THE UNIVERSITY of ADELAIDE

Investigating Alzheimer's disease using zebrafish genetic models

Yang Dong (1661360)

Supervisor: Michael Lardelli

Co-supervisor: Morgan Newman

Department of Genetics and Evolution

School of Biological Sciences

A thesis for the degree of Doctor of Philosophy

The University of Adelaide

Submission Date: 14 October 2020

Table of Contents

The following manuscript submitted for publication is included in this thesis:

Transcriptome analyses of 7-day-old zebrafish larvae possessing a familiar Alzheimer's disease-like mutation in *psen1* indicate effects on oxidative phosphorylation, mcm functions, and iron homeostasis

Yang Dong, Morgan Newman, Stephen Pederson, Nhi Hin, Michael Lardelli

bioRxiv 2020.05.03.075424; doi: https://doi.org/10.1101/2020.05.03.075424

Abstract

Early-onset familial Alzheimer's disease (EOfAD) is promoted by autosomal, dominant mutations, enabling the generation of EOfAD-like mutations in animal models for the study of its pathogenic mechanism. This thesis has succeeded in generating an EOfAD-like mutation and an Acne Inversa-like mutation in *psen1* using the CRISPR/Cas system. The EOfAD-like mutation can be used for investigation of EOfAD-specific transcriptome signatures, while the Acne Inversa-like mutation (non-EOfAD-like) will be used in transcriptome comparisons with EOfAD-like mutations to refine (by exclusion) our definition of the AD-relevant aspects of EOfAD-like brain transcriptome profiles. On the other hand, the Alzheimer's Disease Genetics Laboratory (ADGL) has generated two EOfAD-relevant mutations in zebrafish, $psen1^{K97fs}$ and $psen1^{Q96\_K97del}$, and has performed brain transcriptome analyses on heterozygous 6-month-old and 24-month-old mutants compared to their wild type siblings to investigate gene expression changes induced by the mutations. Work by others in the ADGL had observed significant differences in gene expression due to the mutations at 6 months compared to 24 months. In the work described in this thesis, further analyses were performed on these two mutations. I analysed transcript splicing associated with the $psen1^{K97fs}$ mutation and found that I could not detect equivalent alternative splice products associated with the human mutation upon which this mutation model is based. This indicated that the zebrafish $psen1^{K97fs}$ mutation is probably not a close model of the pathological action of human $PSEN2^{K115fs}$. In addition, I performed transcriptome analysis of heterozygous $psen1^{Q96\_K97del}$ mutant larvae to test their possible utility

in chemical library screening. I found that genes DE in the larvae were not consistent with those identified in 6-month-old heterozygous mutant brains, although similar cellular pathways were affected. Also, some larvae-specific effects were identified. However, these analyses encountered difficulties. I initially aimed to investigate the nature of this age-dependent transition in gene expression using quantitative PCR (qPCR). I intended to monitor the change in expression of these genes month by month. However, variability in the expression of the selected differentially expressed (DE) genes thwarted this analysis. Furthermore, I compared the behaviour of heterozygous *psen1*$^{Q96\_K97del}$ mutant adult fish to their wild type siblings using the free-movement pattern (FMP) Y-maze test to assess impairments in short-term spatial working memory at 6, 12 and 24 months of age. However, no statistically significant differences in behaviour were identified.

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.

Name of Candidate: Yang Dong

Signature: Yang Dong Date: 29/09/2020

Ethics

The research described in this project was carried out under permit S-2017-073 from the Animal Ethics Committee of the University of Adelaide and permit 12210 from the Institutional Biosafety Committee of the University of Adelaide.

Acknowledgements

This research was completed under the guidance and support of my principle supervisor A/Prof. Michael Lardelli and my co-supervisor Dr. Morgan Newman. I sincerely thank Michael and Morgan for guiding and encouraging me to overcome difficulties and persist through my Ph.D. candidature. In particular, when I had a time with a number of unexpected difficulties in my experiments, they did not stress me and gave me helpful suggestions. It was my pleasure and good fortune to work with them, and I learned a lot from them during this time. Also I extremely thankful to Michael for providing suggestions on, and corrections to my writing.

I would also like to thank Dr. Stephen Martin Pederson for his support in bioinformatics. At the outset, I did not have much experience in bioinformatics, but he guided me go through the bioinformatics analysis with great patience.

Finally, I want to express my thanks to my previous and present lab-members, Karissa Barthelson, Nhi Hin, Lachlan Warren William Baer, Tanya Jayne, Haowei Jiang, Jiayu Ruan, Ewan Gerken and Seyyed Hani Moussavi Nik. I always received help and support from them, and feel so pleased to have experienced such a good, positive research environment.

Chapter 1. Literature review

1.1 Introduction

This project focuses on the analysis of two early onset familial Alzheimer's disease (EOfAD)-like zebrafish models, *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$, generated in the previous work of this laboratory, and attempts to generate another two mutations in the zebrafish *psen1* gene to advance the investigation of Alzheimer's disease (AD) pathological mechanisms. Therefore, this Introduction chapter first looks at Alzheimer's disease and hypotheses of the pathological mechanisms that lead to it, followed by an examination of the genetics of this disease, animal models and the previous work of the Alzheimer's Disease Genetics Laboratory.

1.2 The definition of Alzheimer's disease

1.2.1 The classification of AD

AD is a progressive neurodegenerative brain disorder that impairs cognitive functions and eventually develops to dementia. Without efficient treatments available, AD has become a major leading cause of death and shows an increasing disease incidence [1]. AD can be classified by its age of onset and whether or not it shows inheritance in families. Late onset, sporadic AD (sAD), occurs after 65 years of age and is responsible for most cases of AD (>95%) [2]. This form of AD has a very limited inheritance and a family history of affected close relatives is unusual

[2, 3], while it is affected by multiple factors including age, lifestyle, diet, genetic and environmental factors [4]. In contrast, EOfAD is autosomal and dominantly inherited. Most EOfAD-causative mutations have been identified in genes *PSEN1*, *PSEN2* and *APP* [5]. The mean age of EOfAD onset is around 45 years of age [3, 6]. Although EOfAD only contributes less than 5% of all AD cases [6], its dominant mode of inheritance enables genetic modeling in animals for AD studies. sAD and EOfAD share many common pathological features such as the accumulation of amyloid plaques and neurofibrillary tangles (**Figure 1.1**, described later), and there is no evidence showing that they have different pathological progressions. Therefore, in many AD pathology hypotheses, sAD and EOfAD are assumed to share a similar pathological process. Also, in most clinical respects, the rate of disease progression and biomarker profiles of these two forms are comparable [7]. However, as the results of genome-wide association studies (GWAS) for sAD did not identify the major EOfAD loci, *APP* and *PSEN* genes, as susceptibility loci [8], some argue that sAD and EOfAD should probably not be considered together but regarded as different diseases [9]. Nevertheless, analysis of EOfAD should probably provide insight into the genetic basis, etiology and pathology of both AD types, and treatments targeting the pathological development of EOfAD may also be effective in sAD.

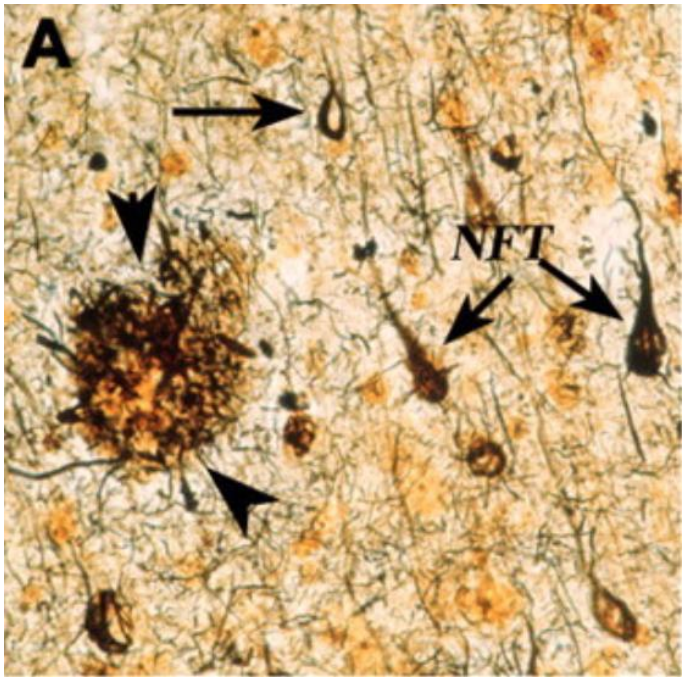


Figure 1.1. The pathological features of AD revealed by the Bielschowsky silver stain (from [10]). Amyloid plaque is indicated by arrowheads while neurofibrillary tangles (NFTs) are indicated by arrows.

1.2.2 The diagnosis of AD

The pathology and diagnosis of AD remains controversial. Accumulation of amyloid plaques (amyloid-β peptides) and neurofibrillary tangles (tau proteins) have been widely regarded as pathological features of AD for over one hundred years [11]. Currently the presence of deposits of these proteins in patients' brains is still used to biologically define AD [12, 13]. However, the clinical diagnosis of AD is based on the observation of a progressive amnestic multidomain cognitive impairment culminating in dementia when other possible causes have been excluded [13]. These two definitions complicate the diagnostic criteria for AD. Most patients with biologically defined AD do not have these clinical symptoms, while some patients clinically diagnosed with AD, upon autopsy, have been shown

to lack these biological features [9, 13]. On the other hand, most current investigations of AD treatments attempt to the prevent or delay the formation of these biological features, but have met with challenges and failure [14]. A study by Jack et al., 2020 [13] suggested that patients with biologically defined AD are more prevalent than those with clinically diagnosed AD at any age, and raised concerns about the specific treatments that are used to prevent biological symptoms onset. Therefore, there is a great need to investigate the cause of AD. In future this will also require a more accurate definition of the disease to improve the reliability of diagnosis and facilitate investigation of treatments.

1.2.3 Cognitive impairment

The cognitive impairment caused by AD is a progressive process with three stages: the preclinical stage with no symptoms, a middle stage of mild cognitive impairment (MCI) and a final stage of dementia [15]. An early clinical feature of AD is often a prominent memory disturbance [16]. With disease progression, other cognitive issues, such as psychosis and depression [17], and behavioural disturbances [18] start to emerge. The clinical evolution of MCI due to AD involves impairments in one or more cognitive domains, typically including memory [19]. The impairment of short-term memory is normally one of the symptoms at the disease early stage. Short-term memory, also called primary or active memory, is a functional storage element that has the ability to hold pieces of information for a relatively short time [20]. Short-term memory is not equivalent to working memory but acts as a component of it, as working memory refers to the structures and

processes used for temporarily storing and manipulating information [20]. In contrast to short-term memory, long-term memory may hold an indefinite amount of information for many years [20]. As Alzheimer's disease progresses, long-term memory trends to become fragmented, and eventually the capacity for memory may be entirely lost during the disease's final stage.

1.3 Hypotheses of Alzheimer's disease

1.3.1 The Amyloid Cascade Hypothesis

The Amyloid Cascade Hypothesis is the most widely known hypothesis for the mechanism underlying AD. It posits that the major cause of AD is related to a pathological feature, the accumulation of a minor form of the amyloid-β (Aβ) peptide, $A\beta_{42}$. (The production of Aβ is discussed in ***section 1.4.2*** below.) The proposed AD progression in this hypothesis is shown in **Figure 1.2**. Aβ normally occurs in both 40 amino acid residue ($A\beta_{40}$) and 42 amino acid residue ($A\beta_{42}$) forms. Dominant missense mutations in EOfAD lead to increased production of the less abundant $A\beta_{42}$, while the failure of Aβ clearance in sAD also results in the accumulation of $A\beta_{42}$. $A\beta_{42}$ accumulates in both sAD and EOfAD brains. An oligomeric form(s) of $A\beta_{42}$ aggregate is assumed to be toxic. The oligomers containing $A\beta_{42}$ aggregate and become plaques through gradual deposition, leading to a series of downstream reactions, such as inflammatory responses and the formation of tangles, finally resulting in dementia [21, 22].

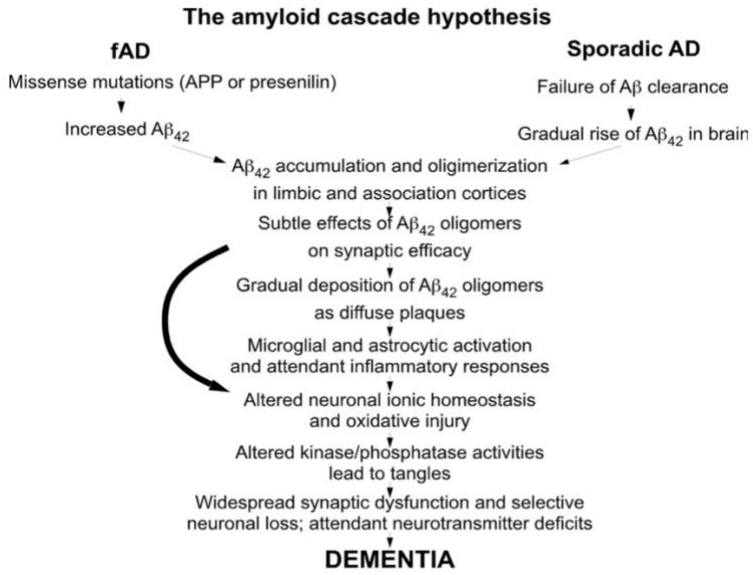


Figure 1.2. The amyloid hypothesis of Alzheimer's disease as described by Herrup [21]. Different etiologies in EOfAD and sAD both lead to the production of an excess of $A\beta_{42}$. The $A\beta_{42}$ peptides are assumed to be toxic and the aggregation of $A\beta_{42}$ triggers a cascade of downstream events, resulting in a series of biological and neurological AD symptoms.

The increased Aβ seen in the brains of those with Down syndrome results in the development of pathological changes resembling AD by 35 years of age [23]. This is due to the extra copy of chromosome 21 in this syndrome. Chromosome 21 contains the *APP* gene, leading to the increase of Aβ production. Therefore, this early onset form of AD in Down syndrome supports that the accumulation of Aβ may contribute to the pathology of AD.

Some experimental data is inconsistent with the Amyloid Hypothesis. A recent study by Sun et al., 2017 [24] examined the ability of 138 reported EOfAD mutations in human *PSEN1* to produce $A\beta_{40}$ and $A\beta_{42}$ *in vitro,* and found that the production of $A\beta_{40}$ and $A\beta_{42}$ was reduced for about 90% of these mutations and

10% of these mutations showed decreased ratios of $A\beta_{42}/A\beta_{40}$ (the $A\beta_{42}/A\beta_{40}$ ratio will be discussed further in ***section 1.4.2***). The decreased $A\beta_{42}$ production and $A\beta_{42}/A\beta_{40}$ ratios are not consistent with the Amyloid Cascade Hypothesis. Also, the Amyloid Cascade Hypothesis fails to explain why there are cognitively normal people (without dementia) who have plaques and tangles in their brains [9]. The impacts of these plaques and tangles on cognitively normal people and AD patients remain unclear.

1.3.2 Tau hypothesis

Another major pathological feature, the accumulation of neurofibrillary tangles, is thought to cause AD in the tau hypothesis. Tau is a highly soluble microtubule-associated protein (MAP) that can interact with tubulin to stabilize microtubule assembly [25]. There are six isoforms of tau transcripts expressed in adult human brains with or without Exons 2, 3 and 10. Inclusion or otherwise of exon 10 results in isoforms of Tau with either four or three microtubule binding domains, ie. 4-repeat (4R) isoforms and 3-repeat (3R) isoforms (**Figure 1.3**, [26]).

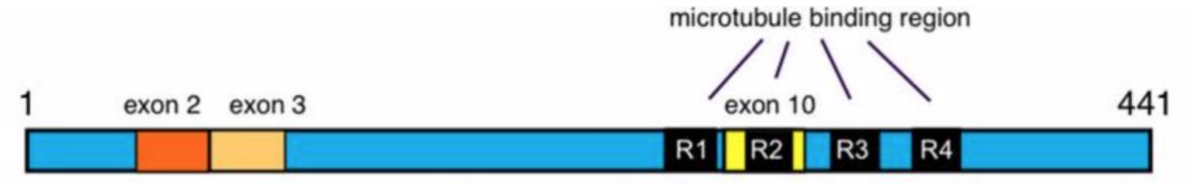


Figure 1.3. The alternative splicing of tau transcripts (from [26]). Six tau transcript isoforms are expressed in adult human brains with or without Exons 2, 3 and 10. There are four repeats of a microtubule binding domain, R1-4. Exon 10 contains the repeat R2.

The abnormal hyperphosphorylation of tau may affect isoform expression and promote the assembly of tau into paired helical filaments (PHFs) or straight filaments (SFs), which are the major components of the intracellular neurofibrillary tangles (NFTs) [25, 27] seen in AD brains. The insoluble structure of NFTs may accumulate and damage neuronal cells, eventually developing into dementia. The tau hypothesis has a similar limitation to the amyloid hypothesis in failing to explain the presence of neurofibrillary tangles in normal brains and their absence in some AD brains. Also, there are no clear explanations for how the genetic factors, such as the EOfAD-causative mutations, induce the tau pathology.

1.3.3 MAM hypothesis

The amyloid cascade hypothesis and the tau hypothesis focus on the major histopathological features of AD, amyloid plaques and neurofibrillary tangles respectively. However, they put little emphasis on other disease-related phenomena that occur years before the accumulation of plaques and tangles [28]. These phenomena include the altered metabolism of fatty acids and phospholipids [29, 30], increased levels of circulating cholesterol [31], the deposition of lipid droplets in cells [32-34], aberrant calcium homeostasis [35], increased stress on the endoplasmic reticulum (ER) [36], and mitochondrial dysfunction [37, 38]. These phenomena are associated with the functions of a specialized subdomain of the ER known as the mitochondria-associated ER membranes (MAM) [28]. MAM, which has the features of a lipid raft, connects mitochondria and ER physically and is critical for the interactions between mitochondria and the ER, including

phospholipid biosynthesis, fatty acid metabolism, cholesterol esterification, calcium homeostasis, mitochondrial dynamics, and communication between the two organelles [39]. The MAM hypothesis focuses on these early features and proposes that they result from increased ER-mitochondrial apposition and perturbed MAM function, which are causative and early events in AD pathology. The accumulation of plaques and tangles is seen as a consequence rather than as a cause of these changes (reviewed in Area-Gomez and Schon, 2017 [28]). Furthermore, MAM is also related to Aβ production [40, 41], and the active forms of PSENs, (in γ-secretase activity) and the γ-secretase cleavage of APP show localisation to the MAM in brains [42, 43]. Increases in MAM functionality are seen in fibroblasts both from sAD patients and when expression of the PSEN genes is reduced [44]. Therefore, a unified explanation for the pathology of both EOfAD and sAD is achieved in the MAM hypothesis.

1.3.4 PSEN holoprotein multimer hypothesis

EOfAD is autosomal and dominantly inherited, and *PRESENILIN* (*PSEN*) genes are major genetic loci of EOfAD [45]. The EOfAD-causative mutations share the common feature of following the "fAD reading frame preservation rule" [5]. These mutations do not change the open reading frame of the transcripts and retain the production of full-length proteins. Based on this feature, it is assumed in the PSEN holoprotein multimer hypothesis (illustrated in **Figure 1.4**) that the production of full-length proteins plays an essential role in disease pathology. In this hypothesis,

the normal PSEN holoprotein can form dimers/multimers that are critical for the maintenance of cellular homeostasis, while EOfAD-mutated PSEN holoprotein may multimerize with wild type holoprotein and inhibit the formation of normal PSEN holoprotein dimers/multimers, leading to the disruption of holoprotein functions and cellular homeostasis such as autophagy or secretion of Aβ [5]. The hypothesis also proposes that, under hypoxia, hypoxia-inducible factor 1, alpha subunit (HIF1α) might promote the endoproteolysis of PSEN holoprotein dimers/multimers to stimulate the formation of active γ-secretase [5]. As a result, EOfAD-mutated PSEN holoprotein would probably lead to increased production of Aβ. Although this PSEN holoprotein multimer hypothesis has been advanced to interpret the action of EOfAD-causative mutations, it remains unknown how this hypothesis relates to sAD. This is a major limitation of this hypothesis and requires further exploration.

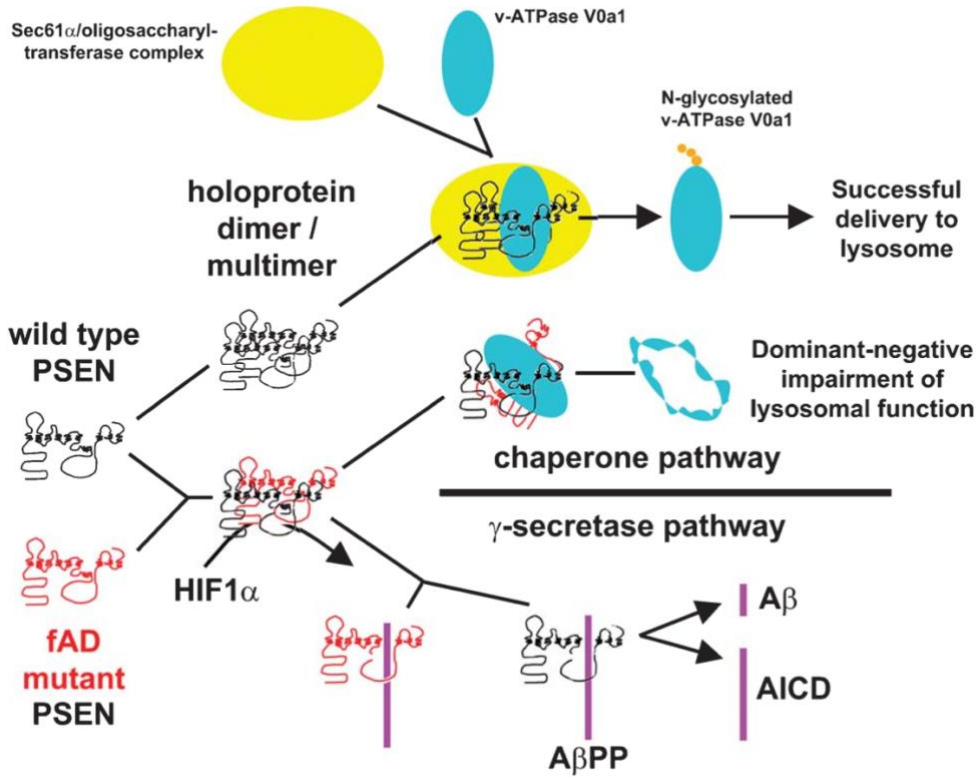


Figure 1.4. PSEN holoprotein multimer hypothesis (from [5]). PSEN holoproteins can form dimers/multimers the may be necessary for the endoproteolysis that activates γ-secretase activity, while mutated PSEN holoprotein may bind to wild type holoprotein and inhibit its activity, influencing cellular homeostasis. On the other hand, under hypoxia, HIF1α may be able to interact with holoprotein dimers/multimers to stimulate the formation of active γ-secretase.

1.3.5 The vascular hypothesis

In addition to genetic factors, age, as a major factor for both EOfAD and sAD, is taken into consideration in many hypotheses [21]. The features of brain degeneration during normal aging likely contribute to the occurrence of AD to some degree. The vascular hypothesis interprets AD progression from the viewpoint of the brain ageing and accumulated damage of vascular and neuronal cells. The

degeneration of vascular cells causes progressive reduction in cerebral blood flow (CBF), hypoperfusion and hypoxia [46]. Furthermore, the physical breakdown of the blood-brain barrier (BBB) caused by pericyte detachment leads to the leakage and accumulation of serum proteins and focal microhemorrhages and the extravasation of red blood cells (RBCs), subsequently resulting in hypoxia, hypoperfusion and the formation of toxic reactive oxygen species (ROS) (**Figure 1.5**, [46]). Also, neurotoxic proteins are able to pass the defective BBB and accumulate in the brain. Therefore, the vascular hypothesis proposes that these factors causing neuronal injury become aggravated during ageing and eventually cause dementia.

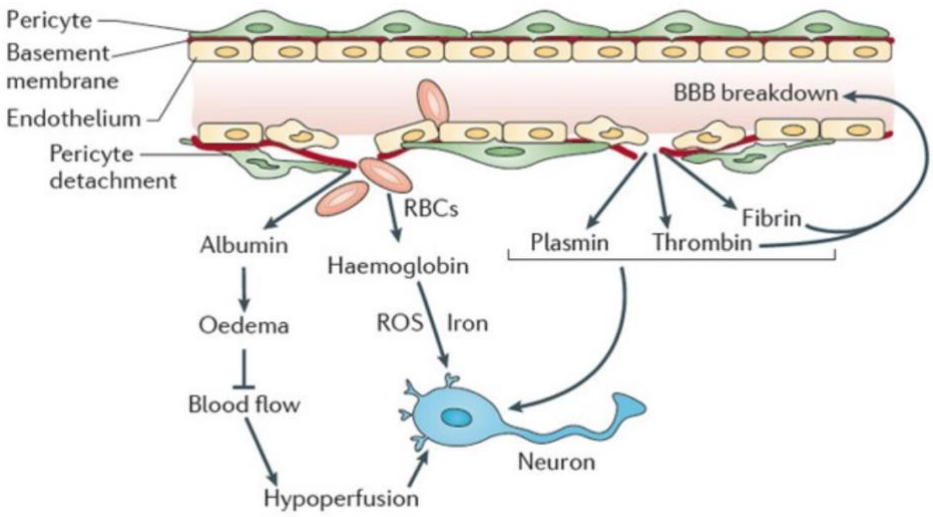


Figure 1.5. Neurodegeneration induced by the breakdown of BBB (from [46]). Pericyte detachment leads to leakage of serum proteins and focal microhaemorrhages, with extravasation of RBCs. Neural injury is aggravated by three proposed mechanisms: 1) Albumin promotes the formation of vascular oedema, which affects blood flow and subsequently induces hypoxia and hypoperfusion. 2) RBCs release haemoglobin, which provides iron to catalyse formation of toxic ROS. 3) The defective BBB allows vasculotoxic and neurotoxic proteins to enter the brain.

Under hypoxic conditions, upregulation of genes involved in Aβ production has been observed, including *APP*, *PSEN1* and *BACE1* [47-49], probably resulting in increased production of Aβ. In addition, due to the degeneration of pericytes and vascular smooth muscle cells, the clearance of Aβ is probably reduced [50]. The accumulation of Aβ is a major pathological feature of AD and may possibly cause other vascular diseases, such as cerebral amyloid angiopathy (CAA) [46], which, co-existing with AD, would significantly increase the risk of dementia [51]. Also, the reduced clearance flow in brains with vascular pathology probably contributes to the accumulation of secreted tau [52].

1.3.6 Change-of-state hypothesis

Another age-dependent hypothesis, the change-of-state hypothesis, has been proposed by Herrup, 2010 (**Figure 1.6** [21]). This hypothesis describes the pathological development of AD involving three key consecutive events: an initiating injury, a chronic inflammatory response and the change of brain cell states. Here, age acts as a risk factor for the initiating injury, which induces an inflammatory response to begin the pathogenic process. The correlation between the neuroinflammatory process and the neuropathology of AD has been demonstrated in several papers [53, 54]. Also, long-term use of nonsteroidal anti-inflammatory drugs has been found associated with a reduced prevalence of AD [54-56]. The inflammatory response causes a chronic stress that drives brain cells from a normal state into a different, pathological state, affecting cell cycle events

and causing defects in autophagy and alterations of tau [21]. Furthermore, in this hypothesis, the amyloid hypothesis has not been subverted, but contributes as a cause of pathogenesis. Inflammatory responses stimulate the production of Aβ, while the aggregation of Aβ also stimulates inflammation, causing an amyloid deposition cycle.

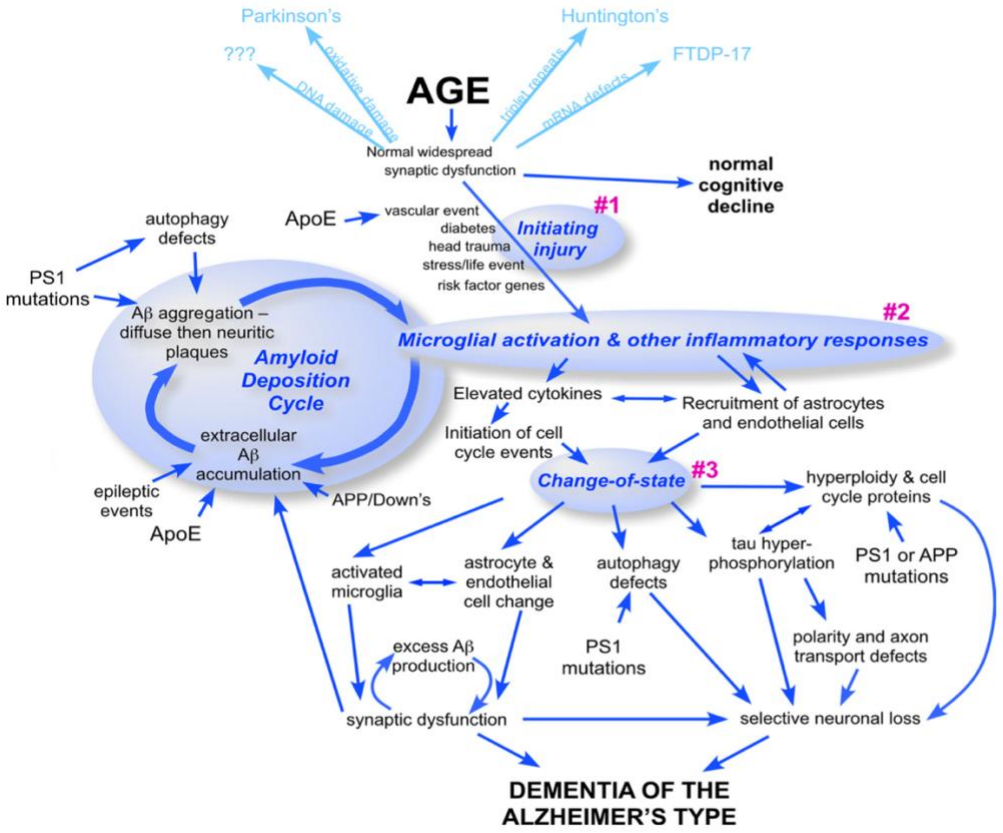


Figure 1.6. The expanded change-of-state hypothesis of AD of Herrup [21]. In this hypothesis, the three steps displayed in blue ovals and numbered are assumed to be key and necessary processes in disease pathology leading to AD. This expanded model also correlates relevant brain cell biology and biomedical processes with AD pathology, indicating the proposed role and involvement of related elements. Aβ accumulation is both a key element and a symptom in AD and is involved in the amyloid deposition cycle.

1.3.7 Summary of AD hypotheses

AD is a complicated neurodegenerative disease, and many AD hypotheses have been proposed focusing on different elements and features observed during disease progression. In the ***section 1.3***, six AD hypotheses selected from among the most popular AD hypotheses have been discussed. Due to the complexity of human brains, most AD hypotheses have limitations in comprehensively explaining AD pathology, and AD treatments based on these hypotheses have faced difficulties [14, 57]. Moreover, AD pathology involves interactions between different pathways, making it even more difficult to investigate AD etiology. As a result, although there is quite a variety of AD hypotheses proposed, it is difficult to determine which hypothesis is most closely in alignment with the reality of AD pathology. Therefore, rather than focusing on AD hypotheses, the ADGL models EOfAD-relevant mutations in zebrafish models and pays attention to transcriptome changes induced by these mutations, to investigate transcriptional signatures of AD and to interpret AD pathology.

1.4 Genetics

1.4.1 Genetic risk factors for sAD

Although sAD shows only limited inheritance, genetic factors remain a strong contributor among the multiple risk factors. More than twenty genetic risk loci of sAD have been identified by large-scale GWAS and meta-analysis (**Table 1.1**, [58]). Among these loci, *APOE* is the major risk locus and is located on chromosome 19. The function of APOE is related to the normal catabolism of triglyceride-rich lipoproteins [7]. In the central nervous system (CNS), APOE is one of the major components of lipoprotein particles, which are responsible for transmitting cholesterol from astrocytes to neurons [59, 60]. A relationship between APOE and sAD has been identified through the observation of immunoreactivity to APOE at amyloid plaques and neurofibrillary tangles [61]. APOE protein has 299 amino acids and has three isoforms in humans: APOE ε2 (Cys112 and Cys158), APOE ε3 (Cys112 and Arg158) and APOE ε4 (Arg112 and Arg158) [7]. Based on population studies, the *APOE ε4* allele increases the risk of developing Alzheimer's disease [62], while the *APOE ε2* allele significantly decreases the risk [63]. In a study by Deane et al., 2008 [64], the role of *APOE* was investigated in *APOE* knock-in mice and it was revealed that the clearance of $A\beta_{40}$ was inhibited in an allele-specific manner, showing the involvement of *APOE* in Aβ clearance.

Gene	Location	SNP	Risk allele frequency controls	OR (95% CI)	Population-attributable fraction (%)	Potential functional variant
APOE (apolipoprotein E)	19q13.32	ε4	0.16	3.78 (2.60–5.48)	30.8[a]	ε4
SORL1 (sortilin-related receptor-1)	11q24.1	rs11218343-T	0.96	1.30 (1.22–1.39)	0.91[b]	Common and rare pathogenic variants[34,35]
BIN1 (bridging integrator 1)	2q14.3	rs6733839-T	0.41	1.22 (1.18–1.25)	8.2[a]	rs59335482, 3 bp insertion[40]
CR1 (complement component (3b/4b) receptor 1)	1q32.2	rs6656401-A	0.20	1.18 (1.14–1.22)	3.5[a]	Intragenic CNV resulting in different CR1 isoforms[41]
CLU (clusterin)	8p21.1	rs9331896-T	0.62	1.16 (1.12–1.19)	5.1[b]	Rare coding and common regulatory variants[30,31]
PICALM (phosphatidylinositol-binding clathrin assembly protein)	11q14.2	rs10792832-G	0.64	1.15 (1.12–1.18)	4.5[b]	—
ABCA7 (ATP-binding cassette transporter A)	19p13.3	rs4147929-A	0.19	1.15 (1.11–1.19)	2.8[a]	Loss-of-function variants[37,38]
FERMT2 (fermitin family member 2)	14q22.1	rs17125944-C	0.09	1.14 (1.09–1.19)	1.2[a]	—
CASS4 (Cas scaffolding protein family member 4)	20q13.31	rs7274581-T	0.92	1.14 (1.09–1.19)	1.0[b]	—
MS4A6A locus (membrane-spanning 4-domains, subfamily A)	11q12.2	rs983392-A	0.60	1.11 (1.09–1.15)	3.8[b]	—
EPHA1 (EPH receptor A1)	7q35	rs11771145-G	0.66	1.11 (1.08–1.14)	3.3[b]	—
HLA-DRB5, HLA-DRB1 locus (major histocompatibility complex, class II, DR beta 5/beta 1)	6p21.32	rs9271192-C	0.28	1.11 (1.08–1.18)	3.0[a]	—
PTK2B (protein tyrosine kinase 2 beta)	8p21.2	rs28834970-C	0.37	1.10 (1.08–1.13)	3.6[a]	—
CD2AP (CD2-associated protein)	6p12.3	rs10948363-G	0.27	1.10 (1.07–1.13)	2.6[a]	—
ZCWPW1 locus (zinc finger, CW type with PWWP domain 1)	7q22.1	rs1476679-T	0.71	1.10 (1.06–1.12)	2.5[b]	—
SLC24A4/RIN3 locus (solute carrier family 24/Ras and Rab interactor 3)	14q32.12	rs10498633-G	0.78	1.10 (1.06–1.14)	1.9[b]	—
INPP5D (inositol polyphosphate-5-phosphatase)	2q37.1	rs35349669-T	0.49	1.08 (1.05–1.11)	3.8[a]	—
MEF2C (myocyte enhancer factor 2C)	5q14.3	rs190982-A	0.59	1.08 (1.05–1.11)	2.8[b]	—
NME8 locus (NME/NM23 family member 8)	7p14.1	rs2718058-A	0.63	1.08 (1.05–1.11)	2.5[b]	—
CELF1 locus (CUGBP, Elav-like family member 1)	11p11.2	rs10838725-C	0.32	1.08 (1.05–1.11)	2.5[a]	—
CD33 (CD33 molecule)	19q13.41	rs3865444-C	0.69	1.06 (1.04–1.1)	1.8[b]	rs12459419 located in a putative SRSF2 splice site of exon 2, leading to alternative splicing of the IgV domain[44]

Table 1.1. Overview of sAD susceptibility loci from GWAS and meta-analysis summarized by Van Cauwenberghe et al. in 2016 (from [58]), with the description of the most significant SNP at each locus. The ranking of significant SNPs is based on their effect size. The population-attributable fraction (PAF) types are indicated with "a" for risk PAF and "b" for preventive PAF.

OR, odds ratio; CI, confidence interval; CNV, copy-number variation.

1.4.2 EOfAD-causative mutation loci

Autosomal dominant genetic mutations leading to EOfAD have been found in the genes encoding *AMYLOID BETA A4 PRECURSOR PROTEIN* (APP), *PRESENILIN 1* (*PSEN1*) or *PRESENILIN 2* (*PSEN2*) [5] and, possibly, *SORTILIN-RELATED RECEPTOR* (*SORL1*) [65, 66]. APP is a type I transmembrane protein, encoded by the *APP* gene located on chromosome 21. APP contains cleavage sites for multiple enzymes (**Figure 1.7**, [5]), but Aβ can only be produced through cleavage by β-secretase and γ-secretase [67, 68]. Presenilin is the catalytic subunit of the γ-secretase enzyme that participates in APP cleavage. The neuronal function of APP remains ill defined, while Aβ seems to be a normal product of APP metabolism and is generated at high levels in neurons [7].

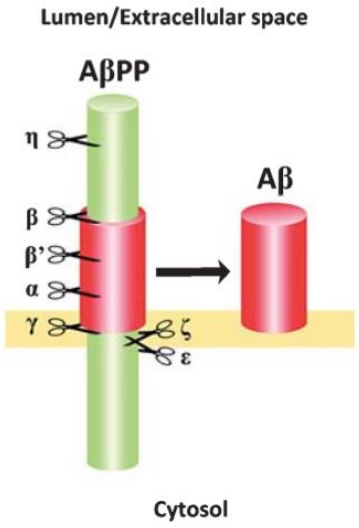


Figure 1.7. Cleavage of APP (from [5]). APP contains multiple cleavage sites, but Aβ can only be produced through cleavage at both the β-site and the γ-site.

