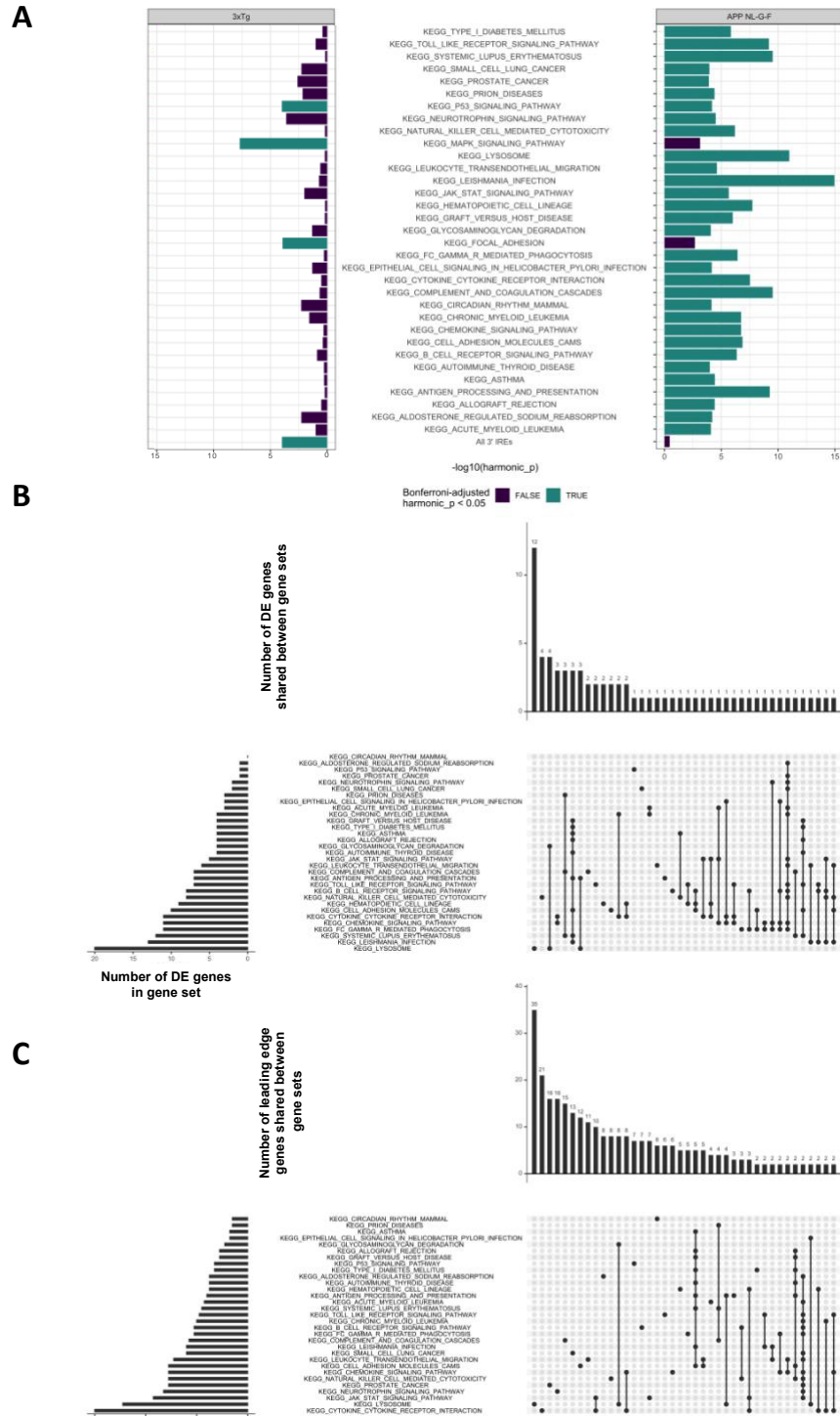


Figure 5: Ranked-list enrichment testing. **A)** Summary of significantly enriched *KEGG* and *IRE* gene sets in *App^{NL-G-F/NL-G-F}* and 3xTg mice. Gene sets are coloured according to whether they were below the threshold of a Bonferroni-adjusted harmonic mean p-value of < 0.05. **B)** Upset plot indicating the overlap of DE genes across the significantly altered gene sets in *App^{NL-G-F/NL-G-F}* mice. **C)** Upset plot indicating the leading edge genes from the *GSEA* algorithm in *App^{NL-G-F/NL-G-F}* mice.

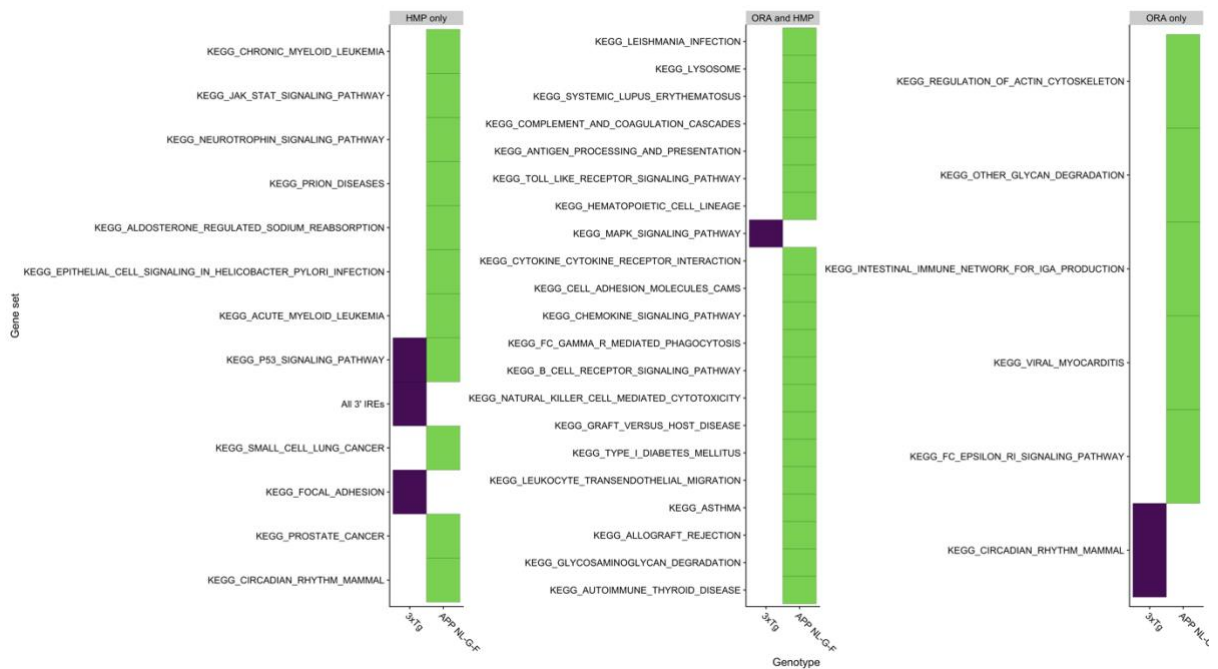


Figure 6: Summary of significantly altered gene sets in *App^{NL-G-F/NL-G-F}* and 3xTg mice. Gene sets only found to be altered by calculation of the harmonic mean p-value (HMP) are shown on the left. Gene sets only found to be altered by over-representation analysis (ORA) using *kegga* are shown on 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.

Changes to cell type proportions are present in *App^{NL-G-F/NL-G-F}* mice

Castillo et al. (2017) noted an increase in expression of marker genes of astrocytes, microglia and oligodendrocytes, while relatively consistent expression

of marker genes of neurons in $App^{NL-G-F/NL-G-F}$ mice. However, only a maximum of 16 genes were used as markers of these cell types. Therefore, we inspected larger sets of genes obtained from (Cahoy et al., 2008) and (Oosterhof et al., 2017) to assess whether cell type proportions are altered in these microarray samples. Using *fry* with a directional hypothesis, we observed statistical evidence for increased expression of marker genes of astrocytes, oligodendrocytes and microglia (FDR < 0.05) in $App^{NL-G-F/NL-G-F}$ mice, suggesting increased abundances of these broad cell types found in the mouse brain (**Figure 7**). Approximately half of the DE genes in the $App^{NL-G-F/NL-G-F}$ mice overlap with marker genes of microglia. Therefore, a significant proportion of the changes to gene expression we observed in this analysis may be artefactual due to increased proportions of microglia in the $App^{NL-G-F/NL-G-F}$ samples.

genotype is indicated by the red diamonds. To assist with visualisation of the increased expression of marker genes in $App^{NL-G-F/NL-G-F}$ 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 $App^{NL-G-F/NL-G-F}$ (left) and 3xTg (right) mice with the cell type marker gene sets.

Discussion and Conclusion

In summary, we draw three main conclusions from our re-analysis of the microarray transcriptomic data from the cortices of 12 month old male $App^{NL-G-F/NL-G-F}$ mice and 3xTg mice:

1. $App^{NL-G-F/NL-G-F}$ mice and 3xTg mice show distinct brain transcriptomic disturbances relative to controls.
2. Brain transcriptomes from $App^{NL-G-F/NL-G-F}$ mice, when compared to transcriptomes wild type mice, show apparent strong inflammatory signals and changes to gene expression implicating the lysosome.
3. A significant factor contributing (artefactually) to apparent disturbances of gene expression observed in $App^{NL-G-F/NL-G-F}$ mice is likely increased proportions of microglia within the cortex samples

Our re-analysis of the microarray dataset described by Castillo et al. (2017) mostly replicated their conclusions. At the single gene level, we found some overlap of DE genes with Castillo et al. (2017), and some distinct DE genes (data not shown). This is likely due to different approaches used in pre-processing of the raw data and testing for differential gene expression. However, at the pathway (i.e. gene set) level, both analyses predict similar processes to be affected in each mouse model.

We also observed highly significant upregulation of genes involved in lysosome function in *App*^{NL-G-F/NL-G-F} mice but not in 3xTg mice. Changes to the lysosomal network have been observed previously in *App*^{NL-G-F/NL-G-F} mice at a young age at the protein level (Whyte et al., 2020), supporting the validity of this result. The Iberian mutation of *APP* present in *App*^{NL-G-F/NL-G-F} mice has been shown to increase levels of the C99 fragment of APP (Guardia-Laguarta et al., 2010), which would impair acidification of the endo-lysosomal system (Jiang et al., 2019), and likely result in an intracellular ferrous iron deficiency, mitochondrial dysfunction and inflammation (Yambire et al., 2019). We did not observe any evidence for mitochondrial dysfunction or iron dyshomeostasis in this dataset. However, we suspect that these signals may be lost in the noise due to sample litter-of-origin issues and the strong inflammatory signal.

From this dataset, it appears that the *App*^{NL-G-F/NL-G-F} mice have a more severe phenotype than the 3xTg mouse at 12 months of age. This is unexpected, as the more extensive genomic differences of the 3xTg mice relative to *App*^{NL-G-F/NL-G-F} mice, would, in a simplistic view, be thought to affect more cellular processes. Indeed, 3xTg mice show cognitive impairment from 4 months age (Billings et al., 2005), while *App*^{NL-G-F/NL-G-F} mice do not show cognitive impairment until 6 months of age (Mehla et al., 2019). However, the memory impairments at 12 months of age appeared to be most severe in *App*^{NL-G-F/NL-G-F} mice (Mehla et al., 2019), which would imply that the effects on cellular processes are also more severe at this age.

Notably, the comparisons made in this microarray analysis between the *App*^{NL-G-F/NL-G-F} and *App*^{+/+} mice observed the compounded effects of homozygosity for 6

mutations simultaneously. An interesting future experiment could entail comparison of the *App*^{NL-G-F/NL-G-F} mice to mice carrying only the humanised A β sequence (i.e. (Serneels et al., 2020)) to assist with identification of changes to the transcriptome due specifically to the EOfAD mutations in this knock-in mouse model. Finally, the inclusion of three EOfAD mutations in *APP* within the same animal is of questionable relevance to the genetic state of human EOfAD. Therefore, the results presented here are not ideal for comparison with the transcriptome changes observed in our heterozygous, single knock-in mutation zebrafish models. The *App*^{NL} mouse model carries only the Swedish *APP* EOfAD mutation within the humanised A β sequence (Saito et al., 2014). A brain transcriptome analysis of young (3 or 6 months of age) heterozygous *App*^{NL} mice, relative to mice carrying only the humanised A β sequence, would not require the generation of new mouse mutants and would provide data more comparable with that already available from our studies using zebrafish.

References:

Barthelson, K., Pederson, S.M., Newman, M., and Lardelli, M. (2020). 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.

Billings, L.M., Oddo, S., Green, K.N., McGaugh, J.L., and LaFerla, F.M. (2005). Intraneuronal Abeta causes the onset of early Alzheimer's disease-related cognitive deficits in transgenic mice. *Neuron* 45, 675-688.

Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A., *et al.* (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain

development and function. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 28, 264-278.

Carvalho, B.S., and Irizarry, R.A. (2010). A framework for oligonucleotide microarray preprocessing. *Bioinformatics* 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.

Dolgalev, I. (2020). msigdb: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format. p. R package.

Guardia-Laguarta, C., Pera, M., Clarimón, J., Molinuevo, J.L., Sánchez-Valle, R., Lladó, A., Coma, M., Gómez-Isla, T., Blesa, R., Ferrer, I., and Lleó, A. (2010). Clinical, Neuropathologic, and Biochemical Profile of the Amyloid Precursor Protein I716F Mutation. *Journal of Neuropathology & Experimental Neurology* 69, 53-59.

Hin, N., Newman, M., Pederson, S.M., and Lardelli, M.M. (2020). Iron Responsive Element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease. *bioRxiv*, 2020.2005.2001.071498.