As APP, PSEN1 and PSEN2 are all involved in the production of Aβ by γ-secretase cleavage of APP, this process is a major research focus in EOfAD. The γ-secretase

enzyme complex performs “endopeptidase” cleavages firstly to release APP’s cytosolic “intracellular domain” (AICD), and subsequently “carboxypeptidase” cleavages to shorten the remaining part of APP so that the hydrophobicity is reduced to release soluble Aβ from the lipid bilayer [5]. A cell-free study by Chavez-Gutierrez et al., 2012 [69] further investigated the γ-secretase changes due to the EOfAD mutations in *PSEN1* and *APP* gene, and found that these mutations have variable effects on the endopeptidase cleavage, indicating the impairment of the endopeptidase cleavage is not necessary for AD pathology. Another study introduced ten different EOfAD mutations into the *PSEN1* gene to investigate their effects on γ-secretase activity using an *in vitro* cleavage assay [70]. The changes in γ-secretase activity were variable, but all these mutations led to increased ratios of $A\beta_{42}/A\beta_{40}$, suggesting that the regulation of γ-secretase activity possibly is not relevant to EOfAD pathology while a relatively increased production of $A\beta_{42}$ ($A\beta_{42}/A\beta_{40}$ ratios) may contribute to disease progression. A similar result was observed in children carrying the mutation *PSEN1* E280A who had increased plasma $A\beta_{42}$ levels and $A\beta_{42}/A\beta_{40}$ ratios compared to non-carriers [71]. However, an inconsistent result was achieved through examining the *in vitro* γ-secretase activity of 138 reported EOfAD mutations in human *PSEN1*, which found decreased production of both $A\beta_{40}$ and $A\beta_{42}$ for about 90% of these mutations and decreased $A\beta_{42}/A\beta_{40}$ ratios for 10% of the mutations [24]. Therefore, the role of γ-secretase activity and $A\beta_{42}/A\beta_{40}$ ratios in AD pathology remains controversial and requires further investigation. Moreover, EOfAD-causative mutations follow a “fAD reading frame preservation rule” that the mutations do not destroy the open reading frame such that mutant transcript are always present encoding a “full-

length" protein [5]. The human EOfAD mutation, *PSEN2*^{K115fs}, which is a frame shift in *PSEN2* leading to a premature termination codon, was previously thought to be an exception to this rule [72]. However, the identification of a new transcript splicing isoform from the *PSEN2*^{K115fs} allele which restores the open reading frame and leads to production of full-length protein, indicated that the human *PSEN2*^{K115fs} mutation also follows this rule [73]. Therefore, it has been argued that production of a full-length protein may be a necessary element in disease pathology, such as for the formation of PSEN1 holoprotein dimers/multimers in the PSEN holoprotein multimer hypothesis ([5], discussed in previous ***section 1.3.3***).

The gene *SORTILIN-RELATED RECEPTOR* (*SORL1*), also known as *LR11*, is another rare EOfAD mutation locus although this status is still subject to debate (reviewed in [66]). The identified EOfAD variants in *SORL1* are illustrated in **Figure 1.8**. *SORL1* encodes a transmembrane receptor with multiple functional domains, which participates in the endosomal sorting of proteins between the trans-Golgi network (TGN), endosomes and the plasma membrane [65]. Unlike the PSEN genes for which hundreds of EOfAD variants have been defined, only a few EOfAD variants have been identified in *SORL1* gene. These are disturbed over different protein domains (**Figure 1.8**). Also, genetic associations between *SORL1* and sAD have been identified in genome-wide association studies [74-76], indicating that *SORL1* is involved in both sAD and EOfAD. SORL1 protein act in complex and essential roles in the trafficking and processing of APP within cells (reviewed in Barthelson et al., 2020 [65]). The overexpression of SORL1 leads to the accumulation of APP in the Golgi and failure to sort APP to late endosomal membranes where some Aβ is produced, while deficiency of SORL1 results in an

increase in Aβ levels [65, 77]. Therefore, the correlation between the expression of SORL1 and the production of Aβ indicates potential roles for SORL1 in Aβ-related pathologies in both sAD and EOfAD.

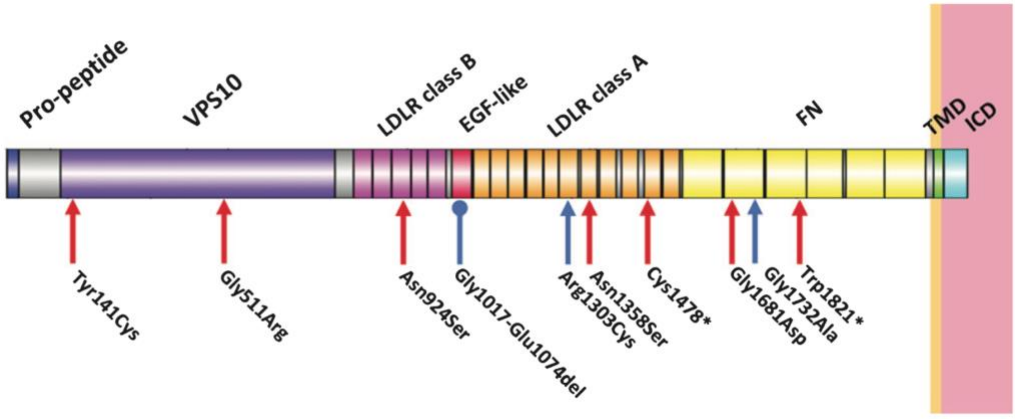


Figure 1.8. A summary of the identified EOfAD variants in the gene *SORL1* (from [65]). It encodes a protein with multiple domains: VPSP10, vacuolar protein sorting 10; LDLR, low density lipoprotein receptor; EGF, epidermal growth factor; FN, fibronectin-type; TMD, transmembrane domain; ICD, intracellular domain. The position of known coding variants are indicated by arrows. Red arrows indicate variants published in Pottier et al., 2012 [78], and blue arrows indicate variants published in Thonberg et al., 2017 [79].

1.4.3 EOfAD-causative mutations in the gene *PSEN1*

PSEN1 gene is the major EOfAD locus, and mutations in *PSEN1* gene contribute to 70%-80% of EOfAD cases [45]. More than 205 different EOfAD-causative mutations in *PSEN1* have been identified with a scattered distribution [80] (**Figure 1.9**). These mutations all follow the "fAD reading frame preservation rule", and missense mutations form a large proportion of these [5]. Interestingly, a known

frame shift mutation in human *PSEN1*, P242LfsX11, does not cause EOfAD but does cause another disease, the skin disease Acne Inversa (Hidradenitis Suppurativa) [43]. This mutation leads to a premature termination codon and may produce a truncated protein lacking γ-secretase activity. The existence and phenotypic effects of this mutation may support the PSEN holoprotein multimer hypothesis raised by Jayne et al., 2016 [5] that EOfAD-causative PSEN mutant holoproteins interact with wild type PSEN holoproteins to inhibit the formation of the normal holoprotein dimers/multimers, while the truncated PSEN proteins may be structurally unable to interact pathologically with wild type PSEN holoproteins.

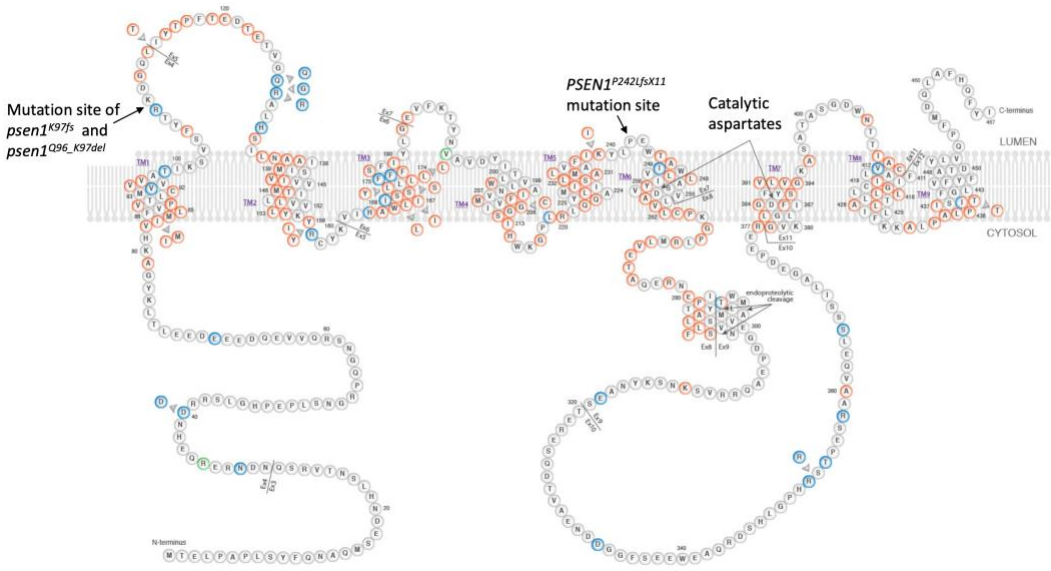


Figure 1.9. The structure of human PSEN1 protein (modified from [81]). The pathogenic EOfAD mutations are labelled in orange, while the non-pathogenic mutations are labelled in green. Mutations with unclear effects are labelled in blue. Two catalytic aspartate residues are critical for its γ-secretase activity.

The major function of PSEN1 in AD pathology is currently thought to be the γ-secretase cleavage of APP to form Aβ. However, the role of γ-secretase in AD

pathology remains controversial (see above), as the increased Aβ production or increased $A\beta_{42}/A\beta_{40}$ ratios proposed to be pathogenic are not observed for all EOfAD mutant forms of PSEN1 [24] while the truncating Acne Inversa mutation that should reduce γ-secretase activity does not cause EOfAD [43]. Therefore, further investigation of the role of γ-secretase in AD pathology is desirable, as the molecular changes caused by changes in γ-secretase activity can be excluded from consideration if the γ-secretase activity does not contribute to AD pathology. Also, other functions of PSEN1 and the pathways it is involved in have been investigated in animal models and require further exploration. It has been found that the *PSEN1* gene is essential for animal development, and *Psen1* knock-out mice (*Psen1*$^{-/-}$) are unable to survive during development [82] while the blockage of *psen1* translation using morpholino antisense oligonucleotides in zebrafish embryos does not kill the zebrafish but leads to the formation of aberrant somites and defects in Notch signaling [83-85]. Furthermore, *psen1* may act as a regulator of histaminergic neuronal development in the brains of *psen1*$^{-/-}$ zebrafish [86]. The brain histaminergic system is a major modulatory system regulating multiple functions including cognition and memory which are closely related to AD [87]. However, our understanding of PSEN1 functions is far from comprehensive, and further investigation is required to figure out what functional changes induced by EOfAD mutations directly contribute to AD pathology.

1.5 Animal models of AD

1.5.1 AD in wild animals

The accumulation of amyloid plaques, which is the major histological sign in AD, has been detected not only in human brains but also in the brains of other nonhuman primates [88]. Even dogs develop amyloid plaques as they age and have been used as a natural model of age-dependent cognitive dysfunction. The impacts of Aβ in dogs' brains on their cognition, such as their ability to learn, have been investigated in several studies [89-91]. However, it is an ongoing debate whether another essential feature of AD, the accumulation of NFTs, is present in these animals' brains [92, 93]. In the study by Schmidt et al., 2015 [93], the presence of NFTs was detected in some extremely aged canines with cognitive dysfunction syndrome, indicating that canines might also develop NFTs but that a longer life span might be required. Also, the presence of both amyloid plaques and NFTs in brains has been found in dolphins, which have a relatively long life span like humans [88]. Therefore, it has been argued that the post-reproductive life span and longevity might contribute to the onset of AD rather than aging [88].

1.5.2 Rodents models in AD studies

Due to the inherited nature of EOfAD, modeling EOfAD in an animal model is an obvious pathway for investigating the molecular changes that occur in this disease *in vivo*. As mammals, rodents are the most common animal models used in AD

research. Compared to other mammals, rodent models are easy to manipulate genetically and have a relatively short life span. The formation of amyloid plaques has been successfully identified in several transgenic models of AD that overexpress the genes associated with EOfAD [94]. However, the majority of AD therapies that have succeeded in preclinical testing using rodent models have failed in clinical trials [94], suggesting that current transgenic EOfAD-like rodent models may not develop aspects of AD pathology and do not reflect closely the disease state of human patients. Hargis and Blalock, 2017 [95] compared brain transcriptional profiles in AD patients and five different transgenic mouse models of AD, and demonstrated that the transcriptional profiles of AD patients were consistent across studies while the transcriptional profiles of transgenic mouse models were not similar to one another or to the transcriptional profiles of AD patients. Also, none of the transgenic mice showed an essential, consistent transcriptional signature identified in AD patients, a downregulation of synaptic and mitochondrial genes. Therefore, due to their failure to reflect the transcriptional changes observed in postmortem human AD brains, the transgenic rodent models probably only reflect specific pathways involved to a limited degree, and are obviously deficient for the study of human AD pathology. Indeed, each animal model has its own advantages and disadvantages and there are no perfect animal models for human AD, while a combination of models might be used to study particular aspects of AD [96]. Therefore, the use of other relevant animal models in AD studies is highly desirable, and may inform us regarding potential molecular changes that are involved in human AD pathology but that do not occur in transgenic rodent models.

1.5.3 The zebrafish as a genetic model

The zebrafish, *Danio rerio*, is a small tropical freshwater fish that is originally from the Ganges River and its tributaries in northern India [97]. The genome of the zebrafish has been fully sequenced and is well annotated [98]. Due to some obvious advantages over mammalian models, the zebrafish has become a popular vertebrate animal model that has been increasingly used in the study of human disease in recent decades. Although the divergence time between humans and zebrafish (teleosts) is approximately 450 million years [99], the zebrafish, as a vertebrate, shares more anatomical, physiological and biological properties with humans than invertebrate models such as *Drosophila melanogaster* and *Caenorhabditis* elegans [100]. The external and transparent embryos of the zebrafish are easily accessible for subtle and complex genetic manipulation and ideal for *in vivo* imaging, and the zebrafish has a short embryo development period and a relatively short generation time. Also, the zebrafish's capacity to produce large families of siblings (100+ per spawning), which can then be raised in the same environment, is desirable for achieving high sample numbers for experimental analysis with limited environmental and genetic noise [101]. Moreover, multiple genes can be manipulated effectively simultaneously in the zebrafish, something which is currently relatively difficult to achieve within rodent models [101].

1.5.4 Zebrafish models of AD

When used as an animal model in the study of AD, the zebrafish should show clear gene orthology to the relevant human genes involved in the disease. The zebrafish orthologs of the three human EOfAD genes *PSEN1*, *PSEN2* and *APP* have been identified and are denoted as *psen1* [102], *psen2* [103], *appa* and *appb* [104] (the latter are duplicate "co-orthologues" of human *APP*). Also, orthologous genes in zebrafish involved with other elements of AD pathology in human include those encoding β-secretase (*bace1* [105] and *bace2* [106]) and γ-secretase complex components (*psenen* [85, 107], *ncstn* [108] and *aph1b* [107]), the genes encoding *tau* (*mapta* and *maptb* co-orthologues in zebrafish [109]) as well as the *apoea* and *apoeb* co-orthologues of *APOE* [110] (reviewed in Newman et al., 2014 [101]). Furthermore, the normal functions of these genes in the zebrafish have been investigated respectively through the inhibition of protein translation, and several unique characteristics of these genes have been identified, which were not observed in knock-out rodent models (reviewed in Newman et al., 2014 [101]). Therefore, zebrafish EOfAD mutation models can reveal different aspects of the functions of these genes compared to rodent models, inducing specific effects on molecular interactions and cellular pathways. The combination of analyses from multiple animal models can contribute to a more thorough understanding of the pathology of human AD.

1.6 Introducing mutations into zebrafish models using genome editing technologies

To use animal models for the study of human disease, mutations equivalent to those causing disease in humans can be introduced into an animal's endogenous genes to generate animals with a genetic state resembling the human disease. Previously, the generation of equivalent mutations could only be achieved in rodent models based on homologous recombination through embryonic stem cell technologies, while in zebrafish models, the technologies available only enabled the temporary blockage of protein function through targeted knockdown approaches [111]. With the discovery of genome editing tools, the possibilities for generating desired mutations in animal models have dramatically improved. In particular, the clustered regularly interspaced short palindromic repeats (CRISPR) /CRISPR-associated gene (Cas) system has made genome editing feasible in almost any organism [112].

All genome editing tools contain two essential elements: a sequence-specific DNA targeting subunit and a double-strand DNA-cleaving nuclease. These two elements introduce double-strand breaks at a sequence of interest, and subsequently initiate innate DNA repair pathways, non-homologous end-joining (NHEJ) or homology-directed repair (HDR) (**Figure 1.10a**, [111]). NHEJ repairs the broken double-stranded DNA without a template, so that small insertions and deletions (indels) are frequently introduced into the breakpoint. Incorrect NHEJ repair of an open reading frame can lead to a frame shift and/or a premature stop codon, resulting in gene knock-out. In contrast, HDR requires a homologous DNA template, and enables the insertion of foreign DNA into a breakpoint to generate a desired mutation [113].

Three major genome editing tools, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and the type II prokaryotic CRISPR/Cas9 system, are frequently used in animal models including zebrafish. In both ZFNs and TALENs, DNA-binding proteins fused to the Fok1 endonuclease are used to target DNA and subsequently cause cleavage (**Figure 1.10b and 1.10c**), while the CRISPR/Cas9 system uses a guide RNA (gRNA) to guide the Cas9 nuclease to the desired genomic locus (**Figure 1.10d**). ZFNs were the first genome editing tool used in zebrafish [114], but the complicated design of zinc finger motifs, time-consuming *in vitro* tests of binding affinity, and their high cost limited the use of this technology in the zebrafish [111]. In contrast to ZFNs, TALENs have a more predictable binding affinity and the design of a transcription activator-like effector (TALE) is technically easier [111]. Also, the high specificity of TALENs to the targeting sequence leads to rare off-targets effects [115, 116], which is preferable for genome editing. The CRISPR/Cas9 system identified from bacteria and archea is built on a different concept, having different technical features. The CRISPR/Cas9 system is the most cost-efficient and technically easiest to use, and the generation of mutations can be achieved easily in most laboratories [111]. Also, the influence of off-target mutations caused by the CRISPR/Cas9 system can be significantly reduced by outcrossing edited zebrafish to wild type fish [111]. Furthermore, the recent discovery of the class II type V CRISPR/Cas12a (Cpf1) system, which is similar to the CRISPR/Cas9 system but uses a different RNA guided endonuclease recognizing a different PAM site, typically 5'-TTTV-3', provides an alternative selection of sequence target sites [117]. Therefore, the use

of CRISPR/Cas systems is an appropriate option for the generation of desired mutations in zebrafish.

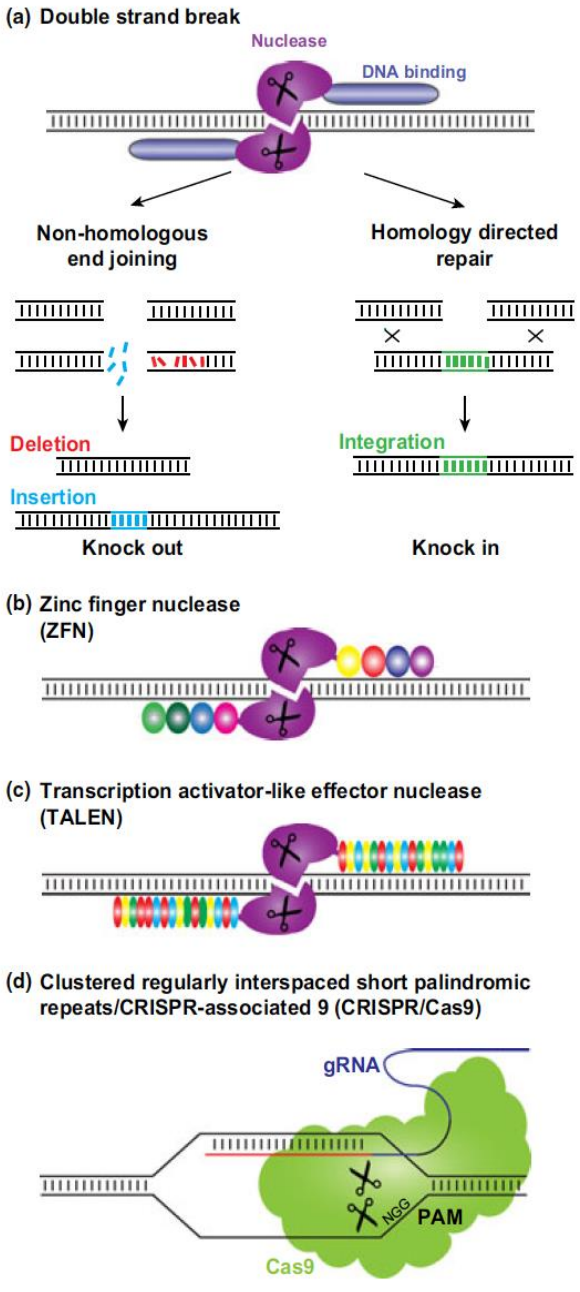


Figure 1.10. Mechanisms of genome editing technologies (from [111]).

(a) Schematic representation of genome editing to generate a double-stranded DNA break and subsequently introduce mutations through DNA repair.

(b) Schematic representation of ZFNs. Left and right ZFNs bind near each other to constitute an active Fok1 nuclease to induce a double-stranded break. Three to four zinc fingers are used on each side, and each zinc finger recognizes three nucleotides.

(c) Schematic representation of TALENs. Left and right TALENs bind near each other to constitute an active Fok1 nuclease to induce a double-stranded break. 16 TALE elements are used on each side, and each TALE recognizes one nucleotide.

(d) Schematic representation of the CRISPR-/Cas9 system. The target site for cleavage is a three-nucleotide NGG motif adjacent to a 20-nucleotide sequence called the protospacer adjacent motif (PAM). A specific guide RNA (gRNA) with a 20-nucleotide sequence complementary to the target and a sequence that recruits the Cas9 nuclease is shown.

1.7 Methods for quantification of gene expression

Assessment of a transcriptome involves description of a set of RNA transcripts and their quantity in a cell, tissue, or whole organism. Through transcriptome analysis, the global regulation genes in a sample can be assessed at a particular time point, and the changes in cellular function can be predicted. Therefore, transcriptome analysis is widely used for the investigation of disease-specific effects at genetic levels. DNA microarrays, a hybridization-based approach, was an early technique for transcriptome analysis. However, it has be limited by its reliance on existing knowledge about genome sequences, high background signals induced by cross-hybridization of probes with similar rather than identical RNA sequences [118], and a limited range of detection due to background and the saturation of signals [119]. With the development of next-generation sequencing, DNA microarrays have been almost replaced by sequenced-based approaches, RNA-seq. A general workflow of RNA-seq is illustrated in **Figure 1.11**. A population of RNA is converted into a library of cDNA fragments through either RNA fragmentation followed by reverse transcription or reverse transcription followed by cDNA fragmentation. With adaptors attached to one or both ends, each fragment is then able to be sequenced from one end (single-end sequencing) or both ends (paired-end sequencing) using high-throughput sequencing technology. The short reads achieved by sequencing can be aligned to a reference genome or to reference transcripts, or assembled *de novo* [119]. Gene quantification can be achieved through counting the reads aligned to each gene.

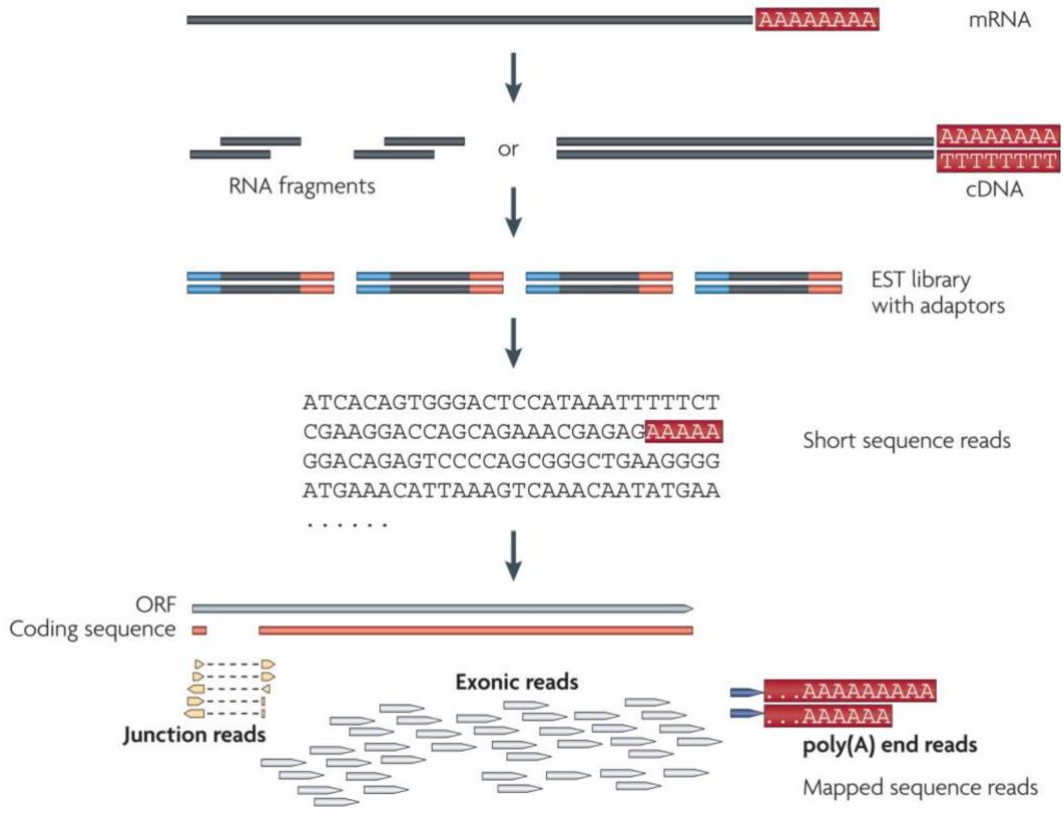


Figure 1.11. The workflow of a typical RNA-Seq experiment (from [119]). mRNAs are converted into a library of cDNA fragments through either RNA fragmentation followed by reverse trancription or reverse trancription followed by cDNA fragmentation. Each cDNA fragment is then attached to a sequencing adaptor (blue). A short sequence read is subsequently achieved from each cDNA through high-throughput sequencing technology.

RNA-seq can provide a comprehensive overview of gene expression and regulation in a cell, tissue, or organism. The expression of genes of interest is subsequently normally assessed by reverse transcription quantitative PCR (RT-qPCR). The relative standard curve method of qPCR is based on the detection of fluorescence to monitor the production of amplification products during each cycle of a PCR monitored in real-time [120]. In this technique, an external standard curve for PCR of a particular gene is generated through performing a series of separate real-time PCRs on a serial dilution of a unknown sample. The standard curve produced can

then be used to estimate the quantity of that gene in unknown samples. The relative quantification of that gene from the standard curve can be normalized to that of a house-keeping gene in the same sample, and the normalized numbers enable the comparison of gene expression between samples.

Digital qPCR (dqPCR) is an alternative technique for gene quantification (**Figure 1.12**). DNA templates are firstly diluted and separated into individual reactions in cells (approximately 1 copy per two cells/reactions), and the presence of PCR products in each cell is detected by fluorescent probes [121]. The gene concentration can be estimated through statistical analysis, as positive signals should be distributed according to Poisson probabilities [122]. Compared to traditional standard curve qPCR, digital qPCR has the advantages of absolute quantification without requiring a standard curve.

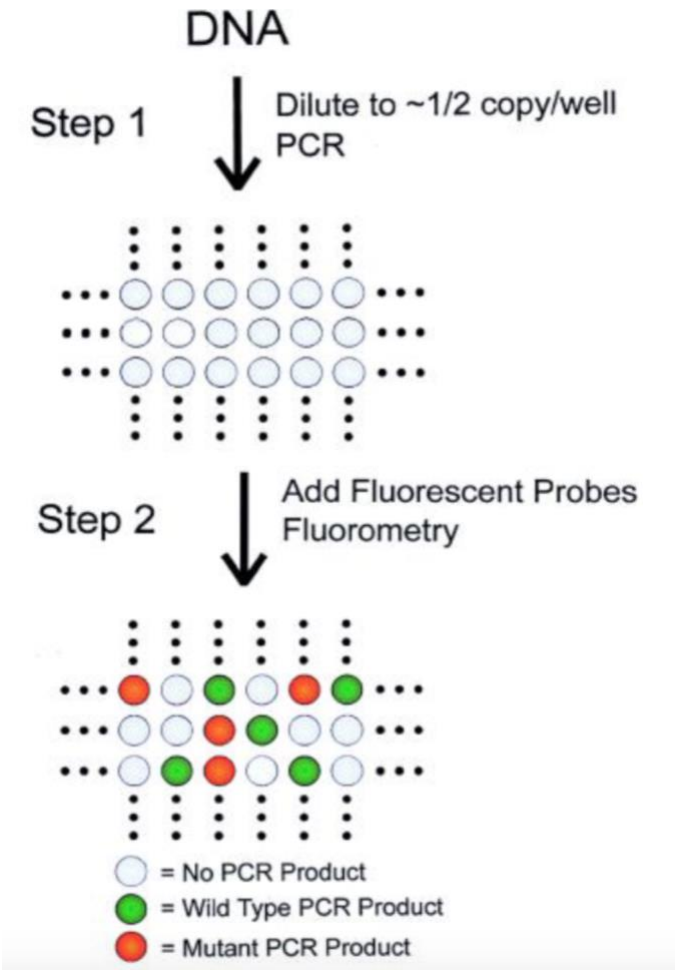


Figure 1.12. Schematic of Digital PCR (from [121]). DNA templates are diluted and separated into individual reaction cells, and the presence of specific PCR products can be detected using corresponding fluorescence probes. Template concentrations are estimated by the subsequent statistical analysis.

1.8 Behavioural Analysis: Testing of short term working memory in animal models using Y-maze

Cognitive impairment, including decline in memory and other executive functions, is the most significant clinical symptom of AD. Working memory involves temporary storage, which is known as short-term memory, and manipulation of information that is used for various cognitive activities [20, 123]. The most current multiple-component model of working memory consists of four subcomponents, the central executive, the phonological loop, the visuospatial sketch pad and the episodic buffer [123]. Spatial memory, which is related to the storage and retrieval of information for the identification and navigation of proximal or distal space [124], participates in the subcomponent visuospatial sketch pad. A study by Guariglia, 2007 [125] used the Corsi block-trapping test to assess short-term spatial working memory in AD patients. Impairments in short-term spatial working memory were observed in AD patients with moderate dementia, while those with mild dementia did not show a significant decline. Therefore, short-term spatial working memory can be used for assessment of cognitive decline in EOfAD models. Although short-term spatial memory in animals cannot be assessed as easily as in humans, the Free-movement pattern (FMP) Y-maze is a rapid method that can be employed for this purpose. The original Y-maze was designed to detect changes in short-term spatial working memory through assessment of spontaneous alternative arm choice based on the idea that animals are driven by an innate curiosity to explore unvisited areas and thus trend to enter a less recently visited arm [126]. No training of the animals is required for this test. A high percentage of entries into consecutive arms is

regarded as indicating superior spatial working memory performance, as the tested animal can recall the arms it has already visited and enter another arm, resulting in a similar proportion of entries to each arm over time (**Figure 1.13**, [127]).

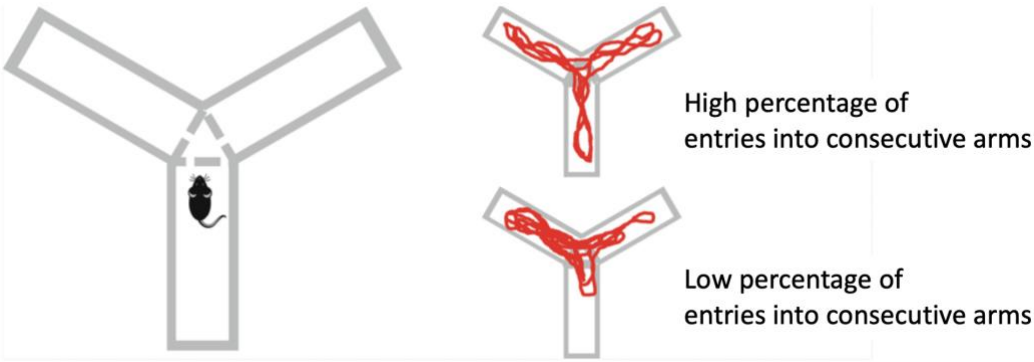


Figure 1.13. Assessment of spontaneous alternative arm choice in the Y-maze. A high percentage of entries into consecutive arms is reflected as a similar proportion of entries into each of the arms (**upper**) indicating good spatial working memory, while a low percentage of entries into consecutive arms is reflected as a higher proportion of repeated entries into the same arm (**lower**) regarded as poor spatial working memory [127].

However, the assessment of spontaneous alternative arm choice using the original Y-maze has limitations. Difficulties in interpreting results have been reported in several studies, especially when the tested models showed hypo- or hyperlocomotion, stereotyped behaviour, or anxiety-related novelty avoidance that could significantly affect the assessment of animals' spontaneous alternative arm choice [128-132]. For example, in a behavioural study of the Tg2576 AD rodent model, these mice showed a perfectly stereotyped behaviour of always making right turns which results in equal entries into each arm during the test period, consequently giving a high percentage of entries into consecutive arms [129]. However, this performance was not due to superior spatial working memory but

was a result of the rodent model's stereotyped behaviour. To overcome these limitations, the FMP Y-maze has been designed by Cleal et al., 2020 [132], which shifts the focus from novelty response to navigational search pattern by recording sequences of animals' left and right turns apportioned into 16 possible tetragrams consisting of 4 turns in a 1-hour runtime of free exploration to identify strategic search patterns. This method is advanced in classifying movement patterns affected by other factors, such as stereotyped behaviour, into different search strategies, which do not overlap with normal search strategies. The search strategies used by vertebrates, including mice, zebrafish and humans, mainly consist of alternations (LRLR and RLRL tetragrams), and impairment of working memory would cause a reduction in these alternations, indicating the potential application of FMP Y-maze in the assessment of working memory [132].

There are several behavioural studies performed on zebrafish to assess cognition. A study by Aoki et al., 2015 [133] performed a 2-hour training session using a specific colour paired with electric shock on adult zebrafish. The zebrafish learned to choose the correct arm at an efficiency of 89%, demonstrating their memory and learning abilities. Another study by Cognato et al., 2012 [134] used a Y-maze memory task with a training session that blocks one arm as the novel arm to assess changes in zebrafish behaviour induced by either MK-801 or scopolamine dissolved in tank water. These drugs produced deficits in memory. Treatment of zebrafish pre-training with either MK-801 or scopolamine or post-training with MK-801 resulted in reduced exploration of the novel arm, highlighting the potential application of the zebrafish model to the study of cognition and memory. As impairment of short-term spatial working memory is one of the clinical symptoms

of AD and has been found in AD patients with moderate dementia [125], it will be a straightforward strategy to determine whether impaired short-term working memory can be observed in EOfAD-like zebrafish mutants through testing in a FMP Y-maze without training. In future, other cognitive abilities, such as learning ability and spatial reference memory, in EOfAD-like zebrafish mutants could be investigated further in Y-maze studies involving specific training treatments.

1.9 Previous work of the Alzheimer's Disease Genetics laboratory on presenilin mutations

1.9.1 The generation of two EOfAD-like mutations in the zebrafish gene *psen1*

To investigate the changes in brain molecular state induced by EOfAD mutations, we have previously introduced two EOfAD-like mutations into the zebrafish *psen1* gene, *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$. These mimic human EOfAD-causative mutations, and the location of the equivalent site on human PSEN1 protein for these mutations is highlighted in **Figure 1.9**.

The *psen1*^{K97fs} mutation is a 2 nucleotide deletion, leading to a reading frame shift and a premature termination codon. This mutation mimics the K115fs mutation of human *PSEN2*, which is a frameshift mutation leading to an early termination codon and was previously thought to represent the only known exception to the "fAD reading frame preservation rule" of EOfAD mutations in *presenilin* genes [72]. (However, the human mutation was subsequently shown to obey this rule, see below). The human *PSEN2* mutation results in transcripts encoding a truncated PSEN2 protein similar to that produced by a naturally occurring alternative transcript of the *PSEN2* gene, PS2V. In PS2V transcripts, exon 5 of PSEN2 is excluded leading to a frameshift and a premature termination codon [135]. The alternative splicing event that generates PS2V is induced by hypoxia and is observed at higher levels in sAD brains [136]. PS2V-like isoforms have been observed in the brains of guinea pigs and are predicted to occur in most other

mammals [137], but alternative PS2V-like splicing has been identified for the zebrafish *psen2* gene under hypoxia [138]. Instead, the Alzheimer's Disease Genetics Laboratory (ADGL) found that, in zebrafish, a PS2V-like isoform is produced from the paralogous *psen1* gene, and has been named PS1IV [136]. Although human PS2V and zebrafish PS1IV share little structural similarity, they both stimulate γ-secretase activity and suppress the unfolded protein response under hypoxia [136]. Therefore, in attempting to model the human $PSEN2^{K115fs}$ mutation in zebrafish, equivalent mutation was introduced into the *psen1* gene (rather than *psen2*) and is designated $psen1^{K97fs}$. In 2019, a new transcript splicing isoform from the human $PSEN2^{K115fs}$ allele was identified that restores the open reading frame [73], which means that this mutation also follows the "fAD mutation reading frame preservation rule" [5] like all other EOfAD-causative mutations in the *presenilin* genes. In this thesis I demonstrate that no such isoform can be identified for transcripts of the $psen1^{K97fs}$ allele, indicating that the $psen1^{K97fs}$ mutation is not a good model of $PSEN2^{K115fs}$ as it does not follow the reading frame preservation rule of EOfAD mutations.

Another mutation of *psen1* that is EOfAD-like, $psen1^{Q96\_K97del}$, has a deletion of 6 nucleotides in *psen1* gene, but maintains the open reading frame, leading to structural and hydrophilicity changes in the first lumenal loop of the translated protein (**Figure 1.9**). This mutation is not an exact equivalent of any currently known human EOfAD mutation, but is similar to numerous EOfAD mutations that distort the first luminal loop of human PSEN1, such as $PSEN1^{L113\_I114insT}$ [139] and $PSEN1^{P117L}$ [140]. Like all the various and widely distributed human EOfAD mutations in the *PRESENILIN* genes, this mutation follows the "fAD mutation

reading frame preservation rule" [5]. Most of the analyses conducted in this thesis use the *psen1*$^{Q96\_K97del}$ mutation in the heterozygous state as a model of EOfAD-like mutation.

1.9.2 Brain transcriptome analysis of two heterozygous EOfAD-like mutants

To investigate the molecular changes induced by the *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$ mutations, brain transcriptome analyses have been performed on young adult zebrafish (6 months of age, recently sexually mature) and aged adult zebrafish (24 months of age) [141, 142]. Numerous differentially expressed (DE) genes have been identified in each mutant strain at both ages. There is a set of common DE genes identified in both young adult and aged adult zebrafish brains heterozygous for the *psen1*^{K97fs} mutation, but showing transcriptional "inversion" (the regulation of these genes was in opposite directions at the two ages; **Figure 2.2**, discussed in following ***sections 2.1 and 2.2.1***) [141]. A similar transcriptional inversion has also been seen in humans in a postmortem transcriptome comparison of mild cognitive impairment (MCI), Alzheimer's disease and age-matched normal brains [143]. This found that the expression of particular genes related to synaptic function, energy metabolism and protein homeostasis were predominantly upregulated in patients with mild cognitive impairment but downregulated in AD brains. As the process driving disease progression remains unknown, we have proposed two possible models by which this change in gene expression could occur. The expression of the

genes undergoing "inversion" could 1) change smoothly during disease progression or 2) it could change rapidly at a certain stage in the progression. Further experiments are required to clarify this issue.

The cellular functions and pathways involved in disease progression in the heterozygous *psen1*^{K97fs} zebrafish brains have been predicted through gene ontology analysis (GO) and gene set enrichment analysis (GSEA). These analyses predict accelerated brain aging and increased glucocorticoid signaling in young adult heterozygous *psen1*^{K97fs} mutant brains, while the inversion of gene expression in aged adult heterozygous *psen1*^{K97fs} mutant brains may reflect glucocorticoid resistance [141]. The GO analysis of young adult heterozygous *psen1*$^{Q96\_K97del}$ mutant brain transcriptomes indicates impairment of mitochondrial function, particularly energy production (ATP synthesis), and lysosomal dysfunction [142].

1.10 Project aims

In our previous work, transcriptome analyses have only been perform on the young adult and aged adult zebrafish brains of two heterozygous EOfAD-like mutations, *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$. The exploration of these two EOfAD-like models is far from comprehensive. Therefore, this thesis focuses mainly on further analyses of these two mutations and attempts to address the following questions:

1) whether the gene expression changes during aging (e.g. the inversion of DE gene expression seen in *psen1*^{K97fs}) occur smoothly or show rapid change at a particular age;

2) whether the *psen1*^{K97fs} mutation promotes alternative splicing of transcripts as seen for human *PSEN2*^{K115fs};
3) whether heterozygous *psen1*$^{Q96\_K97del}$ zebrafish larvae show similar bulk transcriptome signatures to those identified in heterozygous 6-month-old zebrafish adult *psen1*$^{Q96\_K97del}$ brains;
4) whether (and, if so, at what age) impairments of short-term spatial working memory can be detected in heterozygous mutant adult zebrafish by the assessment of FMP Y-maze tests.

In addition, two additional *psen1* mutations were generated to investigate the similarities and differences in their effects on the brain transcriptome from *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$ heterozygous mutants. One is an Acne Inversa-like frame shift mutation, which truncates PSEN1 protein but does not cause EOfAD in humans. This Acne Inversa-like mutation will help exclude irrelevant molecular changes caused by impaired protein functions from the transcriptome "signature" of EOfAD mutations that we are attempting to define. The second mutation will be a missense mutation to alter one of the catalytic aspartate residues in the Psen1 protein. This should cause production of a full-length Psen1 protein lacking γ-secretase activity, which can be used to investigate the involvement of γ-secretase activity in the EOfAD brain transcriptome signature.

1.11 References

1. Gaugler, J., et al., *2019 Alzheimer's disease facts and figures.* Alzheimers & Dementia, 2019. **15**(3): p. 321-387.
2. Kepp, K.P., *Ten Challenges of the Amyloid Hypothesis of Alzheimer's Disease.* J Alzheimers Dis, 2017. **55**(2): p. 447-457.
3. Masters, C.L., et al., *Alzheimer's disease.* Nature Reviews Disease Primers, 2015. **1**.
4. Beydoun, M.A., et al., *Epidemiologic studies of modifiable factors associated with cognition and dementia: systematic review and meta-analysis.* BMC Public Health, 2014. **14**: p. 643.
5. Jayne, T., et al., *Evidence For and Against a Pathogenic Role of Reduced gamma-Secretase Activity in Familial Alzheimer's Disease.* J Alzheimers Dis, 2016. **52**(3): p. 781-99.
6. Barber, R.C., *The genetics of Alzheimer's disease.* Scientifica (Cairo), 2012. **2012**: p. 246210.
7. Masters, C.L., et al., *Alzheimer's disease.* Nat Rev Dis Primers, 2015. **1**: p. 15056.
8. Tosto, G. and C. Reitz, *Genome-wide association studies in Alzheimer's disease: a review.* Curr Neurol Neurosci Rep, 2013. **13**(10): p. 381.
9. Whitehouse, P., *Is There Alzheimer's Disease?* 2015: Journal of Alzheimer's Disease.
10. Nixon, R.A., *Autophagy, amyloidogenesis and Alzheimer disease.* J Cell Sci, 2007. **120**(Pt 23): p. 4081-91.
11. Alzheimer, A., *Über eine eigenartige Erkrankung der Hirnrinde.* Allgemeine Zeitschrift fur Psychiatrie und Psychisch-gerichtliche Medizin, 1907. **64**: p. 146-148.
12. DeTure, M.A. and D.W. Dickson, *The neuropathological diagnosis of Alzheimer's disease.* Mol Neurodegener, 2019. **14**(1): p. 32.
13. Jack, C.R., Jr., et al., *Prevalence of Biologically vs Clinically Defined Alzheimer Spectrum Entities Using the National Institute on Aging-Alzheimer's Association Research Framework.* JAMA Neurol, 2019.

14. Goedert, M. and M.G. Spillantini, *A century of Alzheimer's disease.* Science, 2006. **314**(5800): p. 777-81.
15. Sperling, R.A., et al., *Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.* Alzheimers Dement, 2011. **7**(3): p. 280-92.
16. Kelley, B.J. and R.C. Petersen, *Alzheimer's disease and mild cognitive impairment.* Neurol Clin, 2007. **25**(3): p. 577-609, v.
17. Rosen, J. and G.S. Zubenko, *Emergence of psychosis and depression in the longitudinal evaluation of Alzheimer's disease.* Biol Psychiatry, 1991. **29**(3): p. 224-32.
18. Teri, L., E.B. Larson, and B.V. Reifler, *Behavioral disturbance in dementia of the Alzheimer's type.* J Am Geriatr Soc, 1988. **36**(1): p. 1-6.
19. Albert, M.S., et al., *The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.* Alzheimers Dement, 2011. **7**(3): p. 270-9.
20. Cascella, M. and Y. Al Khalili, *Short Term Memory Impairment*, in *StatPearls*. 2020: Treasure Island (FL).
21. Herrup, K., *Reimagining Alzheimer's disease--an age-based hypothesis.* J Neurosci, 2010. **30**(50): p. 16755-62.
22. Hardy, J. and D.J. Selkoe, *The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics.* Science, 2002. **297**(5580): p. 353-6.
23. Wisniewski, K.E., et al., *Alzheimer's disease in Down's syndrome: clinicopathologic studies.* Neurology, 1985. **35**(7): p. 957-61.
24. Sun, L., et al., *Analysis of 138 pathogenic mutations in presenilin-1 on the in vitro production of Abeta42 and Abeta40 peptides by gamma-secretase.* Proc Natl Acad Sci U S A, 2017. **114**(4): p. E476-E485.
25. Mohandas, E., V. Rajmohan, and B. Raghunath, *Neurobiology of Alzheimer's disease.* Indian J Psychiatry, 2009. **51**(1): p. 55-61.

26. Kametani, F. and M. Hasegawa, *Reconsideration of Amyloid Hypothesis and Tau Hypothesis in Alzheimer's Disease.* Front Neurosci, 2018. **12**: p. 25.
27. Iqbal, K. and I. Grundke-Iqbal, *Alzheimer neurofibrillary degeneration: significance, etiopathogenesis, therapeutics and prevention.* J Cell Mol Med, 2008. **12**(1): p. 38-55.
28. Area-Gomez, E. and E.A. Schon, *On the Pathogenesis of Alzheimer's Disease: The MAM Hypothesis.* FASEB J, 2017. **31**(3): p. 864-867.
29. Fraser, T., H. Tayler, and S. Love, *Fatty acid composition of frontal, temporal and parietal neocortex in the normal human brain and in Alzheimer's disease.* Neurochem Res, 2010. **35**(3): p. 503-13.
30. Pettegrew, J.W., et al., *Brain membrane phospholipid alterations in Alzheimer's disease.* Neurochem Res, 2001. **26**(7): p. 771-82.
31. Stefani, M. and G. Liguri, *Cholesterol in Alzheimer's disease: unresolved questions.* Curr Alzheimer Res, 2009. **6**(1): p. 15-29.
32. Pani, A., et al., *Altered cholesterol ester cycle in skin fibroblasts from patients with Alzheimer's disease.* J Alzheimers Dis, 2009. **18**(4): p. 829-41.
33. Pani, A., et al., *Accumulation of neutral lipids in peripheral blood mononuclear cells as a distinctive trait of Alzheimer patients and asymptomatic subjects at risk of disease.* BMC Med, 2009. **7**: p. 66.
34. Gomez-Ramos, P. and M. Asuncion Moran, *Ultrastructural localization of intraneuronal Abeta-peptide in Alzheimer disease brains.* J Alzheimers Dis, 2007. **11**(1): p. 53-9.
35. Bezprozvanny, I. and M.P. Mattson, *Neuronal calcium mishandling and the pathogenesis of Alzheimer's disease.* Trends Neurosci, 2008. **31**(9): p. 454-63.
36. Lindholm, D., H. Wootz, and L. Korhonen, *ER stress and neurodegenerative diseases.* Cell Death Differ, 2006. **13**(3): p. 385-92.
37. Gillardon, F., et al., *Proteomic and functional alterations in brain mitochondria from Tg2576 mice occur before amyloid plaque deposition.* Proteomics, 2007. **7**(4): p. 605-616.

38. Wang, X., et al., *The role of abnormal mitochondrial dynamics in the pathogenesis of Alzheimer's disease.* J Neurochem, 2009. **109 Suppl 1**: p. 153-9.
39. Hayashi, T., et al., *MAM: more than just a housekeeper.* Trends Cell Biol, 2009. **19**(2): p. 81-8.
40. Puglielli, L., et al., *Role of acyl-coenzyme a: cholesterol acyltransferase activity in the processing of the amyloid precursor protein.* J Mol Neurosci, 2004. **24**(1): p. 93-6.
41. Puglielli, L., et al., *Acyl-coenzyme A: cholesterol acyltransferase modulates the generation of the amyloid beta-peptide.* Nat Cell Biol, 2001. **3**(10): p. 905-12.
42. Area-Gomez, E., et al., *Presenilins are enriched in endoplasmic reticulum membranes associated with mitochondria.* Am J Pathol, 2009. **175**(5): p. 1810-6.
43. Newman, M., 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): p. 602-17.
44. Area-Gomez, E., et al., *Upregulated function of mitochondria-associated ER membranes in Alzheimer disease.* EMBO J, 2012. **31**(21): p. 4106-23.
45. Dai, M.H., et al., *The genes associated with early-onset Alzheimer's disease.* Oncotarget, 2018. **9**(19): p. 15132-15143.
46. Zlokovic, B.V., *Neurovascular pathways to neurodegeneration in Alzheimer's disease and other disorders.* Nat Rev Neurosci, 2011. **12**(12): p. 723-38.
47. Tamagno, E., et al., *Oxidative stress activates a positive feedback between the gamma- and beta-secretase cleavages of the beta-amyloid precursor protein.* J Neurochem, 2008. **104**(3): p. 683-95.
48. Sun, X., et al., *Hypoxia facilitates Alzheimer's disease pathogenesis by up-regulating BACE1 gene expression.* Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18727-32.

49. Guglielmotto, M., et al., *The up-regulation of BACE1 mediated by hypoxia and ischemic injury: role of oxidative stress and HIF1alpha.* J Neurochem, 2009. **108**(4): p. 1045-56.
50. Winkler, E.A., A.P. Sagare, and B.V. Zlokovic, *The pericyte: a forgotten cell type with important implications for Alzheimer's disease?* Brain Pathol, 2014. **24**(4): p. 371-86.
51. Attems, J. and K.A. Jellinger, *The overlap between vascular disease and Alzheimer's disease--lessons from pathology.* BMC Med, 2014. **12**: p. 206.
52. Iliff, J.J., et al., *Impairment of glymphatic pathway function promotes tau pathology after traumatic brain injury.* J Neurosci, 2014. **34**(49): p. 16180-93.
53. Glass, C.K., et al., *Mechanisms underlying inflammation in neurodegeneration.* Cell, 2010. **140**(6): p. 918-34.
54. McGeer, P.L., M. Schulzer, and E.G. McGeer, *Arthritis and anti-inflammatory agents as possible protective factors for Alzheimer's disease: a review of 17 epidemiologic studies.* Neurology, 1996. **47**(2): p. 425-32.
55. Stewart, W.F., et al., *Risk of Alzheimer's disease and duration of NSAID use.* Neurology, 1997. **48**(3): p. 626-32.
56. Vlad, S.C., et al., *Protective effects of NSAIDs on the development of Alzheimer disease.* Neurology, 2008. **70**(19): p. 1672-7.
57. Du, X., X. Wang, and M. Geng, *Alzheimer's disease hypothesis and related therapies.* Transl Neurodegener, 2018. **7**: p. 2.
58. Van Cauwenberghe, C., C. Van Broeckhoven, and K. Sleegers, *The genetic landscape of Alzheimer disease: clinical implications and perspectives.* Genet Med, 2016. **18**(5): p. 421-30.
59. Liu, C.C., et al., *Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy.* Nature Reviews Neurology, 2013. **9**(2): p. 106-118.
60. Vance, J.E., *Dysregulation of cholesterol balance in the brain: contribution to neurodegenerative diseases.* Disease Models & Mechanisms, 2012. **5**(6): p. 746-755.

61. Namba, Y., et al., *Apolipoprotein E immunoreactivity in cerebral amyloid deposits and neurofibrillary tangles in Alzheimer's disease and kuru plaque amyloid in Creutzfeldt-Jakob disease.* Brain Res, 1991. **541**(1): p. 163-6.
62. Farrer, L.A., 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): p. 1349-56.
63. Corder, E.H., et al., *Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families.* Science, 1993. **261**(5123): p. 921-3.
64. Deane, R., et al., *apoE isoform-specific disruption of amyloid beta peptide clearance from mouse brain.* J Clin Invest, 2008. **118**(12): p. 4002-13.
65. Barthelson, K., M. Newman, and M. Lardelli, *Sorting Out the Role of the Sortilin-Related Receptor 1 in Alzheimer's Disease.* J Alzheimers Dis Rep, 2020. **4**(1): p. 123-140.
66. Campion, D., C. Charbonnier, and G. Nicolas, *SORL1 genetic variants and Alzheimer disease risk: a literature review and meta-analysis of sequencing data.* Acta Neuropathol, 2019. **138**(2): p. 173-186.
67. Serneels, L., et al., *Differential contribution of the three Aph1 genes to gamma-secretase activity in vivo.* Proc Natl Acad Sci U S A, 2005. **102**(5): p. 1719-24.
68. Vassar, R., et al., *Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.* Science, 1999. **286**(5440): p. 735-41.
69. Chavez-Gutierrez, L., et al., *The mechanism of gamma-Secretase dysfunction in familial Alzheimer disease.* EMBO J, 2012. **31**(10): p. 2261-74.
70. Bai, X.C., et al., *An atomic structure of human gamma-secretase.* Nature, 2015. **525**(7568): p. 212-7.
71. Quiroz, Y.T., et al., *Brain Imaging and Blood Biomarker Abnormalities in Children With Autosomal Dominant Alzheimer Disease: A Cross-Sectional Study.* JAMA Neurol, 2015. **72**(8): p. 912-9.

72. Jayadev, S., et al., *Alzheimer's disease phenotypes and genotypes associated with mutations in presenilin 2.* Brain, 2010. **133**(Pt 4): p. 1143-54.
73. Braggin, J.E., et al., *Alternative splicing in a presenilin 2 variant associated with Alzheimer disease.* Ann Clin Transl Neurol, 2019. **6**(4): p. 762-777.
74. Miyashita, A., et al., *SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians.* PLoS One, 2013. **8**(4): p. e58618.
75. Wen, Y., et al., *SORL1 is genetically associated with neuropathologically characterized late-onset Alzheimer's disease.* J Alzheimers Dis, 2013. **35**(2): p. 387-94.
76. Meng, Y., et al., *Association between SORL1 and Alzheimer's disease in a genome-wide study.* Neuroreport, 2007. **18**(17): p. 1761-4.
77. Yin, R.H., J.T. Yu, and L. Tan, *The Role of SORL1 in Alzheimer's Disease.* Mol Neurobiol, 2015. **51**(3): p. 909-18.
78. Pottier, C., et al., *High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease.* Mol Psychiatry, 2012. **17**(9): p. 875-9.
79. Thonberg, H., 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): p. 43.
80. Cruts, M., J. Theuns, and C. Van Broeckhoven, *Locus-specific mutation databases for neurodegenerative brain diseases.* Hum Mutat, 2012. **33**(9): p. 1340-4.
81. *MUTATIONS PSEN-1*. n.d. [cited 2020 1 July]; Available from: https://www.alzforum.org/mutations/psen-1.
82. Shen, J., et al., *Skeletal and CNS defects in Presenilin-1-deficient mice.* Cell, 1997. **89**(4): p. 629-39.
83. Nornes, S., et al., *Developmental control of Presenilin1 expression, endoproteolysis, and interaction in zebrafish embryos.* Exp Cell Res, 2003. **289**(1): p. 124-32.

84. Nornes, S., et al., *Interference with splicing of Presenilin transcripts has potent dominant negative effects on Presenilin activity.* Human Molecular Genetics, 2008. **17**(3): p. 402-412.
85. Campbell, W.A., et al., *Zebrafish lacking Alzheimer presenilin enhancer 2 (Pen-2) demonstrate excessive p53-dependent apoptosis and neuronal loss.* J Neurochem, 2006. **96**(5): p. 1423-40.
86. Sundvik, M., Y.C. Chen, and P. Panula, *Presenilin1 regulates histamine neuron development and behavior in zebrafish, danio rerio.* J Neurosci, 2013. **33**(4): p. 1589-97.
87. Kohler, C.A., et al., *Histaminergic mechanisms for modulation of memory systems.* Neural Plast, 2011. **2011**: p. 328602.
88. Gunn-Moore, D., et al., *Alzheimer's disease in humans and other animals: A consequence of postreproductive life span and longevity rather than aging.* Alzheimers & Dementia, 2018. **14**(2): p. 195-204.
89. Borghys, H., et al., *Young to Middle-Aged Dogs with High Amyloid-beta Levels in Cerebrospinal Fluid are Impaired on Learning in Standard Cognition tests.* J Alzheimers Dis, 2017. **56**(2): p. 763-774.
90. Cotman, C.W. and E. Head, *The Canine (Dog) Model of Human Aging and Disease: Dietary, Environmental and Immunotherapy Approaches.* Journal of Alzheimers Disease, 2008. **15**(4): p. 685-707.
91. Studzinski, C.M., J.A. Araujo, and N.W. Milgram, *The canine model of human cognitive aging and dementia: Pharmacological validity of the model for assessment of human cognitive-enhancing drugs.* Progress in Neuro-Psychopharmacology & Biological Psychiatry, 2005. **29**(3): p. 489-498.
92. Youssef, S.A., et al., *Pathology of the Aging Brain in Domestic and Laboratory Animals, and Animal Models of Human Neurodegenerative Diseases.* Vet Pathol, 2016. **53**(2): p. 327-48.
93. Schmidt, F., et al., *Detection and Quantification of beta-Amyloid, Pyroglutamyl Abeta, and Tau in Aged Canines.* J Neuropathol Exp Neurol, 2015. **74**(9): p. 912-23.

94. Drummond, E. and T. Wisniewski, *Alzheimer's disease: experimental models and reality.* Acta Neuropathol, 2017. **133**(2): p. 155-175.
95. Hargis, K.E. and E.M. Blalock, *Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us?* Behav Brain Res, 2017. **322**(Pt B): p. 311-328.
96. Neff, E.P., *Animal models of Alzheimer's disease embrace diversity.* Lab Anim (NY), 2019. **48**(9): p. 255-259.
97. Teame, T., et al., *The use of zebrafish (Danio rerio) as biomedical models.* Anim Front, 2019. **9**(3): p. 68-77.
98. Howe, K., et al., *The zebrafish reference genome sequence and its relationship to the human genome.* Nature, 2013. **496**(7446): p. 498-503.
99. Kumar, S. and S.B. Hedges, *A molecular timescale for vertebrate evolution.* Nature, 1998. **392**(6679): p. 917-20.
100. Lieschke, G.J. and P.D. Currie, *Animal models of human disease: zebrafish swim into view.* Nat Rev Genet, 2007. **8**(5): p. 353-67.
101. Newman, M., E. Ebrahimie, and M. Lardelli, *Using the zebrafish model for Alzheimer's disease research.* Front Genet, 2014. **5**: p. 189.
102. Leimer, U., et al., *Zebrafish (Danio rerio) presenilin promotes aberrant amyloid beta-peptide production and requires a critical aspartate residue for its function in amyloidogenesis.* Biochemistry, 1999. **38**(41): p. 13602-9.
103. Groth, C., et al., *Identification of a second presenilin gene in zebrafish with similarity to the human Alzheimer's disease gene presenilin2.* Dev Genes Evol, 2002. **212**(10): p. 486-90.
104. Musa, A., H. Lehrach, and V.A. Russo, *Distinct expression patterns of two zebrafish homologues of the human APP gene during embryonic development.* Dev Genes Evol, 2001. **211**(11): p. 563-7.
105. Moussavi Nik, S.H., et al., *The BACE1-PSEN-AbetaPP regulatory axis has an ancient role in response to low oxygen/oxidative stress.* J Alzheimers Dis, 2012. **28**(3): p. 515-30.

106. van Bebber, F., et al., *Loss of Bace2 in zebrafish affects melanocyte migration and is distinct from Bace1 knock out phenotypes.* J Neurochem, 2013. **127**(4): p. 471-81.
107. Francis, R., et al., *aph-1 and pen-2 are required for Notch pathway signaling, gamma-secretase cleavage of betaAPP, and presenilin protein accumulation.* Dev Cell, 2002. **3**(1): p. 85-97.
108. Strausberg, R.L., et al., *Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16899-903.
109. Chen, M., R.N. Martins, and M. Lardelli, *Complex splicing and neural expression of duplicated tau genes in zebrafish embryos.* J Alzheimers Dis, 2009. **18**(2): p. 305-17.
110. Woods, I.G., et al., *The zebrafish gene map defines ancestral vertebrate chromosomes.* Genome Res, 2005. **15**(9): p. 1307-14.
111. Schmid, B. and C. Haass, *Genomic editing opens new avenues for zebrafish as a model for neurodegeneration.* J Neurochem, 2013. **127**(4): p. 461-70.
112. Gaj, T., C.A. Gersbach, and C.F. Barbas, 3rd, *ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.* Trends Biotechnol, 2013. **31**(7): p. 397-405.
113. Zu, Y., et al., *TALEN-mediated precise genome modification by homologous recombination in zebrafish.* Nat Methods, 2013. **10**(4): p. 329-31.
114. Doyon, Y., et al., *Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases.* Nat Biotechnol, 2008. **26**(6): p. 702-8.
115. Kim, Y., et al., *A library of TAL effector nucleases spanning the human genome.* Nat Biotechnol, 2013. **31**(3): p. 251-8.
116. Ding, Q., et al., *A TALEN genome-editing system for generating human stem cell-based disease models.* Cell Stem Cell, 2013. **12**(2): p. 238-51.
117. Paul, B. and G. Montoya, *CRISPR-Cas12a: Functional overview and applications.* Biomed J, 2020. **43**(1): p. 8-17.

118. Okoniewski, M.J. and C.J. Miller, *Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations.* BMC Bioinformatics, 2006. **7**: p. 276.
119. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics.* Nat Rev Genet, 2009. **10**(1): p. 57-63.
120. Bustin, S.A. and R. Mueller, *Real-time reverse transcription PCR (qRT-PCR) and its potential use in clinical diagnosis.* Clin Sci (Lond), 2005. **109**(4): p. 365-79.
121. Vogelstein, B. and K.W. Kinzler, *Digital PCR.* Proc Natl Acad Sci U S A, 1999. **96**(16): p. 9236-41.
122. Fazekas de St, G., *The evaluation of limiting dilution assays.* J Immunol Methods, 1982. **49**(2): p. R11-23.
123. Baddeley, A., *Working memory and language: an overview.* J Commun Disord, 2003. **36**(3): p. 189-208.
124. Mehta, M.A., *Spatial Memory in Humans,* in *Encyclopedia of Psychopharmacology*, I.P. Stolerman, Editor. 2010, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 1262-1266.
125. Guariglia, C.C., *Spatial working memory in Alzheimer's disease: A study using the Corsi block-tapping test.* Dement Neuropsychol, 2007. **1**(4): p. 392-395.
126. Lalonde, R., *The neurobiological basis of spontaneous alternation.* Neurosci Biobehav Rev, 2002. **26**(1): p. 91-104.
127. Kraeuter, A.K., P.C. Guest, and Z. Sarnyai, *The Y-Maze for Assessment of Spatial Working and Reference Memory in Mice.* Methods Mol Biol, 2019. **1916**: p. 105-111.
128. Heredia-Lopez, F.J., et al., *An automated Y-maze based on a reduced instruction set computer (RISC) microcontroller for the assessment of continuous spontaneous alternation in rats.* Behav Res Methods, 2016. **48**(4): p. 1631-1643.
129. Stewart, S., F. Cacucci, and C. Lever, *Which Memory Task for My Mouse? A Systematic Review of Spatial Memory Performance in the Tg2576*

Alzheimer's Mouse Model. Journal of Alzheimers Disease, 2011. **26**(1): p. 105-126.

130. Hughes, R.N., *The value of spontaneous alternation behavior (SAB) as a test of retention in pharmacological investigations of memory.* Neuroscience and Biobehavioral Reviews, 2004. **28**(5): p. 497-505.
131. King, D.L. and G.W. Arendash, *Behavioral characterization of the Tg2576 transgenic model of Alzheimer's disease through 19 months.* Physiology & Behavior, 2002. **75**(5): p. 627-642.
132. Cleal, M., et al., *The Free-movement pattern Y-maze: A cross-species measure of working memory and executive function.* Behav Res Methods, 2020.
133. Aoki, R., T. Tsuboi, and H. Okamoto, *Y-maze avoidance: an automated and rapid associative learning paradigm in zebrafish.* Neurosci Res, 2015. **91**: p. 69-72.
134. Cognato Gde, P., et al., *Y-Maze memory task in zebrafish (Danio rerio): the role of glutamatergic and cholinergic systems on the acquisition and consolidation periods.* Neurobiol Learn Mem, 2012. **98**(4): p. 321-8.
135. Sato, N., et al., *A novel presenilin-2 splice variant in human Alzheimer's disease brain tissue.* J Neurochem, 1999. **72**(6): p. 2498-505.
136. Moussavi Nik, S.H., et al., *Alzheimer's disease-related peptide PS2V plays ancient, conserved roles in suppression of the unfolded protein response under hypoxia and stimulation of gamma-secretase activity.* Hum Mol Genet, 2015. **24**(13): p. 3662-78.
137. Sharman, M.J., et al., *The Guinea Pig as a Model for Sporadic Alzheimer's Disease (AD): The Impact of Cholesterol Intake on Expression of AD-Related Genes.* PLoS One, 2013. **8**(6): p. e66235.
138. Moussavi Nik, S.H., M. Newman, and M. Lardelli, *The response of HMGA1 to changes in oxygen availability is evolutionarily conserved.* Exp Cell Res, 2011. **317**(11): p. 1503-12.
139. De Jonghe, C., et al., *Aberrant splicing in the presenilin-1 intron 4 mutation causes presenile Alzheimer's disease by increased Abeta42 secretion.* Hum Mol Genet, 1999. **8**(8): p. 1529-40.

140. Wisniewski, T., et al., *A novel Polish presenilin-1 mutation (P117L) is associated with familial Alzheimer's disease and leads to death as early as the age of 28 years.* Neuroreport, 1998. **9**(2): p. 217-21.
141. Hin, N., et al., *Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2.* PLoS One, 2020. **15**(1): p. e0227258.
142. Newman, M., et al., *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): p. 43.
143. Berchtold, N.C., et al., *Brain gene expression patterns differentiate mild cognitive impairment from normal aged and Alzheimer's disease.* Neurobiol Aging, 2014. **35**(9): p. 1961-72.

Chapter 2: Investigating changes in transcriptome state during disease progression in heterozygous *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$ mutant brains

2.1 Introduction

AD is thought to take decades to develop before disease onset [1]. The process of disease progression in AD is a controversial research issue. There are two broad possibilities for describing this process (**Figure 2.1**): Either brain cellular damage accumulates during ageing and eventually degrades brain function sufficiently to cause AD cognitive change, or increasing stress and other factors in ageing brains cause cells to shift into an alternative, "stable" pathological state where cognition progressively worsens. Importantly, if the onset of AD is caused by a state shift of the brain cells, it may be possible to find drugs that can revert the pathological state back to normal [2].

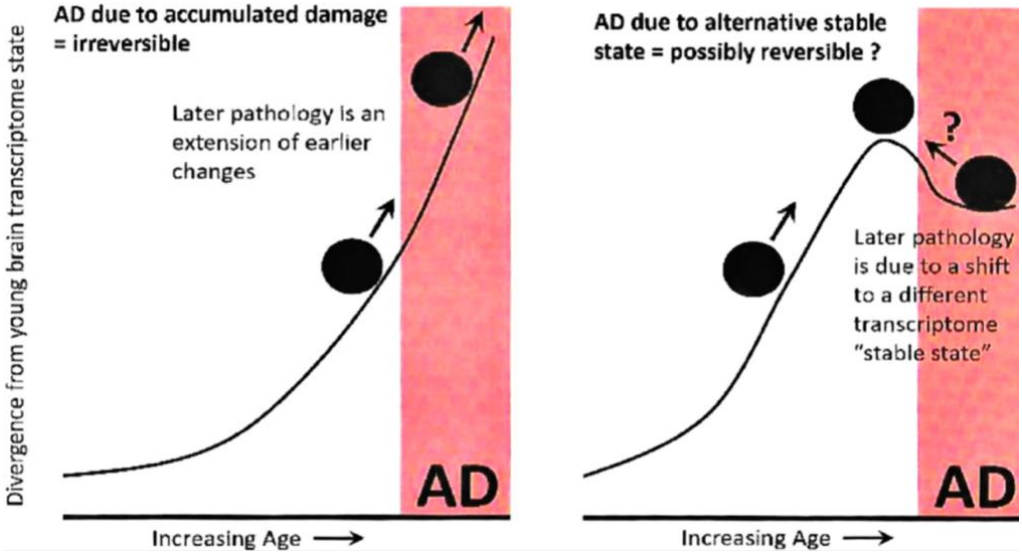


Figure 2.1. Two proposed processes of gene expression change from a normal state to a pathological state. With increasing age, the state of the brain transcriptome during progression to AD may reflect accumulated damage until it irreversibly passes a pathological threshold (**left**), or a change in state at a certain disease stage and shift into an alternative stable state (**right**) which may possibly be reversed.

In previous work in the ADGL, bulk transcriptome analyses were performed on the entire brains of young adult zebrafish (6 months of age, recently sexually mature) and aged adult zebrafish (24 months of age) for two heterozygous EOfAD-like mutations, *psen1*^{K97fs} [3] and *psen1*$^{Q96\_K97del}$ [4] (discussed in previous ***section 1.9.1 and section 1.9.2***). In heterozygous *psen1*^{K97fs} mutant brains, a group of DE genes was identified showing an obvious expression inversion between young adult and aged adult zebrafish ([3], **Figure 2.2**). A similar transcriptional inversion has been seen in humans in a postmortem transcriptome comparison of mild cognitive impairment (MCI), Alzheimer's disease and age-matched normal brains, which found the expression of particular genes was predominantly upregulated in patients with MCI but downregulated in AD brains [5]. However, we do not know whether the expression inversion represents a gradual change with age consistent with

steady, incremental accumulation of damage as in **Figure 2.1**, or a more rapid shift to an alternative pathological state as also illustrated in that figure. Therefore, we designed a month-by-month study monitoring transcriptional state during aging from 6 months to 24 months. Through this month-by-month study, we expected to determine both the rate of gene expression change during aging and whether all DE genes invert their expression simultaneously or whether this occurs separately for different genes at different ages. This work could further inform us on the process of AD progression and the likelihood that interventions could reverse this pathological process.

Initially, I aimed to select a set of highly DE genes identified from among those identified in the transcriptome analyses of heterozygous *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$ brains as potential biomarkers to represent brain genetic states. The expression of these DE genes could then be assessed by qPCR. The brains of heterozygous *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$ mutant zebrafish with their wild type siblings were collected every month from 6 months to 24 months. Through tracking these highly DE genes month by month, I expected to clarify the rapidity of the expression change leading to the inversion of expression between 6 months and 24 months of age.

Digital quantitative PCR (dqPCR) can be employed to measure the expression of representative genes in zebrafish brains collected each month. Before conducting this month-by-month study, I first needed to be validate the changes in gene expression measured by dqPCR corresponded to the changes that had been detected in the brain transcriptome analyses. Therefore, I performed dqPCRs on selected

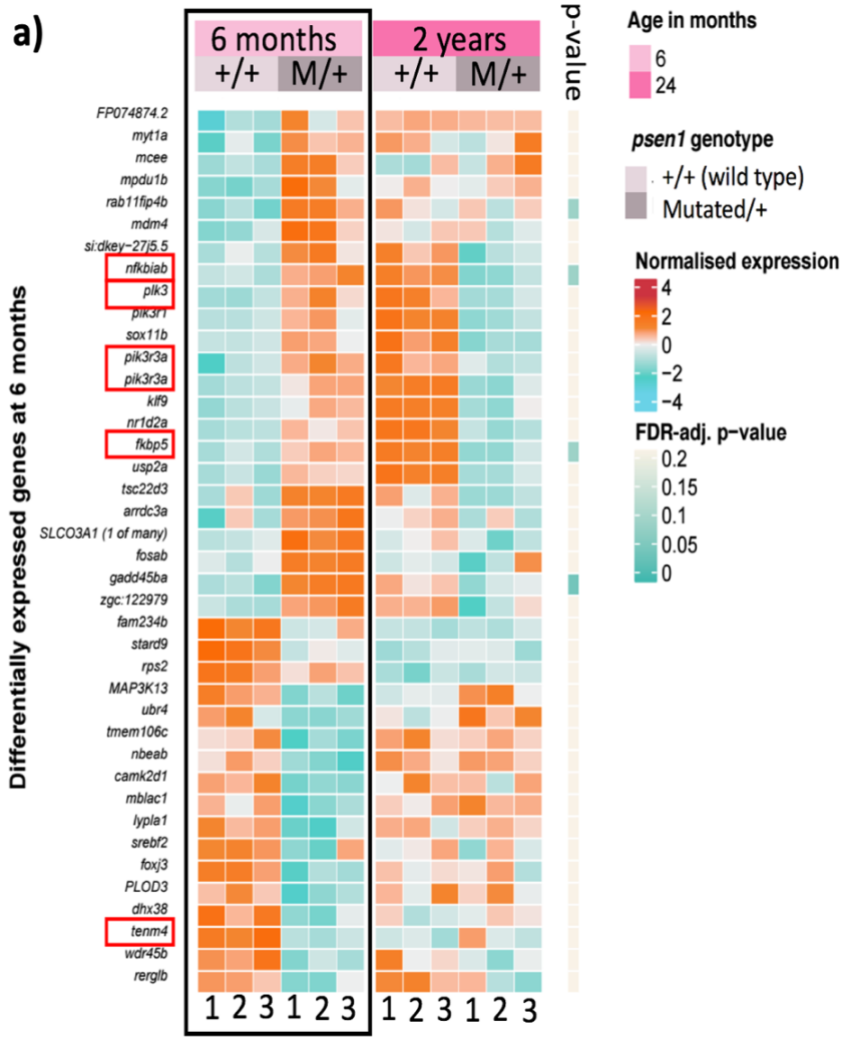
highly DE genes to validate the brain transcriptome analyses of both the heterozygous *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$ mutants. The results of these dqPCRs are showed in this chapter and were inconsistent with the transcriptome analyses, even when using cDNA synthesized from the same RNA as was used in a transcriptome analysis. The reason for this inconsistency remains unclear but reflects variability in gene expression seen between the brains of different individuals. (Also, this contrasts with the consistency of our dqPCR and transcriptome analysis results when analyzing pools of individuals in ***Chapter 4***.) Since the dqPCR results were in apparent conflict with the transcriptome results, the bulk of the month-by-month study was, ultimately, abandoned.

2.2 Results

2.2.1 dqPCR assessment of supposedly DE genes in heterozygous *psen1*^{K97fs} mutant brains

The DE genes identified in the transcriptome analysis of 6-month-old and 24-month-old heterozygous *psen1*^{K97fs} mutant brains are displayed in **Figure 2.2**. **Figure 2.2A** displays the DE genes identified at 6 months and **Figure 2.2B** displays the DE genes identified at 24 months. Interestingly, there was a set of common DE genes identified in both 6-month-old and 24-month-old mutant brains, but the regulation of these genes was in opposite directions (i.e. their differential expression was "inverted"). These genes were upregulated at 6 months of age but showed downregulation at 24 months. Among these genes, five genes, *nfkbiab*, *plk3*,

pik3r3a and *fkbp5*, which showed more significant fold changes in gene expression, were selected to represent this group and were subjected to dqPCR to validate the changes seen in the RNA-seq analyses. Also, two DE genes, *tenm4 and nr3c2*, showing upregulation in only one age group, were also assessed by dqPCRs.



(The duplication of *pik3r3a* gene in **Figure 2.2a** is due to the genome reference used in mapping, which contains multiple visions of a gene.)