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.

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- β CTF (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.

Klaus, B., and Reisenauer, S. (2018). An end to end workflow for differential gene expression using Affymetrix microarrays.

Mehla, J., Lacoursiere, S.G., Lapointe, V., McNaughton, B.L., Sutherland, R.J., McDonald, R.J., and Mohajerani, M.H. (2019). Age-dependent behavioral and biochemical characterization of single APP knock-in mouse (APPNL-G-F/NL-G-F) model of Alzheimer's disease. *Neurobiology of Aging* 75, 25-37.

Oddo, S., Caccamo, A., Shepherd, J.D., Murphy, M.P., Golde, T.E., Kaye, 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, 409-421.

Oosterhof, N., Holtman, I.R., Kuil, L.E., van der Linde, H.C., Boddeke, E.W.G.M., Eggen, B.J.L., and van Ham, T.J. (2017). Identification of a conserved and acute neurodegeneration-specific microglial transcriptome in the zebrafish. *Glia* 65, 138-149.

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.

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.

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

Serneels, L., T'Syen, D., Perez-Benito, L., Theys, T., Holt, M.G., and De Strooper, B. (2020). Modeling the β -secretase cleavage site and humanizing amyloid-beta precursor protein in rat and mouse to study Alzheimer's disease. *Molecular Neurodegeneration* 15, 60.

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., and Mesirov, J.P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* 102, 15545.

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

Whyte, L.S., Hassiotis, S., Hattersley, K.J., Hemsley, K.M., Hopwood, J.J., Lau, A.A., and Sargeant, T.J. (2020). Lysosomal Dysregulation in the Murine AppNL-G-F/NL-G-F Model of Alzheimer's Disease. *Neuroscience* 429, 143-155.

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* 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.

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.

Supplemental File 3: APOE-TR RNA-seq re-analysis

Possession of the $\epsilon 4$ allele of apolipoprotein E (*APOE*) is the greatest genetic risk factor for development of sporadic, late-onset Alzheimer's disease (LOAD).

Homozygosity for the $\epsilon 4$ allele increases an individual's risk at least 9-fold relative to the common, $\epsilon 3$ allele (depending on sex and ethnicity) (Farrer et al., 1997).

There is also an $\epsilon 2$ allele of *APOE*, which is protective against development of AD (Farrer et al., 1997). To understand the effects of the different alleles of *APOE*, targeted replacement mice have been developed expressing humanised $\epsilon 2$, $\epsilon 3$, or $\epsilon 4$ alleles from the mouse *Apoe* locus (APOE-TR) (Sullivan et al., 1997).

Zhao et al. (2020) performed an RNA-seq experiment investigating the effect of *APOE* genotype, age (3, 12 and 24 months of age) and sex in APOE-TR homozygous mice. In that analysis, pairwise comparisons between the $\epsilon 2$, or $\epsilon 4$ alleles relative to the $\epsilon 3$ 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 was reported. We are interested in the cellular processes implicated as affected by AD-causative mutations in young brains.

Therefore, here we will re-analyse the RNA-seq dataset from Zhao et al. to ask which processes are affected by homozygosity for the $\epsilon 2$, or $\epsilon 4$ alleles relative to the $\epsilon 3$ allele of *APOE* in the brains of three and twelve month old mice.

Throughout this analysis, we refer to these homozygous mice as "APOE2", "APOE3" and "APOE4".

RNA-seq data processing

We first 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. The gene expression counts matrix was generated 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* (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. The effect of filtering lowly expressed genes is found in **Figure 1** below.

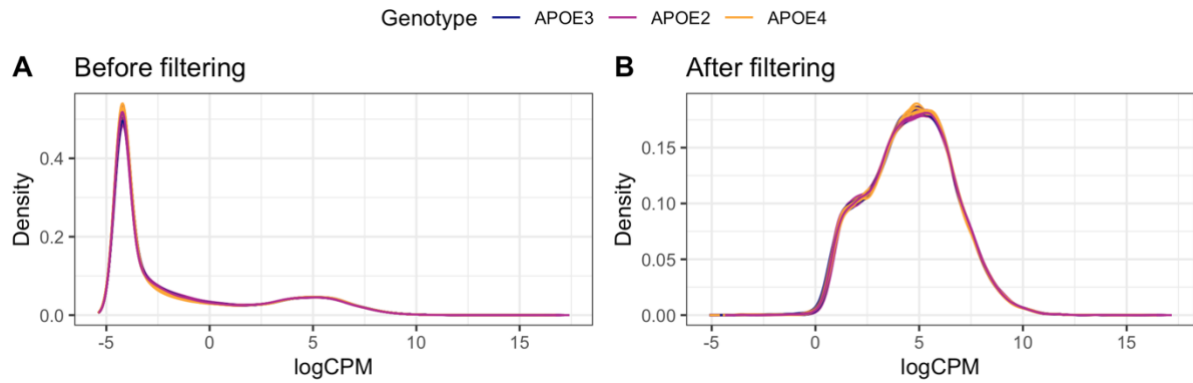


Figure 1: The density of the log counts per million (logCPM) values detected in 3 month old mouse brain samples is shown before filtering in **A**, then after omitting samples with a logCPM of < 2 in at least one third of the samples in **B**.

Assessment of expression of genes from the Y-chromosome to confirm sex.

We then confirmed the sex of each mouse by examination of the expression of genes from the Y-chromosome. 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 (**Figure 2**). This was subsequently corrected for the rest of the analysis.

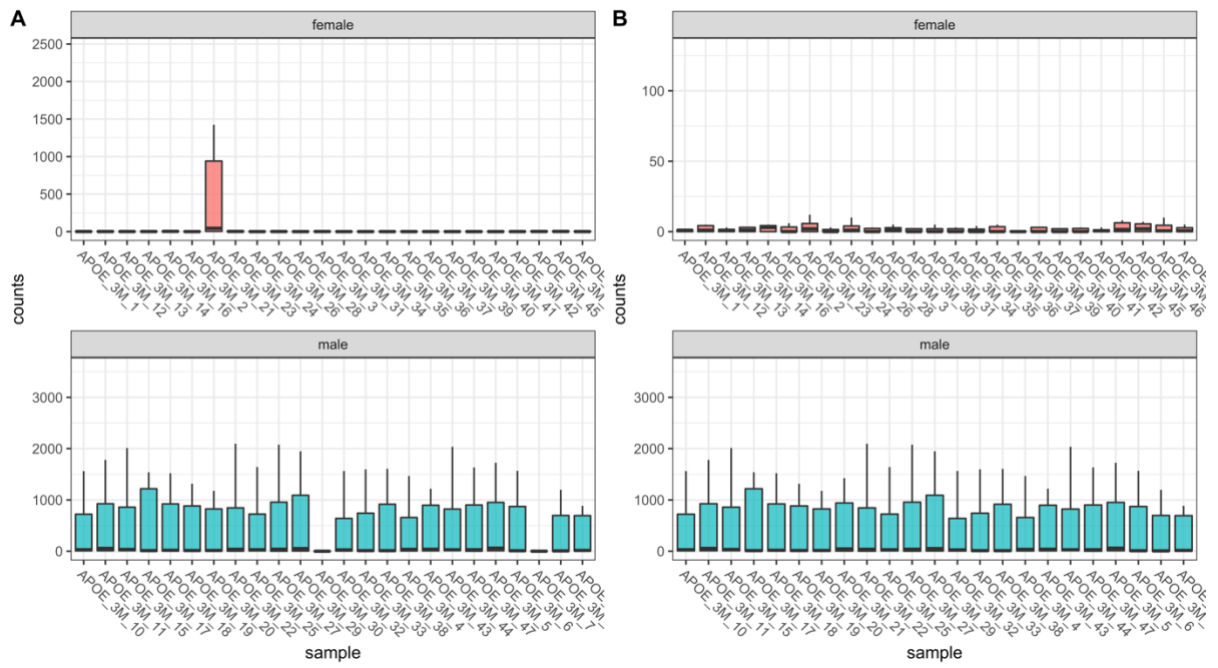


Figure 2: A) Boxplots showing the distribution of expression of male-specific genes (located on the Y-chromosome) in the cortex of 3 month old APOE-TR mice, grouped by initial sex. **B)** Expression of male-specific genes after correcting the sex of the samples.

Principal component analysis

We next performed principal component analysis (PCA) to explore whether *APOE* genotype results in stark differences to the brain transcriptome at 3 months of age. A plot of principal component 1 (PC1) against principal component 2 (PC2) revealed that samples separated into two distinct clusters of sex across PC2 (**Figure 3A**). 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 distinct cluster 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 arise mostly from single litters, as implied from the date-of-birth of each sample (**Figure 3B**).

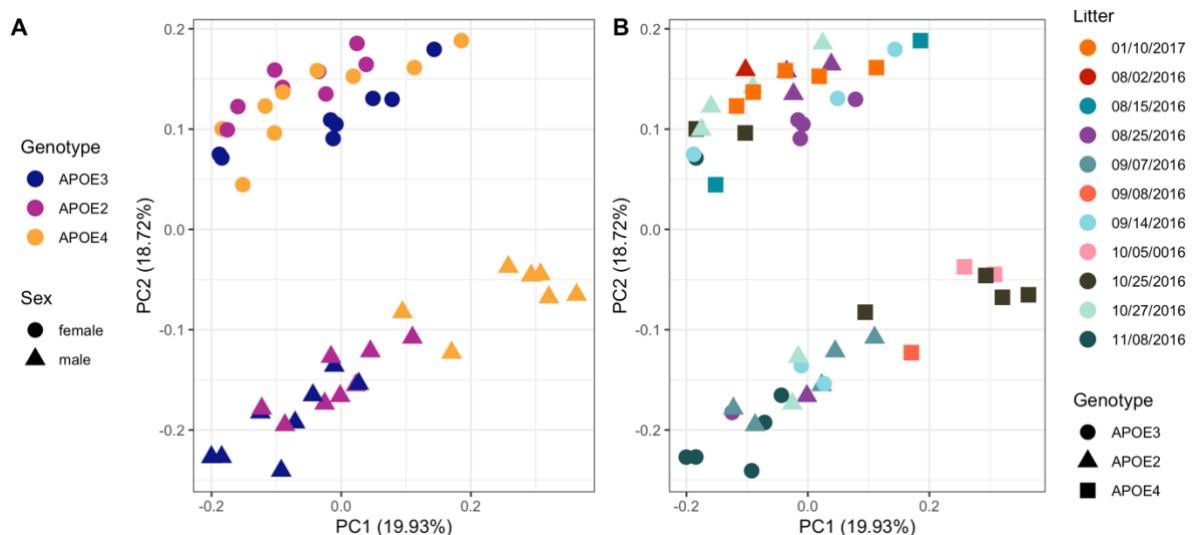


Figure 3: Principal component analysis. **A** shows principal component 1 (PC1) against PC2. The numbers between parentheses indicate the percentage of variation in the dataset that PC explains. Each point represents a sample, which is coloured by *APOE* genotype, and shaped by sex. **B** also shows PC1 against PC2. However, each point is coloured by litter (implied from the date-of-birth of each mouse), and shaped by *APOE* genotype.

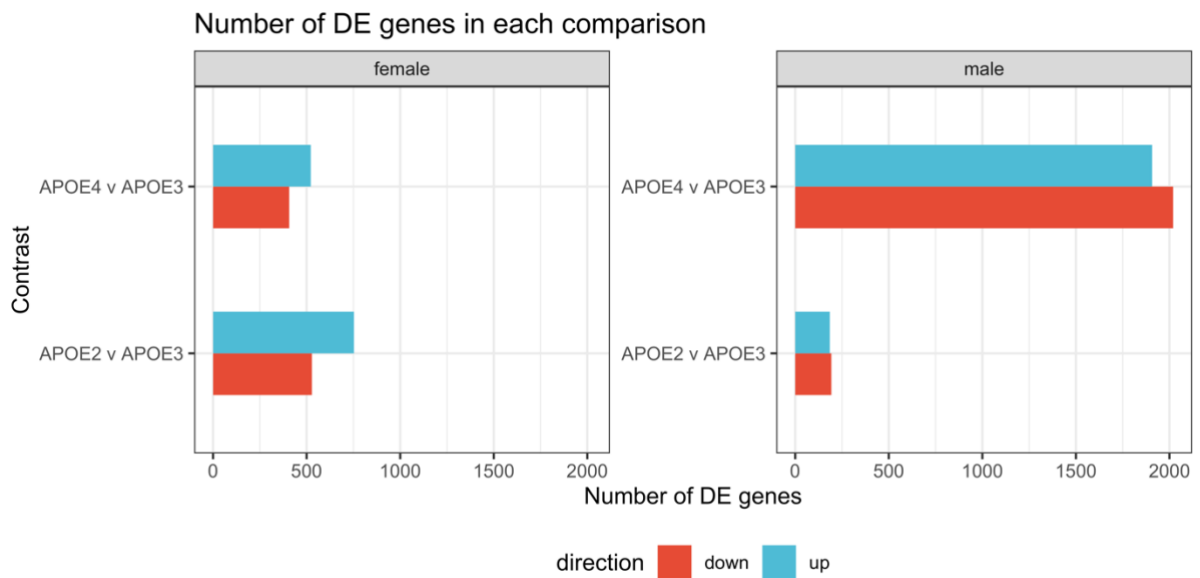
From **Figure 3B**, we observed that some experimental groups of samples (i.e. the male *APOE3* and male *APOE4* samples) appeared to arise mostly from single litters of mice. This is confounding with the effect of genotype and complicates interpretation of whether any effects observed in a pairwise contrast between male *APOE3* and *APOE4* mice are due to *APOE* genotype or litter-of-origin (or, most likely, both). Indeed, χ^2 tests for independence revealed that there is a highly significant dependence of *APOE* genotype and litter across the entire 3-month-old dataset ($\chi^2= 82.7$, $df = 20$, $p\text{-value} = 1.4\text{e-}09$), only within male samples ($\chi^2= 43.1$, $df = 14$, $p\text{-value} = 8.2\text{e-}05$) and only within female samples ($\chi^2= 39.3$, $df = 14$, $p\text{-value} = 3.0\text{e-}04$). 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 assuming that the effect of litter is negligible.