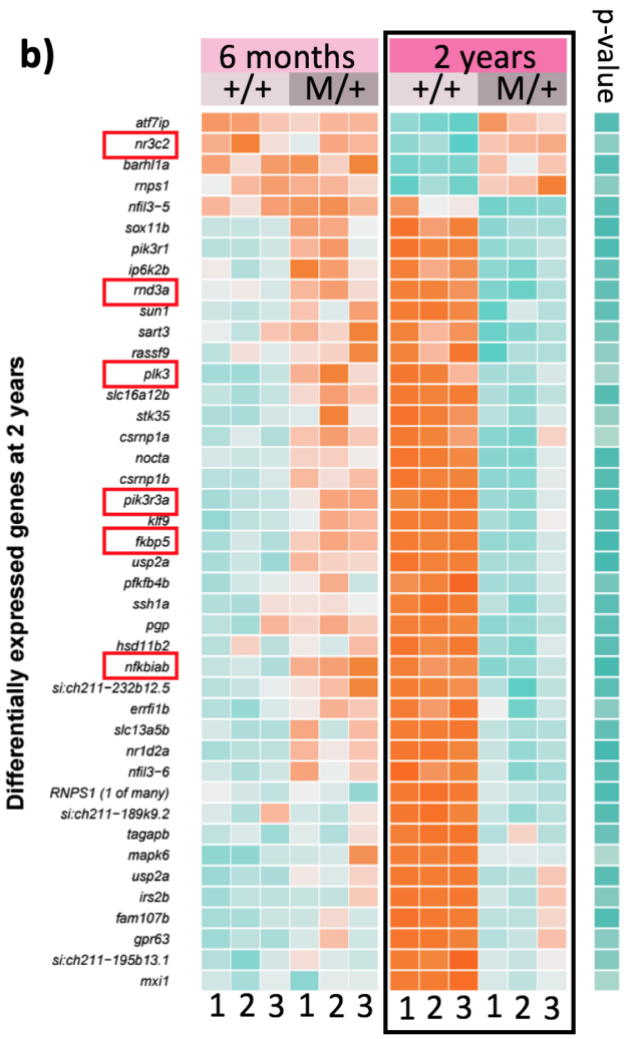
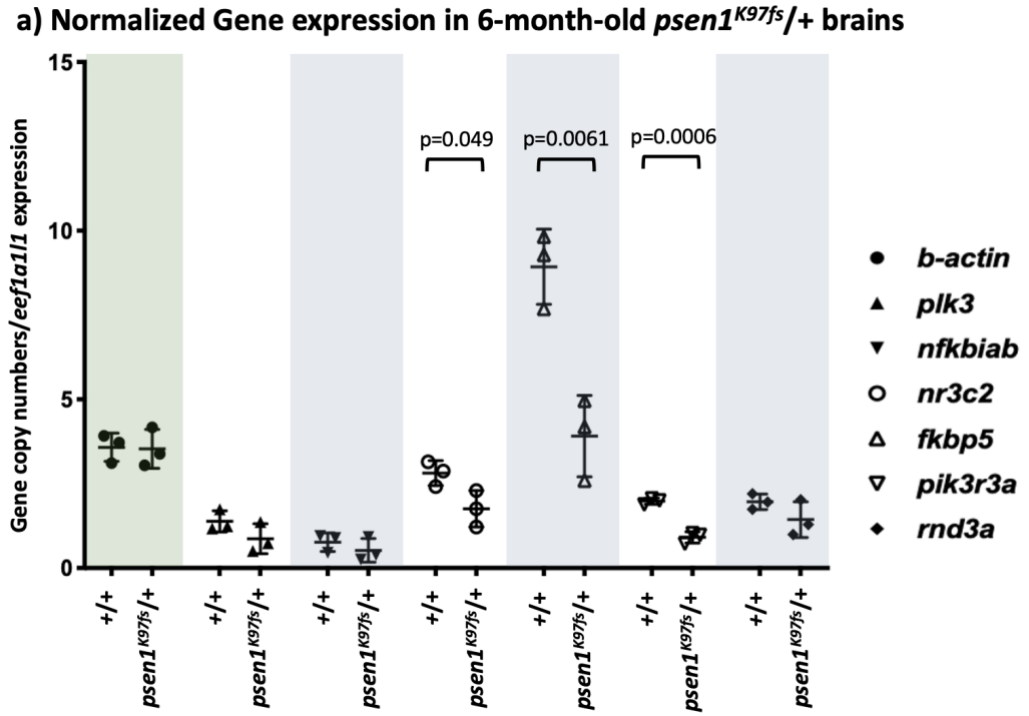


Figure 2.2. DE genes in heterozygous *psen1*^{K97fs} mutant zebrafish brains compared to their wild type siblings at 6 months and 24 months of age (the heatmap was produced by Nhi Hin prior to the publication of [3] in PLoS ONE, therefore it is slightly different from the final published figure). Figure 1a and 1b presents the relative expression of DE genes identified in 6-month and 24-month brains respectively (in black boxes), which have FDR-adjusted p-values < 0.05 and $\log_2$fold change > 0.5. The p-value bar shows the FDR-adjusted p-value of the genes in 24-month brains. The individual zebrafish are numbered below. The genes tested by dqPCR are highlighted in red boxes.

dqPCRs for the gene indicated in **Figure 2.2** were first performed on cDNA generated from the brains of individual 6-month-old and 24-month-old heterozygous *psen1*^{K97fs} mutants. The results are displayed in **Figure 2.3**. Gene expression levels were normalized using the "house-keeping" gene *eef1a1l1*. *Beta-actin* was another house-keeping gene used as a control. In **Figure 2.3**, the normalized expression of *beta-actin* was at the same level between mutant brains and their wild type siblings, supporting that the expression level of this house-keeping gene was stable among samples.

The normalized expression of the genes measured by dqPCR was inconsistent with their differential expression detected by transcriptome analysis. The expression of the genes *nfkbiab*, *nr3c2* and *fkbp5* genes was downregulated in 6-month-old mutant brains (**Figure 2.3A**), which was opposite to the results observed in the transcriptome analysis (**Figure 2.2A**). There were no significant differences in normalized gene expression in 24-month-old mutant brains (**Figure 2.3B**). Notably, the expression of some genes differed greatly between individuals, such as the expression of *fkbp5* gene in 24-month-old brains, indicating that individual differences might have strong effects on gene expression level.



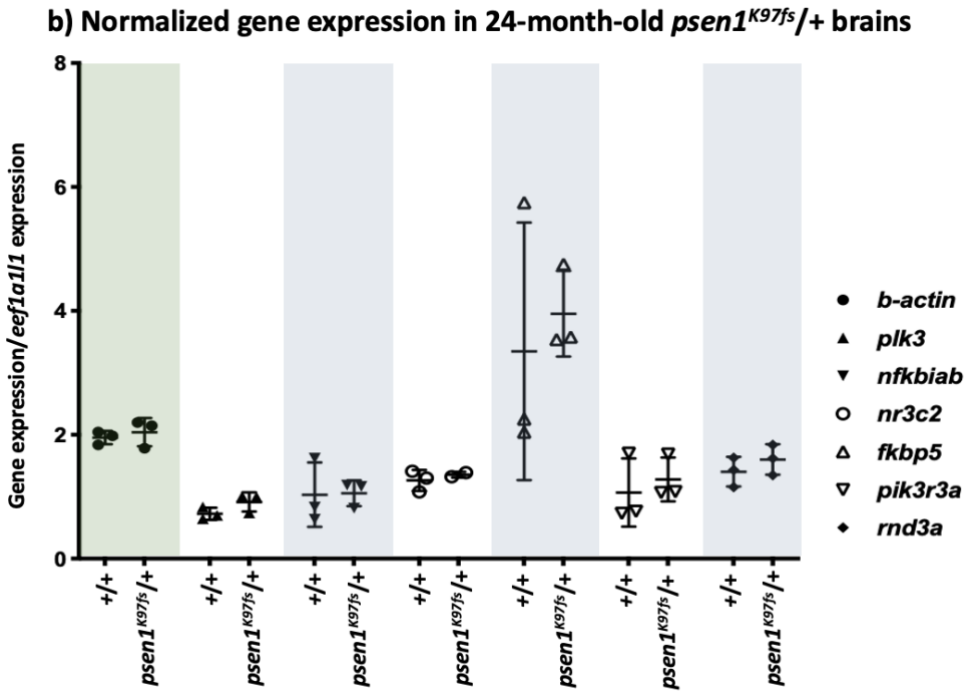
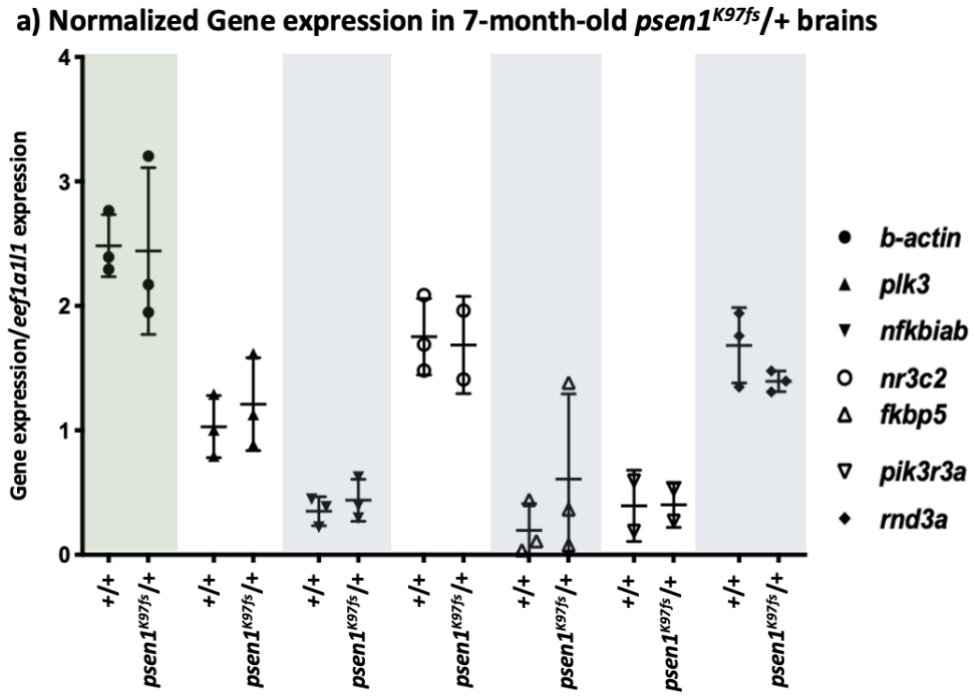
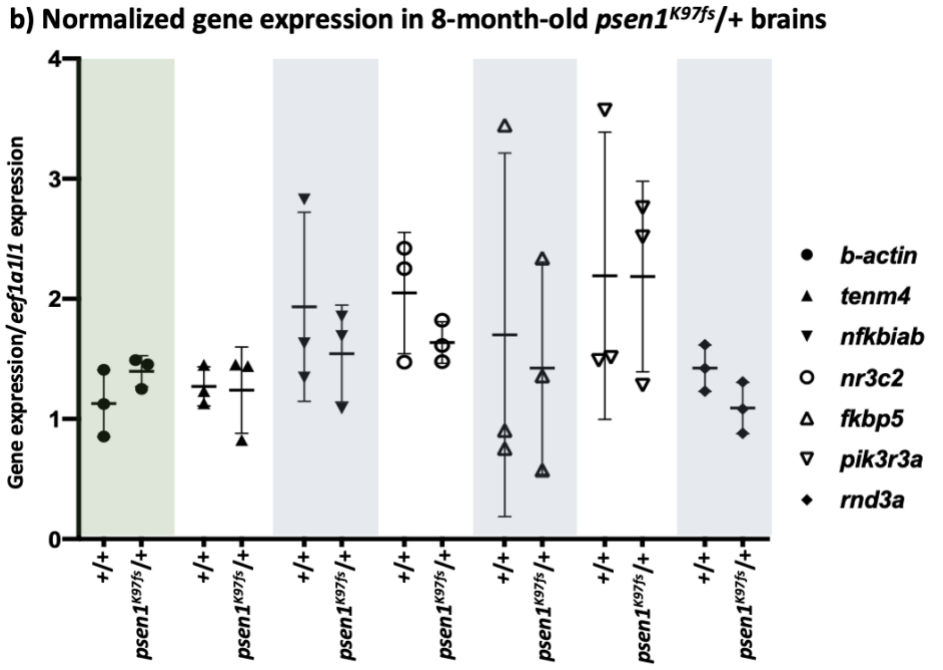


Figure 2.3. Assessing the expression of the DE genes by dqPCR in 6-month-old and 24-month-old heterozygous *psen1*^{K97fs} mutant brains. The gene expression levels were normalized to the expression level of house-keeping gene *eef1a1l1*. The p-values where genes were differentially expressed between genotypes are indicated if less than 0.05.

One possibility for the discrepancy between the results from the dqPCRs and the transcriptome analysis was that different tanks/families of the same zebrafish genotype might be aging at different rates. Therefore, dqPCRs were performed on older heterozygous *psen1*^{K97fs} mutant and wild type sibling brains at 7, 8 and 15 months of age **(Figure. 2.4a-c**). However, none of the dqPCR results showed significant gene expression differences between the mutant brains and their wild type siblings.

As dqPCRs failed to support the differences in gene expression identified in the transcriptome analysis, we could not use these genes as biomarkers to represent brain genetic states in the intended month-by-month study.





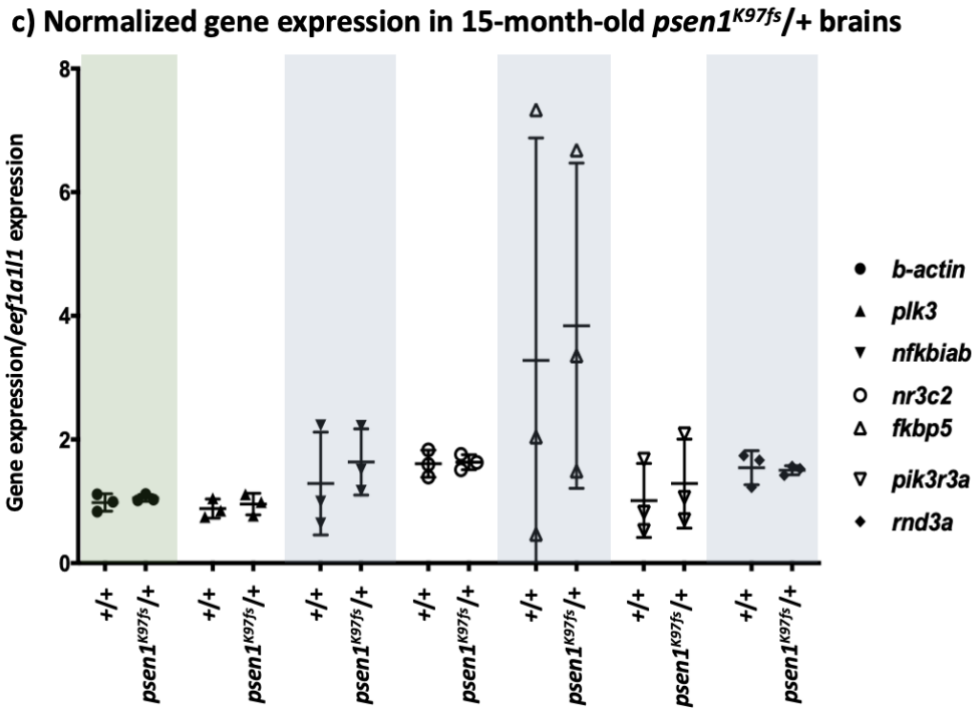


Figure 2.4. Assessing the expression of the DE genes by dqPCRs in 7-month-old, 8-month-old and 15-month-old heterozygous *psen1*^{K97fs} mutant brains. Instead of *plk3*, the gene *tenm4* was assayed in 8-month-old brains. No genes showed statistically significantly differential expression between genotypes (i.e. p-value was always greater than 0.05).

Interestingly, there is large variation in the expression of the gene *fkbp5* between samples of the same genotype, whether wild type or mutant, in the various qdPCR tests. The gene *fkbp5* encodes FK506 binding protein 5 (FKBP5), which acts as a negative-feedback regulator of glucocorticoid receptor (GR) [6, 7]. Glucocorticoids are the major stress hormones, and overexpression of FKBP5 interacts with GRs and promotes resistance to glucocorticoid [8]. The large spreads of *fkbp5* expression shown here probably reflect environmental impacts on these tested zebrafish. However, these zebrafish were treated similarly during euthanasia and brain excision. The zebrafish were collected together and euthanised for brain removal

around 11:00 a.m. after feeding. Therefore, the differences in *fkbp5* expression were not likely due to circadian influences, which regulate the release of glucocorticoid [9]. Neither were there notable external environmental stresses on the fish. Probably zebrafish individuals responded differently to environmental changes, resulting in the variability in *fkbp5* expression.

2.2.2 dqPCR assessment of supposedly DE genes in heterozygous *psen1*$^{Q96\_K97del}$ mutant brains

Transcriptome analysis (by RNA-Seq) had also been performed on 6-month-old and 24-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant brains to investigate the genetic changes induced by this EOfAD-like mutation [4]. While numerous DE genes were identified in this study, a difference compared to the *psen1*^{K97fs} brain analysis was that no genes were found to be DE at both 6 months and 24 months (and so "inverted" patterns of gene expression were not observed, see ***section 2.2.1***). The transcriptome DE genes analysed by dqPCR are shown in **Figure 2.5**. In 6-month-old *psen1*$^{Q96\_K97del}$ mutant brain transcriptomes, the expression of gene *abcd3a, pid1* and *kalrnb* was upregulated and the expression of gene *clcn3* was downregulated; in 24-month-old *psen1*$^{Q96\_K97del}$ mutant brains, gene *spsb1* was upregulated and genes *BX469889.4* and *mtor* genes were downregulated.

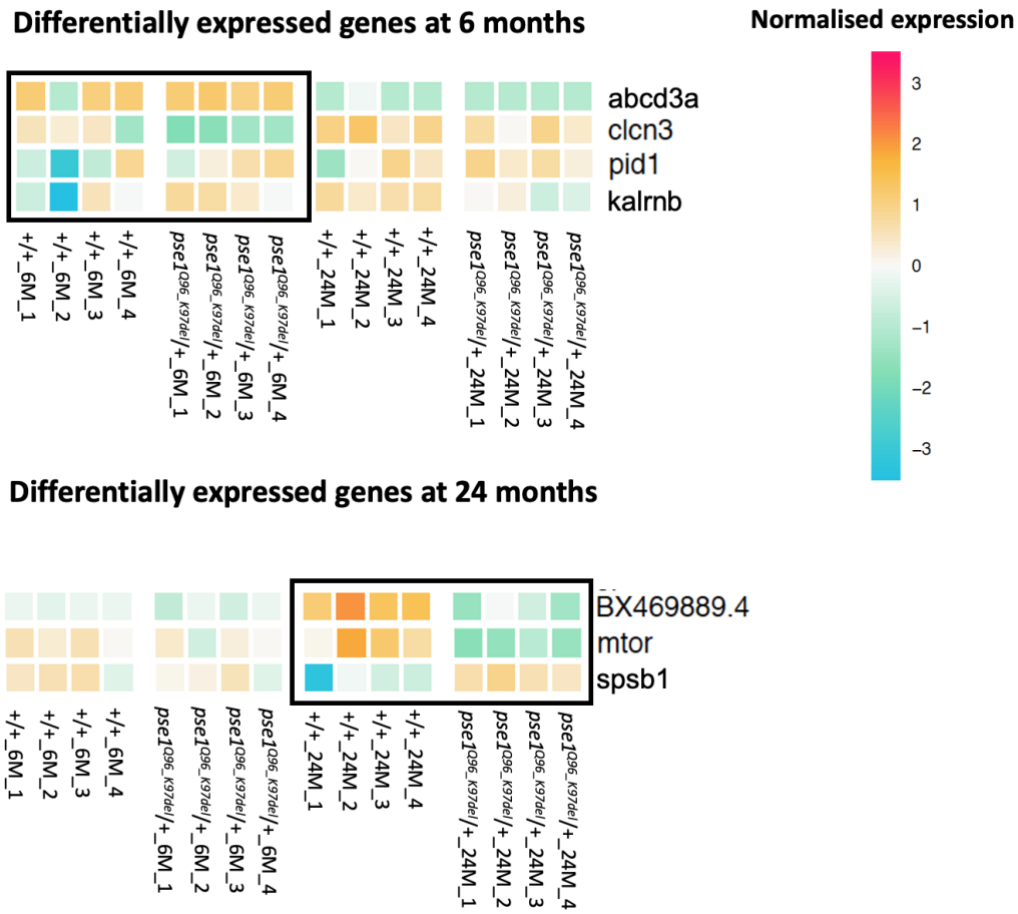


Figure 2.5. DE genes in heterozygous *psen1*$^{Q96\_K97del}$ mutant brains compared to their wild type siblings at 6 months and 24 months of age (data generated by Nhi Hin). As the DE gene list is too long, only the genes used in our dqPCR analyses are shown here.

dqPCRs were performed on the cDNA generated from 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant brain samples to assess the expression of the 7 DE genes selected from transcriptome analysis. The dqPCR results are shown in **Figure 2.6**. Gene expression levels were normalized to that of house-keeping gene *eef1a1l1*. Similar to the results seen from *psen1*^{K97fs} mutant brains, the normalized expression levels of these DE genes were similar between heterozygous *psen1*$^{Q96\_K97del}$ mutant

brains and the wild type siblings. No statistically significant differences in gene expression were detected.

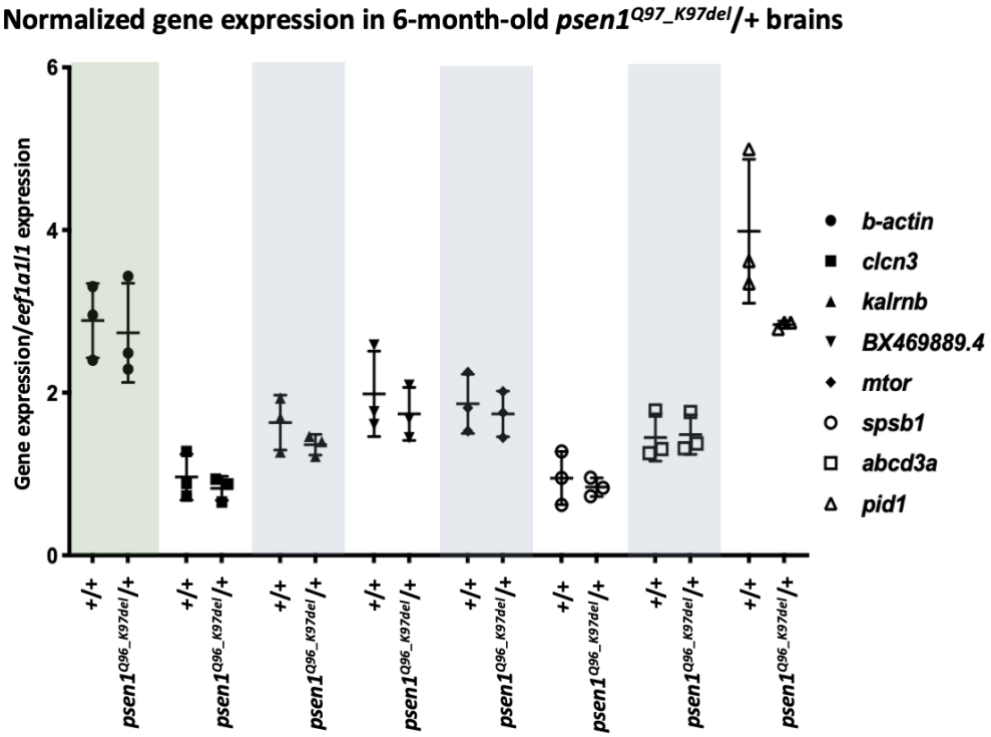


Figure 2.6. Assessing the expression of the DE genes by dqPCR in 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant brains. No genes showed statistically significantly differential expression between genotypes.

Due to the failure to detect the above genes as DE using dqPCR, they could not be used as biomarkers to monitor the brain genetic state in the month-by-month analysis. As the expression level of some of the genes assayed varied considerably between individuals of the same genotype, this represented one possible explanation for the failure to detect them as DE. To counter the problem of sample variability between those brains used to generate mRNA for RNA-Seq in the transcriptome analysis, and those brains used for dqPCR, I then performed dqPCR on cDNAs synthesized from the same RNAs that were used for the transcriptome

analysis (**Figure 2.7A**). dqPCRs performed on genes *clcn3*, *abcd3a* and *kalrnb* still did not detected differences in expression, indicating that between-sample variability did not seem to be the cause of the inconsistency. Another possibility for the inconsistency between the transcriptome and dqPCR results was that the dqPCR technique itself was somehow invalid. To support that dqPCR produces reliable results, I compared data produced by this technique to that produced by another quantitative PCR technique, relative standard curve qPCR. Relative standard curve qPCR analyses were also performed on cDNA synthesized from the RNAs used in the transcriptome analysis. The results are shown in **Figure 2.7B**. Although still inconsistent with the results of the transcriptome analyses, the expression of *clcn3* and *kalrnb* measured by relative standard curve qPCRs was similar to their expression measured in dqPCRs, indicating that both quantitative PCR techniques performed appropriately.

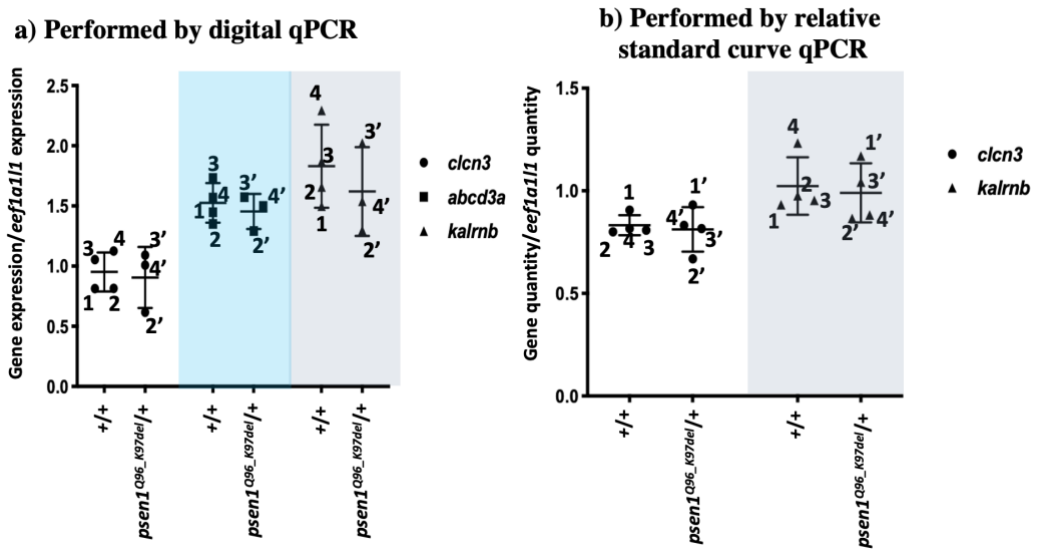


Figure 2.7. Gene expression in 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant brains achieved by **a)** digital qPCR and **b)** relative standard curve qPCR. Each data point is labeled with a sample number, which corresponds with an individual fish in the transcriptome analysis. No genes showed statistically significantly differential expression between genotypes. (The cDNA of Sample 1' of heterozygous *psen1*$^{Q96\_K97del}$ mutant was too degraded in dqPCRs, so Sample 1' was not shown in Figure 2.7A)

The analyses do not support that sample variability or the quantitative PCR technique used contribute to the inconsistency between the results of transcriptome (RNA-Seq) analysis and the dqPCRs. As for the month-by-month study of the *psen1*^{K97fs} heterozygous mutant brains, the month-by-month study of the *psen1*$^{Q96\_K97del}$ heterozygous mutant brains was suspended.

2.3 Discussion

The brain transcriptome analyses of heterozygous *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$ mutants should provide insight into the molecular changes induced by EOfAD-like mutations. However, the differential expression of the DE genes identified in these analyses could not be validated by the qPCRs performed. Some attempts were made to test possible reasons for this inconsistency. The possibility of different rates of aging rates between separate tanks/families was tested through assessing the DE genes identified from 6-month-old heterozygous *psen1*^{K97fs} mutant brains in the brains of fish at older ages (7 months, 8 months and 15 months of age). However, none of these genes showed similar differences in gene expression between heterozygous mutant and wild type sibling fish as those seen in the transcriptome analysis. Furthermore, the influences of variability between samples and the reliability of the dqPCR technique were also assessed through preforming two quantitative RT-PCR techniques in parallel, dqPCR and relative standard curve qPCR, on cDNA synthesized from the actual RNA used in the transcriptome analysis. The two qPCR techniques detected similar patterns of differential gene expression (i.e. no differential expression) which were still inconsistent with the transcriptome analysis, supporting that both qPCR techniques appeared to work appropriately and that variability between samples was not responsible for the inconsistency with the transcriptome data. Some other possibilities were then considered, such as that the observations of differential expression of some genes in the transcriptome analysis were an artefact of the RNA-Seq technique or the bioinformatics analysis that followed it, the bias induced in RNA-seq process or the

following bioinformatic analysis. However, RNA-Seq analyses are expensive and, therefore, difficult to replicate. Nevertheless, two facts are noteworthy:

1) The ADGL did attempt to replicate the RNA-seq analysis of 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant brains compared to wild type siblings using a less expensive non-domestic service. However, the RNA-seq data produced was of lower quality, possibly because of the additional treatment of the RNA required for it to survive export from Australia. Ultimately, little consistency was found between the two sets of RNA-seq data (Lachlan Baer, Honours thesis 2019).

2) In ***Chapter 4*** I describe how relative standard curve qPCR was used successfully to validate the expression patterns of the DE genes identified in transcriptome analysis of heterozygous *psen1*$^{Q96\_K97del}$ mutant larvae (see ***section 4.3, Supplementary data 2*** for detailed explanation). In those experiments, 40 sibling larvae were pooled for each sample (which would "average out" differences between individuals) and the qPCRs were conducted on the same RNA samples as were used in the RNA-seq analysis. Nevertheless, the large differences observed between samples in the differential expression of the various genes tested is remarkable.

During the early era of transcriptome studies using mainly microarrays, it was a common requirement to validate findings of differential expression by qPCR tests. Indeed, qPCR often found larger differences in expression of a gene between samples than were indicated by microarray hybridisation signals [10, 11]. The same tests are not commonly applied to RNA-seq-based studies today.

As gene expression can be diverse between samples, a limited sample size used in a RNA-Seq analysis might misidentify genes as DE, or fail to identify DE genes, both leading to unreliable conclusions. In contrast, using an excessive number of samples can waste time and resources. Thus, a suitable samples size is critical for RNA-seq analysis to achieve confident results with high efficiency. Since the *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$ analyses were performed, power calculations have been performed using the strategy from Zhao et al., 2018 [12], and have showed that n=6 provides a power of ~70% for detection of fold-change > 2 at a false discovery rate (FDR) of 0.05 across the vast majority of expressed transcripts in zebrafish brain transcriptomes (Stephen Pederson, unpublished results). However, n=3 and n=4 were used in *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$ analyses respectively, which were relatively low sample sizes for identification of DE genes. Due to the genetic diversity between siblings, the low sample sizes may have led to misidentification of some genes as DE, which is a major limitation of our transcriptome analyses. Therefore, the sample numbers used in the above brain transcriptome analyses was suboptimal. It is for the reason that a sample size of 6 was used in the transcriptome analysis described in ***Chapter 4***.

2.4 Methods

Excision of zebrafish brains

The *psen1*^{K97fs} and *psen1*$^{Q96\_K97del}$ families used for month-by-month transcriptome monitoring were initially generated through mating an individual heterozygous mutant zebrafish with a wild type individual to produce a family of ~50% heterozygous and ~50% wild type siblings. Three wild type and three heterozygous mutant zebrafish brains were extracted each month from 6 months to 24 months after the family was spawned. Zebrafish were euthanised in a loose icy slurry before brain removal. The extracted brains were placed into 400 µl RNA*later*™ (Invitrogen, Camarillo, California, USA) overnight at 4°C for RNA stabilization. RNA*later*™ solution was then removed and brain samples were stored at -80°C.

RNA purification and cDNA synthesis

Total brain RNA was extracted using the QIAGEN RNeasy mini Kit (QIAGEN, Hilden, Germany). DNase treatment was performed on RNA to remove remaining genomic DNA using the DNA-free™ Kit (Ambion, Life Technologies, Thermo Fisher Scientific, Waltham, MA, USA). Purified RNA (1000 ng) was then used to synthesise 20 µL of first-strand cDNA by reverse transcription (SuperScript III kit, Invitrogen, Camarillo, California, USA).

QuantStudio 3D Digital PCR

Digital quantitative PCR was performed on a QuantStudio™ 3D Digital PCR System (Life Technologies, Thermo Fisher Scientific, Waltham, MA, USA). Each reaction mix contained 1X QuantStudio™3D digital PCR Master Mix, SYBR® green dye, 0.2 μM of specific primers and a specific amount of cDNA for each gene (primer sequences and cDNA amount used in each dqPCR reactions are shown in **2.5 Appendix**), to give a final product concentration in the reactions within a range of 200 to 2000 copy numbers per microliter. The reaction mixture (14.5 μl) was loaded onto a QuantStudio™3D digital PCR 20 K chip using an automatic chip loader following the manufacturer's instructions. Loaded chips underwent thermo-cycling on the Gene Amp 9700 thermo-cycling system under the following conditions: 96°C for 10 min, 39 cycles of 60°C for 2 min and 98°C for 30 sec, followed by a final extension step at 60°C for 2 min. The chips were imaged on a QuantStudio™ 3D instrument, which assesses raw data and calculates the estimated concentration of the nucleic acid sequence targeted by the SYBR® green dye by Poisson distribution [13]. The concentration of tested genes was normalized to the concentration of house-keeping gene *eef1a1l1* to correct for sample to sample variations in RT-PCR efficiency and errors in sample quantification.

Relative standard curve qPCR

Relative standard curve qPCR was used to repeat the quantification on two DE genes to ensure the reliability and reproducibility of dqPCR. Each 25 μL qPCR reaction contained 20 ng cDNA, 0.2 μM of each PCR primers and Power SYBR green master mix PCR solution (Applied Biosystems, Thermo Fisher Scientific Inc., Waltham, MA, USA). The standard curve was generated by a serial dilution, having 50 ng, 25 ng, 12.5 ng and 6.25 ng of wild type cDNA per reaction. The qPCR was performed on an ABI 7000 Sequence Detection System (Applied Biosystems) using a 96-well plate. The amplification consisted of a holding stage and a cycling stage. The holding stage performed 50°C for 2 mins and 95°C for 10mins, and the cycling stage had 40 cycles of 95°C for 15 s and 60°C for 1 min. Three technical PCR replicates were performed on each cDNA sample and their mean was used to represent the quantity. The quantities of the selected genes were calculated relative to the quantities of house-keeping gene *eef1a1l1*, which were then compared using an unpaired t-test.

2.5 References

1. Villemagne, V.L., et al., *Longitudinal assessment of Abeta and cognition in aging and Alzheimer disease.* Ann Neurol, 2011. **69**(1): p. 181-92.
2. Wagner, A., et al., *Drugs that reverse disease transcriptomic signatures are more effective in a mouse model of dyslipidemia.* Mol Syst Biol, 2015. **11**(3): p. 791.
3. Hin, N., et al., *Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2.* PLoS One, 2020. **15**(1): p. e0227258.
4. Newman, M., et al., *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): p. 43.
5. Berchtold, N.C., et al., *Brain gene expression patterns differentiate mild cognitive impairment from normal aged and Alzheimer's disease.* Neurobiol Aging, 2014. **35**(9): p. 1961-72.
6. Binder, E.B., et al., *Association of FKBP5 polymorphisms and childhood abuse with risk of posttraumatic stress disorder symptoms in adults.* JAMA, 2008. **299**(11): p. 1291-305.
7. Li, H., et al., *The glucocorticoid receptor-FKBP51 complex contributes to fear conditioning and posttraumatic stress disorder.* J Clin Invest, 2020. **130**(2): p. 877-889.
8. Szyf, M., *How do environments talk to genes?* Nature Neuroscience, 2013. **16**(1): p. 2-4.
9. Dickmeis, T., *Glucocorticoids and the circadian clock.* J Endocrinol, 2009. **200**(1): p. 3-22.
10. Yuen, T., et al., *Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays.* Nucleic Acids Res, 2002. **30**(10): p. e48.
11. Dallas, P.B., et al., *Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR -- how well do they correlate?* BMC Genomics, 2005. **6**: p. 59.
12. Zhao, S., et al., *RnaSeqSampleSize: real data based sample size estimation for RNA sequencing.* BMC Bioinformatics, 2018. **19**(1): p. 191.

13. Fazekas de St, G., *The evaluation of limiting dilution assays.* J Immunol Methods, 1982. **49**(2): p. R11-23.

2.6 Appendix: Primer information

Ensembl gene ID	Target gene	Primer sequences	Tm (°C)	cDNA amount in dqPCR (ng)
House-keeping genes				
ENSDARG00000020850	*eef1a1l1*	5'-CCAACTTCAACGCTCAGGTCA-3' 5'-CAAACTTGCAGGCGATGTGA-3'	60	0.5
ENSDARG00000037746	*b-actin*	5'- TGGAGAAGAGCTATGAGCTGC-3' 5'- CATACCCAGGAAGGAAGGCTG-3'	60	0.5
DE genes identified in heterozygous psen1^{K97fs} brains				
ENSDARG00000038754	*plk3*	5'-GAACCTAATGGAGGGTGGAGAT-3' 5'-AGGGTTCGGTTGTCAAAGAGC-3'	60	12.5
ENSDARG00000007693	*nfkbiab*	5'-GACACGTATCTACATCTCGCCA-3' 5'-AGTGCAGTCTGTCTCTGGTTG-3'	60	12.5
ENSDARG00000102082	*nr3c2*	5'-GCAGTCGCAACACTGGGAAT-3' 5'-CCTGGGTCATTTGATAAACCGTA-3'	60	25
ENSDARG00000028396	*fkbp5*	5'-TACCTAAAGCCCAAGTACGCCT-3' 5'-ACAGCACGTTCCAGTTTCTCT-3'	60	25
ENSDARG00000103038	*pik3r3a*	5'-CCACTGTAACACGGATCTACCT-3' 5'-TGCCGTCTTCTGAAAAACCCT-3'	60	25
ENSDARG00000076799	*rnd3a*	5'-TCCCTGAGAATTATGTCCCCAC-3' 5'-ACGAGAGTGGTCTGACGTTG-3'	60	25
ENSDARG00000105088	*tenm4*	5'-GAGACCAGAAACATAGCAAAGCA-3' 5'-CAGTACAAAGGCGAGGTCCC-3'	60	50
DE genes identified in heterozygous psen1^{Q96K97del} brains				
ENSDARG00000110064 (ENSDARG00000003269)	*clcn3*	5'-CATAATGGGATTGTATTGGGCA-3' 5'-AAGGAGGCGTCAAGGGCT-3'	60	50
ENSDARG00000111084	*kalrnb*	5'-CCAAGTTCACCTTTGAGACCAG-3' 5'-CATACTCCAGTGTTCCCTCCAGA-3'	60	25
ENSDARG00000112239	*BX469889.4*	5'-GTGGTCTGCTATCGGGTTTATTG-3' 5'-CATTGATGTAGAAGAAGGGGCG-3'	60	25
ENSDARG00000053196	*mtor*	5'-CAGCATCAGAGAGTTTCTTTTCCA-3' 5'-CCACCGAGAGCCACAACAATC-3'	60	25
ENSDARG00000110015	*spsb1*	5'-TATTTAGGGGTTGCTTTTAGAGGAC-3' 5'-AATGGCAGAGGTTCAGGATCA-3'	60	25
ENSDARG00000110888	*abcd3a*	5'-CAACCTTCAAAAAACTGGTGGAT-3' 5'-AAAATACTTGGCGATAATGCTGT-3'	60	25
ENSDARG00000114639	*pid1*	5'-AATGGCAGAGGTTCAGGATCA-3' 5'-CACACATACAGGATACAAGGTGACG-3'	60	25

Chapter 3: Do zebrafish with the *psen1*^{K97fs} mutation show the alternative splicing seen in human *PSEN2*^{K115fs} carriers?

3.1 Introduction

EOfAD is autosomal and dominantly inherited, and hundreds of EOfAD-causative mutations are located in the three major genetic loci *PSEN1*, *PSEN2* and *APP* [1]. Interestingly, these mutations have a common feature of following a "fAD reading frame preservation rule" specifying that these EOfAD mutations must not change the reading frame and so must allow translation of "full-length" proteins [2]. *PSEN2*^{K115fs}, a frame shift EOfAD-causative mutation in human *PSEN2* gene, was previously thought to be a unique exception to this rule [3]. As, supposedly, the only frame shift mutation of the *PREESENLIN* genes causing EOfAD, the *PSEN2*^{K115fs} mutation had been studied to investigate the differences between this mutation and other frame shift mutations and how this mutation could induce AD pathology. It has been proposed that decreased PSEN function may contribute to AD pathogenesis [3], or that the truncated PSEN2 protein encoded by the mutated coding sequence may function like the hypoxia-induced truncated PSEN2 isoform PS2V, which shows increased expression under hypoxia and in sAD brains and can stimulate γ-secretase activity and suppress the unfolded protein response [4]. However, a recent study in 2019 identified an alternative transcript splicing isoform from the *PSEN2*^{K115fs} allele in humans that includes 77-nucleotides of intronic sequence leading to the restoration of the open reading frame [5]. This means that,

in fact, the *PSEN2*^{K115fs} mutation also follows the "fAD mutation reading frame preservation rule" like all other EOfAD-causative mutations in the *presenilin* genes and is not an exception to this rule.