Initial differential gene expression analysis

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 both males and females. In this analysis, we considered a gene to be differentially expressed (DE) if the FDR-adjusted p-value was < 0.05 . Many genes were found to be DE in each comparison, particularly in male *APOE4* mice (**Figure 4A**). Additionally, the bias for GC content and gene length for differential expression noted by Zhao et al. in the original analysis was also apparent (**Figure 4A, B**).

A



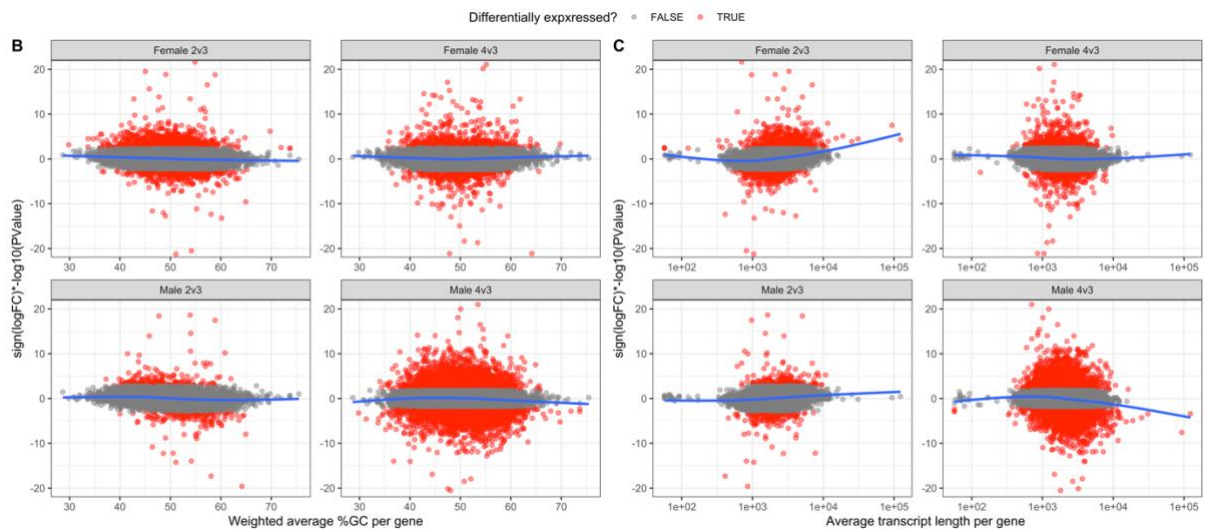


Figure 4: Initial differential expression analysis. A) Bar chart showing the number of differentially expressed genes (DE) in each comparison of the APOE4 or APOE2 genotype to APOE3. **B)** A ranking statistic per gene was calculated as the sign of the logFC multiplied by the negative log10 of the p-value 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.

Conditional quantile normalisation

A gene length bias for differential expression has been shown to influence the results of gene set enrichment analyses (Mandelbroum et al., 2019). Therefore, to correct for the observed bias for gene length (and GC content), we used 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. In these tests, many genes were identified as DE and the bias for %GC and gene

length had improved. A gene length bias was still present in male APOE4 and female APOE2 comparisons to APOE3. However, these were only observed due to a small number of long gene transcripts and likely can be ignored.

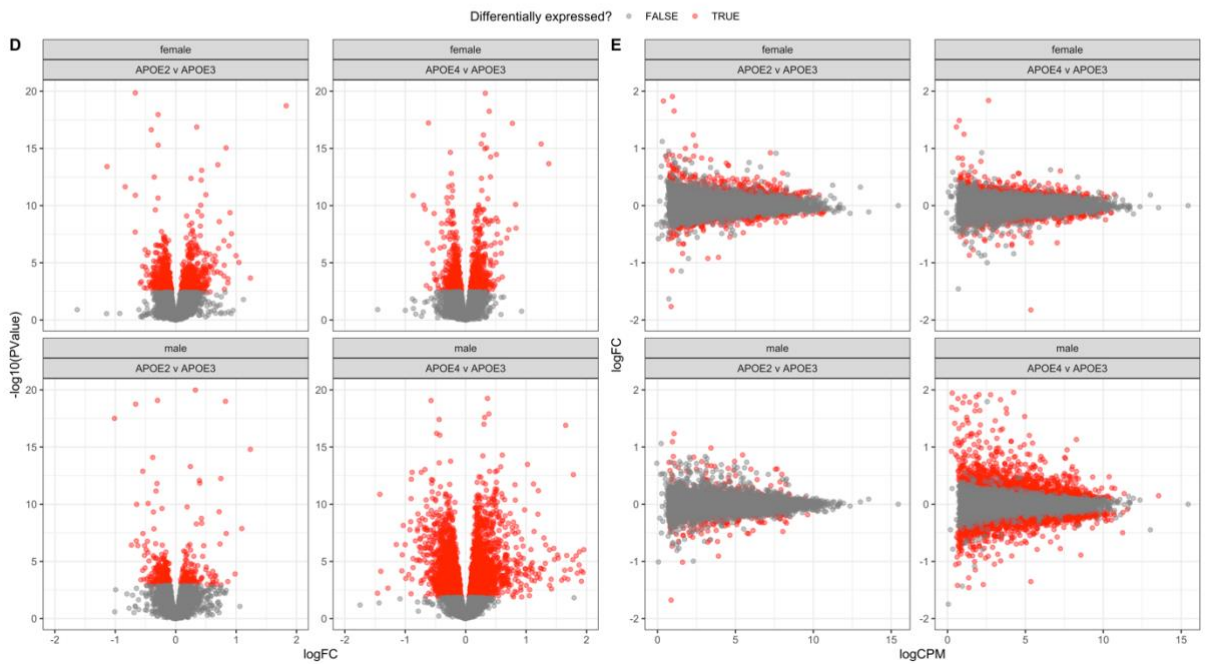
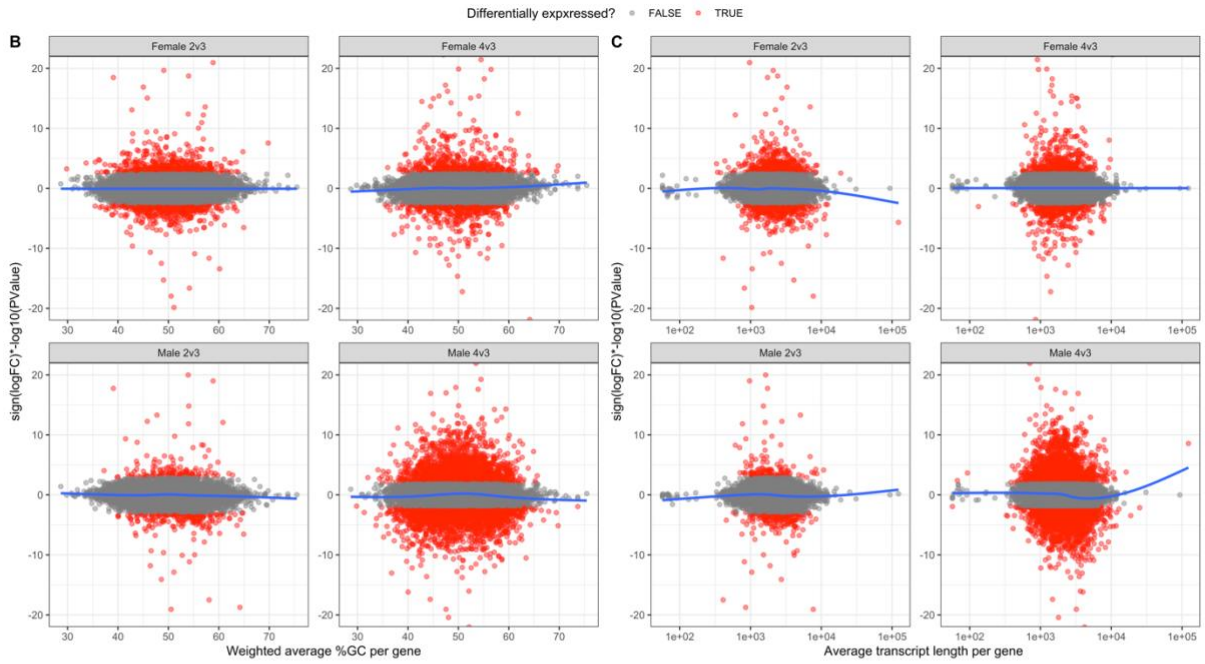
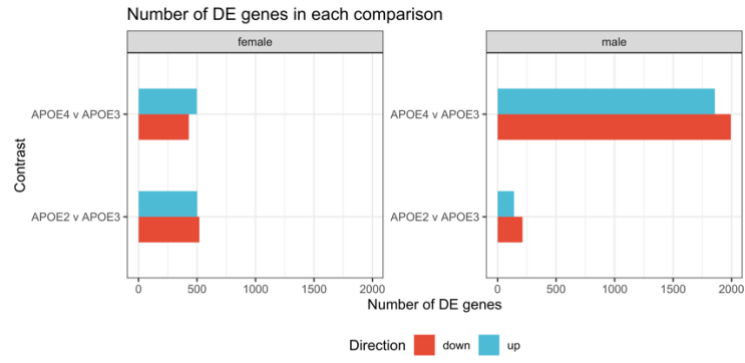
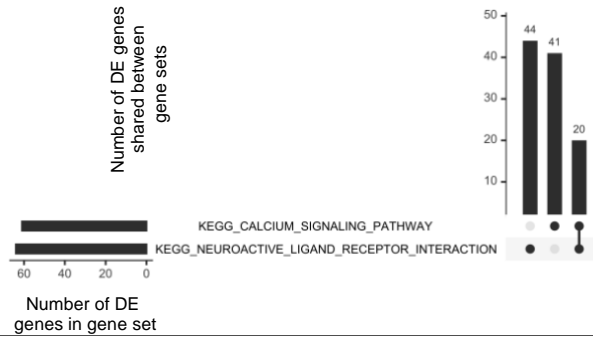
A

Figure 5: 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. **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.

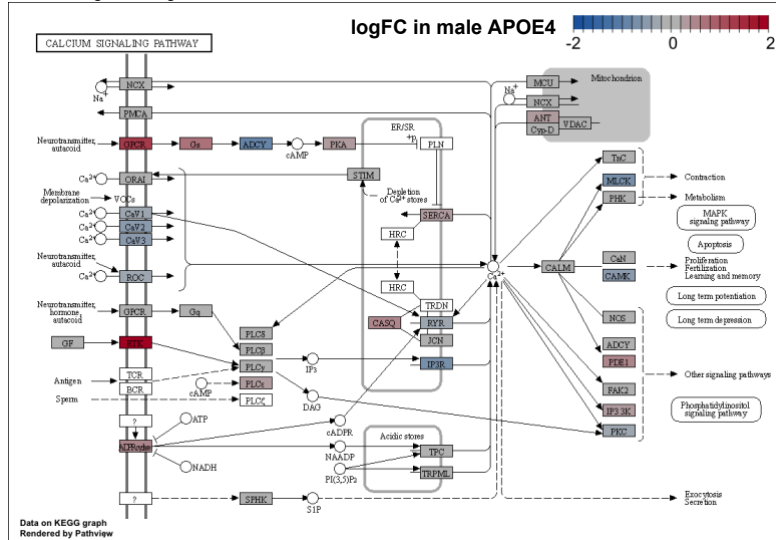
Enrichment analysis

We next tested for over-representation of the *KEGG* (Kanehisa and Goto, 2000) and *IRE* (Hin et al., 2020) gene sets within the DE gene lists using *goseq* (Young et al., 2010), using the average transcript length per gene to calculate the probability weighting function (PWF). After correction for multiple testing using the FDR, we observed a total of three gene sets to be significantly enriched in any of the DE gene lists. Two of these *KEGG* gene sets were enriched in the DE genes for male APOE4 mice: *KEGG_CALCIIUM_SIGNALING_PATHWAY* ($p_{\text{FDR}} = 0.03$) and *KEGG_NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION* ($p_{\text{FDR}} = 0.04$). Approximately one third of the DE genes in each of these gene sets are shared, indicating that the enrichments of these gene sets are driven partially by the same genes (**Figure 6**).

A



B



C

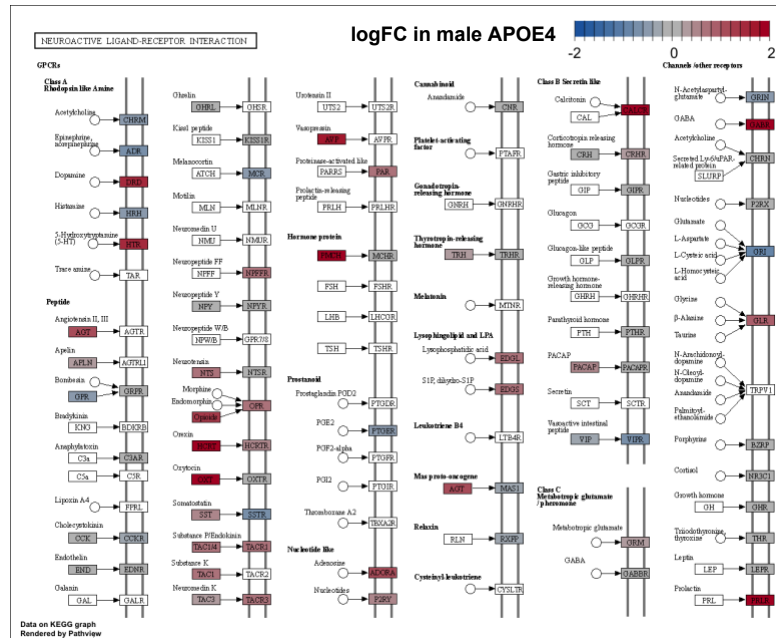


Figure 6: Enrichment analysis within the lists of differentially expressed genes. **A)** Upset plot indicating the overlap of DE genes in male APOE4 samples for the significantly enriched gene sets. **B)** Pathview (Luo et al., 2017) visualisation of the logFC in male APOE4 samples for the *KEGG_CALCIIUM_SIGNALING_PATHWAY* gene set and **C)** *KEGG_NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION* gene set.

Additionally, the gene set *KEGG_STEROID_BIOSYNTHESIS* was found to be significantly enriched among the female APOE2 DE genes ($p_{FDR} = 1e-5$). This gene set appears to be downregulated (**Figure 7**).

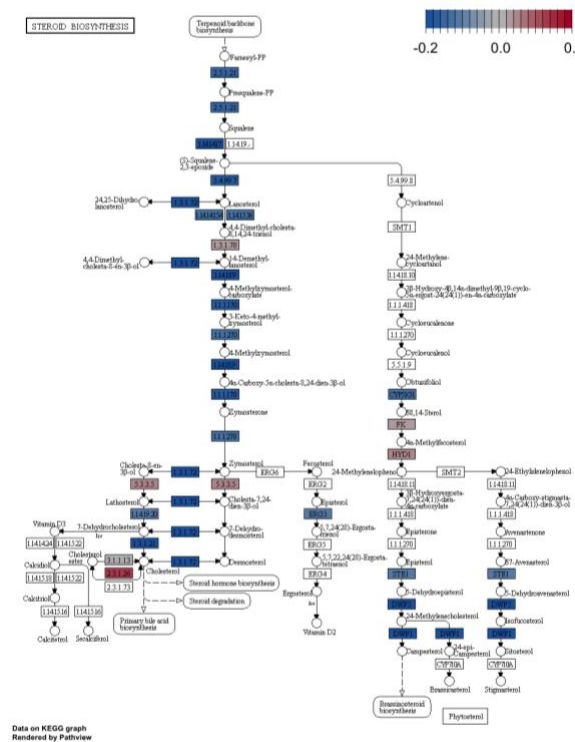
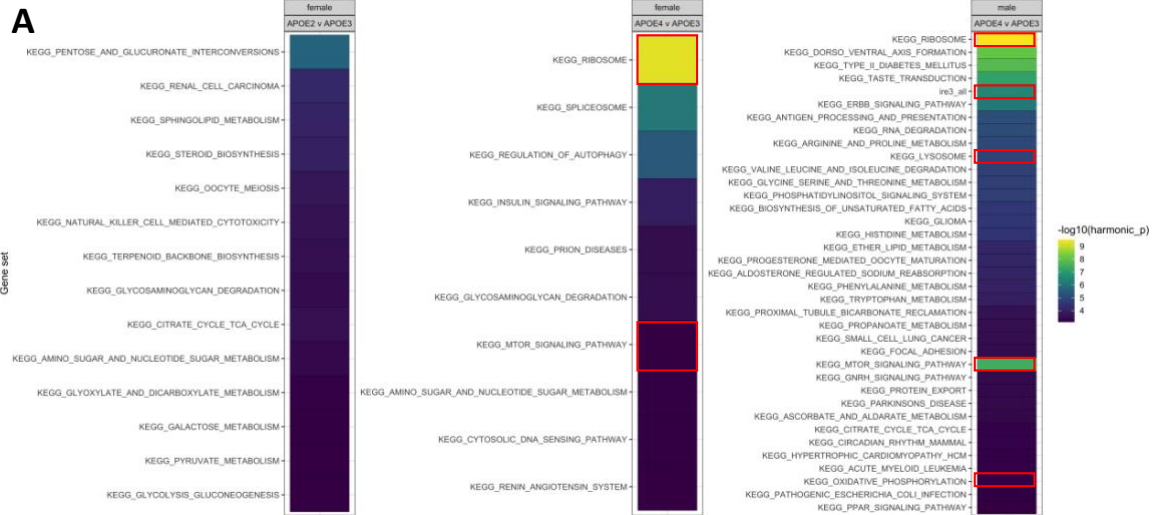


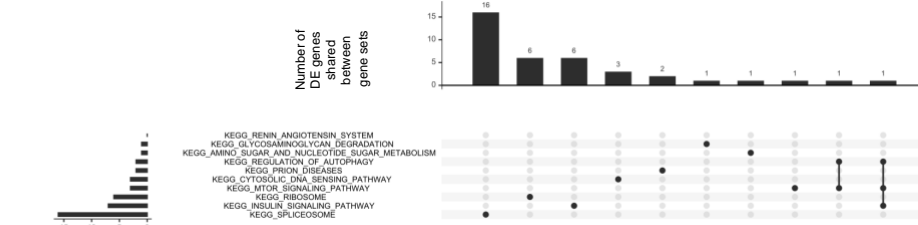
Figure 7: Pathview visualisation indicating the logFC of genes in the *KEGG_STEROID_BIOSYNTHESIS* gene set in female APOE2 mice.

Over-representation analysis using *goseq* relies on a hard threshold for a gene to be classified as DE. Therefore, information may be missed, i.e. genes which are highly ranked in terms of differential expression, but do not reach the threshold of

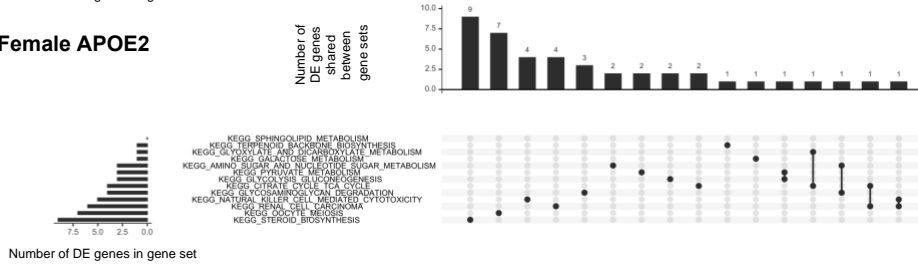
FDR < 0.05. Therefore, to obtain a more complete view of the changes to gene expression in APOE4 or APOE2 mice relative to APOE3 mice, we performed enrichment analysis on the entire list of detectable genes using three methods of rank-based gene set enrichment analysis with different statistical methodologies: *fry* (Wu et al., 2010), *camera* (Wu and Smyth, 2012) and *GSEA* (Subramanian et al., 2005) (implemented in the *fgsea* R package (Sergushichev, 2016)). We then combined raw p-values from each method to obtain an overall significance value by calculation of the harmonic mean p-value (Wilson, 2019) (a method which has been shown specifically to be robust to combining dependent p-values). After FDR adjustment of the harmonic mean p-value, we observed 73 gene sets to be significantly altered (FDR adjusted harmonic mean p-value < 0.05) in male APOE4 mice, 5 gene sets in male APOE2 mice, 30 gene sets in female APOE4 mice, and 30 gene sets in female APOE2 mice. This appears to be a considerable number of significantly altered gene sets, especially for young mice carrying mutations associated with LOAD. However, these effects are likely driven by both *APOE* genotype and litter-of-origin (**Figure 5**). Therefore, to simplify interpretation, we considered a gene set to be altered significantly if the FDR-adjusted harmonic mean p-value was < 0.01, which represents the most significantly altered pathways (as these are likely to have the largest effect of genotype). The statistical significance values of the gene sets are mostly driven by distinct expression signals, as shown by the lack of overlap of the DE genes found in the significantly altered gene sets (with the exception of the *KEGG* gene sets for oxidative phosphorylation and Parkinson's disease, which share 14 DE genes).



B Female APOE4



Female APOE2



Male APOE4

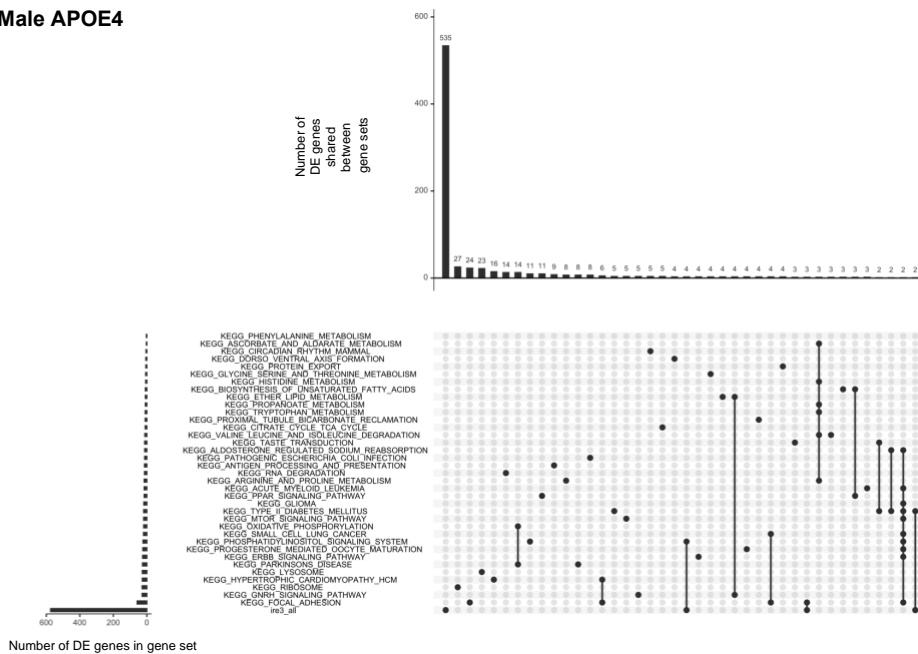


Figure 8: Ranked-list gene set enrichment testing. **A)** Heatmap indicating gene sets with a FDR-adjusted harmonic mean p-value < 0.01 in APOE-TR mice at 3 months of age. Gene sets of interest are highlighted with a red box. Note that no gene sets were found to contain an FDR-adjusted harmonic mean p-value of < 0.01 in male APOE2 mice. **B)** 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 APOE-TR mice.

As described in the main text, the *KEGG_OXIDATIVE_PHOSPHORYLATION* gene set was the only gene set to be affected by all the EOfAD-like mutations in our zebrafish models and not by non AD-relevant mutations. In APOE-TR mice, statistical evidence was observed for this gene set to be altered in male APOE4 mice (**Table 1**). The overall direction of logFC of genes in the *KEGG_OXIDATIVE_PHOSPHORYLATION* gene set was up in the male APOE4 mice (although some genes of the gene set were downregulated) (**Figure 9**).

Table 1: 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

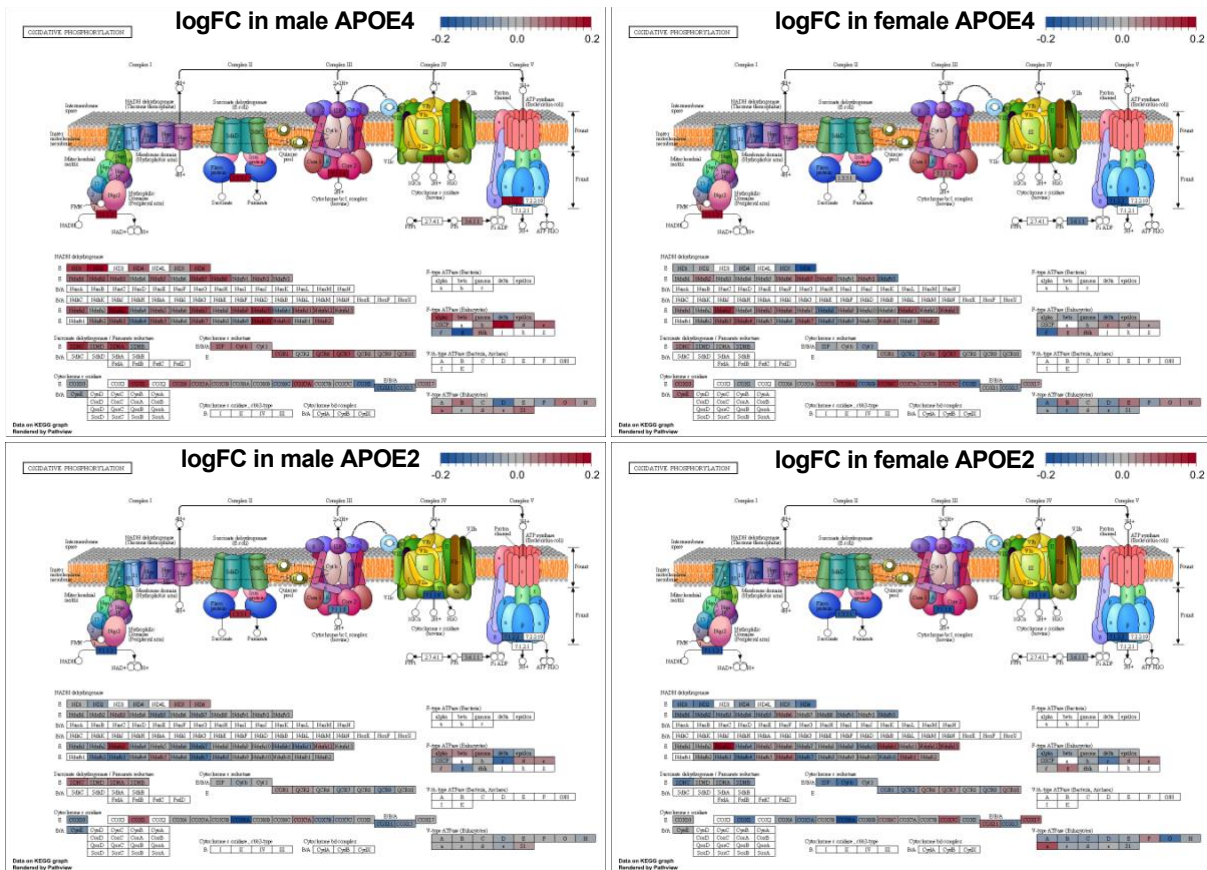


Figure 9: Changes to the *KEGG_OXIDATIVE_PHOSPHORYLATION* gene set in young APOE-TR mice.