Previously, we modeled the human *PSEN2*^{K115fs} mutation in zebrafish by introducing an equivalent 2 nucleotide deletion into the coding sequence of the *psen1* gene, *psen1*^{K97fs} (This mutation was introduced into *psen1*, rather than *psen2* because, in contrast to humans, a PS2V-related isoform is produced in zebrafish from its *psen1* gene rather than from *psen2*, details in ***section 1.9.1***). However, the recent identification of a low-abundance, alternatively spliced isoform from the human *PSEN2* gene that restores the open reading frame in *PSEN2*^{K115fs} mutants raised the question of whether the *psen1*^{K97fs} mutation was influenced by an equivalent isoform and so still a suitable model of the human *PSEN2*^{K115fs} mutation. Therefore, we sought to determine whether such an isoform occurs in zebrafish. This chapter describes an effort to detect alternative splicing in heterozygous *psen1*^{K97fs} mutant brains by using RT-PCR to amplify a region of cDNA spanning the exon containing the mutation and multiple neighbouring exons. However, no alternatively spliced transcript isoforms were identified, indicating that the *psen1*^{K97fs} mutation is likely not a good model of *PSEN2*^{K115fs}.

3.2 Results

A low-abundance, alternative splicing isoform of human *PSEN2* restores the open reading frame in *PSEN2*^{K115fs} mutants through insertion of 77-nucleotides of intronic sequence upstream of the exon in which the *PSEN2*^{K115fs} mutation residues [5]. The *psen1*^{K97fs} mutation is located in zebrafish *psen1* exon 3. As *psen1*^{K97fs} is a model of human *PSEN2*^{K115fs}, we needed to detect whether there is a similar alternatively spliced isoform from the *psen1*^{K97fs} allele that inserts intronic sequence near exon 3 and restores open reading frame. Therefore, two primer pairs for RT-PCR covering the mutation region were designed for the detection of aberrant splicing. Their binding sites are showed in **Figure 3.1**.

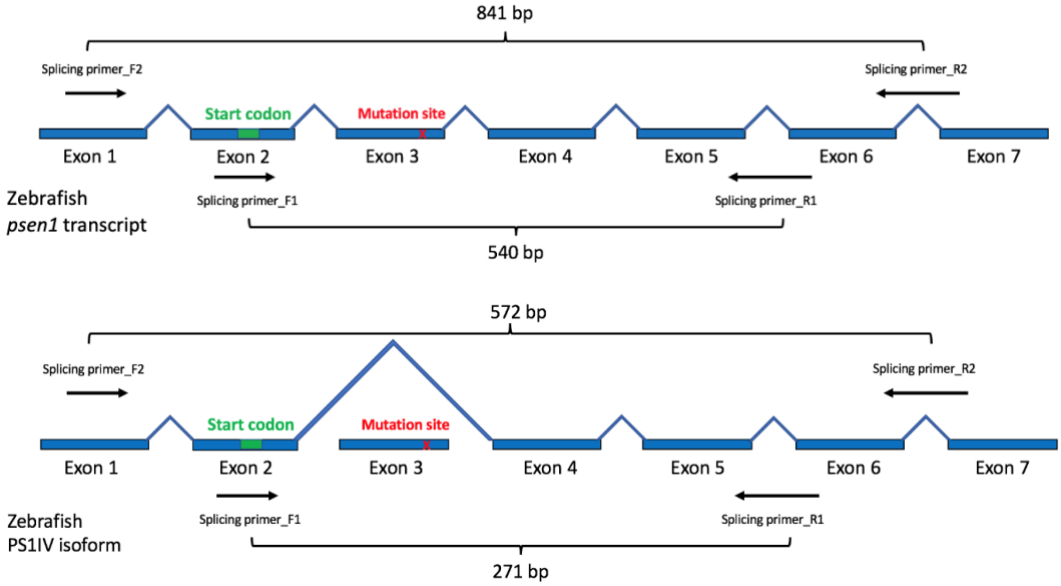


Figure 3.1. The design of RT-PCR "splicing" primer pairs for detection of alternative splicing in *psen1*$^{K97fs/+}$ mutants. These primer pairs should amplify cDNA derived from the normal *psen1* transcript (**Upper**) and the PS1IV isoform (**Lower**) which results from hypoxia-induced exclusion of exon 3 [4]. Under normoxia, the expression of PS1IV is at a low level and can be difficult to detect.

Note that the exons shown above have the following ID codes in the Ensembl database based on transcript psen1-201 with Ensembl ID: ENSDART00000149864.3:

***psen1* exon number**	**Ensembl ID**
1	ENSDARE00001050629
2	ENSDARE00001003996
3	ENSDARE00000951505
4	ENSDARE00000109023
5	ENSDARE00000086874
6	ENSDARE00000138525
7	ENSDARE00000957691

PCRs were performed on cDNA generated from 24-month-old heterozygous *psen1*^{K97fs} mutant brains, and the results are shown in **Figure 3.2**. In the tests of both primer pairs, all wild-type and mutant samples showed a single clear band at the expected product size of normal *psen1* transcripts. No products of a size indicating the presence of PS1IV transcripts were seen, probably because the expression of PS1IV transcripts was at a very low level under normoxia. No aberrant splicing products were identified. Therefore, this result informed us that our zebrafish *psen1*^{K97fs} mutation probably is not a very informative model of the human *PSEN2*^{K115fs} mutation, as it does not appear to produce an alternative transcript that restores the open reading frame.

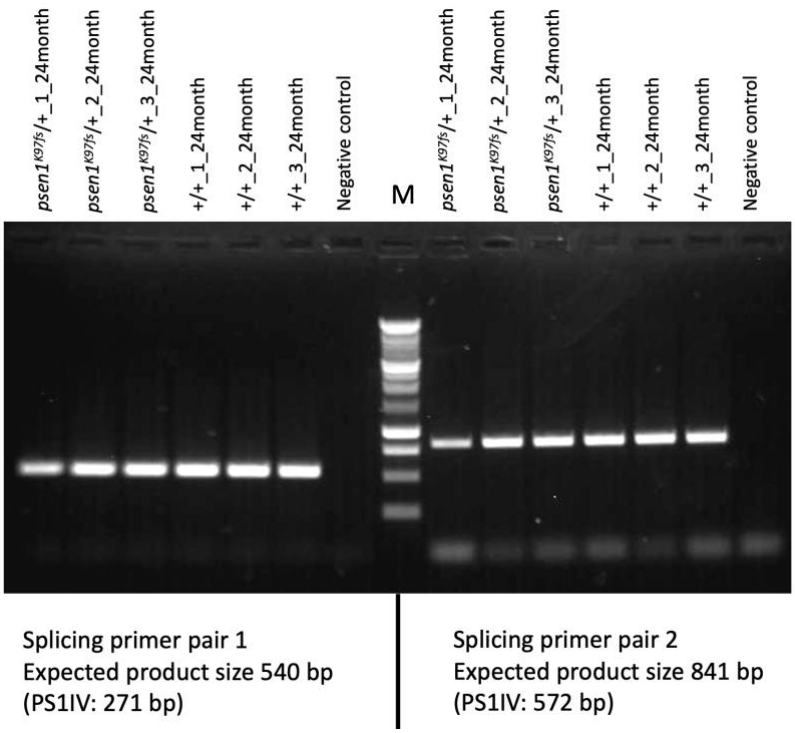


Figure 3.2. PCR tests for detection of alternative splicing on cDNA generated from 24-month-old heterozygous *psen1*^{K97fs} mutant brains. "Splicing primer pair 1" is Splicing_Primer_F1 with Splicing_Primer_R1, and "Splicing primer pair 2" is Splicing_Primer_F2 with Splicing_Primer_R2, corresponding to the primer binding sites illustrated in **Figure 3.1**.

3.3 Discussion

The identification of an alternatively spliced isoform in human *PSEN2*, which restores the open reading frame in the *PSEN2*^{K115fs} mutant allele, indicates that the pathological mechanism through which this mutation drives EOfAD likely does not need to involve mimicry of the naturally truncated isoform PS2V. Instead, like other EOfAD-causative mutations, translation of a "full-length" protein still appears to be an essential element in the pathogenesis of EOfAD. As no alternatively spliced isoform from the *psen1*^{K97fs} allele has been identified that restores this gene's open reading frame, the zebrafish *psen1*^{K97fs} mutation is probably not a close model of the pathological action of human *PSEN2*^{K115fs}. Consequently, the DE genes and predicted pathways identified in transcriptome analysis of zebrafish heterozygous *psen1*^{K97fs} mutant brains may not reflect the molecular genetic state of EOfAD brains. Therefore, based on our current understanding, *psen1*^{K97fs} mutants were not regarded as an EOfAD-like model in my following experiments but analysis of them may help to exclude irrelevant effects induced by reduced PSEN2 expression. A more EOfAD-like mutant, *psen1*^{Q96K97del}, was subsequently used to investigate EOfAD-like effects.

3.4 Methods

RNA purification and cDNA synthesis

Three heterozygous *psen1*^{K97fs} mutant zebrafish and three their wild type siblings at 24 months of age were euthanised in a loose icy slurry before brain extraction. Total RNA was then extracted from each zebrafish brain using the QIAGEN RNeasy mini Kit (QIAGEN, Hilden, Germany). 250 ng of total RNA from each brain was then used to synthesise 20 μL of first-strand cDNA by reverse transcription (SuperScript III kit, Invitrogen, Camarillo, California, USA).

RT-PCR for alternative splicing products

One unit of Phusion high-fidelity DNA polymerase (New England Biolabs, Ipswich, Massachusetts) was used in each 25 μL PCR reaction containing 10 ng of a cDNA sample, 0.2 mM of each deoxyribonucleotide triphosphate (dNTP, each at a final concentration of 0.2 mM), and 0.4 μM of each PCR primer. PCR cycling was performed with 35 cycles of a denaturation temperature of 95°C for 30 s, then an annealing temperature of 60°C for 30 s and then an extension temperature of 72°C for 2 minutes. PCR products were electrophoresed through a 1.5% agarose gel in 1×TAE buffer for separation and identification.

3.5 Appendix: Primer information

Primers for detection of alternative splicing

Ensembl gene ID	Target gene	Primer sequences	Annealing temperature (°C)
ENSDARG00000004870	*psen1* (pair1)	5'-CACAGTTCCGATGGCTGATTTAG-3'	60
		5'-GTCTTGAACACTTCCCCCAAA-3'	
ENSDARG00000004870	*psen1* (pair2)	5'-CACTCTATCGGGCTTTTGTCAG-3'	60
		5'-ACACTGCCAGAAGATCGTAGACT-3'	

3.6 References

1. Cruts, M., J. Theuns, and C. Van Broeckhoven, *Locus-specific mutation databases for neurodegenerative brain diseases.* Hum Mutat, 2012. **33**(9): p. 1340-4.
2. Jayne, T., et al., *Evidence For and Against a Pathogenic Role of Reduced gamma-Secretase Activity in Familial Alzheimer's Disease.* J Alzheimers Dis, 2016. **52**(3): p. 781-99.
3. Jayadev, S., et al., *Alzheimer's disease phenotypes and genotypes associated with mutations in presenilin 2.* Brain, 2010. **133**(Pt 4): p. 1143-54.
4. Moussavi Nik, S.H., et al., *Alzheimer's disease-related peptide PS2V plays ancient, conserved roles in suppression of the unfolded protein response under hypoxia and stimulation of gamma-secretase activity.* Hum Mol Genet, 2015. **24**(13): p. 3662-78.
5. Braggin, J.E., et al., *Alternative splicing in a presenilin 2 variant associated with Alzheimer disease.* Ann Clin Transl Neurol, 2019. **6**(4): p. 762-777.

Chapter 4: Transcriptome analyses of 7-day-old heterozygous *psen1*$^{Q96\_K97del}$ zebrafish larvae

4.1 Summary

Zebrafish larvae can be used for screening of chemical libraries in drug discovery. A 2015 paper by Wagner et al. [1] showed that the most effective drugs in an animal model (of dyslipidemia) were those able to revert disease transcriptome signatures back to normal. In accordance with this philosophy, we investigated whether we could use our EOfAD-like heterozygous *psen1*$^{Q96\_K97del}$ mutant larvae for the discovery of drugs that can reverse their presumably pathological transcriptome state to normal. Although the previous brain transcriptome analysis of 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutants indicated defects in mitochondrial function and lysosomal acidification, similar transcriptome effects would need to be seen in 7 day post fertilization (dpf) mutant larvae if we were use them in a drug screen based on restoring normality to the transcriptome. Therefore, we performed transcriptome analysis on clutches of 7 dpf wild type and heterozygous *psen1*$^{Q96\_K97del}$ larvae using a specific paired-mating strategy. We performed transcriptome analysis on whole larvae rather than on larval brains, although the magnitude of some transcriptome changes only occurring in nervous tissues might be diluted or masked by the inclusion of transcriptomes from other cells of whole larvae. The major reason is that larval brains are so tiny that large numbers of larval brains would be required to generate sufficient RNA for RNA-seq. Physically isolating this tissue from larvae would also be difficult technically. Also, any high-

throughput drug screen would be designed to use whole larvae, as it would be extremely technically challenging to generate sufficient RNA from individual larval brains for assessment of brain transcriptome changes. In larval transcriptome analysis, 228 differentially expressed (DE) genes were identified, and gene set enrichment analysis (GSEA) and Goseq analysis were used to predict the effects of the mutation on cellular functions. Due to the lack of correspondence between DE genes identified in heterozygous *psen1*$^{Q96\_K97del}$ mutant 7 dpf larvae and in heterozygous mutant 6-month-old brains as well as the variability of gene expression level between larvae clutches of the same genotype, *psen1*$^{Q96\_K97del}$ larvae appear to be unsuitable for the discovery of drugs to suppress the transcriptome changes seen in *psen1*$^{Q96\_K97del}$ mutant brains. However, we still observed similar affected cellular pathways between the larvae and the brains, including oxidative phosphorylation, mitochondrial function, lysosomal acidification, and iron ion transport although the larvae lacked detectable enrichment of the prevalence of transcripts containing iron responsive elements (IREs) in their 3'UTRs. Also, some apparently larva-specific effects of the mutation were identified, including effects on extracellular matrix (ECM) and minichromosome maintenance protein complex (mcm) function.

References

1. Wagner, A., et al., Drugs that reverse disease transcriptomic signatures are more effective in a mouse model of dyslipidemia. Mol Syst Biol, 2015. 11(3): p. 791.

4.2 Transcriptome analyses of 7-day-old zebrafish larvae possessing a familial Alzheimer's disease-like mutation in *psen1* indicate effects on oxidative phosphorylation, ECM and mcm functions, and iron homeostasis

Statement of authorship

Statement of Authorship

Title of Paper	Transcriptome analyses of 7-day-old zebrafish larvae possessing a familial Alzheimer's disease-like mutation in psen1 indicate effects on oxidative phosphorylation, ECM and mom functions, and iron homeostasis
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Has been submitted to BMC genomics and is under peer review.

Principal Author

Name of Principal Author (Candidate)	Yang Dong		
Contribution to the Paper	Research plan, most of the experimentation and data analysis, drafting 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	Yang Dong	Date	21/Aug/2020

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 in 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	Supervision in laboratory		
Signature		Date	21/08/2020

Name of Co-Author	Stephen Martin Pederson		
Contribution to the Paper	Supervision of data analysis and editing of manuscript		
Signature		Date	21/08/2020

Name of Co-Author	Nhi Hin		
Contribution to the Paper	Provided IRE data sets for analysis		
Signature		Date	21/08/2020

Name of Co-Author	Michael Lardelli		
Contribution to the Paper	Research plan, supervision, editing and revision of manuscript drafts		
Signature		Date	21/08/2020

Transcriptome analyses of 7-day-old zebrafish larvae possessing a familial Alzheimer's disease-like mutation in *psen1* indicate effects on oxidative phosphorylation, ECM and mcm functions, and iron homeostasis

Yang Dong[1,*], Morgan Newman[1], Stephen M. Pederson[2], Nhi Hin[1,2] & 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

* Corresponding author. Email: a1661360@adelaide.edu.au

Abstract

Background: Early-onset familial Alzheimer's disease (EOfAD) is promoted by dominant mutations, enabling the study of Alzheimer's disease (AD) pathogenic mechanisms through generation of EOfAD-like mutations in animal models. In a previous study, we generated an EOfAD-like mutation, *psen1*$^{Q96\_K97del}$, in zebrafish and performed transcriptome analysis comparing entire brains from 6-month-old wild type and heterozygous mutant fish. We identified predicted effects on mitochondrial function and endolysosomal acidification. Here we aimed to determine whether similar effects occur in 7 day post fertilization (dpf) zebrafish larvae that might be exploited in screening of chemical libraries to find ameliorative drugs.

Results: We generated clutches of wild type and heterozygous *psen1*$^{Q96\_K97del}$ 7 dpf larvae using a paired-mating strategy to reduce extraneous genetic variation before performing a comparative transcriptome analysis. We identified 228 differentially expressed genes and performed Goseq analysis and gene set enrichment analysis (GSEA) to predict cellular functions.

Conclusions: Our analyses predicted a significant effect on oxidative phosphorylation, consistent with our earlier observations of predicted effects on ATP synthesis in adult heterozygous *psen1*$^{Q96\_K97del}$ brains. The dysregulation of minichromosome maintenance protein complex (MCM) genes strongly contributed to predicted effects on DNA replication and the cell cycle and may explain earlier observations of genome instability due to *PSEN1* mutation. The upregulation of *crystallin* gene expression may be a

response to defective activity of mutant Psen1 protein in endolysosomal acidification. Genes related to extracellular matrix (ECM) were downregulated, consistent with previous studies of EOfAD mutant iPSC neurons and postmortem late onset AD brains. Also, changes in expression of genes controlling iron ion transport were observed without identifiable changes in the prevalence of transcripts containing iron responsive elements (IREs) in their 3' untranslated regions (UTRs). These changes may, therefore, predispose to the apparent iron dyshomeostasis previously observed in 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ EOfAD-like mutant brains.

Background

Alzheimer's disease (AD) is a progressive neurodegenerative brain disorder that eventually develops into dementia. AD is a serious worldwide health issue and shows a trend of increasing disease incidence [1]. AD may be classified in numerous ways. Late onset, sporadic AD, occurs after 65 years of age and is the most common form, contributing to more than 95% of AD cases [2]. This form of AD is affected by multiple factors, including age, diet, life style, genetic and environmental factors [3]. Therefore, it has been difficult to model in animals. An early onset, familial form of AD (EOfAD) shows autosomal, dominant inheritance and contributes less than 5% of all AD cases [4]. As both AD forms share similar pathologies [2], many researchers model EOfAD through genetic manipulation of animals to study AD ontology and pathology in general.

Rodent models are the most commonly used in AD research. However, current transgenic rodent models used in EOfAD studies do not reflect closely the disease state of human patients. In 2017, Hargis and Blalock [5] summarized brain transcriptional profiles in human AD, and compared five transgenic mouse models of AD to human AD profiles. All of these mouse models failed to model the most consistent transcriptional signature of human AD, a downregulation of neuronal and mitochondrial genes. Also, the focus of most AD studies is on the pathologies of the advanced disease, such as the accumulation of amyloid-β peptide and tau protein, and on identification of new biomarkers for early diagnosis. Our laboratory seeks deeper insight into the early molecular states of AD brains to explore disease etiology and molecular mechanisms. We have modeled EOfAD-like mutations in another popular vertebrate animal model, the zebrafish. The zebrafish has a fully sequenced and well annotated genome [6], and has the advantages of rapid development with a relatively short generation time. It is easily manipulated genetically and has the capacity to produce large families of siblings which can then be raised together in the same environment to limit the effects of environmental and genetic noise in molecular analyses [7]. Moreover, zebrafish possess orthologs of the human genes mutated in EOfAD. Most recognized EOfAD-causative mutations have been found in the genes *PSEN1*, *PSEN2* and *APP* [8]. (The majority of these mutations, ~63%, occur in the gene *PSEN1* [9].) The zebrafish orthologs of these genes have been identified as *psen1* [10], *psen2* [11], *appa* and *appb* [12]. Therefore, zebrafish have the potential to model EOfAD mutations for the study of the molecular pathological processes of AD. The zebrafish is also a versatile model for drug screening

as its tiny larvae can be obtained in large numbers and arrayed into microtitre plates for molecular, developmental, or behavioural analyses [13].

One EOfAD-like mutation we have generated is *psen1*$^{Q96\_K97del}$, a deletion of 6 nucleotides in the zebrafish *psen1* gene. This mutation deletes 2 codons but maintains the open reading frame, leading to structural and hydrophilicity changes in the first lumenal loop of the translated protein. Although this mutation is not the exact equivalent of any currently known human EOfAD mutation, there are numerous similar EOfAD mutations that distort the first luminal loop of human *PSEN1* (e.g. *PSEN1*$^{L113\_I114insT}$ [14], *PSEN1*^{P117L} [15]) and, like all the many various and widely distributed EOfAD mutations in the *PRESENILIN* genes, it follows the "fAD mutation reading frame preservation rule" [8].

Like human EOfAD mutations, *psen1*$^{Q96\_K97del}$ has dominant effects when heterozygous. We have observed that the brains of 6-month-old (young, recently sexually mature adult) zebrafish heterozygous for *psen1*$^{Q96\_K97del}$ show transcriptome alterations consistent with disturbances in energy production (ATP synthesis) and lysosomal dysfunction [16]. These may represent the initial stresses that, after decades in humans, lead to AD.

Could our heterozygous mutant zebrafish be used to identify drugs that suppress these molecular defects and so prevent the pathological progression to AD? A 2015 paper by Wagner et al. [17] showed that the most effective drugs in an animal model (of dyslipidemia) were those that best caused reversion of the transcriptome disease signature to normal. In accordance with this philosophy, zebrafish mutants could be used to screen for AD-preventative drugs based on the drugs' ability to revert transcriptome signatures of ATP synthesis disruption and lysosomal dysfunction back to wild type. Therefore, as a first step in assessing the viability of this idea, we were interested to observe whether the transcriptome signatures evident in 6-month-old zebrafish *psen1*$^{Q96\_K97del}$ heterozygous adult mutant brains were discernable in whole zebrafish larvae.

Our previous analysis of *psen1*$^{Q96\_K97del}$ heterozygous adult mutant brain transcriptomes was facilitated by the ability to perform bulk RNA-seq on the entire ~7 mg brains of individual mutant zebrafish and their wild type siblings. While an individual zebrafish larva at 7 days post fertilization (dpf, when feeding would normally begin) is too small

to provide sufficient RNA for bulk RNA-seq analysis without some form of amplification, we can produce clutches of uniformly heterozygous larvae by crossing a homozygous mutant parent fish with a wild type parent. Analysis of pooled RNA from multiple individuals also reduces between-genotype variability due to "averaging" of the mRNA expression levels contributed by each larva in the pool. Also, using a single male fish to produce both a heterozygous mutant clutch and a wild type clutch of larvae (through mating with a single homozygous mutant or wild type female fish respectively) further reduces genetic variability in the analysis (see Figure 3).

In this paper we describe a transcriptome analysis on clutches of 7 dpf heterozygous mutant and wild type larvae structured as described above to minimize genetic variation. This identified 228 potentially differentially expressed (DE) genes. Bioinformatic predictive analysis identified probable significant changes in DNA replication and cell cycle processes, to which changes in the regulation of genes related to the minichromosome maintenance protein complex (MCM) were the main contributors. In addition, effects on iron ion transport were identified, suggesting a potential early disruption of iron homeostasis components that might lead, ultimately, to mitochondrial dysfunction including disruption of ATP synthesis.

Results

Our previous study examined the effects of heterozygosity for the *psen1*$^{Q96\_K97del}$ mutation on the transcriptome of 6-month-old zebrafish brains. The changes in gene expression observed were predicted to affect ATP synthesis and lysosomal acidification [16]. Here we sought to identify the changes present in entire, heterozygous 7 dpf larvae to assess whether these larvae might be a suitable system in which to screen drug libraries for compounds ameliorating the young adult brain ATP synthesis and lysosomal acidification effects. The mating scheme described in Figure 3 was employed to generate n=6 pairs of heterozygous mutant and wild type clutches of larvae. (Power calculations performed since our first publication indicated that n=6 provides a power of approximately 70% for detection of fold-change > 2 at a false discovery rate (FDR) of 0.05 across the vast majority of expressed transcripts in zebrafish brain transcriptomes [18], data not shown.) RNA-seq was performed on RNA purified from these clutches followed by a comparative transcriptome analysis to identify differentially expressed genes and explore potential functional effects caused by the mutation.

Differentially expressed genes (DE genes)

Gene expression differences between wild type and heterozygous *psen1*$^{Q96\_K97del}$/+ clutches were calculated through a design matrix considering each pair of clutches as a factor and genotypes as the common difference.

228 significantly DE genes were identified (Supplementary data 1), although most of these genes showed only minor fold-change differences in expression. Comparison of the significantly DE genes identified from heterozygous mutant 7 dpf larvae with those seen in heterozygous mutant 6-month brains [16], revealed only one gene, *lgals8b*, as common between the two datasets. It is upregulated in both.

To support the accuracy and reliability of the RNA-seq data, relative standard curve quantitative PCRs (qPCRs) were performed for four of the most statistically significantly DE genes that showed relatively large fold-changes in expression. The qPCRs were performed using cDNA synthesized from the same preparations of RNA that were used in the RNA-seq analysis. Three of the four genes were seen to be differentially expressed ($p < 0.05$, Supplementary data 2).

GOseq analysis of pathways and GO terms

To predict the cellular functions affected by heterozygosity for the *psen1*$^{Q96\_K97del}$ mutation, we analysed the DE genes using the Hallmark, KEGG, and Wiki pathway databases and the Gene Ontology database. Different pathway databases may contain different representations of similar biological pathways. Hallmark gene sets summarize well-defined biological states or processes built on the overlapping of several gene set collections, and so are useful to achieve an overall view [19]. The KEGG and Wiki gene sets are two popular pathway databases allowing examination of high-level functions. Different pathway databases might show low between-database consistency due to the incomprehensive gene sets and gene interactions in each category [20]. Therefore, to generate a more comprehensive result, we used both KEGG and Wiki pathway databases for pathway analysis.

Pathway and GO analysis were performed using Goseq, which weighted DE genes and calculated each category's significance amongst DE genes to identify significantly changed pathways or GO terms. Goseq analysis only focuses on the proportions of DE genes in each category but does not consider gene expression fold change and pathway regulation direction. Table 1 shows the Goseq results with a FDR cutoff of 0.05 in the analysis of Hallmark, KEGG and Wiki pathways (Table 1) and of GO terms (Figure 1). In the Hallmark pathway (Table 1), G2M_CHECKPOINT contains genes critical for cell division cycle progression, and E2F_TARGETS includes numerous genes that play essential rolls in the cell cycle and DNA replication [21]. Therefore, the Goseq results of the Hallmark, KEGG and Wiki pathway analyses (Table 1) show significant changes in DNA replication and cell cycle control. Among the DE genes in these two categories, most are members of the minichromosome maintenance (MCM) protein family. Downregulation of the genes *mcm2, mcm3, mcm4, mcm5, mcm6* and *mcmbp* and upregulation of the gene *mcm7* were observed in the heterozygous mutant larvae.

Pathway	DE genes	Genes in category	FDR
Hallmark pathway			
G2M_CHECKPOINT	19	182	1.37E-10
E2F_TARGETS	13	174	2.53E-05
KEGG pathway			
DNA_REPLICATION	7	34	4.17E-06
CELL_CYCLE	8	109	5.29E-04
Wiki pathway			
DNA Replication	6	31	4.26E-05
Cell cycle	7	71	2.08E-04
G1 to S cell cycle control	6	49	2.24E-04

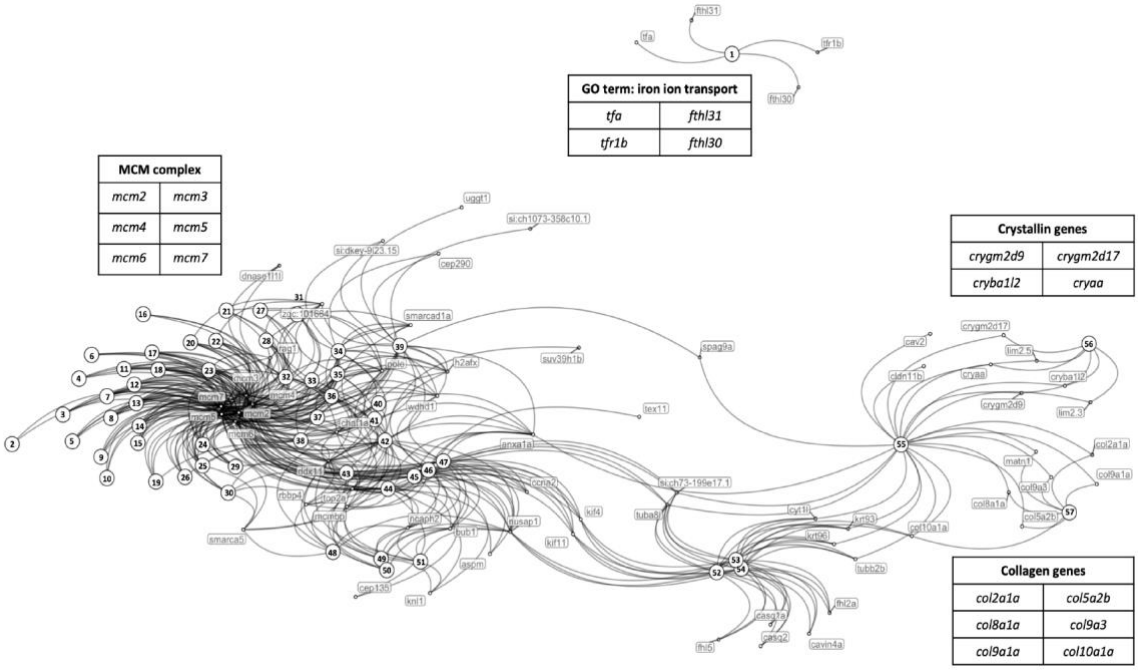
Table 1. Significantly-changed pathways in the Goseq analysis of Hallmark, KEGG and Wiki pathways filtered by a FDR cutoff of 0.05.

In GO analysis, one DE gene can contribute to several related GO terms. The network shown in Figure 1 illustrates how the DE genes are shared between GO terms. Similar to the pathway analyses, most of the GO terms showing significant enrichment for DE genes are related to the cell cycle and DNA replication. In the network, these GOs cluster around the *MINICHROMOSOME MAINTENANCE (mcm)* genes. The network also illustrates how numerous genes can form a functionally related cluster contributing to only one or a few GOs. This is seen for the significantly upregulated *CRYSTALLIN* genes that contribute to eye lens structure (GO: *Structural constituent of eye lens*) but also function in lysosomal acidification (not revealed here, see Discussion). In contrast, the four genes included in the GO *Iron ion transport* show significantly changed regulation. This includes downregulation of the genes *tfa* and *tfr1b* that act to import iron via the endolysosomal pathway [22]. The *ferritin heavy chain like* genes *fthl30* and *fthl31* are upregulated and downregulated respectively, presumably influencing the storage of ferric iron within cells.

We recently published an analysis using a novel method of transcriptome analysis to detect differences in ferrous iron (Fe^{2+}) status in cells [23]. Using this technique, we detected for the first time, that young (6-month-old) adult brains from *psen1*$^{Q96\_K97del}$/+ zebrafish are likely deficient for ferrous iron. Therefore, we were very interested to see evidence of iron ion transport gene expression changes in the 7 dpf *psen1*$^{Q96\_K97del}$/+ larvae. To confirm these changes in gene expression we performed qPCRs for the genes

tfa, *tfr1b*, and *fthl31* on cDNA made from the same mRNA samples that were subjected to RNA-seq (see Supplementary data 2, *fthl31* was not examined because its expression level is particularly low). The qPCRs for these three genes were consistent with the RNA-seq results.

When ferrous iron is deficient in cells, Iron Regulatory Proteins bind to Iron-Responsive elements in the 3' untranslated regions (3'UTRs) of mRNAs encoding proteins that function to increase ferrous iron levels (such as human TFR1 [24] or zebrafish Tfr1b [25]). To detect ferrous iron dyshomeostasis in transcriptome data, we looked for enrichment of a large set of gene mRNAs that include putative IREs in their 3' UTRs. We did not see enrichment of this gene set in the 7 dpf *psen1*$^{Q96\_K97del}$/+ zebrafish larvae, likely indicating that the apparent ferrous iron deficiency of young adult *psen1*$^{Q96\_K97del}$/+ brains requires time to develop (Supplementary data 4).



Number	GO ID	Terms	Number	GO ID	Terms	Number	GO ID	Terms
1	GO:0006826	iron ion transport	20	GO:0000725	recombinational repair	39	GO:0033554	cellular response to stress
2	GO:1990518	single-stranded 3'-5' DNA helicase activity	21	GO:0022616	DNA strand elongation	40	GO:0032508	DNA duplex unwinding
3	GO:0043142	single-stranded DNA helicase activity	22	GO:0000724	double-strand break repair via homologous recombination	41	GO:0032392	DNA geometric change
4	GO:1902969	mitotic DNA replication	23	GO:0065004	protein-DNA complex assembly	42	GO:0051276	chromosome organization
5	GO:0017116	single-stranded DNA helicase activity	24	GO:0071824	protein-DNA complex subunit organization	43	GO:0071103	DNA conformation change
6	GO:0043138	3'-5' DNA helicase activity	25	GO:0004386	helicase activity	44	GO:0022402	cell cycle process
7	GO:0003697	single-stranded DNA binding	26	GO:0003678	DNA helicase activity	45	GO:1903047	mitotic cell cycle process
8	GO:0006270	DNA replication initiation	27	GO:0006271	DNA strand elongation involved in DNA replication	46	GO:0007049	cell cycle
9	GO:0006268	DNA unwinding involved in DNA replication	28	GO:0006302	double-strand break repair	47	GO:0000278	mitotic cell cycle
10	GO:1902975	mitotic DNA replication initiation	29	GO:0042555	MCM complex	48	GO:0006323	DNA packaging
11	GO:0000727	double-strand break repair via break-induced replication	30	GO:0008094	DNA-dependent ATPase activity	49	GO:0007059	chromosome segregation
12	GO:0033260	nuclear DNA replication	31	GO:1902315	nuclear cell cycle DNA replication initiation	50	GO:0000819	sister chromatid segregation
13	GO:0044786	cell cycle DNA replication	32	GO:0140097	catalytic activity, acting on DNA	51	GO:0098813	nuclear chromosome segregation
14	GO:0003688	DNA replication origin binding	33	GO:0006310	DNA recombination	52	GO:0099512	supramolecular fiber
15	GO:1902299	pre-replicative complex assembly involved in cell cycle DNA replication	34	GO:0006974	cellular response to DNA damage	53	GO:0099081	supramolecular polymer
16	GO:1902292	cell cycle DNA replication initiation	35	GO:0006281	DNA repair	54	GO:0099080	supramolecular complex
17	GO:0006267	pre-replicative complex assembly involved in nuclear cell cycle DNA replication	36	GO:0006259	DNA metabolic process	55	GO:0005198	structural molecule activity
18	GO:0036388	pre-replicative complex assembly	37	GO:0006261	DNA-dependent DNA replication	56	GO:0005212	structural constituent of eye lens
19	GO:0004003	ATP-dependent DNA helicase activity	38	GO:0006260	DNA replication	57	GO:0005201	extracellular matrix structural constituent

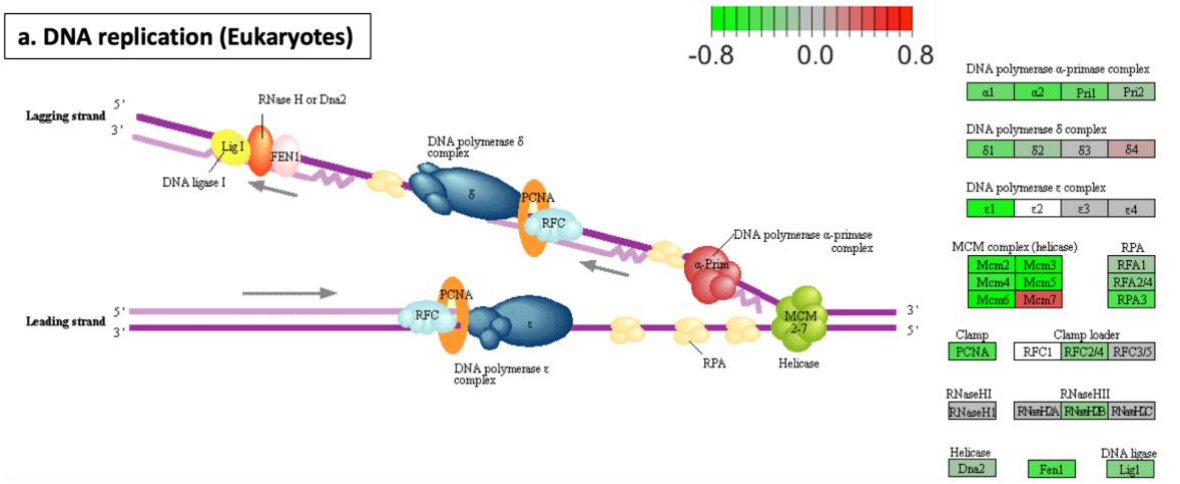
Figure 1. Network of relationships between DE genes and significantly-changed GO terms in the Goseq analysis. Dots represent DE genes and are labelled with gene names., Numbered circles represent those GO terms showing significant enrichment for the DE genes. The table below the network indicates the GO represented by each number.