Additionally, the *KEGG_RIBOSOME* gene set was found to be altered in common by EOfAD-like (and non-EOfAD-like) mutations in our zebrafish models. This gene set was found to be highly significantly altered in both male and female APOE4 mice and not in APOE2 mice (**Figure 8A**). The genes in this gene set appear to be mostly upregulated in APOE4 mice, and are both upregulated or downregulated in APOE2 mice.

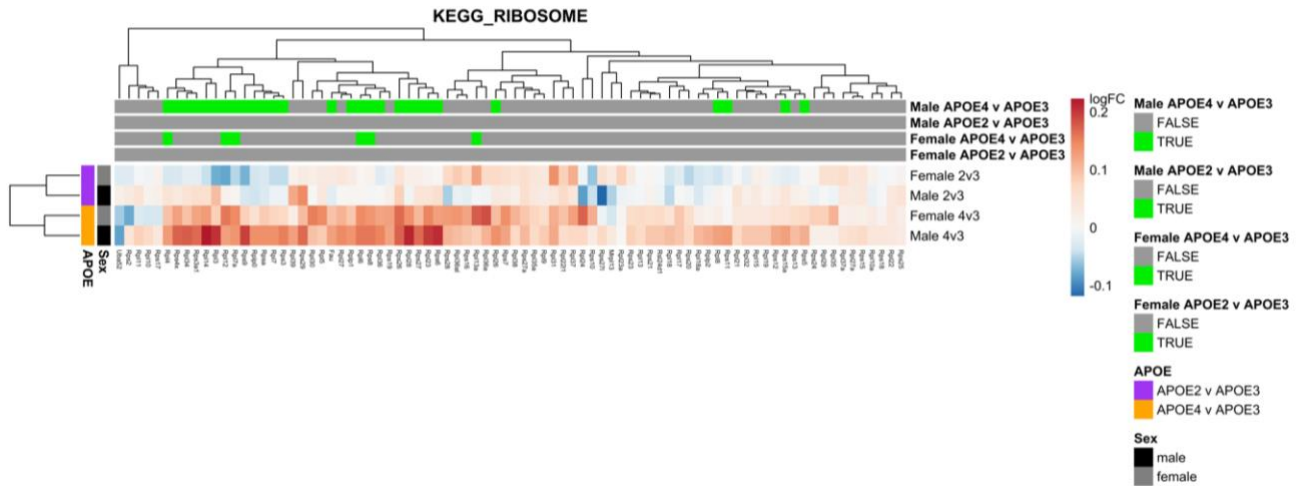


Figure 10: The columns in the heatmap represent a gene in the *KEGG_RIBOSOME* gene set, while the rows indicate the comparisons between APOE-TR mice in the differential expression analysis. The colour of a cell represents the logFC of a particular gene, and the genes are labelled in green above if they were classified as differentially expressed (FDR < 0.05) in the differential gene expression analysis in a particular comparison.

The *KEGG_MTOR_SIGNALING_PATHWAY* gene set was observed to be significantly altered in both male and female APOE4 mice. No clear direction of change is evident, as both up- and down-regulation of genes in this pathway is observed (**Figure 11**).

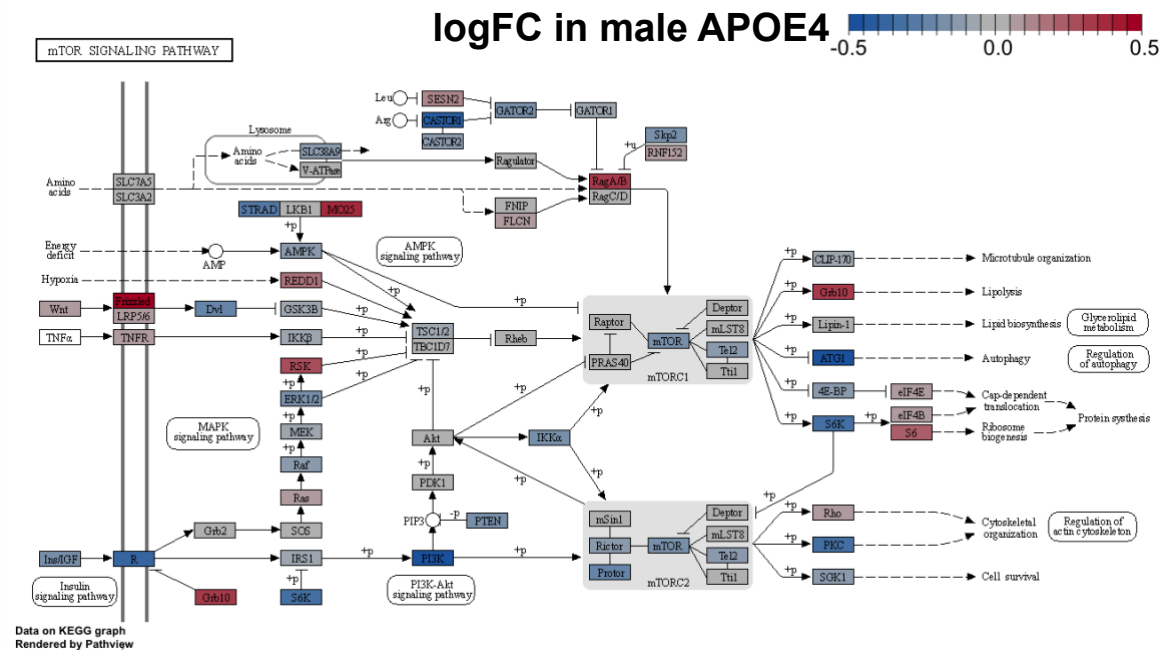
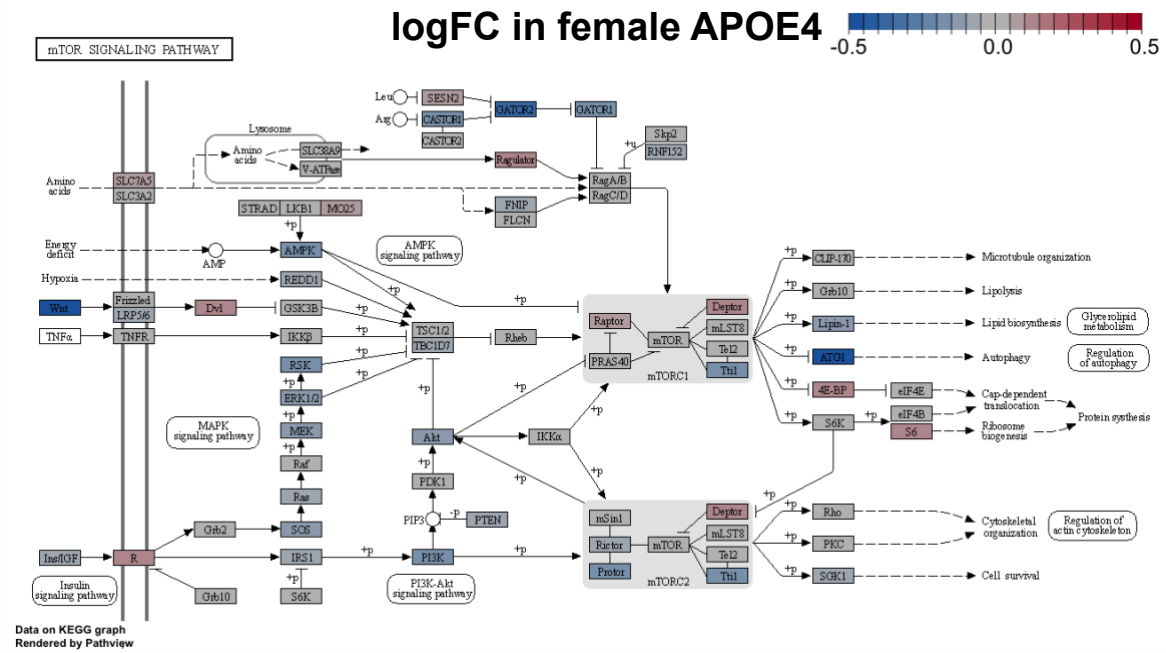


Figure 11: Changes to the *KEGG_MTOR_SIGNALING_PATHWAY* gene set in young APOE4 mice.

Cell type proportions

We next assessed whether changes to gene expression observed in young APOE-TR mice are due to changes to cell type proportions as described in

Supplementary File 2. We observed a slight decrease in expression values of neuronal marker genes in female APOE2 mice, and increased expression of marker genes of oligodendrocytes and astrocytes in male APOE4 mice, suggesting that the overall changes to gene expression observed in these mice may be driven partially by differences in cell type proportions (**Figure 12**).

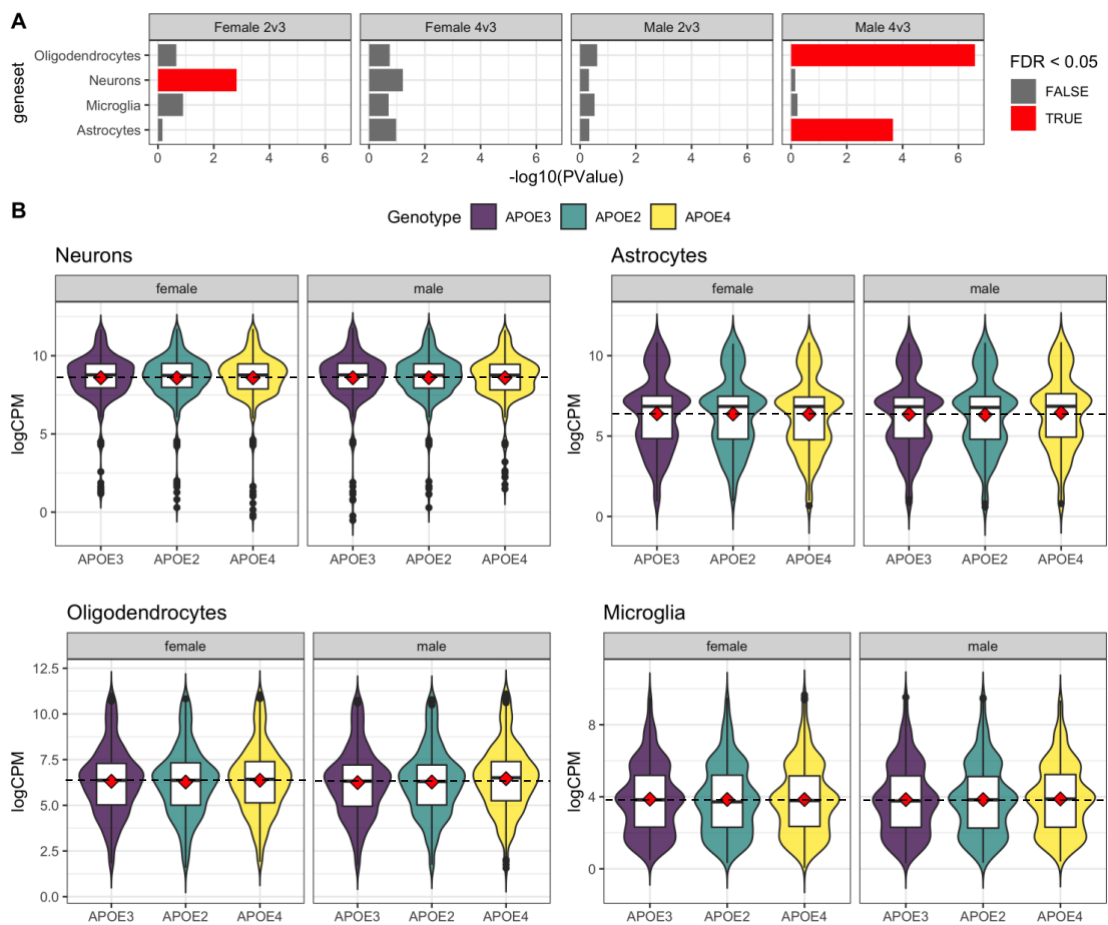


Figure 12: A) Significance of gene set testing from *fry* with a directional hypothesis of gene sets consisting of marker genes of neurons, astrocytes and oligodendrocytes from (Cahoy et al., 2008) and microglia from (Oosterhof et al., 2017). **B)** Distribution of expression (logCPM) of these cell type marker genes in APOE-TR mice. Boxplots show the summary statistics, while the violin plots summarise the density of logCPM expression values. The mean of each experimental group is shown by the red diamonds, and the mean of the APOE3 expression values is shown by black dashed lines.

Discussion and conclusion

In conclusion, our re-analysis of the comprehensive study of Zhao et al. (2020) predicts many cellular processes to be affected in young APOE-TR mice.

However, sources of external variation may be contributing to the observed effects (i.e. litter-of-origin, changes to cell type proportions and possibly gene length). Importantly, the specific changes to gene expression observed in our zebrafish models of EOfAD are observed to be altered in APOE4 mice

(*KEGG_RIBOSOME* in both males and females, and

KEGG_OXIDATIVE_PHOSPHORYLATION only in males). However, replication

of this study is desirable, particularly with modifications to compensate for the

litter-of-origin confounding effect, to confirm that these effects are due to *APOE* genotype.

References

Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A., *et al.* (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 28, 264-278.

- 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* 29, 15-21.
- 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, 1349-1356.
- 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.
- Hin, N., Newman, M., Pederson, S.M., and Lardelli, M.M. (2020). Iron Responsive Element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease. *bioRxiv*, 2020.2005.2001.071498.
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* 28, 27-30.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923-930.
- 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.
- Mandelbom, S., Manber, Z., Elroy-Stein, O., and Elkon, R. (2019). Recurrent functional misinterpretation of RNA-seq data caused by sample-specific gene length bias. *PLoS Biol* 17, e3000481-e3000481.
- 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.
- Oosterhof, N., Holtman, I.R., Kuil, L.E., van der Linde, H.C., Boddeke, E.W.G.M., Eggen, B.J.L., and van Ham, T.J. (2017). Identification of a conserved and acute neurodegeneration-specific microglial transcriptome in the zebrafish. *Glia* 65, 138-149.
- 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* 26, 139-140.
- 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.

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., and Mesirov, J.P. (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.

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

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* 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.

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.e726.

Supplemental File 4: Additional Discussion

Energy production changes in zebrafish and mouse models of genetic variation driving AD

The majority of heterozygous EOfAD-like mutations we have studied in zebrafish show overall (majority) downregulation of the oxidative phosphorylation gene set in young adult brains relative to brains of the wild type genotype. 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 and so it was not included in this paper. Transheterozygosity for mutations in *sor11* 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. 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 changes in water quality, microbiome, handling etc.. Also, 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 sensitivity for detection of statistically significant differences.)

Importantly, male mice homozygous for the late onset AD (LOAD) risk allele, APOE4, 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. This demonstrates the similarity, at the molecular level, between the cellular effects of genetic variants causing EOfAD and promoting LOAD, and supports the validity of analysing early molecular events in AD pathogenesis using knock-in models in zebrafish. By extension, our results also support that 'omics analyses of the brains of mouse knock-in models of EOfAD mutations will yield valuable information on AD pathogenesis. Although numerous such models were created 15-20 years ago, and some showed subtle cognitive effects (Guo et al., 1999; Kawasumi et al., 2004), their relevance to understanding AD was apparently dismissed since they failed to reproduce the A β plaque and neurofibrillary tangle phenotypes of the human disease. Consequently, the molecular state of their brains has never been analysed in detail.

In humans, it is thought that A β -related abnormalities precede metabolic defects (Jack et al., 2013; Leclerc and Abulrob, 2013). However, metabolic abnormalities are generally measured *in vivo* using FDG-PET imaging, which is likely not sensitive enough to detect subtle changes. Nevertheless, reduced glucose uptake has been observed by FDG-PET in the brains of living subjects before the onset of dementia (Mosconi et al., 2008; Mosconi et al., 2009). Additionally, changes to the expression of genes in the oxidative phosphorylation pathway has been observed 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). While neurons derived from a patient carrying the *PSEN1*^{S170F} EOfAD mutation also show mitochondrial abnormalities (Li et al., 2020). Together, these results support that changes to mitochondrial function are an early AD pathology.

Our findings regarding EOfAD mutations in zebrafish were not consistent with findings from our analysis of transcriptome data from in 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 for the A β region and three EOfAD mutations), motivated by the idea that the more aggregation-prone human A β 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 this model at a younger age. However, the expression of genes involved in lysosomal function (*KEGG_LYSOSOME*) was observed to be highly significantly upregulated in the homozygous *App*^{NL-G-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 product of APP (Jiang et al., 2019)). Increased β -CTF has been observed in the brains of *App*^{NL-G-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

(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)), mitochondrial dysfunction and inflammation (Yambire et al., 2019). Additionally, the *App^{NL-G-F}* mouse model shows increased levels of A β from a young age (Saito et al., 2014), and the deposition of A β into plaques was shown to be associated with the increased expression of genes in the complement system in the comprehensive spatial transcriptomics study with aging in the *App^{NL-G-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-G-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-G-F}* mice to those in our zebrafish models and in APOE4 mice. However, their subtle signs 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 sex and litter-of-origin effects).

mTOR signalling can regulate ribosomal gene set expression

In the majority of our zebrafish mutants (all except for zebrafish transheterozygous for EOfAD-like mutations in *sort1*) and in the APOE4 mouse data we have observed changes to the expression of the set of genes encoding components of the ribosomal subunits. 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 cellular processes which are also implicated in AD pathogenesis, such as autophagy and metabolism (reviewed in (Saxton and Sabatini, 2017)).

The mTOR proteins reside 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 A β (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. (2019) observed that inhibition of mTOR signalling (specifically, mTORC1) restores auto- and mito-phagy defects in fibroblasts from Down's syndrome individuals. 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 *sor11* 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 we could only detect significant changes in the transcriptional response to altered mTOR signalling in the other mutants.

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

In a highly sensitive analysis method such as RNA-seq, external sources of variation must be minimised. 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 sources of 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 small litters of 5-10 pups, making it particularly difficult to obtain sufficient samples which are synchronous siblings (particularly when a genotype of interest is homozygous). 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 of whether *App*^{NL-G-F} mice were littermates was not available, and this required us to assume that the effect of litter was 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 difference due to sex, while this has a negligible effect on bulk brain transcriptomes from zebrafish (Barthelson et al., 2020a; Barthelson et al., 2020b; Barthelson et al., 2020c), 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-G-F} mice, a

phenomenon that can create the artefactual appearance of gene expression change. We have not observed cell-type proportion differences in young or “middle aged” (24 month old) zebrafish (Barthelson et al., 2021; Barthelson et al., 2020a; Barthelson et al., 2020b; Barthelson et al., 2020c; Hin et al., 2020), 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 from underwent a 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 A β , 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 A β is causative of AD pathology or a consequence of it continues as a matter of debate in the AD research community (reviewed in (Morris et al., 2018)). If zebrafish cannot produce A β , then the changes we have observed in the brains of our zebrafish models may illuminate A β -independent effects of EOfAD mutations.

References

Barthelson, K., Dong, Y., Newman, M., and Lardelli, M. (2021). 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. *bioRxiv*, 2021.2001.2026.428321.

Barthelson, K., Pederson, S., Newman, M., and Lardelli, M. (2020a). Transcriptome analysis of a protein-truncating mutation in sortilin-related receptor 1 associated with early-onset familial Alzheimer's disease indicates effects on mitochondrial and ribosome function in young-adult zebrafish brains. *bioRxiv*, 2020.2009.2003.282277.

Barthelson, K., Pederson, S.M., Newman, M., Jiang, H., and Lardelli, M. (2020b). Frameshift and frame-preserving mutations in zebrafish presenilin 2 affect different cellular functions in young adult brains. *bioRxiv*, 2020.2011.2021.392761.

Barthelson, K., Pederson, S.M., Newman, M., and Lardelli, M. (2020c). 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.

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 research* 27, 121-130.

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. *Biochem J* 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. *The Journal of biological chemistry* 285, 13107-13120.

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.e919.

Contino, S., Suelves, N., Vrancx, C., Vadukul, D.M., Payen, V.L., Stanga, S., Bertrand, L., and Kienlen-Campard, P. (2021). Presenilin-Deficient Neurons and Astrocytes Display Normal Mitochondrial Phenotypes. *Frontiers in Neuroscience* 14, 1419.

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, 101-106.

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.* (2020). Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2. *PLOS ONE* 15, e0227258.

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). HIF α 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., Knopman, D.S., Jagust, W.J., Petersen, R.C., Weiner, M.W., Aisen, P.S., Shaw, L.M., Vemuri, P., Wiste, H.J., Weigand, S.D., *et al.* (2013). Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *The Lancet Neurology* 12, 207-216.

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- β CTF (C99). *The Journal of Neuroscience* 39, 5255.

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. *The European journal of neuroscience* 19, 2826-2838.

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.

Leclerc, B., and Abulrob, A. (2013). Perspectives in Molecular Imaging Using Staging Biomarkers and Immunotherapies in Alzheimer's Disease. *The Scientific World Journal* 2013, 589308.

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, 1146-1158.

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.

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.

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.

Meyer, A., and Van de Peer, Y. (2005). From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *BioEssays : news and reviews in molecular, cellular and developmental biology* 27, 937-945.

Moore, D.B., Gillentine, M.A., Botezatu, N.M., Wilson, K.A., Benson, A.E., and Langeland, J.A. (2014). Asynchronous evolutionary origins of A β and BACE1. *Molecular biology and evolution* 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.

Mosconi, L., De Santi, S., Li, J., Tsui, W.H., Li, Y., Boppana, M., Laska, E., Rusinek, H., and de Leon, M.J. (2008). Hippocampal hypometabolism predicts cognitive decline from normal aging. *Neurobiol Aging* 29, 676-692.

Mosconi, L., Mistur, R., Switalski, R., Tsui, W.H., Glodzik, L., Li, Y., Pirraglia, E., De Santi, S., Reisberg, B., Wisniewski, T., and de Leon, M.J. (2009). FDG-PET changes in brain glucose metabolism from normal cognition to pathologically verified Alzheimer's disease. *Eur J Nucl Med Mol Imaging* 36, 811-822.

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.

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.

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

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.

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.

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.

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.

Chapter 10: Final discussion and conclusion

Despite over 100 years of research into the pathological mechanisms of AD, effective therapeutics remain to be discovered. Once symptoms of AD begin to show, damage to the brain is considerable and likely cannot be reversed. Therefore, a deep understanding of the early cellular/molecular stresses on brains destined to suffer AD, occurring decades before symptom onset, is required in order to develop preventative treatments. This idea has formed the basis for the work presented in this thesis, where I have introduced into zebrafish, and analysed, the effects of EOfAD-like mutations in their brains. These effects were assessed mostly at 6 months of age, equivalent to recently sexually mature humans when symptom onset is still decades away. The majority of the work presented in this thesis assessed the genome-wide effects of these mutations on bulk brain transcriptomes. Additionally, some non-transcriptomic effects were assessed.

Changes to brain vasculature were not observed in zebrafish *psen1* mutant brains

In **Chapter 3**, I performed an analysis to investigate whether heterozygosity for previously generated *psen1* mutations in zebrafish (*psen1*^{Q96_K97del} and *psen1*^{K97fs}) resulted in any morphological abnormalities in their brain vasculature with aging. If brain vasculature was damaged as a result of the *psen1* mutations, delivery of oxygen and nutrients to the brain could be impaired. This could explain the apparent hypoxia transcriptional response previously observed due to heterozygosity for these mutations in 6 month old zebrafish brains under normoxia [1], and would provide evidence supporting the vascular hypothesis of AD pathogenesis [2]. No statistically significant

differences between *psen1* mutant genotypes were observed in the forebrains of young adult or aged fish, and vessels appeared morphologically similar in each *psen1* genotype. In the manuscript presented in **Chapter 3** [3], I proposed two other mechanisms which could explain the apparent hypoxic response: changes to γ -secretase activity or intracellular ferrous iron deficiency.

Hypoxia-inducible factor 1, alpha subunit (HIF1- α) interacts with PSEN1 protein [4], and interaction of HIF1- α and the γ -secretase complex (contains PSEN1) increases γ -secretase activity [5], resulting in a feed-forward mechanism of increased HIF1- α stability involving p75^{NTR} [6]. Since PSEN1 is the catalytic core of the γ -secretase complex, it is assumed that the *psen1*^{Q96_K97del} and *psen1*^{K97fs} mutations disrupt γ -secretase activity in zebrafish. Indeed, changes to the Notch signalling gene set have been observed in aged *psen1*^{K97fs/+} zebrafish brains, and in both young adult and aged (24 month) *psen1*^{Q96_K97del/+} zebrafish brains relative to non-mutant siblings in previous RNA-seq experiments performed by the Alzheimer's Disease Genetics Laboratory (ADGL) [7, 8]. This is suggestive of altered γ -secretase activity. Similarly, the two mutations in *psen1* analysed in **Chapter 8**: T428del (EOfAD-like) and W233fs (familial acne-inversa-like), also appear to alter γ -secretase activity. The T428del and W233fs mutations of *psen1* were initially expected to decrease γ -secretase activity. The human EOfAD mutation modelled by T428del (*PSEN1*^{T440del} [9]), was shown to contain effectively no γ -secretase activity (for cleavage of an APP fragment) in a cell-free assay [10]. The W233fs mutation is predicted to result in a grossly truncated protein without the critical catalytic aspartate residues required for γ -secretase activity. However,

heterozygosity for either of these mutations was observed to increase the expression of the transcriptional targets of Notch signalling pathway, suggestive of *increased* γ -secretase activity. Why? Unfortunately, no antibodies against zebrafish presenilin proteins currently exist (despite multiple attempts by the ADGL to generate these), meaning this is a difficult question to explore using zebrafish. Future exploration of the Q96_K97del and K97fs brain transcriptome datasets should involve the assessment of expression of the downstream targets of Notch signalling (such as performed in the research described in **Chapter 8**) to assess their overall direction of change, to determine whether these mutations may be increasing or decreasing γ -secretase activity. Additionally, work in the ADGL is underway to confirm these apparent increases of γ -secretase activity using molecular, rather than bioinformatic techniques.