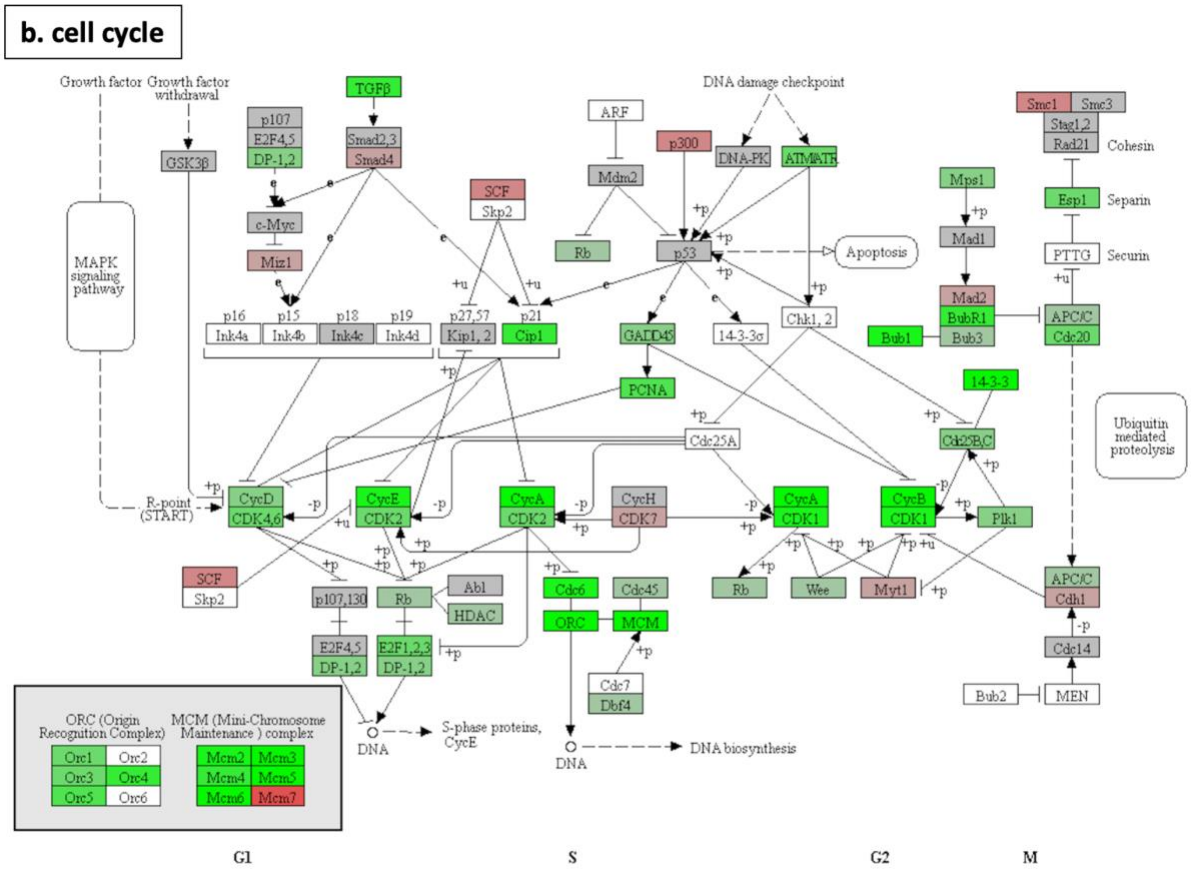
Gene set enrichment analysis (GSEA)

Goseq analysis only focuses on significantly DE genes and predicts affected pathways based on DE gene numbers in each GO. In contrast, GSEA ranks all genes based on fold change and P-value, and then estimates their contributions to each pathway. Therefore, GSEA can show pathway regulation direction, and provides a complementary view of gene sets.

Pathway	ES	NES	size	padj
Hallmark pathway				
OXIDATIVE_PHOSPHORYLATION	-0.555804281	-1.899952116	202	6.10E-04
MTORC1_SIGNALING	-0.532824812	-1.8177652	198	6.12E-04
XENOBIOTIC_METABOLISM	-0.524048494	-1.783121213	193	6.14E-04
GLYCOLYSIS	-0.534085982	-1.81565734	191	6.15E-04
EPITHELIAL_MESENCHYMAL_TRANSITION	-0.53480945	-1.814489884	187	6.17E-04
G2M_CHECKPOINT	-0.715387358	-2.420487445	182	6.20E-04
E2F_TARGETS	-0.687885552	-2.316955128	174	6.24E-04
FATTY_ACID_METABOLISM	-0.564815428	-1.883230276	157	6.32E-04
CHOLESTEROL_HOMEOSTASIS	-0.655693592	-1.997376931	76	6.96E-04
INTERFERON_GAMMA_RESPONSE	-0.524159178	-1.741314463	152	0.004444049
ESTROGEN_RESPONSE_LATE	-0.478794934	-1.631035008	195	0.009199858
COAGULATION	-0.542963825	-1.750486063	117	0.010511102
MYC_TARGETS_V1	-0.475280378	-1.620439179	197	0.011637736
MYOGENESIS	-0.464741292	-1.580492418	192	0.035661584
KEGG pathway				
ECM_RECEPTOR_INTERACTION	-0.687182381	-2.073168142	71	0.002601799
FATTY_ACID_METABOLISM	-0.753821419	-2.097193508	43	0.002766581
BETA_ALANINE_METABOLISM	-0.834990432	-2.046674914	22	0.002964805
GLUTATHIONE_METABOLISM	-0.729274008	-2.020247611	42	0.005551576
CELL_CYCLE	-0.581949911	-1.860528972	109	0.007407899
DNA_REPLICATION	-0.732012688	-1.95378306	34	0.014165169
PYRIMIDINE_METABOLISM	-0.583500782	-1.816436301	89	0.022831735
BUTANOATE_METABOLISM	-0.741595286	-1.931869094	30	0.025837321
FOCAL_ADHESION	-0.484232114	-1.642520802	186	0.032170787
OXIDATIVE_PHOSPHORYLATION	-0.537913914	-1.751874325	127	0.036339481
WIKI pathway				
Cell cycle	-0.699655808	-2.108960687	71	0.001162888
G1 to S cell cycle control	-0.736907821	-2.097483391	49	0.001214409
DNA Replication	-0.80698873	-2.117293192	31	0.001272343
Cholesterol Biosynthesis	-0.837022118	-1.894160937	15	0.020452081

Table 2. Significantly-changed pathways in the GSEA analysis of Hallmark, KEGG and Wiki pathways filtered by a Bonferroni correction P-value of 0.05. ES and NES indicate enrichment score and normalized enrichment score respectively. NES is generated through normalizing enrichment score to mean enrichment score of random samples. Size presents the numbers of genes contributing to each pathway.


a. DNA replication (Eukaryotes)
-0.8
0.0
0.8
Lagging strand
Leading strand
RNase H or Dna2
Lig I
FEN1
DNA ligase I
DNA polymerase δ complex
PCNA
RFC
DNA polymerase α-primase complex
α-Prim
MCM 2-7
Helicase
RPA
DNA polymerase ε complex
DNA polymerase α-primase complex
α1 α2 Pri1 Pri2
DNA polymerase δ complex
δ1 δ2 δ3 δ4
DNA polymerase ε complex
ε1 ε2 ε3 ε4
MCM complex (helicase)
Mcm2 Mcm3
Mcm4 Mcm5
Mcm6 Mcm7
RPA
RFA1
RFA2/4
RPA3
Clamp
PCNA
Clamp loader
RFC1 RFC2/4 RFC3/5
RNaseHI
RNaseH1
RNaseHII
Helicase
Dna2
Fen1
DNA ligase
Lig1


b. cell cycle
Growth factor
Growth factor withdrawal
TGFβ
GSK3β
MAPK signaling pathway
p107
E2F4,5
DP-1,2
Smad2,3
Smad4
c-Myc
Miz1
SCF
Skp2
ARF
Mdm2
DNA damage checkpoint
p300
DNA-PK
ATM&ATR
Rb
p53
Apoptosis
Chk1, 2
GADD45
14-3-3σ
p16 Ink4a
p15 Ink4b
p18 Ink4c
p19 Ink4d
p27,57 Kip1, 2
p21 Cip1
PCNA
Cdc25A
R-point (START)
CycD CDK4,6
CycE CDK2
CycA CDK2
CycH CDK7
CycA CDK1
CycB CDK1
Plk1
p107,130
Rb
Abl
HDAC
E2F4,5
DP-1,2
E2F1,2,3
DP-1,2
Cdc6
Cdc45
ORC
MCM
Cdc7
Dbf4
Rb
Wee
Myt1
Mps1
Mad1
Mad2
BubR1
Bub3
Bub1
Smc1 Smc3
Stag1,2
Rad21
Cohesin
Esp1
Separin
PTTG
Securin
APC/C Cdc20
14-3-3
Cdc25B,C
Ubiquitin mediated proteolysis
APC/C Cdh1
Cdc14
Bub2
MEN
ORC (Origin Recognition Complex)
Orc1 Orc2
Orc3 Orc4
Orc5 Orc6
MCM (Mini-Chromosome Maintenance) complex
Mcm2 Mcm3
Mcm4 Mcm5
Mcm6 Mcm7
DNA
S-phase proteins, CycE
DNA
DNA biosynthesis
G1
S
G2
M

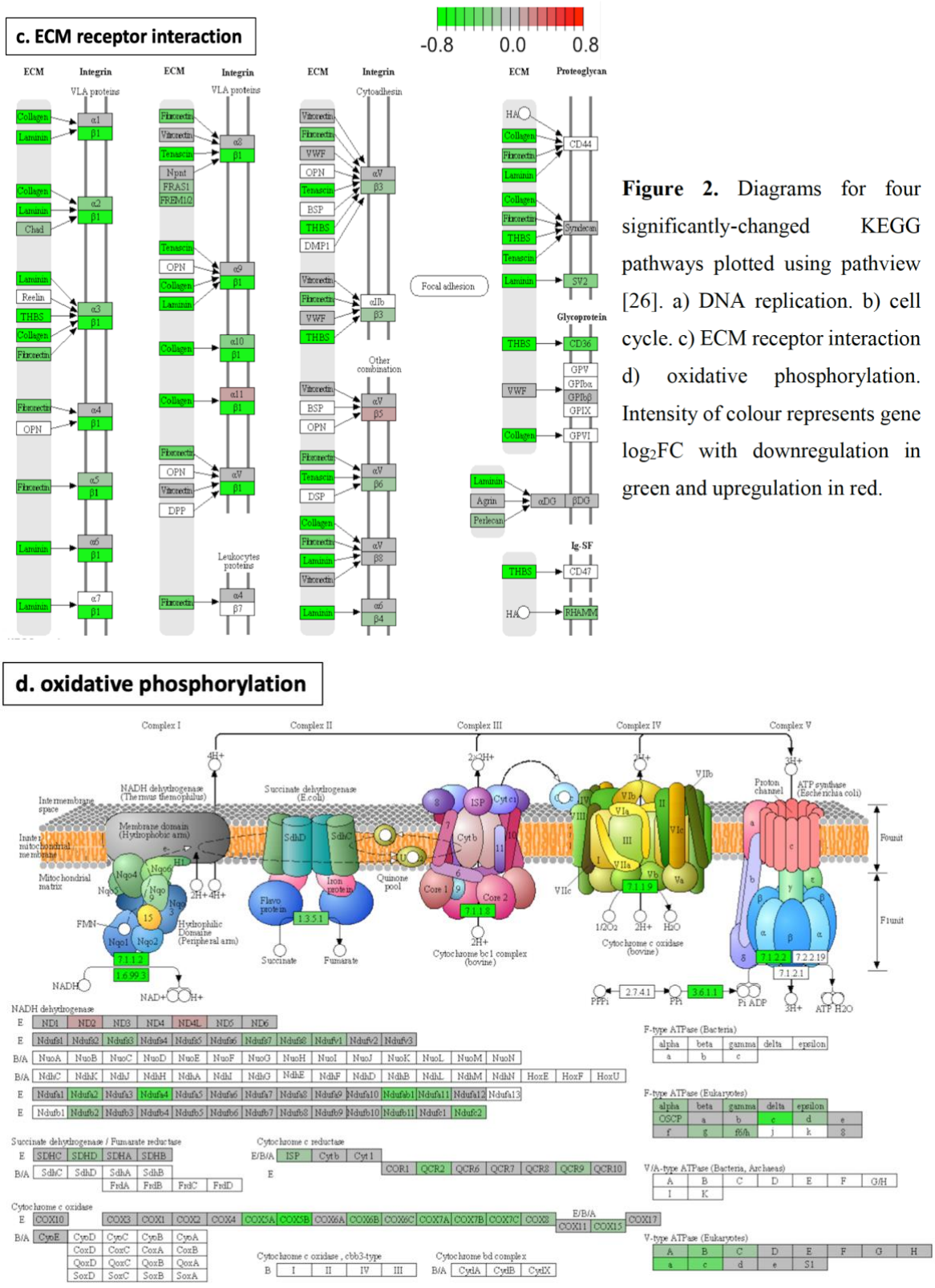


Figure 2. Diagrams for four significantly-changed KEGG pathways plotted using pathview [26]. a) DNA replication. b) cell cycle. c) ECM receptor interaction d) oxidative phosphorylation. Intensity of colour represents gene $\log_2$FC with downregulation in green and upregulation in red.

We applied GSEA using the Hallmark, KEGG and Wiki pathway databases. Several significantly-changed pathways were identified in each analysis (Table 2), including pathways previously identified in the Goseq pathway analysis. Four of the significantly-changed KEGG pathways are illustrated in Figure 2. *DNA replication* (Figure 2a) and *cell cycle* (Figure 2b) were the most significantly affected pathways identified in the Goseq pathway analysis and the GO analysis. Regulation of the MCM complex plays essential roles in both pathways. The MCM complex forms a DNA helicase, which cooperates with replication protein A (RPA) to unwind duplex parental DNA before DNA synthesis (Figure 2a, [27]). Dysregulation of the MCM complex would influence the DNA replication and might cause replication stress leading to genomic instability [28]. The pathways *ECM receptor interaction* (Figure 2c) and *oxidative phosphorylation* (OXPHOS, Figure 2d) were also significantly changed in 7 dpf *psen1*$^{Q96\_K97del}$/+ zebrafish larvae. *ECM receptor interaction* was the most significantly changed pathway in KEGG pathway analysis (the lowest P-value), and most genes involved were downregulated (Figure 2c), including the *COLLAGEN* gene group identified in the previous GO analysis. The KEGG pathway *ECM receptor interaction* plays important roles in control of cellular activities, including functioning to provide cell structural support and to regulate cell-cell and cell matrix interactions [29]. In developing brains, *ECM receptor interaction* participates in cell migration and the guidance of growing axons, having crucial effects on neural cells. This has implicated *ECM receptor interaction* in processes underlying many central nervous system diseases such as AD, schizophrenia and Parkinson's disease [30]. OXPHOS (Figure 2d), as well as *fatty acid metabolism* (shown in Table 2), contribute to the fundamentally important function of energy production. In our previous GO analysis of 6-month-old *psen1*$^{Q96\_K97del}$/+ zebrafish brains, we saw very significant apparent effects on ATP synthesis [16]. The analysis here suggests that that energy production capacity is downregulated in the mutant larvae and this is expected to include ATP synthesis. Furthermore, *Beta-alanine metabolism, glutathione metabolism, pyrimidine metabolism, butanoate metabolism* and *focal adhesion* are also identified as significantly-changed pathways (Table 2). The interpretation of these pathway changes requires further investigation. KEGG diagrams for the statistically significantly affected pathways not shown in Figure 2 are given in Supplementary data 3.

Discussion

No suitable transcriptome biomarkers identified for drug screening

Zebrafish larvae represent a powerful system for screening of chemical libraries in drug discovery [13]. However, the lack of consistency in transcriptome changes seen between heterozygous mutant *psen1*$^{Q96\_K97del}$ 7 dpf larvae compared to heterozygous mutant 6-month-old brains would appear to preclude the use of these mutant larvae to find drugs to suppress the transcriptome changes seen in the brains. Also, although the RNAs used for transcriptome analyses were extracted from clutches of larvae rather than individuals, the expression levels of the DE genes were, nevertheless, quite variable between clutches of the same genotype, either wild type or mutant (Supplementary data 2). We were able to identify significantly changed cellular pathways in common between the larvae and the brains. These related to oxidative phosphorylation, mitochondrial function and lysosomal acidification, indicating that similar stresses/biological effects caused by the presence of the *psen1*$^{Q96\_K97del}$ mutation likely exist during the entirety of zebrafish development from larvae to young adult. However, analysis of transcriptomes at this level would be unsuitable for the massively parallel screening of chemical libraries. Therefore, use of the *psen1*$^{Q96\_K97del}$ EOfAD-like mutation for discovery of AD-preventative drugs remains infeasible until a suitable biomarker can be identified.

MCM complex dysregulation may drive DNA replication stress

Comparison of the transcriptomes of pools of 7 dpf heterozygous *psen1*$^{Q96\_K97del}$ mutant larvae to their wild type siblings revealed highly significant regulatory effects on genes involved with DNA replication and the cell cycle. These were identified by both Goseq analysis and GSEA. The majority of DE genes contributing to these two terms are related to the minichromosome maintenance (MCM) protein family. The eukaryotic MCM complex functions as a DNA helicase essential for DNA replication and cell division. The complex is comprised of the protein products of six genes, MCM2-7 (Figure2a, [31]). We observed a downregulation of zebrafish genes *mcm2, mcm3, mcm4, mcm5, mcm6* and *mcmbp* and an upregulation of *mcm7*. If this gene dysregulation phenomenon affects other vertebrates, including mammals, it may clarify a hitherto unexplained mutagenic effect of EOfAD mutations in, specifically, *PSEN1*. In 2002, Chan et al. [32] showed that forced expression of EOfAD mutant *PSEN1*, but not wild type *PSEN1*, increased the sensitivity of Rat pheochromocytoma (PC12) cells to DNA damage by etoposide. Responses included greater than normal increases in p53 protein levels and

phosphorylation. In 2010, Michelsen et al. [33], studied the effects in mouse brains of transgenes expressing EOfAD mutant forms of the genes *APP*, or *PSEN1*, or both simultaneously. They observed an increase in the number of single-strand DNA breaks occurring in hippocampal granule cells of the dentate gyrus and hippocampal pyramidal cells in areas CA1/2 when brains expressed only the *PSEN1* mutant transgene. Interestingly, simultaneous expression of EOfAD mutant forms of both *APP* and *PSEN1* reduced the single stand break rate so that it was similar to that seen for only *APP* mutant transgene expression alone. In mice, reduced function of the *Mcm4* gene led to susceptibility to chromosome breaks induced by a DNA replication inhibitor, aphidicolin [34]. Notably, in 2011, Yurov et al. [28] suggested a DNA replication stress hypothesis of AD which proposes that replication stress caused by incomplete DNA replication leads to DNA damage or improper repair, subsequently resulting in the accumulation of genomic instabilities in AD brains. Genomic instabilities are associated with neurodegeneration in other aging-related diseases [35, 36]. Our observations, those of others (above), and Yurov et al.'s ideas suggest that associations between *PSEN1* functions and DNA integrity are an interesting area for further exploration and may give us greater insight into cellular stresses driving AD pathologies.

Are crystallin genes upregulated due to disturbance of lysosomal acidification?

The discovery that DE genes were enriched under the GO term *structural constituent of eye lens*, focused our attention on the crystallin genes *crygm2d9*, *crygm2d17*, *cryba1l2*, and *cryaa*. Crystallin genes have not previously been linked with AD or *PSEN* functions. However, a paper by Valapala et al., 2013 [37] reported that the loss of the βA3/A1-crystallin gene in rat (*Cryba1*) reduces endolysosomal acidification, leading to reduced γ-secretase-mediated release of Notch intracellular domain (NICD) and impaired lysosomal-mediated degradation of Notch. Overexpression of *Cryba1* in a *Cryba1*$^{-/-}$ knockout mouse apparently was able to rescue the deficient function of the vacuolar ATPase (v-ATPase) responsible for endolysosomal acidification. EOfAD mutations of human *PSEN1* are thought to reduce lysosomal acidification due to a requirement for the PSEN1 holoprotein for correct N-glycosylation of the v-ATPase subunit, V0a1 [38]. Our transcriptome analysis of the effect of heterozygosity for the *psen1*$^{Q96\_K97del}$ mutation on 6-month-old zebrafish brains also implied effects on lysosomal acidification [16]. While lysosomal acidification was not revealed as an affected GO in our analysis of 7 dpf

heterozygous *psen1*$^{Q96\_K97del}$ larvae, we speculate that the upregulation of crystallin genes observed may be a homeostatic response to cope with disturbed lysosomal acidification.

Interestingly, Notch signaling has been seen to control expression of the protein products of the human *mcm* gene orthologues, *MCM2* and *MCM6* in cells where Notch signaling suppresses proliferation [39]. However, since Notch signaling was seen to downregulate these proteins, loss of Notch signaling due to increased lysosomal pH would not appear to explain the downregulation of *mcm2* and *mcm6* gene expression we observed in the heterozygous mutant zebrafish larvae.

Does downregulation of ECM related genes increase the risk of AD?

The most significant KEGG pathway identified in GSEA was *ECM (extracellular matrix) receptor interaction*. Most genes in this pathway, including those encoding collagen, laminin, tenascin and thrombospondin, were significantly downregulated (Figure 2c). *ECM receptor interaction* is not currently a focus in AD studies, although altered regulation of these genes has previously been observed in some AD relevant research. In 2019, Kwart et al. [40] identified 1,515 overlapping DE genes from three human iPSC lines carrying EOfAD mutations in *APP* and *PSEN1*. These DE genes were used to perform a enrichment analysis, and *ECM receptor interaction* was the second most statistically significant KEGG pathway identified. However, the relationships between EOfAD mutations and *ECM receptor interaction* have not been further explored. Another study by Conejero-Goldberg et al., 2014 [41] analysed human postmortem cortex and identified upregulation of ECM-related gene transcripts in carriers of the AD-protective ε2 allele of the gene APOE (i.e. APOE2). Based on this observation, they assumed increased ECM expression would reduce amyloid-β secretion or excitotoxicity. Thus, it appears that increased ECM gene expression is associated with decreased AD risk while EOfAD mutation-associated decrease of ECM gene expression promotes AD pathology. This may indicate the potential for ECM gene expression to act as an AD risk biomarker. More attention should be focused on ECM gene expression in future studies of AD.

Mitochondrial dysfunction is an early effect of the *psen1*$^{Q96\_K97del}$ mutation

In addition to the pathways mentioned above, *fatty acid metabolism*, *oxidative phosphorylation* (*OX PHOS*) and *cholesterol biosynthesis* were identified in the GSEAs of KEGG and Wiki pathways. These pathways represent linked systems that control ATP production to meet cellular energy demand and in response to oxygen availability. Acetyl-CoA, as the starting point of *cholesterol biosynthesis*, is produced by oxidation reactions including oxidative decarboxylation of pyruvate and β-oxidation of fatty acids [42], and is a substrate in the tricarboxylic acid cycle (TCA cycle) to drive oxidative phosphorylation [43]. The AD brain is hypometabolic [44] and mitochondrial dysfunction is associated with oxidative stress in AD neuropathology through reduced ATP production [45]. Reductions in oxidative phosphorylation enzyme activities and functions have been identified in AD and other neurodegenerative processes [46]. A study by Manczak et al. [47] examined the expression of oxidative phosphorylation genes in AD patients and found downregulation of mitochondrial genes coding for electron transport chain (ETC) complex I, which is consistent with the gene expression in our mutant larvae (Figure 2d). However, Manczak et al. saw increased mRNA expression for components of complexes III and VI in contrast to our observations in mutant larvae. A more recent analysis by Mastroeni et al. [48] saw downregulation of nuclear-encoded ETC genes in AD but increased expression of these relative to age-matched controls in mild cognitive impairment (MCI). The downregulation of *fatty acid metabolism*, *oxidative phosphorylation* (*OX PHOS*) and *cholesterol biosynthesis* genes observed in mutant larvae indicates that the implied impairment of energy production by the EOfAD-like *psen1*$^{Q96\_K97del}$ mutation occurs early in life.

Iron homeostasis

We have previously proposed that cellular iron dyshomeostasis may represent a unifying effect-in-common of the EOfAD mutations in *APP*, *PSEN1* and *PSEN2* [22] since APP was thought to stabilize the iron export protein FERROPORTIN [49], while endolysosomal acidification (affected by EOfAD mutations in *PSEN1* [38]) is important for import of iron into cells [50]. Recently, the role of APP in stabilization of FERROPORTIN has been challenged [51, 52]. However, it has been revealed that EOfAD mutations in APP, (like those in *PSEN1*), also affect acidification of the endolysosomal pathway [53] and so would be expected to affect cellular iron homeostasis. Our identification that the GO *iron ion transport* is affected in 7 dpf *psen1*$^{Q96\_K97del}$/+ larvae, particularly with downregulation of the *tfa* and *tfr1b* genes

required for importation of iron, supports that EOfAD mutations in *PSEN1* disturb ferrous iron homeostasis. The fact that we have seen transcriptome evidence for such dyshomeostasis in 6-month-old *psen1*$^{Q96\_K97del}$/+ brains [16] but were unable to detect stabilization of mRNAs containing IREs in their 3'UTRs in 7 dpf larvae, suggests that any disruption of ferrous iron homeostasis begins subtly. Therefore, we propose that gene dysregulation promoting ferrous iron dyshomeostasis occurs by 7 dpf, but that the iron dyshomeostasis requires time to develop before it becomes apparent as ferrous iron dyshomeostasis in young adult *psen1*$^{Q96\_K97del}$/+ brains.

Methods

Ethics

The research described in this paper was carried out under permit S-2017-073 issued to members of the Alzheimer's Disease Genetics Laboratory by the Animal Ethics Committee of the University of Adelaide.

Zebrafish pair-mating breeding strategy

The zebrafish genetic lines used in this study were bred as stocks within the zebrafish facility of the Alzheimer's Disease Genetics Laboratory of the University of Adelaide. Adult fish were returned to their families after mating. A female zebrafish homozygous for the *psen1*$^{Q96\_K97del}$ allele was crossed to a male wild-type fish to generate a clutch of heterozygous *psen1*$^{Q96\_K97del}$/+ larvae. The same male wild-type fish was then crossed to a female wild-type fish to generate a group of wild-type embryos (Figure 3). The use of the male fish as the common parent in this mating scheme avoids the possibility that maternal effects might obscure transcriptome differences between the larval genotypes [54]. These two clutches of larvae from an individual male wild-type fish were labeled a "pair" (Figure 3). From previous work analysing the transcriptomes of *psen1*$^{Q96\_K97del}$/+ whole zebrafish brains compared to the brains of wild type siblings [16], we found that n = 6 would provide a power of ~70% for detection of fold-change > 2 at a false discovery rate (FDR) of 0.05, across the vast majority of expressed transcripts [18]. Therefore, for this analysis involving pooled larvae, we analysed n=6 pairs of larval clutches. Note that no individual adult fish contributed to more than one larval clutch pair.

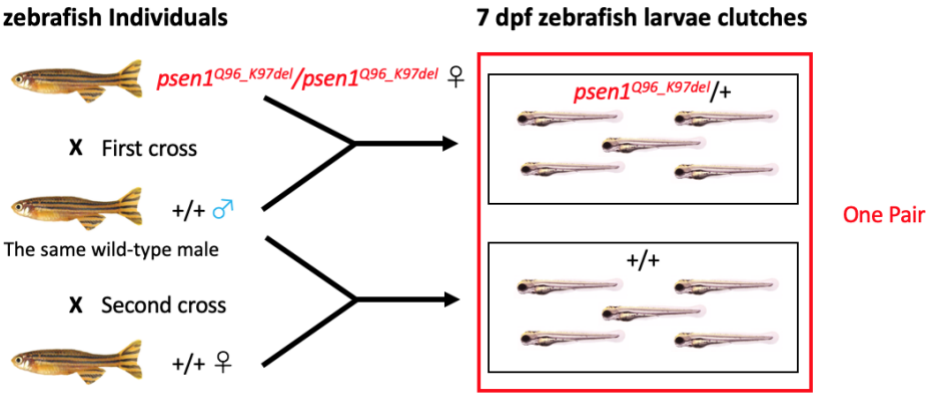


Figure 3. Mating scheme to generate pairs of 7 dpf zebrafish larval clutches.

Sample preparation and RNA purification

Six pairs of larval clutches were generated as described above (12 clutches in total). The parents of 4 pairs of clutches were approximately 5 months old, while the parents of the other two pairs of clutches were approximately 12 months old. The larvae were raised in E3 medium for 7 days, allowing complete larval yolk absorption but not food intake. This was done to minimize the influence of environmental factors on gene expression. As individual larvae are too small to generate sufficient RNA for bulk RNA-seq analysis, 40 larvae were pooled from each clutch (i.e. a total of 480 larvae were consumed in this experiment). Larvae were cooled on ice for anaesthesia before their aqueous medium was replaced with 400 μL of Invitrogen RNA*later*™ Stabilization Solution (Thermo Fisher Scientific, Waltham, MA, USA) and then kept at 4°C for 24 hours before storage at -80°C until lysis for total RNA extraction (below).

Total RNA was extracted from pooled larvae using the mirVanaTM miRNA isolation Kit (Ambion, Life Technologies, Thermo Fisher Scientific, Waltham, MA, USA). DNase treatment was performed on RNA to remove remaining genomic DNA using the DNA-free™ Kit (Ambion, Life Technologies, Thermo Fisher Scientific, Waltham, MA, USA). Purified total RNA was then delivered to the Genomics Facility at the South Australian Health and Medical Research Institute (SAHMRI, Adelaide, Australia) for RNA sequencing. Demultiplexed libraries were provided by the sequencing centre at SAHMRI as 75 bp single-end reads, after using polyA amplification. All libraries were sequenced to a minimum depth of 30 million reads, across two NextSeq lanes which were subsequently merged. The data has been deposited in NCBI Gene Expression Omnibus (GEO) database [55] and are accessible through GEO accession number GSE148631.

Transcriptome data treatment and DE gene identification

Reads from each sample were trimmed using AdapterRemoval [56] setting a minimum quality of 30 and a minimum length of 35. Trimmed reads were aligned to the *Danio rerio* genome Ensembl Release 96 (GRCz11) using STAR v2.5.3a [57], and the aligned reads were assigned to each gene using featureCount [58], specifying that only unique alignments which strictly overlapped exonic regions were counted. The 19,396 genes that received more than 1 count per million reads (CPM) in at least 6 samples, were retained for further analysis. The remaining 12,661 genes were discarded as undetectable genes, giving library sizes ranging between 44,498,373 and 52,779,608 reads. A design

matrix was specified with an intercept for each pair and with genotype as the common difference. An initial Differential Gene Expression (DE) analysis was then performed using the glmLRT method as implemented in edgeR [59, 60]. The 5000 lowest ranked genes from this analysis (i.e. with the highest p-value) were selected as unchanged negative control genes and passed to RUVg [61], setting k = 1. After RUV treatment, the same design matrix with the addition of the offset term from RUV was used to perform a new DE analysis. A FDR cutoff of 0.01 was then used to identify a gene as differentially expressed (DE).

Goseq analysis and Gene set enrichment analysis (GSEA)

The mappings linking zebrafish genes to Hallmark/KEGG pathways, Wiki pathways and GO terms were achieved by msigdbr [19], rWikipathways [62] and org.Dr.eg [63] respectively. To look for enrichment of gene-sets within our defined set of DE genes, a Probability Weighting Function (PWF) was calculated based on the set DE genes using gene lengths as the bias data. Then goseq analysis [64] was performed using this PWF and gene-sets were considered as significantly enriched in the set of DE genes using a FDR threshold of 0.05.

In order to look at the more complete list of all expressed genes, we then used Gene Set Enrichment Analysis (GSEA, [65]) as implemented in the fgsea [66] package. A ranked list was formed using sign($\log_2$FC) * (-$\log_{10}$PValue) as the ranking statistic and the same gene-sets as for goseq were tested, using n = 10^5 iterations. A Bonferroni corrected p-value < 0.05 was used to identify significantly-altered pathways. Pathway Diagrams for detected KEGG pathways were plotted by pathview [26] using $\log_2$FC .

Iron responsive elements (IRE) enrichment analysis

Genes containing iron responsive elements (IRE) in either their 5’ or 3’ UTRs were identified using searching for IREs (SIREs) [67]. Both goseq [64] and GSEA were performed to detect the enrichments of the genes containing IREs in their 5’UTR or 3’UTR respectively. A FDR and a Bonferroni corrected p-value < 0.05 were used to identify significant enrichments in goseq analysis and GSEA respectively.

Abbreviations

CPM: Count per million reads

DE: Differentially expressed

Dpf: Day post fertilization

ECM: Extracellular matrix

EOfAD: Early-onset familial Alzheimer's disease

ETC: Electron transport chain

FDR: False discovery rate

GO: Gene Ontology

GSEA: Gene set enrichment analysis

IRE: Iron responsive elements

MCI: Mild cognitive impairment

MCM: Minichromosome maintenance protein complex

NICD: Notch intracellular domain

OX PHOS: Oxidative phosphorylation

PWF: Probability Weighting Function

qPCRs: Quantitative PCRs

UTRs: Untranslated regions

Declarations

Ethics approval and consent to participate

The research described in this paper was carried out under permit S-2017-073 from the Animal Ethics Committee of the University of Adelaide.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The sequencing data has been deposited in NCBI Gene Expression Omnibus (GEO) database and are accessible through GEO accession number GSE148631.
All data generated or analysed during this study are included in this published article and its supplementary information files.

Funding

YD is supported by an Adelaide Graduate Research Scholarship from the University of Adelaide. ML and MN were both supported by grants GNT1061006 and GNT1126422 from the National Health and Medical Research Council of Australia, https://www.nhmrc.gov.au/. ML and SP are employees of the University of Adelaide. NH is supported by an Australian Government Research Training Program (RTP) Scholarship. No funding body played any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

YD and ML conceived the project. MN provided the homozygous parental *psen1*$^{Q96\_K97del}$ fish. YD conducted fish mating and larvae collection, RNA purification and qPCR validation of RNA-seq data. YD performed analysis of RNA-seq data under the guidance of SP except that NH performed the IRE enrichment analysis. YD drafted

the manuscript that was subsequently edited by SP and ML. All authors read and approved the final manuscript.

Acknowledgements

The authors wish to thank Lachlan Warren William Baer for his technical assistance in construction of the GO plot.

References

1. Gaugler, J., et al., *2019 Alzheimer's disease facts and figures.* Alzheimers & Dementia, 2019. **15**(3): p. 321-387.
2. Masters, C.L., et al., *Alzheimer's disease.* Nat Rev Dis Primers, 2015. **1**: p. 15056.
3. Beydoun, M.A., et al., *Epidemiologic studies of modifiable factors associated with cognition and dementia: systematic review and meta-analysis.* BMC Public Health, 2014. **14**: p. 643.
4. Barber, R.C., *The genetics of Alzheimer's disease.* Scientifica (Cairo), 2012. **2012**: p. 246210.
5. Hargis, K.E. and E.M. Blalock, *Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us?* Behav Brain Res, 2017. **322**(Pt B): p. 311-328.
6. Howe, K., et al., *The zebrafish reference genome sequence and its relationship to the human genome.* Nature, 2013. **496**(7446): p. 498-503.
7. Newman, M., E. Ebrahimie, and M. Lardelli, *Using the zebrafish model for Alzheimer's disease research.* Front Genet, 2014. **5**: p. 189.
8. Jayne, T., et al., *Evidence For and Against a Pathogenic Role of Reduced gamma-Secretase Activity in Familial Alzheimer's Disease.* J Alzheimers Dis, 2016. **52**(3): p. 781-99.
9. Pottier, C., et al., *High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease.* Mol Psychiatry, 2012. **17**(9): p. 875-9.
10. Leimer, U., et al., *Zebrafish (Danio rerio) presenilin promotes aberrant amyloid beta-peptide production and requires a critical aspartate residue for its function in amyloidogenesis.* Biochemistry, 1999. **38**(41): p. 13602-9.
11. Groth, C., et al., *Identification of a second presenilin gene in zebrafish with similarity to the human Alzheimer's disease gene presenilin2.* Dev Genes Evol, 2002. **212**(10): p. 486-90.
12. Musa, A., H. Lehrach, and V.A. Russo, *Distinct expression patterns of two zebrafish homologues of the human APP gene during embryonic development.* Dev Genes Evol, 2001. **211**(11): p. 563-7.
13. Parng, C., et al., *Zebrafish: A preclinical model for drug screening.* Assay and Drug Development Technologies, 2002. **1**(1): p. 41-48.
14. De Jonghe, C., et al., *Aberrant splicing in the presenilin-1 intron 4 mutation causes presenile Alzheimer's disease by increased Abeta42 secretion.* Hum Mol Genet, 1999. **8**(8): p. 1529-40.
15. Wisniewski, T., et al., *A novel Polish presenilin-1 mutation (P117L) is associated with familial Alzheimer's disease and leads to death as early as the age of 28 years.* Neuroreport, 1998. **9**(2): p. 217-21.
16. Newman, M., et al., *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): p. 43.
17. Wagner, A., et al., *Drugs that reverse disease transcriptomic signatures are more effective in a mouse model of dyslipidemia.* Mol Syst Biol, 2015. **11**(3): p. 791.
18. Zhao, S., et al., *RnaSeqSampleSize: real data based sample size estimation for RNA sequencing.* BMC Bioinformatics, 2018. **19**(1): p. 191.
19. Liberzon, A., et al., *The Molecular Signatures Database Hallmark Gene Set Collection.* Cell Systems, 2015. **1**(6): p. 417-425.
20. Soh, D., et al., *Consistency, comprehensiveness, and compatibility of pathway databases.* BMC Bioinformatics, 2010. **11**: p. 449.
21. Bracken, A.P., et al., *E2F target genes: unraveling the biology.* Trends in Biochemical Sciences, 2004. **29**(8): p. 409-417.
22. Lumsden, A.L., et al., *Dysregulation of Neuronal Iron Homeostasis as an Alternative Unifying Effect of Mutations Causing Familial Alzheimer's Disease.* Frontiers in Neuroscience, 2018. **12**.
23. Hin, N., et al., *Iron Responsive Element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease.* bioRxiv, 2020: p. 2020.05.01.071498.
24. Anderson, G.J. and D.M. Frazer, *Current understanding of iron homeostasis.* Am J Clin Nutr, 2017. **106**(Suppl 6): p. 1559S-1566S.
25. Ye, H. and T.A. Rouault, *Erythropoiesis and iron sulfur cluster biogenesis.* Adv Hematol, 2010. **2010**.
26. Luo, W. and C. Brouwer, *Pathview: an R/Bioconductor package for pathway-based data integration and visualization.* Bioinformatics, 2013. **29**(14): p. 1830-1.
27. Waga, S. and B. Stillman, *The DNA replication fork in eukaryotic cells.* Annu Rev Biochem, 1998. **67**: p. 721-51.