In addition to changes in γ -secretase activity, the HIF1-mediated transcriptional response to hypoxia requires not only oxygen, but also ferrous iron [11, 12]. Previous work by the ADGL [8] provided evidence that heterozygosity for each *psen1* mutation may result in iron dyshomeostasis in 6 month old zebrafish brains. This provides support that a deficiency of intracellular ferrous iron, and therefore, stabilisation of the HIF1 complex (through HIF1- α), could be driving the increased basal expression of the hypoxia responsive genes under normoxia.

In summary, we observed no evidence supporting that changes to brain vasculature drive the apparent hypoxic stress in young adult *psen1* mutant brains. However, other evidence from the ADGL support that the apparent

hypoxic stress may be driven partially by iron dyshomeostasis and/or changes to the interaction of the Psen1 and Hif1- α proteins in zebrafish.

The work presented in **Chapter 3** does have limitations regarding its conclusions. Due to financial constraints and microscope availability, I performed this work using confocal laser scanning microscopy. Ideally, these brains should be imaged using light sheet microscopy, which has the benefit of increased resolution and greater fluorescence intensity in deeper focal planes (i.e. on the ventral side of the brain which was opposite to the detector in the confocal imaging). Imaging time is greatly reduced in light sheet relative to confocal microscopy. This would have allowed increased numbers of biological replicates to be analysed, which would increase statistical power to detect the likely subtle morphological differences which may be present in zebrafish modelling the early AD disease state. Additionally, labelling of vessel lumens was recently shown to outperform labelling vessel walls in terms of accurate detection of vessels, and in contrasting between signal and background fluorescence in mouse brains [13]. Finally, inter-family variation makes comparisons between the two *psen1* mutants difficult. An analysis of a family of fish arising from a pair mating of $Tg(fli1:GFP);psen1^{Q96_K97del/+}$ and $Tg(fli1:GFP);psen1^{K97fs/+}$ would give both heterozygous *psen1* mutants and their wild type siblings in the same family (i.e. as in the RNA-seq experiments of **Chapters 8** and **9**), allowing direct comparison between these two mutations without confounding batch effects, and with the added benefit of only imaging wild type brains once. Therefore, future work may include labelling of the vessel lumen in a single family of fish as just described, rather than labelling of endothelial cells (i.e. in vessel walls) by the *fli1::GFP* transgene, and obtaining

of images for 3D reconstruction using light-sheet microscopy to allow more precise comparisons.

Putative evidence for impaired spatial working memory in 12-month-old *psen1* EOfAD-like mutants

The primary pathology observed in AD is cognitive impairment. Therefore, assessment of cognitive abilities in any animal model of AD is an obvious experiment to determine model relevance. In **Chapter 4**, I performed an initial characterisation of whether heterozygosity for the *psen1*^{Q96_K97del}, *sorl1*^{V1482Afs}, or *sorl1*^{W1818*} mutations in zebrafish result in alterations of spatial working memory. To achieve this, I utilised the free movement pattern (FMP) Y-maze task [14], a simple test of spatial working memory. Statistical evidence was observed in the re-analysis of the previously generated raw data for *psen1*^{Q96_K97del/+} mutant fish at 12 months of age. However, this finding was not replicated when analysing a different family with the same *psen1* genotypes. No statistical evidence was observed for behavioural differences for the *sorl1* mutants at 12 months of age. Future work should involve replication of these experiments to confirm whether the observed (or not observed) impairments of spatial working memory are reproducible and, potentially, testing of our other EOfAD-like mutant zebrafish lines. The sample sizes used for this analysis were based on a the previously performed calculation by Cleal and Parker [15], who calculated that n = 14 fish per treatment would provide 80% power to detect statistically significant differences (alpha = 0.05) due to a range of pharmacological treatments (effect size of 1.2). The effect of heterozygosity for EOfAD-like mutations in young adult brains is likely less disruptive to brain function than pharmacological treatments. Therefore, the expected sample size

to detect a smaller effect size would be larger. In **Chapter 4**, statistical analysis was performed using generalised linear mixed effect models, as the type of data generated in this kind of analysis is count data, and both fixed and random effects were included in the statistical model. For these types of models, a power calculation can be performed using Monte Carlo simulations (i.e. [16]), which allow both fixed and random effects to be accounted for. To confirm the necessary number of zebrafish required to determine the likely subtle influence of an EOfAD-like mutation on spatial working memory, a Monte Carlo simulation should be performed in the future.

Disturbance of the oxidative phosphorylation pathway is a transcriptomic signature of knock-in models of early-onset familial Alzheimer's disease.

In **Chapters 5-8**, I performed bulk RNA-seq analysis of EOfAD-like mutations in the zebrafish orthologues of *SORL1*, *PSEN1* and *PSEN2*. Comparisons of the effects of these mutations, of other mutations previously generated by our group, and with knock-in mouse models of EOfAD and LOAD mutations, has revealed that changes to the oxidative phosphorylation pathway appear to be an early effect in AD-mutation carrier brains. As discussed in **Chapter 9**, both up- and down-regulation of the oxidative phosphorylation gene set was observed in different mutants for, as yet, unknown reasons.

One possible explanation for the differences in direction of changes to gene expression observed in the different EOfAD-like mutants could involve the maintenance of homeostasis. Cellular processes fluctuate in their activity to maintain an overall steady state [17]. If a cellular function is impaired, such as

due to an environmental stress, homeostatic mechanisms will attempt to correct for this stress by upregulating that function. This could explain a potential explanation for why we observe both up- and down-regulation of the oxidative phosphorylation genes across our different EOfAD mutation model zebrafish. Overall, the mutations may result in impaired mitochondrial function. Therefore, the oxidative phosphorylation genes may be upregulated as an attempt to compensate for the impairment. Subsequently, if the upregulation exceeds the steady-state levels, the genes must then be downregulated. The regulation of the oxidative phosphorylation genes may resemble a waveform, where it oscillates within a “normal” range. Cyclic patterns of gene expression are typical of feedback mechanisms, in which a delay in the response of the system can result in oscillating patterns of gene expression (reviewed in [18, 19]). In transcriptome analysis, we only observe a “snapshot” of the transcriptome state. Therefore, we may be observing the regulation of the oxidative phosphorylation genes between the peaks and troughs of this theoretical waveform as their expression is regulated to maintain homeostasis. However, this is highly speculative. Replication of the transcriptome analyses ideally should be performed to confirm whether each mutation results in a consistent direction of change to the oxidative phosphorylation gene set. A “targeted transcriptomics” approach such as BART-Seq [20], where only the expression of the oxidative phosphorylation genes can be analysed, could provide a more cost-effective method to achieve this. Additionally, the reality of energy production by mitochondria could be explored *in vivo*. Fluorescent sensors for ATP have previously been utilised in the living brains of mice [21], and application of this in zebrafish is likely feasible. Additionally, Mandal et al.

[22] have described an *in vivo* fluorescent reporter which can measure the ratio of ATP to ADP utilising a ratiometric, fluorescent sensor, PercevalHR [23], in zebrafish larvae. However, the changes to gene expression observed in our zebrafish models were relatively subtle (i.e. the magnitudes of the fold changes were generally modest, and statistical significance was only detected at the gene set, not single gene level). Therefore, exploring these changes at the molecular level could prove difficult *in vivo*, as the true signals may be lost in background noise.

Potential therapeutics for the treatment of AD

The brain transcriptome analyses presented in this thesis suggest that changes to mitochondrial function are an early stress in AD mutation-carrier brains. Therefore, targeting of mitochondria for therapeutic development, which is already being performed, is recommended. Naturally occurring compounds such as curcumin (derived from the spice turmeric) have antioxidant and anti-inflammatory properties which have been utilised to potentially alleviate complications due to dysfunctional mitochondria in AD. Treatment of curcumin in a transgenic mouse model of AD reduced oxidative damage, inflammation and plaque pathology relative to controls [24]. However, this was not translated in human clinical trials and the AD-related protective effects were not observed, likely due to the lack of uptake into the brain (reviewed in [25]). Alternate methods of increasing the bioavailability of curcumin are currently being explored (reviewed in [26]). Other compounds have been explored and these are recently reviewed in detail in [27, 28]. However, the overall consensus is that variable success has been observed in human clinical trials utilising

mitochondrial-based therapeutics. Notably, at the animal model testing stage, almost all of these therapeutics were tested using transgenic mouse models of AD. Future therapeutics should be tested on knock-in animal models, as these may more reliably indicate therapeutic efficacy than transgenic models (which have not yet led to successful therapeutics for AD).

Additionally, an individual's level of physical activity and diet is known to have AD-protective effects which have properties which act at the mitochondrial level. Regular physical exercise is associated with neuro-protective effects such as increase blood flow to the brain [29], reduced oxidative stress [30], increased brain mitochondrial biogenesis [31], as well as association with improved cognitive abilities [29, 32, 33]. Diets which consist of high intake of foods naturally high in antioxidants and unsaturated fats, with minimal consumption of processed foods and added sugar (i.e. the Mediterranean diet) is associated with reduced risk of AD [34], as well as a lighter load of AD biomarkers [35]. Additionally, caloric restriction is associated with decreased oxidative stress as well as increased mitochondrial respiration [36, 37]. The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) [38] (which has led to a global network of similar studies [39]), is an investigation of whether a range of lifestyle modifications decrease the incidence of dementia. These lifestyle interventions span diet, exercise and cognitive training. In the initial study in the Finnish cohort, small protective effects were observed in the intervention group after two years, supporting that lifestyle intervention can reduce the incidence of dementia (including AD) [38]. A second follow-up analysis is due to be performed after seven years post-intervention, where larger effect sizes may be observed. A meta-analysis of all

of the FINGER-type studies occurring globally (summarised in [39]) will provide a comprehensive investigation of the efficacy of lifestyle modifications on the prevention of cognitive decline in humans.

Alternatively, mitochondrial dysfunction may be a secondary event in AD pathogenesis. As discussed multiple times throughout this thesis, failure to acidify the endolysosomal system (such as due to *PSEN1* mutations [40], increased expression of the C99 fragment of APP [41] and the presence of the $\epsilon 4$ allele of *APOE* [42]) can result in mitochondrial abnormalities [43]. In promising work from the group of Professor Ralph Nixon, $\beta 2$ -adrenergic receptor agonists (such as the heart medication isoproterenol) were shown to rescue lysosomal acidification in three different cells lines with loss of *PSEN1* function [44]. Lysosomes of *PSEN1* deficient cells showed not only decreased acidification relative to controls, but also reduced levels of chloride ions (Cl^-) [44]. Chloride ions normally act as counter ions for the protons concentrated in lysosomes by the action of the v-ATPase proton pump (reviewed in [45]). Treatment of *PSEN1* deficient cells with isoproterenol rescued lysosomal pH by facilitating the delivery of chloride channel-7 (ClC-7) from the endoplasmic reticulum to lysosomes, subsequently restoring lysosomal proteolysis, calcium homeostasis and autophagic flux to wild type levels [44]. This work has uncovered a novel phenotype in *PSEN1* mutant cells, implicating a broader role of PSEN1 in maintenance of lysosomal ion homeostasis. Could a similar phenotypic rescue occur for the impairment of lysosomal acidification associated with C99, or the $\epsilon 4$ allele of *APOE*? Would these drugs be effective *in vivo*? Does treatment of $\beta 2$ -adrenergic receptor agonists rescue ferrous iron deficiency, mitochondrial abnormalities and inflammation associated with

impaired acidification of the endo-lysosomal system? These questions must be addressed in future work. However, this novel therapeutic strategy presents a promising avenue for the treatment of AD.

Cell-type specific effects should be explored in future experimentation

Analysis of bulk RNA-seq data gives information on global changes of the transcriptomic state of the brain, where gene expression values become averaged across the brain's various cell types. The next logical step in understanding the effects of these mutations in young adult, mutation carrier brains is to perform cell-type specific analyses, i.e. single-cell or single nuclear RNA-seq (scRNA-seq and snRNA-seq respectively). These types of experiments have recently been performed on human AD samples [46-49], and on transgenic/knock-out mouse models [50-52] and have revealed cell-type- (and gender-) specific changes to gene expression with relevance to AD pathogenesis. For example, Mathys et al. [48] profiled > 80,000 nuclear transcriptomes from the prefrontal cortex of 48 individuals of various AD pathology stages. A comparison of the snRNA-seq data with comparable bulk RNA-seq data was performed, and this revealed that the changes to gene expression observed in the bulk RNA-seq data were driven by gene expression changes in mainly two cell types: excitatory neurons and oligodendrocytes. Additionally, *APOE* was strongly upregulated in microglia, and downregulated (to a lesser extent) in astrocytes, an interesting detail which would be missed in bulk RNA-seq data. Some of the strongest cell-type specific changes to gene expression were observed when comparing samples arising from individuals with "early" AD pathology relative to controls. Conversely, changes to gene

expression which occur in samples derived from patients with late-stage AD pathology were relatively consistent across cell types. This highlights the power of sc/snRNA-seq for understanding the early effects of AD mutations *in vivo*. Future work should include single cell/nuclei analysis of the zebrafish models described in this thesis to tease out the cell-type-specific changes occurring due to the EOfAD-like mutations.

Conclusion

The research presented in this thesis has harnessed the power of zebrafish as a model organism to reveal that an early transcriptomic signature of EOfAD is oxidative phosphorylation. To our knowledge, this is the first time brain transcriptome analyses have been performed in *young, knock-in* models of EOfAD in any species. Comparison of the transcriptomes of our zebrafish models with publicly available brain transcriptome data from mouse models of LOAD revealed similar processes affected, supporting that EOfAD and LOAD are the same disease. This definition of an early EOfAD transcriptomic signature, avoiding previous assumptions regarding pathogenic mechanism, provides a step towards development of a successful treatment for AD.

References

1. Newman M, Nik HM, Sutherland GT, Hin N, Kim WS, Halliday GM, Jayadev S, Smith C, Laird AS, Lucas CW, et al: **Accelerated loss of hypoxia response in zebrafish with familial Alzheimer's disease-like mutation of presenilin 1.** *Human molecular genetics* 2020, **29**:2379-2394.
2. de la Torre JC, Mussivan T: **Can disturbed brain microcirculation cause Alzheimer's disease?** *Neurological Research* 1993, **15**:146-153.

3. Barthelson K, Newman M, Nowell CJ, Lardelli M: **No observed effect on brain vasculature of Alzheimer's disease-related mutations in the zebrafish presenilin 1 gene.** *Molecular Brain* 2021, **14**:22.
4. De Gasperi R, Sosa MAG, Dracheva S, Elder GA: **Presenilin-1 regulates induction of hypoxia inducible factor-1 α : altered activation by a mutation associated with familial Alzheimer's disease.** *Molecular Neurodegeneration* 2010, **5**:38-38.
5. Villa JC, Chiu D, Brandes AH, Escorcía FE, Villa CH, Maguire WF, Hu CJ, de Stanchina E, Simon MC, Sisodia SS, et al: **Nontranscriptional role of Hif-1 α in activation of gamma-secretase and notch signaling in breast cancer.** *Cell Replication* 2014, **8**:1077-1092.
6. Le Moan N, Houslay DM, Christian F, Houslay MD, Akassoglou K: **Oxygen-dependent cleavage of the p75 neurotrophin receptor triggers stabilization of HIF-1 α .** *Molecular cell* 2011, **44**:476-490.
7. Hin N, Newman M, Kaslin J, Douek AM, Lumsden A, Nik SHM, Dong Y, Zhou X-F, Mañucat-Tan NB, Ludington A, et al: **Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2.** *PLOS ONE* 2020, **15**:e0227258.
8. Hin N, Newman M, Pederson SM, Lardelli MM: **Iron Responsive Element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease.** *bioRxiv* 2020:2020.2005.2001.071498.
9. Ishikawa A, Piao Y-S, Miyashita A, Kuwano R, Onodera O, Ohtake H, Suzuki M, Nishizawa M, Takahashi H: **A mutant PSEN1 causes dementia with lewy bodies and variant Alzheimer's disease.** *Annals of Neurology* 2005, **57**:429-434.
10. Sun L, Zhou R, Yang G, Shi Y: **Analysis of 138 pathogenic mutations in presenilin-1 on the in vitro production of A β 42 and A β 40 peptides by γ -secretase.** *Proceedings of the National Academy of Sciences* 2017, **114**:E476-E485.
11. Wang GL, Jiang BH, Rue EA, Semenza GL: **Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O₂ tension.** *Proceedings of the National Academy of Sciences* 1995, **92**:5510-5514.
12. Wang GL, Semenza GL: **Desferrioxamine induces erythropoietin gene expression and hypoxia-inducible factor 1 DNA-binding activity: implications for models of hypoxia signal transduction.** *Blood* 1993, **82**:3610-3615.
13. Di Giovanna AP, Tibo A, Silvestri L, Müllenbroich MC, Costantini I, Allegra Mascaro AL, Sacconi L, Frasconi P, Pavone FS: **Whole-Brain Vasculature Reconstruction at the Single Capillary Level.** *Scientific Reports* 2018, **8**:12573.
14. Cleal M, Fontana BD, Ranson DC, McBride SD, Swinny JD, Redhead ES, Parker MO: **The Free-movement pattern Y-maze: A cross-species measure of working memory and executive function.** *Behavior Research Methods* 2020.

15. Cleal M, Parker MO: **Moderate developmental alcohol exposure reduces repetitive alternation in a zebrafish model of fetal alcohol spectrum disorders.** *Neurotoxicology and Teratology* 2018, **70**:1-9.
16. Green P, MacLeod CJ: **SIMR: an R package for power analysis of generalized linear mixed models by simulation.** *Methods in Ecology and Evolution* 2016, **7**:493-498.
17. Cannon WB: **The wisdom of the body.** 1939.
18. Nelson DE, Sée V, Nelson G, White MRH: **Oscillations in transcription factor dynamics: a new way to control gene expression.** *Biochemical Society Transactions* 2004, **32**:1090-1092.
19. Brandman O, Meyer T: **Feedback Loops Shape Cellular Signals in Space and Time.** *Science* 2008, **322**:390.
20. Uzbas F, Opperer F, Sönmezer C, Shaposhnikov D, Sass S, Krendl C, Angerer P, Theis FJ, Mueller NS, Drukker M: **BART-Seq: cost-effective massively parallelized targeted sequencing for genomics, transcriptomics, and single-cell analysis.** *Genome Biology* 2019, **20**:155.
21. Kitajima N, Takikawa K, Sekiya H, Satoh K, Asanuma D, Sakamoto H, Takahashi S, Hanaoka K, Urano Y, Namiki S, et al: **Real-time in vivo imaging of extracellular ATP in the brain with a hybrid-type fluorescent sensor.** *eLife* 2020, **9**:e57544.
22. Mandal A, Pinter K, Drerup CM: **Analyzing Neuronal Mitochondria in vivo Using Fluorescent Reporters in Zebrafish.** *Frontiers in Cell and Developmental Biology* 2018, **6**:144.
23. Tantama M, Martínez-François JR, Mongeon R, Yellen G: **Imaging energy status in live cells with a fluorescent biosensor of the intracellular ATP-to-ADP ratio.** *Nature Communications* 2013, **4**:2550.
24. Lim GP, Chu T, Yang F, Beech W, Frautschy SA, Cole GM: **The Curry Spice Curcumin Reduces Oxidative Damage and Amyloid Pathology in an Alzheimer Transgenic Mouse.** *The Journal of Neuroscience* 2001, **21**:8370.
25. Belkacemi A, Doggui S, Dao L, Ramassamy C: **Challenges associated with curcumin therapy in Alzheimer disease.** *Expert Reviews in Molecular Medicine* 2011, **13**:e34.
26. Serafini MM, Catanzaro M, Rosini M, Racchi M, Lanni C: **Curcumin in Alzheimer's disease: Can we think to new strategies and perspectives for this molecule?** *Pharmacological Research* 2017, **124**:146-155.
27. Cenini G, Voos W: **Mitochondria as Potential Targets in Alzheimer Disease Therapy: An Update.** *Frontiers in Pharmacology* 2019, **10**:902.
28. Cunnane SC, Trushina E, Morland C, Prigione A, Casadesus G, Andrews ZB, Beal MF, Bergersen LH, Brinton RD, de la Monte S, et al: **Brain energy rescue: an emerging therapeutic concept for neurodegenerative disorders of ageing.** *Nature Reviews Drug Discovery* 2020, **19**:609-633.

29. Chapman S, Aslan S, Spence J, DeFina L, Keebler M, Didehbani N, Lu H: **Shorter term aerobic exercise improves brain, cognition, and cardiovascular fitness in aging.** *Frontiers in Aging Neuroscience* 2013, **5**:75.
30. Wang Q, Xu Z, Tang J, Sun J, Gao J, Wu T, Xiao M: **Voluntary exercise counteracts A β 25-35-induced memory impairment in mice.** *Behavioural Brain Research* 2013, **256**:618-625.
31. Steiner JL, Murphy EA, McClellan JL, Carmichael MD, Davis JM: **Exercise training increases mitochondrial biogenesis in the brain.** *Journal of applied physiology (Bethesda, Md : 1985)* 2011, **111**:1066-1071.
32. Adlard PA, Perreau VM, Pop V, Cotman CW: **Voluntary Exercise Decreases Amyloid Load in a Transgenic Model of Alzheimer's Disease.** *The Journal of Neuroscience* 2005, **25**:4217.
33. Winchester J, Dick MB, Gillen D, Reed B, Miller B, Tinklenberg J, Mungas D, Chui H, Galasko D, Hewett L, Cotman CW: **Walking stabilizes cognitive functioning in Alzheimer's disease (AD) across one year.** *Archives of Gerontology and Geriatrics* 2013, **56**:96-103.
34. Scarmeas N, Stern Y, Tang M-X, Mayeux R, Luchsinger JA: **Mediterranean diet and risk for Alzheimer's disease.** *Annals of neurology* 2006, **59**:912-921.
35. Hill E, Goodwill AM, Gorelik A, Szoeka C: **Diet and biomarkers of Alzheimer's disease: a systematic review and meta-analysis.** *Neurobiology of Aging* 2019, **76**:45-52.
36. Amigo I, Menezes-Filho SL, Luévano-Martínez LA, Chausse B, Kowaltowski AJ: **Caloric restriction increases brain mitochondrial calcium retention capacity and protects against excitotoxicity.** *Aging Cell* 2017, **16**:73-81.
37. Redman LM, Smith SR, Burton JH, Martin CK, Il'yasova D, Ravussin E: **Metabolic Slowing and Reduced Oxidative Damage with Sustained Caloric Restriction Support the Rate of Living and Oxidative Damage Theories of Aging.** *Cell Metabolism* 2018, **27**:805-815.e804.
38. Kivipelto M, Solomon A, Ahtiluoto S, Ngandu T, Lehtisalo J, Antikainen R, Bäckman L, Hänninen T, Jula A, Laatikainen T, et al: **The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER): Study design and progress.** *Alzheimer's & Dementia* 2013, **9**:657-665.
39. Kivipelto M, Mangialasche F, Snyder HM, Allegri R, Andrieu S, Arai H, Baker L, Belleville S, Brodaty H, Brucki SM, et al: **World-Wide FINGERS Network: A global approach to risk reduction and prevention of dementia.** *Alzheimer's & Dementia* 2020, **16**:1078-1094.
40. Lee JH, Yu WH, Kumar A, Lee S, Mohan PS, Peterhoff CM, Wolfe DM, Martinez-Vicente M, Massey AC, Sovak G, et al: **Lysosomal proteolysis and autophagy require presenilin 1 and are disrupted by Alzheimer-related PS1 mutations.** *Cell* 2010, **141**:1146-1158.

41. Jiang Y, Sato Y, Im E, Berg M, Bordi M, Darji S, Kumar A, Mohan PS, Bandyopadhyay U, Diaz A, et al: **Lysosomal Dysfunction in Down Syndrome Is APP-Dependent and Mediated by APP- β CTF (C99).** *The Journal of Neuroscience* 2019, **39**:5255.
42. Prasad H, Rao R: **Amyloid clearance defect in ApoE4 astrocytes is reversed by epigenetic correction of endosomal pH.** *Proceedings of the National Academy of Sciences* 2018, **115**:E6640.
43. Yambire KF, Rostosky C, Watanabe T, Pacheu-Grau D, Torres-Odio S, Sanchez-Guerrero A, Senderovich O, Meyron-Holtz EG, Milosevic I, Frahm J, et al: **Impaired lysosomal acidification triggers iron deficiency and inflammation in vivo.** *Elife* 2019, **8**.
44. Lee J-H, Wolfe DM, Darji S, McBrayer MK, Colacurcio DJ, Kumar A, Stavrides P, Mohan PS, Nixon RA: **β 2-adrenergic Agonists Rescue Lysosome Acidification and Function in PSEN1 Deficiency by Reversing Defective ER-to-lysosome Delivery of CLC-7.** *Journal of Molecular Biology* 2020, **432**:2633-2650.
45. Bose S, He H, Stauber T: **Neurodegeneration Upon Dysfunction of Endosomal/Lysosomal CLC Chloride Transporters.** *Frontiers in Cell and Developmental Biology* 2021, **9**:323.
46. Del-Aguila JL, Li Z, Dube U, Mihindikulasuriya KA, Budde JP, Fernandez MV, Ibanez L, Bradley J, Wang F, Bergmann K, et al: **A single-nuclei RNA sequencing study of Mendelian and sporadic AD in the human brain.** *Alzheimer's Research & Therapy* 2019, **11**:71.
47. Grubman A, Chew G, Ouyang JF, Sun G, Choo XY, McLean C, Simmons RK, Buckberry S, Vargas-Landin DB, Poppe D, et al: **A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation.** *Nat Neurosci* 2019, **22**:2087-2097.
48. Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, Menon M, He L, Abdurrob F, Jiang X, et al: **Single-cell transcriptomic analysis of Alzheimer's disease.** *Nature* 2019, **570**:332-337.
49. Lau S-F, Cao H, Fu AKY, Ip NY: **Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease.** *Proceedings of the National Academy of Sciences* 2020, **117**:25800.
50. Zhong S, Wang M, Zhan Y, Zhang J, Yang X, Fu S, Bi D, Gao F, Shen Y, Chen Z: **Single-nucleus RNA sequencing reveals transcriptional changes of hippocampal neurons in APP23 mouse model of Alzheimer's disease.** *Bioscience, Biotechnology, and Biochemistry* 2020, **84**:919-926.
51. Keren-Shaul H, Spinrad A, Weiner A, Matcovitch-Natan O, Dvir-Szternfeld R, Ulland TK, David E, Baruch K, Lara-Astaiso D, Toth B, et al: **A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease.** *Cell* 2017, **169**:1276-1290.e1217.
52. Zhou Y, Song WM, Andhey PS, Swain A, Levy T, Miller KR, Poliani PL, Cominelli M, Grover S, Gilfillan S, et al: **Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease.** *Nature Medicine* 2020, **26**:131-142.