28. Yurov, Y.B., S.G. Vorsanova, and I.Y. Iourov, *The DNA replication stress hypothesis of Alzheimer's disease.* ScientificWorldJournal, 2011. **11**: p. 2602-12.
29. Van der Ven, A.T., A. Vivante, and F. Hildebrandt, *Novel Insights into the Pathogenesis of Monogenic Congenital Anomalies of the Kidney and Urinary Tract.* JASN, 2018. **29**(1): p. 36-50.
30. Berezin, V., et al., *Targeting of ECM molecules and their metabolizing enzymes and receptors for the treatment of CNS diseases.* Prog Brain Res, 2014. **214**: p. 353-88.
31. Li, N., et al., *Structure of the eukaryotic MCM complex at 3.8 A.* Nature, 2015. **524**(7564): p. 186-91.
32. Chan, S.L., et al., *Presenilin-1 mutations sensitize neurons to DNA damage-induced death by a mechanism involving perturbed calcium homeostasis and activation of calpains and caspase-12.* Neurobiol Dis, 2002. **11**(1): p. 2-19.
33. Michelsen, K.A., et al., *Presenilin 1-related alterations in DNA integrity in a transgenic mouse model of Alzheimer's disease.* Brain Research, 2010. **1316**: p. 139-144.
34. Shima, N., et al., *A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice.* Nature Genetics, 2007. **39**(1): p. 93-98.
35. Iourov, I.Y., et al., *Increased chromosome instability dramatically disrupts neural genome integrity and mediates cerebellar degeneration in the ataxia-telangiectasia brain.* Hum Mol Genet, 2009. **18**(14): p. 2656-69.
36. Nussenzweig, A., *Causes and consequences of the DNA damage response.* Cell Cycle, 2007. **6**(19): p. 2339-2340.
37. Valapala, M., et al., *Impaired endolysosomal function disrupts Notch signalling in optic nerve astrocytes.* Nature Communications, 2013. **4**.
38. Lee, J.H., et al., *Lysosomal proteolysis and autophagy require presenilin 1 and are disrupted by Alzheimer-related PS1 mutations.* Cell, 2010. **141**(7): p. 1146-58.
39. Noseda, M., et al., *Notch-dependent cell cycle arrest is associated with downregulation of minichromosome maintenance proteins.* Circ Res, 2005. **97**(2): p. 102-4.
40. Kwart, D., et al., *A Large Panel of Isogenic APP and PSEN1 Mutant Human iPSC Neurons Reveals Shared Endosomal Abnormalities Mediated by APP beta-CTFs, Not Abeta.* Neuron, 2019. **104**(5): p. 1022.
41. Conejero-Goldberg, C., et al., *APOE2 enhances neuroprotection against Alzheimer's disease through multiple molecular mechanisms.* Molecular Psychiatry, 2014. **19**(11): p. 1243-1250.
42. Bhagavan, N.V. and C.E. Ha, *Lipids I: Fatty Acids and Eicosanoids*, in *Essentials of Medical Biochemistry (Second Edition)*. 2015, Academic Press. p. 269-297.
43. Martinez-Reyes, I. and N.S. Chandel, *Mitochondrial TCA cycle metabolites control physiology and disease.* Nat Commun, 2020. **11**(1): p. 102.
44. Mosconi, L., A. Pupi, and M.J. De Leon, *Brain Glucose Hypometabolism and Oxidative Stress in Preclinical Alzheimer's Disease.* Mitochondria and Oxidative Stress in Neurodegenerative Disorders, 2008. **1147**: p. 180-195.
45. Zhang, C., R.A. Rissman, and J. Feng, *Characterization of ATP alternations in an Alzheimer's disease transgenic mouse model.* J Alzheimers Dis, 2015. **44**(2): p. 375-8.
46. Shoffner, J.M., *Oxidative phosphorylation defects and Alzheimer's disease.* Neurogenetics, 1997. **1**(1): p. 13-9.
47. Manczak, M., et al., *Differential expression of oxidative phosphorylation genes in patients with Alzheimer's disease: implications for early mitochondrial dysfunction and oxidative damage.* Neuromolecular Med, 2004. **5**(2): p. 147-62.
48. Mastroeni, D., et al., *Nuclear but not mitochondrial-encoded oxidative phosphorylation genes are altered in aging, mild cognitive impairment, and Alzheimer's disease.* Alzheimers & Dementia, 2017. **13**(5): p. 510-519.
49. Wong, B.X., et al., *beta-Amyloid precursor protein does not possess ferroxidase activity but does stabilize the cell surface ferrous iron exporter ferroportin.* PLoS One, 2014. **9**(12): p. e114174.
50. Yambire, K.F., et al., *Impaired lysosomal acidification triggers iron deficiency and inflammation in vivo.* Elife, 2019. **8**.
51. Dlouhy, A.C., et al., *Fluorescence resonance energy transfer links membrane ferroportin, hephaestin but not ferroportin, amyloid precursor protein complex with iron efflux.* Journal of Biological Chemistry, 2019. **294**(11): p. 4202-4214.

52. Ji, C.Y., et al., *The Ferroxidase Hephaestin But Not Amyloid Precursor Protein is Required for Ferroportin-Supported Iron Efflux in Primary Hippocampal Neurons.* Cellular and Molecular Neurobiology, 2018. **38**(4): p. 941-954.
53. Jiang, Y., et al., *Lysosomal Dysfunction in Down Syndrome Is APP-Dependent and Mediated by APP-beta CTF (C99).* Journal of Neuroscience, 2019. **39**(27): p. 5255-5268.
54. Pelegri, F., *Maternal factors in zebrafish development.* Dev Dyn, 2003. **228**(3): p. 535-54.
55. Edgar, R., M. Domrachev, and A.E. Lash, *Gene Expression Omnibus: NCBI gene expression and hybridization array data repository.* Nucleic Acids Res, 2002. **30**(1): p. 207-10.
56. Lindgreen, S., *AdapterRemoval: easy cleaning of next-generation sequencing reads.* BMC Res Notes, 2012. **5**: p. 337.
57. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
58. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.* Bioinformatics, 2014. **30**(7): p. 923-30.
59. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-140.
60. McCarthy, D.J., Y.S. Chen, and G.K. Smyth, *Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation.* Nucleic Acids Research, 2012. **40**(10): p. 4288-4297.
61. Risso, D., et al., *Normalization of RNA-seq data using factor analysis of control genes or samples.* Nature Biotechnology, 2014. **32**(9): p. 896-902.
62. Slenter, D.N., et al., *WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research.* Nucleic Acids Research, 2018. **46**(D1): p. D661-D667.
63. Carlson, M., *org.Dr.eg.db: Genome wide annotation for Zebrafish*. 2017: R package version 3.5.0.
64. Young, M.D., et al., *Gene ontology analysis for RNA-seq: accounting for selection bias.* Genome Biology, 2010. **11**(2).
65. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.
66. Sergushichev, A. *An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation*. bioRxiv, 2016. DOI: 10.1101/060012.
67. Campillos, M., et al., *SIREs: searching for iron-responsive elements.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W360-7.

4.3 Supplementary data

4.3.1 Supplementary data 1: DE gene list

Ensembl Gene ID	Gene Name	log_2FC	log_2CPM	P-Value	FDR	Rank statistics
ENSDARG00000039682	si:ch211-121a2.2	-3.195357333	1.235549038	1.02E-19	1.97E-15	-18.99272019
ENSDARG00000093024	si:ch211-213a13.2	1.024518193	2.754478867	4.23E-17	4.11E-13	16.37330497
ENSDARG00000040781	sult3st4	2.138069863	0.454818767	8.44E-16	5.46E-12	15.07353918
ENSDARG00000077068	si:ch211-11p18.6	-0.864354589	3.602471014	1.20E-15	5.80E-12	-14.92244699
ENSDARG00000099511	CABZ01034698.2	1.543570078	1.421034013	2.11E-15	8.19E-12	14.67552448
ENSDARG00000079589	si:dkeyp-73d8.6	1.443060699	1.352980164	2.62E-15	8.48E-12	14.58124549
ENSDARG00000070011	si:ch211-167j6.4	-1.908088181	0.767768566	1.91E-14	5.29E-11	-13.7187776
ENSDARG00000078995	tex11	1.597973265	0.050168575	2.32E-13	5.63E-10	12.63444414
ENSDARG00000091150	mki67	-0.808754804	4.228953545	3.96E-13	8.53E-10	-12.40271393
ENSDARG00000103849	mdh1ab	-0.997592103	4.619204283	7.00E-13	1.36E-09	-12.15487997
ENSDARG00000057128	hadhaa	-0.75378883	4.163723502	1.32E-12	2.32E-09	-11.88042279
ENSDARG00000090600	si:ch211-213a13.1	-0.976482451	3.064074262	4.45E-12	7.20E-09	-11.35129904
ENSDARG00000055618	acta1b	-1.006861962	8.137870503	6.27E-12	9.35E-09	-11.20292308
ENSDARG00000097973	si:ch1073-190k2.1	-3.505690905	0.796352288	7.11E-12	9.85E-09	-11.14811482
ENSDARG00000002991	cep135	1.421043604	3.223492835	1.05E-11	1.35E-08	10.97994542
ENSDARG00000101331	tekt1	1.777045206	1.441434581	2.59E-11	3.13E-08	10.58742462
ENSDARG00000039266	rbm25a	0.791212475	4.697857144	8.63E-11	9.84E-08	10.06407143
ENSDARG00000057911	zgc:86709	-0.829827607	3.052393611	2.11E-10	2.28E-07	-9.675355646
ENSDARG00000052905	zgc:165423	-0.871362117	5.364876864	2.26E-10	2.31E-07	-9.64516326
ENSDARG00000100584	ccdc40	0.936664349	2.876472155	2.54E-10	2.36E-07	9.594642588
ENSDARG00000055644	prss60.2	-0.993640356	2.054457212	2.67E-10	2.36E-07	-9.573680566
ENSDARG00000099470	muc5.3	-0.836508789	5.136950151	2.67E-10	2.36E-07	-9.572796952
ENSDARG00000053558	rtkn2a	1.631287552	1.380797037	4.56E-10	3.84E-07	9.341240402
ENSDARG00000087046	si:ch211-222k6.2	-0.99726875	1.384573132	1.14E-09	9.23E-07	-8.942485719
ENSDARG00000098754	si:ch211-156p11.1	1.190011799	0.592942087	1.23E-09	9.47E-07	8.910371332
ENSDARG00000044976	krt93	-0.70878209	2.99783341	1.27E-09	9.47E-07	-8.896342655
ENSDARG00000115545	clec19a	-1.084395275	1.641603661	2.09E-09	1.50E-06	-8.679414234
ENSDARG00000103019	gstp2	-1.270901121	2.320425211	2.44E-09	1.69E-06	-8.612332837
ENSDARG00000023861	dnase1l1l	0.718054558	3.960688269	2.84E-09	1.90E-06	8.547159082
ENSDARG00000045089	si:ch211-270n8.1	-1.4986985	0.57880414	3.88E-09	2.51E-06	-8.411536327
ENSDARG00000074531	spag9a	0.556940205	5.063666392	6.00E-09	3.76E-06	8.221646887
ENSDARG00000060622	si:ch73-14h1.2	-0.800851248	4.222779945	7.60E-09	4.60E-06	-8.119394682
ENSDARG00000026726	anxa1a	-0.911330132	6.883579699	1.62E-08	9.50E-06	-7.791620686

ENSDARG00000104899	CR931813.2	0.797276959	2.222911283	2.28E-08	1.30E-05	7.64203466
ENSDARG00000093699	si:ch73-27e22.4	-1.11597696	0.776462666	2.83E-08	1.57E-05	-7.547667417
ENSDARG00000042708	tuba8l	-0.595215461	6.540256941	3.46E-08	1.86E-05	-7.461509512
ENSDARG00000054753	col10a1a	-0.86519819	5.992846267	4.17E-08	2.19E-05	-7.379646882
ENSDARG00000002192	aspn	-0.796058659	3.202490712	4.47E-08	2.28E-05	-7.350010695
ENSDARG00000111240	dhrs13a.2	-0.69005071	3.063465613	5.08E-08	2.53E-05	-7.293895761
ENSDARG00000095082	ITLN1 (1 of many)	1.221736969	2.517845631	6.37E-08	3.09E-05	7.19584603
ENSDARG00000090946	si:zfos-364h11.2	-1.064785007	2.008884056	1.25E-07	5.92E-05	-6.902753457
ENSDARG00000061196	emilin2a	-0.773607167	4.174377486	1.65E-07	7.63E-05	-6.782149367
ENSDARG00000077620	cdca7a	-1.020105533	1.69086497	1.75E-07	7.91E-05	-6.755867887
ENSDARG00000074844	CABZ01061495.1	-0.58476051	3.508221205	1.93E-07	8.35E-05	-6.715442107
ENSDARG00000002403	nusap1	-0.795198538	2.050108206	1.94E-07	8.35E-05	-6.713068228
ENSDARG00000007377	odc1	-0.672830165	2.791019939	2.18E-07	9.16E-05	-6.662173734
ENSDARG00000095643	si:dkey-253d23.3	-1.044948874	2.061783969	2.22E-07	9.16E-05	-6.653924496
ENSDARG00000016771	tfa	-0.671676383	8.554705029	2.47E-07	0.000100008	-6.606437353
ENSDARG00000097080	si:ch73-181m17.1	-1.65835602	2.097257516	2.63E-07	0.000104035	-6.580335285
ENSDARG00000040004	si:ch211-244a23.1	-0.692229586	2.752512157	2.86E-07	0.000111123	-6.542939515
ENSDARG00000053315	tmprss3a	-1.070024307	1.929990633	3.35E-07	0.000127454	-6.474789887
ENSDARG00000040944	ntd5	-0.468030637	4.548711917	4.30E-07	0.000159063	-6.366749628
ENSDARG00000054292	si:ch211-14a17.11	-0.807462069	2.276822992	4.48E-07	0.000159063	-6.34827596
ENSDARG00000115701	crygm2d9	0.599624682	6.269020854	4.49E-07	0.000159063	6.348077012
ENSDARG00000114577	si:dkey-159n16.2	0.633189257	4.05784447	4.51E-07	0.000159063	6.34577921
ENSDARG00000014803	cryba1l2	0.519039708	5.570475475	5.55E-07	0.000191717	6.256040925
ENSDARG00000019995	bmp4	0.524759243	3.440304714	5.63E-07	0.000191717	6.249176993
ENSDARG00000042018	fhl2a	-0.46009233	4.931139513	6.33E-07	0.000211738	-6.198485901
ENSDARG00000035700	zgc:101664	-0.78608191	1.547410463	7.22E-07	0.000237207	-6.141733201
ENSDARG00000040192	nenf	-0.983845052	1.784594465	7.50E-07	0.000242531	-6.124794204
ENSDARG00000069817	crygm2d17	0.802494829	7.428088002	8.22E-07	0.000260853	6.08488259
ENSDARG00000052122	rag1	-0.812963532	2.882403277	8.42E-07	0.000260853	-6.074611713
ENSDARG00000062152	chaf1a	-0.589901411	3.40881466	8.47E-07	0.000260853	-6.071975734
ENSDARG00000104600	CABZ01070527.1	0.75093988	4.101177283	9.34E-07	0.000283074	6.029632389
ENSDARG00000071347	aftphb	0.566251958	3.348876842	9.64E-07	0.000287552	6.016082274
ENSDARG00000092243	BX005065.1	3.110361261	0.767626336	9.82E-07	0.000288531	6.007975389
ENSDARG00000100311	si:dkey-165e24.1	-1.115812313	1.185592335	1.06E-06	0.000307334	-5.974026734
ENSDARG00000094210	fthl31	0.721720727	4.52370205	1.21E-06	0.000340598	5.917673361
ENSDARG00000071604	si:ch211-156p11.1	-0.750495506	1.643388314	1.21E-06	0.000340598	-5.916620921
ENSDARG00000039422	fuom	-0.923275979	0.393504689	1.43E-06	0.000397583	-5.843186044
ENSDARG00000056386	TMC1	-1.284697882	0.500523184	1.49E-06	0.000405801	-5.82814103
ENSDARG00000104403	oscp1a	1.09236325	0.393697183	1.56E-06	0.00042042	5.806695993
ENSDARG00000067520	zgc:158482	-0.621573732	2.832616145	1.74E-06	0.000457334	-5.758632634

ENSDARG00000052779	zgc:153932	-1.039601293	3.871096157	1.74E-06	0.000457334	-5.758247353
ENSDARG00000092890	si:ch73-44m9.5	3.475227963	1.38806443	1.81E-06	0.000467173	5.743173218
ENSDARG00000093997	si:dkey-9i23.15	-0.640748862	2.046458609	1.91E-06	0.00048633	-5.719967404
ENSDARG00000024488	top2a	-0.689341995	3.45397341	2.01E-06	0.000506099	-5.696986266
ENSDARG00000038668	gbp1	-0.915790111	1.995418233	2.15E-06	0.000528743	-5.667701139
ENSDARG00000039051	hhatla	-0.504417271	6.919728817	2.18E-06	0.000528743	-5.661969186
ENSDARG00000071543	si:dkey-42i9.7	1.307596316	0.755250411	2.18E-06	0.000528743	5.661377194
ENSDARG00000086917	crygm2d2	0.533488034	8.871350006	2.34E-06	0.000559655	5.631306798
ENSDARG00000024278	adh8b	-0.699520736	5.923896572	2.41E-06	0.000571005	-5.617258462
ENSDARG00000105183	CT009487.2	0.891864324	2.831933755	2.46E-06	0.000575673	5.608458016
ENSDARG00000062618	kcnj12b	-0.577582883	3.408560592	2.61E-06	0.00060271	-5.583324669
ENSDARG00000090847	si:ch211-209l18.4	-0.905488007	2.617969336	2.64E-06	0.000603548	-5.577581568
ENSDARG00000024847	col5a2b	-0.823672363	2.875421324	2.72E-06	0.000613074	-5.565700708
ENSDARG00000093628	s100a11	-0.735886765	2.519790467	2.84E-06	0.000630335	-5.547264023
ENSDARG00000117693	CR318673.1	0.630086321	3.096319409	2.86E-06	0.000630335	5.543657852
ENSDARG00000015998	wdhd1	-0.780646227	1.011028119	3.01E-06	0.000655056	-5.52204392
ENSDARG00000053323	zgc:112285	0.552984827	6.528180422	3.16E-06	0.000680667	5.500535173
ENSDARG00000014805	fhl5	-0.966360641	1.129981443	3.47E-06	0.000732596	-5.460239767
ENSDARG00000069846	zgc:162944	-0.576344774	4.013066855	3.47E-06	0.000732596	-5.459059645
ENSDARG00000055270	si:ch1073-358c10.1	-0.849574258	1.415830401	3.52E-06	0.000733341	-5.453922973
ENSDARG00000096905	si:ch73-23l24.1	-0.78693341	1.628948068	4.29E-06	0.00088487	-5.367704884
ENSDARG00000058285	cpt1b	-0.798461743	3.016533095	4.36E-06	0.00089104	-5.360091607
ENSDARG00000036139	lctla	0.472389158	5.821464535	4.49E-06	0.000906347	5.348146365
ENSDARG00000030215	matn1	-0.562516104	5.954454526	4.88E-06	0.000975883	-5.311542474
ENSDARG00000074546	si:ch211-213a13.2	-0.720514193	1.68968825	4.93E-06	0.000976143	-5.306972728
ENSDARG00000009844	dusp23a	-0.475972548	4.253820595	5.19E-06	0.001008038	-5.28512206
ENSDARG00000078419	filip1a	-1.006270345	0.606602049	5.20E-06	0.001008038	-5.284235076
ENSDARG00000069559	muc13a	-0.541359769	4.559822655	5.29E-06	0.001015369	-5.276766854
ENSDARG00000033757	ncaph2	-0.94369771	0.565800288	5.74E-06	0.0010921	-5.240849629
ENSDARG00000052734	hmgcra	-0.87452332	3.414099804	5.95E-06	0.001120243	-5.225562618
ENSDARG00000095615	si:dkeyp-86h10.3	-0.479758572	3.248050203	6.54E-06	0.001217807	-5.18417276
ENSDARG00000038716	casq1a	-0.779606314	5.862824768	6.59E-06	0.001217807	-5.180944441
ENSDARG00000040306	otomp	-0.443774273	4.184549758	6.86E-06	0.001254837	-5.163818994
ENSDARG00000104823	znf648	-0.827036777	3.962059604	7.06E-06	0.001279242	-5.151375846
ENSDARG00000094466	si:ch73-199e17.1	-0.573899315	2.547311741	8.02E-06	0.00143975	-5.096001466
ENSDARG00000101495	ugt5b2	-1.220029149	1.350745811	8.35E-06	0.00148559	-5.078386838
ENSDARG00000058966	zgc:112332	0.837936727	3.402575014	8.65E-06	0.001524377	5.063227097
ENSDARG00000037613	lgals8b	0.679196419	1.279105177	9.32E-06	0.001628137	5.030698219
ENSDARG00000101180	mcm7	0.484027692	6.247258419	9.81E-06	0.001689977	5.008208849
ENSDARG00000105443	si:ch211-63p21.8	0.718399505	3.432737805	9.85E-06	0.001689977	5.006752922

ENSDARG00000077372	tfr1b	-0.531823196	4.874578002	1.01E-05	0.001726522	-4.993635259
ENSDARG00000070140	RETSAT	0.838124313	0.363779272	1.06E-05	0.001774807	4.976540887
ENSDARG00000028664	ahsa1a	-0.653234178	1.87197908	1.06E-05	0.001774807	-4.974102965
ENSDARG00000074749	abca12	-0.454219232	6.486104629	1.09E-05	0.001804261	-4.963227038
ENSDARG00000052039	caspb	-0.589137419	2.783999382	1.11E-05	0.001824512	-4.954683508
ENSDARG00000077360	fthl30	-0.845820443	1.35832497	1.15E-05	0.001871454	-4.939985991
ENSDARG00000003311	pank2	0.43819455	6.752264535	1.19E-05	0.001925239	4.924046313
ENSDARG00000086103	slc37a1	-0.579442767	2.395649313	1.30E-05	0.002082787	-4.886282032
ENSDARG00000024204	mcm3	-0.782118733	2.849975251	1.34E-05	0.002129828	-4.873007717
ENSDARG00000052000	cav2	0.433401574	4.060659341	1.36E-05	0.002139985	4.867396287
ENSDARG00000073720	si:dkey-250l23.4	0.609534896	2.736974186	1.37E-05	0.002149592	4.861934549
ENSDARG00000019507	mcm5	-0.925367859	2.847134848	1.50E-05	0.00232581	-4.824227839
ENSDARG00000052113	hexa	0.463508612	4.461687782	1.53E-05	0.002355283	4.81529856
ENSDARG00000019752	rom1a	0.434083015	4.635254278	1.55E-05	0.002369383	4.809273258
ENSDARG00000079145	CABZ01077217.1	-1.036107287	1.398289198	1.58E-05	0.002401764	-4.799971919
ENSDARG00000052288	zmp:0000000634	-1.330947602	-0.219842759	1.64E-05	0.002470597	-4.784320635
ENSDARG00000052057	pcolceb	-0.604547676	2.224035147	1.66E-05	0.002470941	-4.780906373
ENSDARG00000052437	mia	-0.693337732	2.122799642	1.74E-05	0.002555561	-4.760494573
ENSDARG00000114670	ube2c	-1.026576353	0.2743903	1.75E-05	0.002555561	-4.758002832
ENSDARG00000102050	MCOLN3	-0.920846017	0.748878872	1.75E-05	0.002555561	-4.756374231
ENSDARG00000021086	znf367	-0.770521938	0.556173729	1.91E-05	0.00277151	-4.717890884
ENSDARG00000029058	rbbp4	-0.558348587	3.508495036	1.99E-05	0.002819741	-4.700788661
ENSDARG00000075707	nid2a	-0.479884109	4.350436462	1.99E-05	0.002819741	-4.700501036
ENSDARG00000099355	BX890548.1	0.45130522	3.555172292	2.00E-05	0.002819741	4.699991208
ENSDARG00000102798	mcm2	-0.66867466	3.600919282	2.01E-05	0.002819741	-4.696808587
ENSDARG00000026771	tmem41ab	-0.666128118	3.530667336	2.02E-05	0.002819741	-4.694488102
ENSDARG00000087164	crygm2d4	0.48382435	7.829751638	2.06E-05	0.002842833	4.686677563
ENSDARG00000038156	dusp11	-0.553300866	2.778590783	2.07E-05	0.002842833	-4.684741772
ENSDARG00000091918	BX548073.1	0.812541603	0.508383523	2.09E-05	0.002852079	4.680262293
ENSDARG00000104721	si:dkey-203a12.9	-1.009461717	0.646226571	2.14E-05	0.002895781	-4.669609164
ENSDARG00000062727	cep290	0.43088837	6.052947754	2.15E-05	0.002895781	4.66758394
ENSDARG00000070239	knl1	-0.801767444	1.318446228	2.23E-05	0.002972651	-4.65211634
ENSDARG00000037845	col9a3	-0.658972232	6.100689257	2.24E-05	0.002972651	-4.650215325
ENSDARG00000057074	rpgrb	0.513603171	4.867581488	2.26E-05	0.00298425	4.645303478
ENSDARG00000030254	cbln9	-0.871377861	1.015030745	2.28E-05	0.00298425	-4.642615194
ENSDARG00000058533	pole	-0.603929278	1.992628356	2.40E-05	0.003130243	-4.618947872
ENSDARG00000061416	c2cd4a	0.559234714	4.758865837	2.50E-05	0.003208046	4.602541442
ENSDARG00000038658	cavin4a	-0.432620139	5.063362768	2.50E-05	0.003208046	-4.602494649
ENSDARG00000103744	hacd1	-0.417952662	5.241863198	2.64E-05	0.003373923	-4.577733376
ENSDARG00000069093	col2a1a	-0.61082304	9.257959447	2.68E-05	0.003396925	-4.571934809

ENSDARG00000036239	gatm	-0.817409366	6.042497757	2.76E-05	0.00345868	-4.558478556
ENSDARG00000098950	si:dkeyp-92c9.4	-0.913597935	0.157446222	2.76E-05	0.00345868	-4.558470073
ENSDARG00000034063	unm_sa911	-0.792175134	1.673333752	2.79E-05	0.003472785	-4.553909672
ENSDARG00000053502	cryaa	0.457328425	6.95995161	2.85E-05	0.003522545	4.544955988
ENSDARG00000077029	bub1	-0.674879796	1.277637322	2.93E-05	0.00359882	-4.532894968
ENSDARG00000026759	ldlrb	-0.481905048	3.332563054	2.98E-05	0.003637537	-4.525507663
ENSDARG00000045808	rlbp1b	-0.70352991	3.402360079	3.03E-05	0.003673809	-4.518475602
ENSDARG00000030723	cldn11b	0.414332185	4.000552431	3.12E-05	0.003760725	4.505614721
ENSDARG00000021404	zgc:110319	0.404499666	4.276119365	3.31E-05	0.003965078	4.479945393
ENSDARG00000103277	cyp24a1	0.420580436	8.36185793	3.34E-05	0.003979636	4.475681264
ENSDARG00000038281	natd1	0.426554579	5.256276176	3.38E-05	0.003994797	4.471283722
ENSDARG00000014041	smarcad1a	-0.683564654	1.579195159	3.40E-05	0.003994797	-4.468733499
ENSDARG00000098591	tubb2b	-0.427661245	6.976508032	3.57E-05	0.004173969	-4.446760348
ENSDARG00000062817	crym	-0.80851753	1.487359795	3.59E-05	0.004173969	-4.444446468
ENSDARG00000040041	mcm4	-0.590566707	3.977590377	3.74E-05	0.00429947	-4.426573573
ENSDARG00000023587	kcnk5a	-0.830636307	1.508119253	3.75E-05	0.00429947	-4.426410572
ENSDARG00000105511	BX248521.2	0.921371001	1.116761537	3.88E-05	0.004423142	4.411532327
ENSDARG00000096920	si:zfos-375h5.1	0.459151211	3.354551175	3.91E-05	0.004437406	4.407586902
ENSDARG00000094077	si:dkey-21e2.16	0.554659101	2.80155962	3.96E-05	0.004470649	4.401813153
ENSDARG00000010948	kif11	-0.641351805	2.053896762	4.13E-05	0.004635277	-4.383590391
ENSDARG00000026629	gmds	-0.408708627	4.373027833	4.21E-05	0.004695769	-4.375456182
ENSDARG00000100954	wars	-0.544076214	2.328006465	4.34E-05	0.004811284	-4.362413118
ENSDARG00000071083	si:dkeyp-34c12.1	0.599143599	2.865601086	4.48E-05	0.004935496	4.348868666
ENSDARG00000103754	aspm	-0.594377687	2.718018205	4.58E-05	0.005023664	-4.338718289
ENSDARG00000004748	zgc:100868	-0.524062321	5.482403449	4.63E-05	0.005047765	-4.334193048
ENSDARG00000093323	si:ch211-114l13.1	0.628793996	2.25132992	4.88E-05	0.005272815	4.311541766
ENSDARG00000068431	si:ch211-195h23.3	1.469329276	4.461700867	4.89E-05	0.005272815	4.31039711
ENSDARG00000055753	suv39h1b	-0.799687991	0.251437576	4.94E-05	0.005276249	-4.306408584
ENSDARG00000052348	smarca5	-0.415370117	4.608576662	4.95E-05	0.005276249	-4.305315526
ENSDARG00000055314	mcmbp	-0.433850967	3.612137046	5.00E-05	0.005300081	-4.300978612
ENSDARG00000073699	col9a1a	-0.443331149	6.062570951	5.03E-05	0.005301124	-4.298526385
ENSDARG00000092033	si:dkey-239h2.3	0.404465442	6.004973016	5.46E-05	0.005720189	4.263130053
ENSDARG00000097825	BX465227.1	0.434870817	3.215065311	5.68E-05	0.005892225	4.245844891
ENSDARG00000041295	lim2.5	0.462360561	4.792372145	5.72E-05	0.005892225	4.242645521
ENSDARG00000101670	kntc1	-0.638677789	1.858140436	5.73E-05	0.005892225	-4.242140101
ENSDARG00000079347	zgc:194659	0.582281571	5.341517537	5.75E-05	0.005892225	4.240428608
ENSDARG00000096988	CR589944.1	0.977171665	1.909412314	5.77E-05	0.005892225	4.238679264
ENSDARG00000110961	zgc:110045	0.575023831	3.510295679	5.89E-05	0.005911272	4.230205235
ENSDARG00000029406	h2afx	-0.751857654	1.040403266	5.91E-05	0.005911272	-4.228456946
ENSDARG00000014179	pfkma	-0.415850527	5.937842326	5.92E-05	0.005911272	-4.227388574

ENSDARG00000079245	si:dkey-73p2.2	-0.990247659	2.868716424	5.96E-05	0.005911272	-4.224478313
ENSDARG00000037402	lim2.3	0.373060962	5.144101081	5.98E-05	0.005911272	4.223492093
ENSDARG00000042876	abracl	-0.499141376	2.994246545	6.01E-05	0.005911272	-4.220858565
ENSDARG00000063195	pxk	0.375150905	6.027453402	6.05E-05	0.005911272	4.218527802
ENSDARG00000011094	ccna2	-0.747639605	1.451212255	6.06E-05	0.005911272	-4.217779277
ENSDARG00000002986	gda	-0.510740933	4.174478014	6.06E-05	0.005911272	-4.217178153
ENSDARG00000091260	mylk4a	-0.445329476	4.814196135	6.17E-05	0.00597961	-4.210009332
ENSDARG00000093494	si:ch211-217k17.9	-0.55064334	3.785132109	6.24E-05	0.006014323	-4.20496645
ENSDARG00000077403	col8a1a	-0.692800371	2.36446439	6.26E-05	0.006014323	-4.203174075
ENSDARG00000044685	nr0b2a	0.715383776	4.781100335	6.45E-05	0.006159909	4.190641812
ENSDARG00000057683	mcm6	-0.766793	3.161765769	6.54E-05	0.006218099	-4.1844244
ENSDARG00000095147	krt96	0.683397345	5.212867833	6.72E-05	0.006358072	4.172632896
ENSDARG00000053068	cyp8b1	-0.642855268	3.19002589	6.95E-05	0.006547502	-4.157769333
ENSDARG00000004232	dlb	-0.521416038	2.483284408	7.04E-05	0.006594415	-4.15256553
ENSDARG00000062702	ankmy1	0.788949229	0.697934556	7.62E-05	0.007087448	4.118195776
ENSDARG00000036832	cyt1l	-0.461112249	8.226281012	7.64E-05	0.007087448	-4.117075994
ENSDARG00000051880	kcnj11	-0.617937174	1.986641382	8.29E-05	0.007652435	-4.081693215
ENSDARG00000008982	casq2	-0.393529917	5.079445292	8.52E-05	0.007832735	-4.069399752
ENSDARG00000076830	si:dkey-65b12.6	-0.788841214	2.531839128	8.56E-05	0.007832735	-4.067462891
ENSDARG00000051879	abcc8	0.387248666	5.109168371	8.77E-05	0.007966141	4.057246933
ENSDARG00000011072	ddx11	-0.694392704	0.69733658	8.79E-05	0.007966141	-4.056050399
ENSDARG00000095048	si:dkey-250k15.7	0.574092494	2.015510223	8.91E-05	0.00804196	4.049911799
ENSDARG00000090232	clpb	0.41586575	5.393156067	9.13E-05	0.008200855	4.039399282
ENSDARG00000099873	zgc:163083	-0.716395363	1.083635739	9.23E-05	0.008246957	-4.034958705
ENSDARG00000093402	si:dkey-286j17.4	-1.094567707	2.222106568	9.27E-05	0.008248745	-4.032867782
ENSDARG00000038066	kpna2	-0.636640614	2.02478688	9.50E-05	0.008417618	-4.022078852
ENSDARG00000062943	taco1	-0.622466487	2.268344237	9.85E-05	0.008681344	-4.006702511
ENSDARG00000042285	atp8b3	-0.521477384	2.264896863	9.98E-05	0.008759846	-4.000823454
ENSDARG00000054746	uggt1	-0.461280339	3.760335472	0.00010353	0.009045524	-3.984925475
ENSDARG00000087994	si:ch211-114c17.1	0.610780904	2.59634886	0.00010598	0.0092178	3.974780053
ENSDARG00000078425	oat	-0.579462651	5.409791643	0.00010836	0.009350453	-3.965123832
ENSDARG00000035018	thy1	-0.50665046	2.6239244	0.00010847	0.009350453	-3.964697022
ENSDARG00000005462	kif4	-0.593412791	2.055927611	0.0001119	0.009603498	-3.951174296
ENSDARG00000100915	si:ch211-255f4.7	0.42801666	3.737811127	0.00011529	0.009835525	3.938204603
ENSDARG00000095559	si:dkey-246j7.1	-0.606413561	1.175049085	0.00011562	0.009835525	-3.936979785

Notes

Rank statistics uses sign(log_2FC) * (-log_{10}PValue).

Only genes with a FDR less than 0.01 are shown.

"FC" indicates fold change.

"CPM" indicates counts per million.

Positive values under "log_2FC" indicate increased expression in the mutant.

Negative values under "log_2FC" indicate decreased expression in the mutant.

4.3.2 Supplementary data 2: qPCR validation of RNA-seq data

Supplementary data 2: qPCR validation of RNA-seq data

To support the accuracy and reliability of RNA-seq data, we performed relative standard curve quantitative PCRs (qPCRs) on four of the most statistically significantly DE genes with a relatively large fold-change in expression. Also, qPCRs were performed on three of the significantly changed genes identified in *GO iron ion transport* to confirm these results. The cDNAs used in qPCRs were synthesized from the same RNA samples as were used in the RNA-seq.

Method

Relative standard curve quantitative PCR

1000 ng of the total RNA from each sample remaining after RNA-seq was used to synthesise 20 μL of first-strand cDNA by reverse transcription (SuperScript III kit, Invitrogen, Camarillo, California, USA). Standard curves were generated using a dilution series having 40 ng, 20 ng, 10 ng and 5 ng of wild type cDNA per reaction. Each 25 μL qPCR reaction contained 20 ng of cDNA, 0.2 μM of each PCR primer and Power SYBR green master mix PCR solution (Applied Biosystems, Thermo Fisher Scientific Inc., Waltham, MA, USA). The qPCR was performed on an ABI 7000 Sequence Detection System (Applied Biosystems) using a 96-well plate. The amplification consisted of a holding stage and a cycling stage. The holding stage was performed at 50°C for 2 mins and then 95°C for 10 mins, and the cycling stage had 40 cycles of 95°C for 15 s and 60°C for 1 min. Each reaction was conducted in triplicate. The quantities of amplified product in each reaction were determined from the standard curve. The mean value of three technical replicates was used to represent the quantity. The quantities of transcripts of the genes of interest were then calibrated to the quantities of transcripts of the house-keeping gene *rpl13* for relative quantification, enabling quantitative comparison between samples.

The gene expression measured by RNA-seq analysis was plotted as $\log_2$CPM (see the left half of the following figures in Results), while the qPCR results were plotted as quantities relative to *rpl13* (see the right half of the following figures in Results) with P-values calculated using a paired t-test.

Primer Information

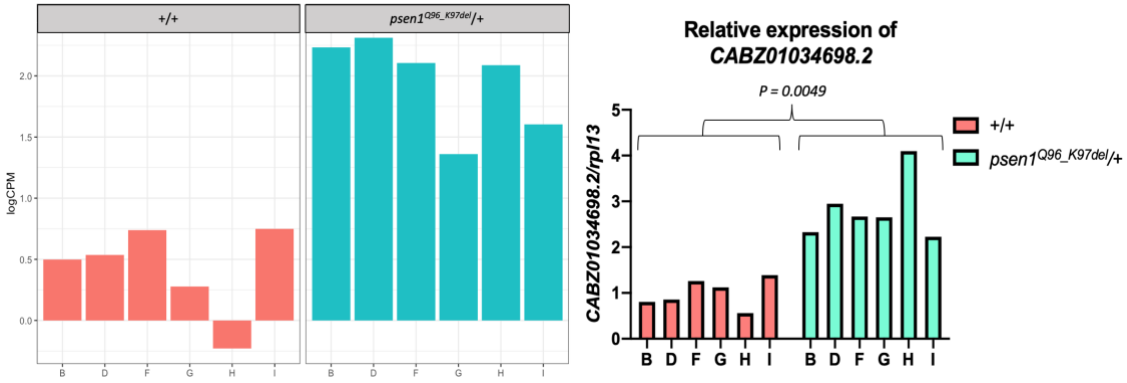
Ensembl Gene ID	Primer name	Primer sequence
House-keeping gene		
ENSDARG00000099380	rpl13_F	5'- CGCTAAGGACGGAGTGAACAAC-3'
	rpl13_R	5'- CTCTCTTCTGCCAGTCTTTATGA-3'
Statistically significantly DE genes		
ENSDARG00000099511	CABZ01034698.2_F	5'- CTTATTGACGGAAAGATAGGAGAA-3'
	CABZ01034698.2_R	5'- CTAGTCTGTGTCTAAATCTCTCATCGTG-3'
ENSDARG00000103849	mdh1ab_F	5'- GCTTCCTCACAGGTGGCATT-3'
	mdh1ab_R	5'- GGCAGTCTCGTCCTTCCCTT-3'
ENSDARG00000077068	si:ch211-11p18.6_F	5'- TGCTGATACTCTGCTAAATCAACTG-3'
	si:ch211-11p18.6_R	5'- CCTGTTCACTGCCACCCTGAG-3'
ENSDARG00000093024	si:ch211-213a13.2_F	5'- CTCCAAAAGAGCCTTGTTAAATGC-3'
	si:ch211-213a13.2_R	5'- TTACTGAGCCCCAATATCTCCAA-3'
Genes identified in the GO **iron ion transport**		
ENSDARG00000016771	tfa_F	5'- GCAATCCCAGAGAGTGAGAGG-3'
	tfa_R	5'- CTGCTTCAGGACATCATAAACAT-3'
ENSDARG00000077372	tfr1b_F	5'- CCGTATTATGAAGGTTGAGCA-3'
	tfr1b_R	5'- CAGGTCTGTGGCGGTTTGC-3'
ENSDARG00000094210	fthl31_F	5'- ATAGTAAAACCTGCCGAGATGGA-3'
	fthl31_R	5'- TAAAATAGTGAGCCATGGAGGTG-3'

Results

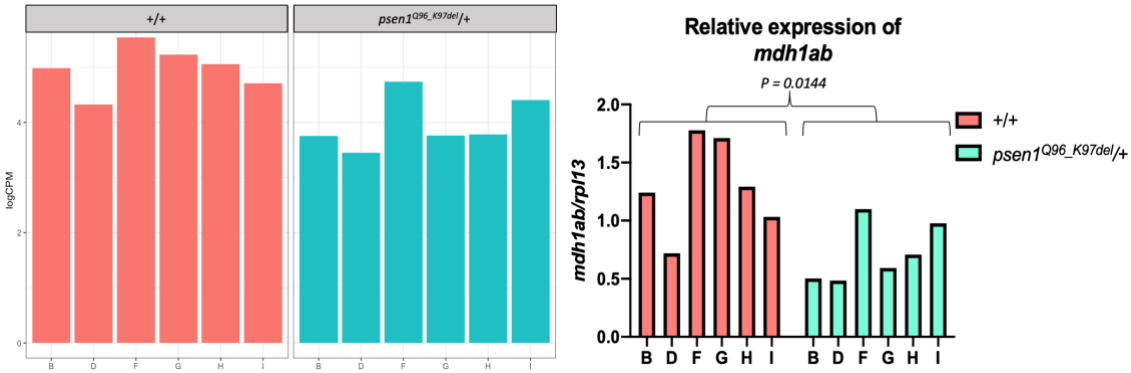
Validation of statistically significantly DE genes

QPCRs were performed on transcripts of the genes *CABZ01034698.2*, *mdh1ab*, *si:ch211-11p18.6* and *si_ch211-213a13.2*. The relative quantifications of *CABZ01034698.2*, *mdh1ab* and *si:ch211-11p18.6* indicated significant differences (P-value < 0.05) between wild type and heterozygous mutant larvae with the same regulation direction as observed in RNA-seq. However, the relative quantification of *si_ch211-213a13.2* was different from the RNA-seq result, probably due to mis-mapping or misidentification of reads in the RNA-seq analysis. Therefore, three of four statistically significantly DE genes identified in RNA-seq were validated by qPCR, supporting the accuracy of most RNA-seq results.

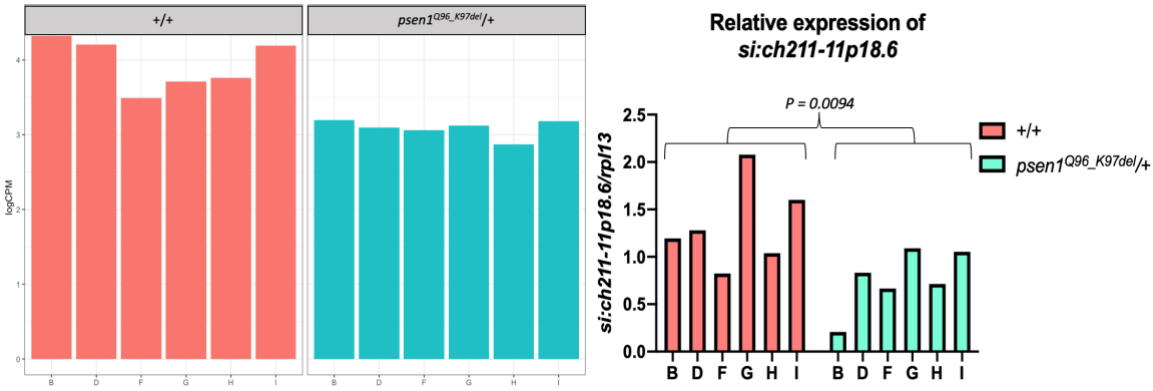
CABZ01034698.2



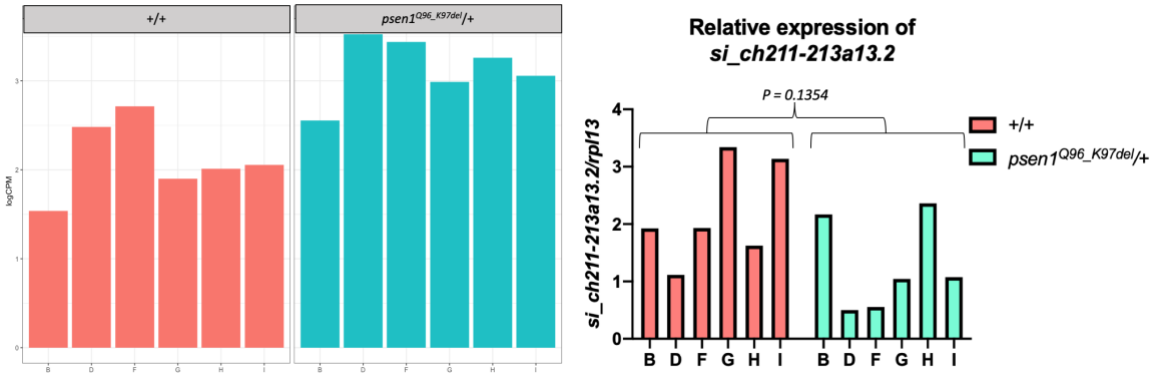
mdh1ab



si:ch211-11p18.6



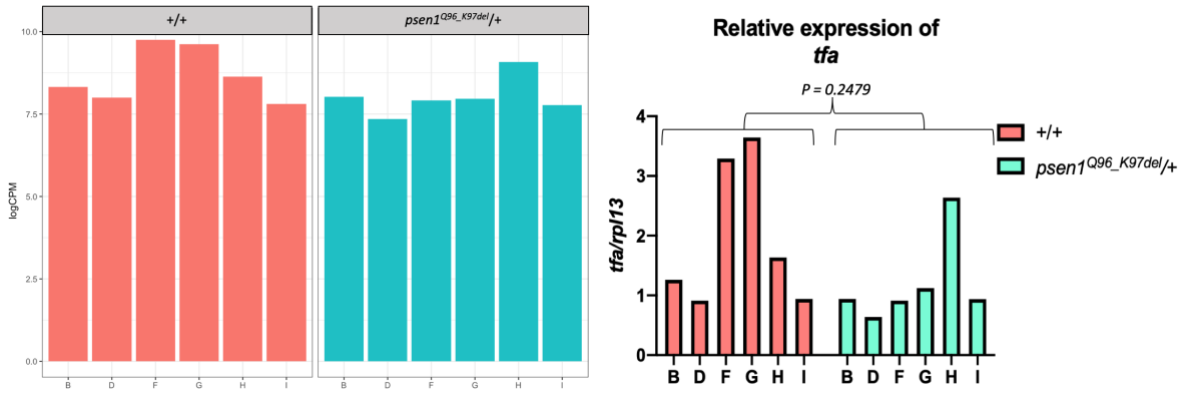
si_ch211-213a13.2



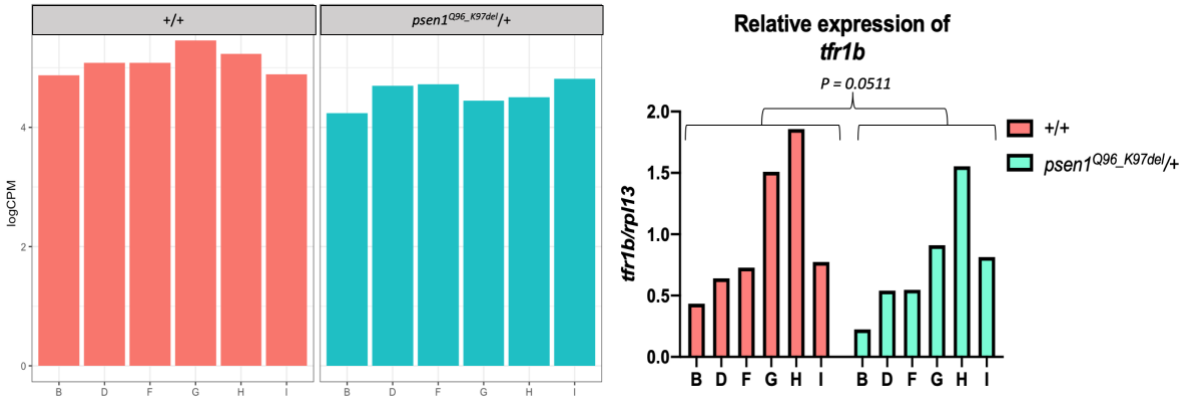
Validation of the genes identified in the GO *iron ion transport*

The same qPCR approach as above was applied to the genes *tfa*, *tfr1b* and *fthl31* identified in the GO *iron ion transport*. Only the difference in relative expression level of *fthl31* validated the RNA-seq result at the statistical significance threshold chosen (P-value < 0.05) in a paired t-test. However, although the differences in relative expression levels for *tfa* and *tfr1b* did not show statistical significance, the relative expression level in each sample generally corresponded with the relative $\log_2$CPM level in RNA-seq analysis, indicating that the qPCR results supported the reliability of the RNA-seq results. Also, detection of statistically significantly DE genes by RNA-seq included all the detected genes of the transcriptome in estimating variance, while qPCR only examined variation in the gene of interest across samples and so was less sensitive in detecting statistical significance.

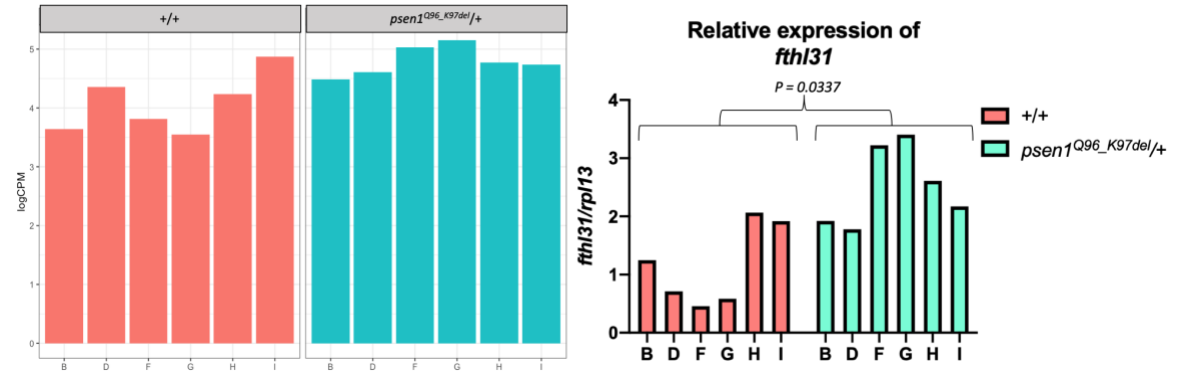
tfa



tfr1b



fthl31



Both the RNA-seq plots and the qPCR plots display obvious variability in gene expression within clutches of larvae of the same genotype. For example, although the expression levels of *fthl31* in wild-type clutches H and I were similar to the expression levels in mutant clutches B and D, when analyzed in pairs, the expression levels of *fthl31* in the mutants were obviously

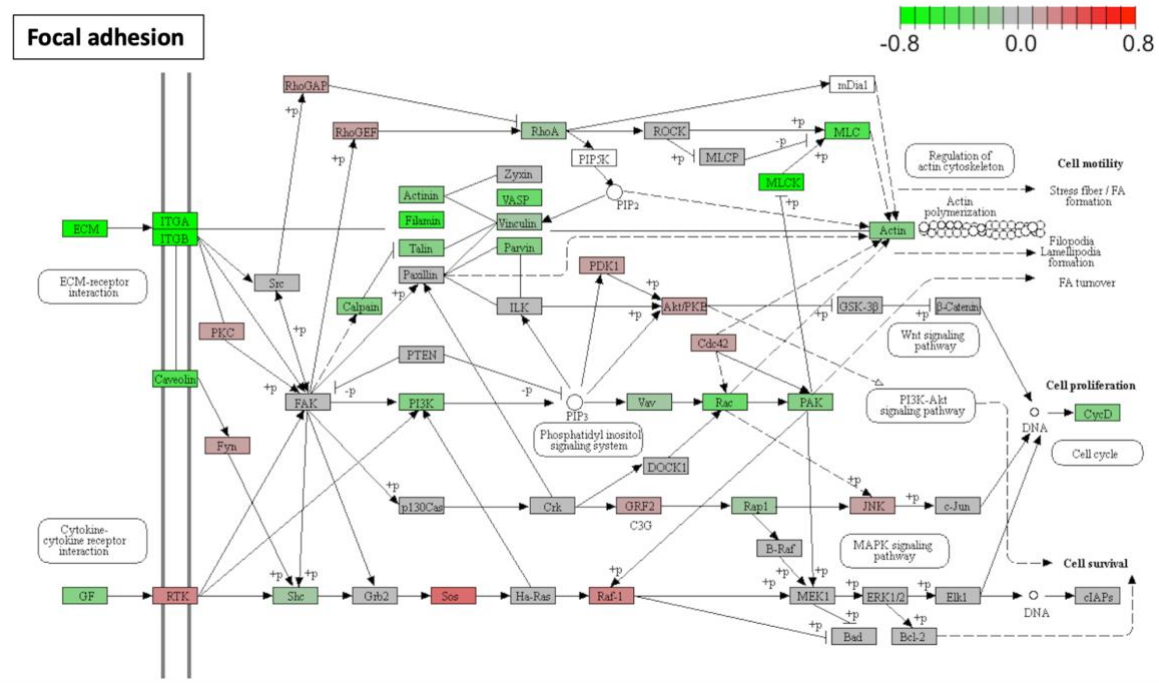
higher than in the wild-type clutches. In this sense, our paired-mating strategy (Figure 3 of the paper) aided in DE gene identification.

4.3.3 Supplementary data 3: KEGG pathway diagrams

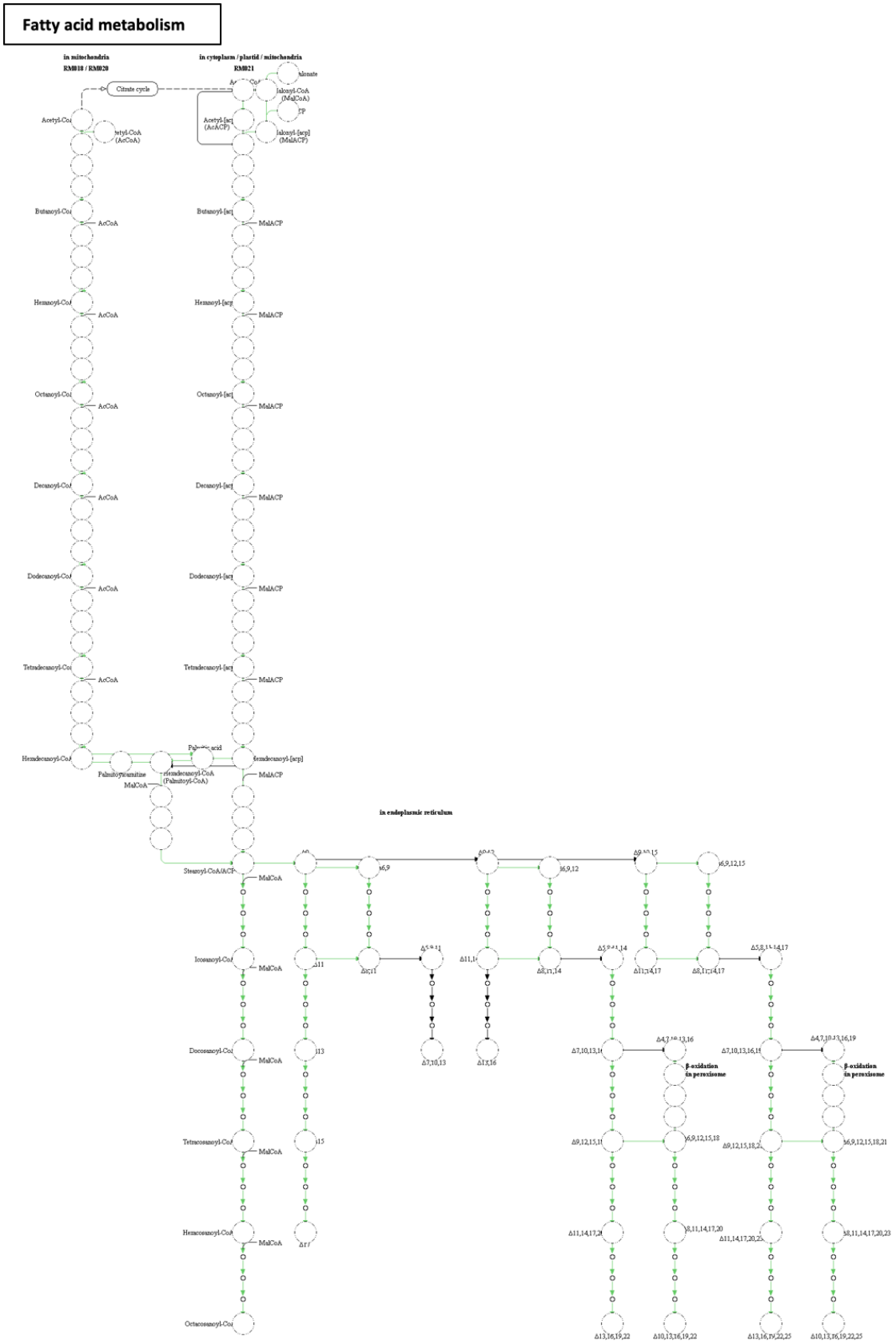
Supplementary data 3: KEGG pathway diagrams

Ten significantly-changed KEGG pathways were identified by GSEA. The KEGG diagrams of *DNA replication*, *cell cycle*, *ECM receptor interaction* and *oxidative phosphorylation* are shown in the main paper. The diagrams of the other 6 significantly changed KEGG pathways are listed below. Intensity of colour in diagrams represents gene log_2FC ranging from -0.8 to 0.8 with downregulation in green and upregulation in red.

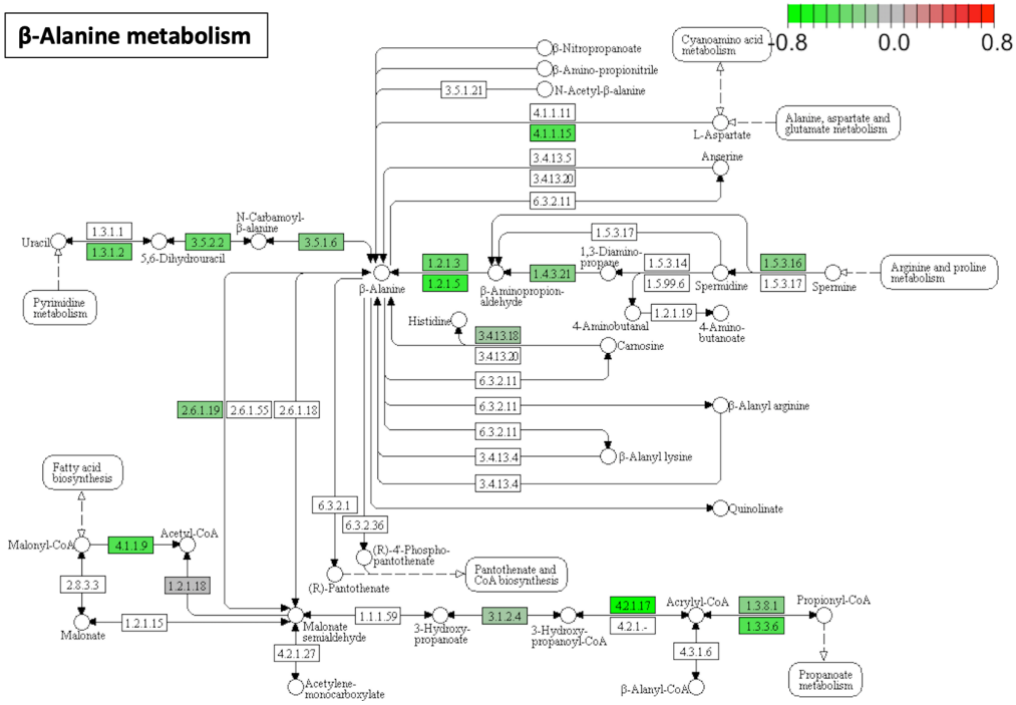
Focal adhesion



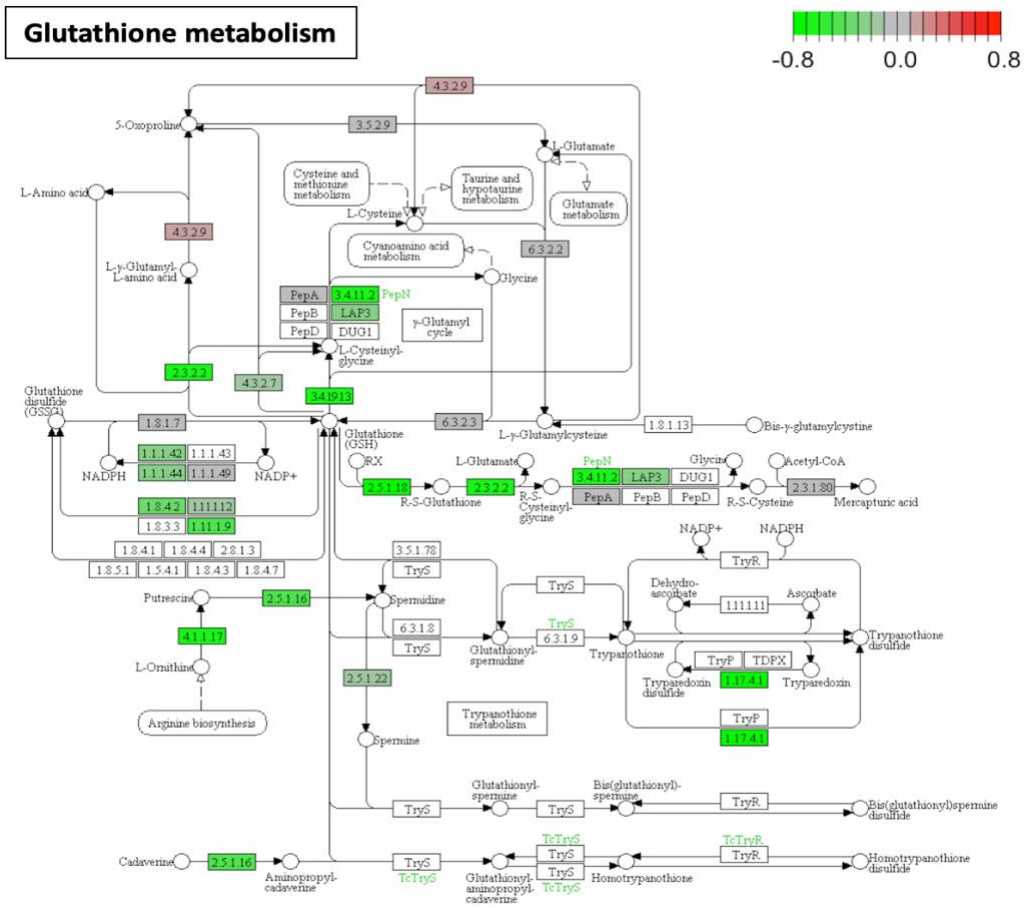
Fatty acid metabolism



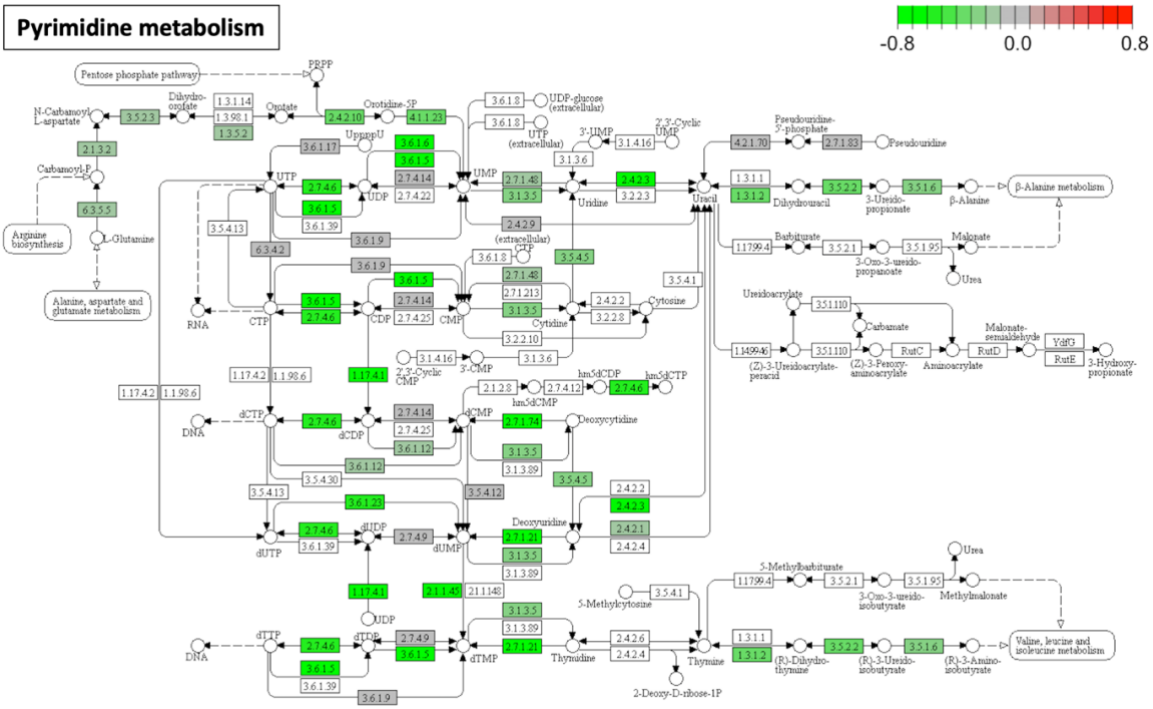
Beta-Alanine metabolism



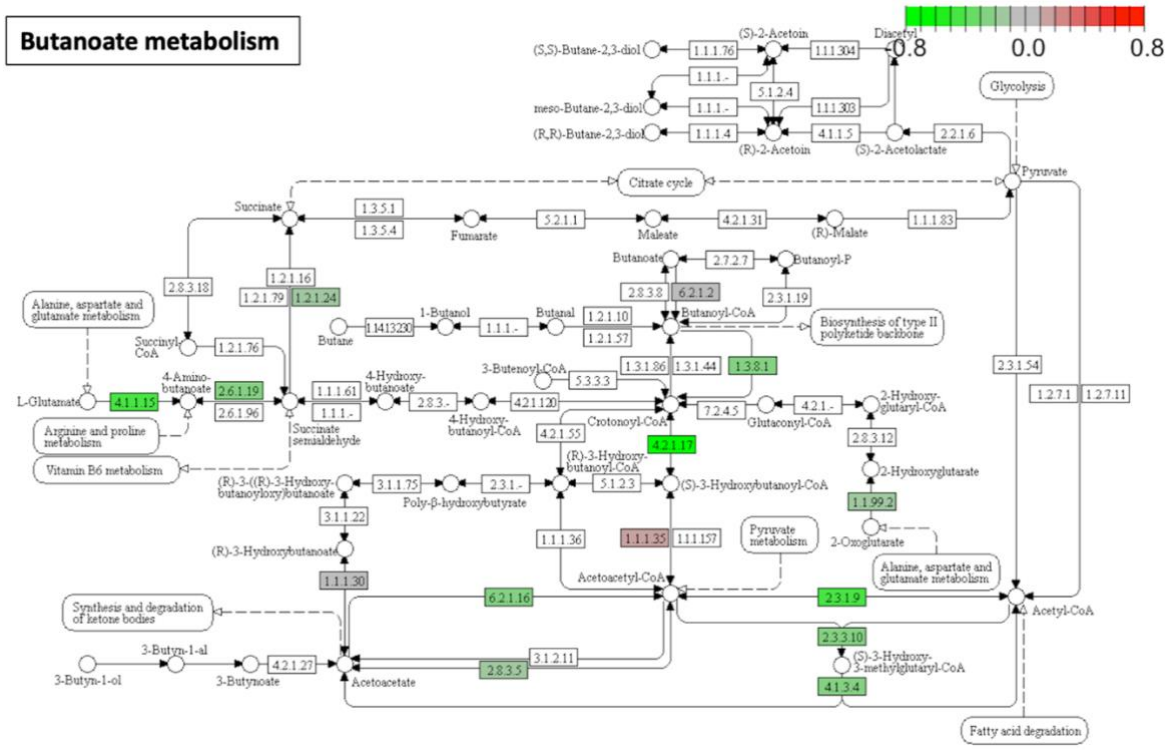
Glutathione metabolism



Pyrimidine metabolism



Butanoate metabolism



4.3.4 Supplementary data 4: IRE enrichment analysis

Supplementary data 4: IRE enrichment analysis

The GO *iron ion transport* was identified as significantly changed in heterozygous *psen1*$^{Q96\_K97del}$ mutant larvae, indicating the influence of the *psen1*$^{Q96\_K97del}$ mutation on iron homeostasis. Here, in order to detect whether ferrous iron dyshomeostasis effects are evident in the transcriptome data, we performed enrichment analysis of the gene sets containing IREs (iron-responsive elements) in either their 5' or 3' UTRs as defined by Hin et al., 2020 [1] using Goseq analysis (Supplementary data 4, Table 1) and GSEA (Supplementary data 4, Table 2). Genes containing high confidence IRE predictions were also selected as independent categories for enrichment analysis.

Category	DE genes	Genes in category	FDR
5'UTR_All	6	306	0.729629735
5'UTR_High	1	39	0.816242185
3'UTR_High	1	72	0.818415842
3'UTR_All	9	883	0.947236787

Supplementary data 4, Table 1. Enrichment analysis of the sets of genes containing IREs using Goseq analysis. "All" indicates all genes with such predicted IREs, while "High" contains only genes with such IREs predicted with high confidences (see [1]). FDR, false discovery rate.

Category	ES	NES	size	padj
3'UTR_High	-0.407666115	-1.231491999	72	0.605622962
5'UTR_High	0.365554576	1.114862574	39	1
3'UTR_All	-0.257651669	-0.96281786	883	1
5'UTR_All	-0.261396244	-0.925624336	306	1

Supplementary data 4, Table 2. Enrichment analysis of the sets of genes containing IREs using GSEA. ES and NES indicate enrichment score and normalized enrichment score respectively. "Size" presents the numbers of genes contributing to each category. padj, adjusted P-value.

In both Goseq analysis and GSEA, none of the gene sets showed significant enrichment (P-value < 0.05), indicating that the ferrous iron dyshomeostasis identified in young adult *psen1*$^{Q96\_K97del}$/+ brains was not detectable by this method at the transcriptome level in 7 dpf zebrafish larvae.

References

1. Hin, N., et al., *Iron Responsive Element (IRE)-mediated responses to iron dyshomeostasis in Alzheimer's disease.* bioRxiv, 2020: p. 2020.05.01.071498.

Chapter 5. Assessing short term spatial working memory in heterozygous *psen1*$^{Q96\_K97del}$ zebrafish adults of different ages

5.1 Introduction

Transcriptome analysis of 7 pdf whole heterozygous *psen1*$^{Q96\_K97del}$ mutant larval zebrafish showed changes due to the mutation inconsistent with those identified in 6-month heterozygous mutant brains, although changes in similar cellular pathways were predicted, including oxidative phosphorylation, mitochondrial function, lysosomal acidification and iron ion transport. It is likely that some gene expression changes seen in 6-month-old mutant brains required time to develop and so are not seen in larvae (e.g. iron regulatory element-containing transcript abundance changes). Also, a greater number of tissue types is found in whole larvae than in adult brains. In addition, gene expression levels showed obvious variability between larval clutches of the same genotype as, indeed, was seen for gene expression in brains in **Chapter 2** and this can make it difficult to identify a selection of gene expression changes to represent the pathological genetic state of EOfAD. Therefore, it did not appear viable to use mutant zebrafish larvae for drug screening through the assessment of gene expression states.

Cognitive impairment is a major clinical symptom of AD. Zebrafish larvae also show behavioural responses to drugs similar to adult zebrafish and other vertebrates [1]. Zebrafish larvae show an obvious locomotor response to photic stimuli. They

move to a source of light when illumination is reduced, and become transiently hyperactive in complete darkness before adopting a resting state [2]. Also, the influence of lighting on larvae's locomotion was studied by MacPhail et al., 2009 [3]. Locomotion was tested in an initial 10-min period of darkness followed by either a 40-min period of illumination or darkness. Locomotion in darkness (including the initial darkness) initially increased to a maximum at 4 minutes and then decreased to a stable low level after 20 minutes, while locomotion during light was initially low and gradually increased to a stable, high level after 20 minutes. However, when exposed to the stimulation of alternating light and dark periods of around 10 minutes each, zebrafish larvae showed obvious changes in locomotor activity where they were stably inactive in light but extremely active in darkness. These features have been increasingly employed in high-throughput screening for neuroactive drugs [4]. To observe whether the behaviour of heterozygous *psen1*$^{Q96\_K97del}$ mutant larvae might be useful in drug screening, their response to simulation by repeating intervals of 3-min light and 1-min dark, was assessed by Morgan Newman (unpublished data). However, no significant differences in activity patterns between wild type and mutant larvae were observed. This is not surprising as AD is an age-dependent disease, and so cognitive impairment is unlikely to occur at very early ages.

As heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish larvae did not show differences in behaviour compared to wild type larvae, we sought to observe behavioural changes in older, adult zebrafish. As a model of human disease, *psen1*$^{Q96\_K97del}$ mutant zebrafish are expected to show equivalent behavioural changes, such as decline in memory. As impairments in short-term spatial working memory have

been observed in AD patients with moderate dementia using the Corsi block-trapping test [5], the assessment of an animal model's short-term spatial working memory might reveal progressive decline in cognition. The Free-movement pattern (FMP) Y-maze test designed by Cleal et al., 2020 [6] investigates animals' search strategies by recording direction choices (left, L, or right, R) during an hour of free exploration. Impacts on working memory can be assessed by quantifying differences in "alternation tetragrams" (LRLR and RLRL), which are the dominant search strategy used in vertebrates (see details in ***section 1.8***). Therefore, as part of an initial attempt to investigate behavioural differences in adult mutant zebrafish compared to their wild type siblings, the FMP Y-maze was employed to assess working memory by quantification of alternation tetragrams.

An assessment of alternation tetragrams using the FMP Y-maze was previously performed by Jiayu Ruan (an M.Sc. candidate in the ADGL) on heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish and their wild type siblings at 6, 12, and 24 months of age. The 6- and 24-month-old mutant zebrafish did not show any statistically significant differences in alternation tetragrams while a slight decrease in alternation tetragrams was identified in 12-month-old mutant zebrafish, indicating possible impairments of spatial working memory (**Figure 5.1**, [7]).

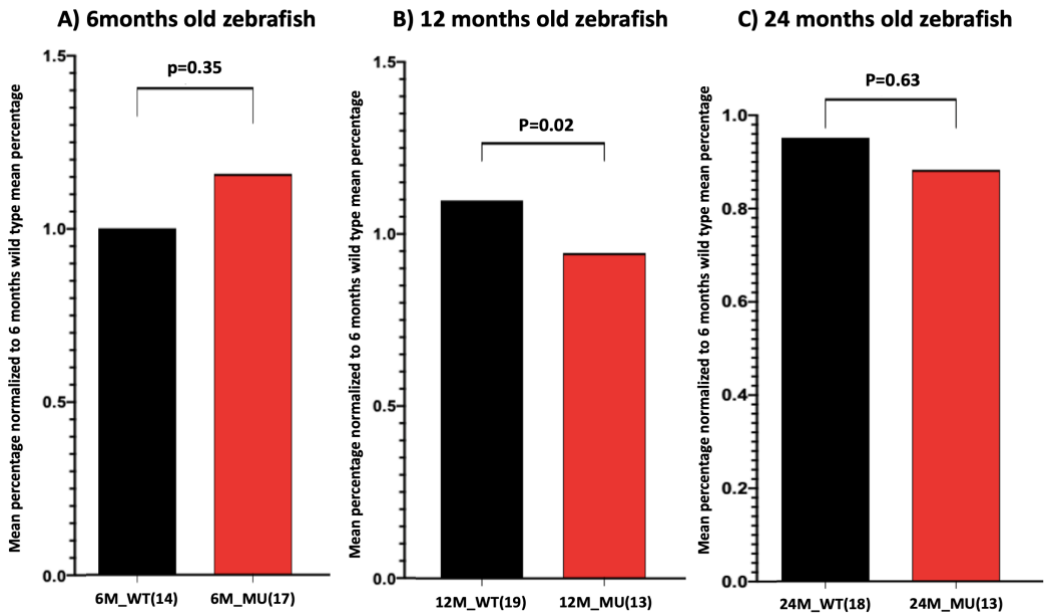


Figure 5.1. The normalized percentage of alternation tetragrams in zebrafish of A) 6 months of age, B) 12 months of age and C) 24 months of age (The figures were taken from the thesis of Jiayu Ruan [7]). The mean percentages of alternation tetragrams are normalized to that of 6-month-old zebrafish. The number in brackets indicates the zebrafish numbers used in tests. WT, wild type; MU, mutant.

However, the limited sample size and the high variability between samples of the same genotype in these tests undermined our confidence in the validity of the cognitive difference at 12 months. Therefore, for the work described in this chapter, we re-performed the assessment of alternation tetragram behaviour in heterozygous *psen1*$^{Q96\_K97del}$ mutant adult zebrafish of 6 months, 12 months as well as 24 months of age respectively compared to their wild type siblings to investigate changes in short-term spatial working memory during adult aging. No statistically significant differences in alternation behaviour were identified at any age.

The sample sizes used in FMP Y-tests exceeded the required n=14/per group, which was calculated using effect size [d] =1.2, power=0.8, alpha=0.05 [6, 8]. Therefore, the sample sizes should have been sufficiently large to detect differences with

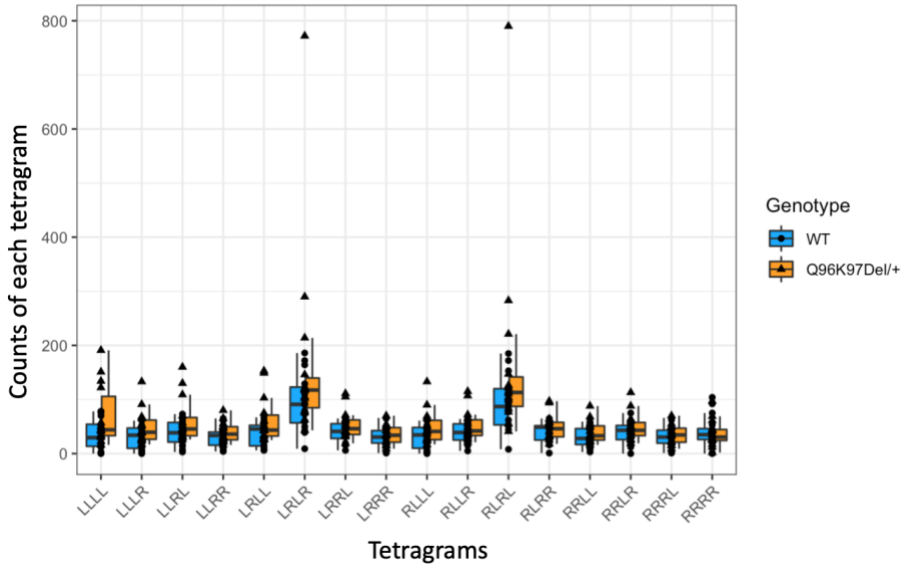
statistical significance. The failure to identify differences in alternation behaviour at any age may mean that any difference observed previously was artefactual. Alternatively, the failure to detect any difference between genotypes may have been due to other factors such as increased variability of behaviour e.g. due to environmental stressors that were not accounted for. Also, the movement patterns of zebrafish might not reflect the search strategy on which the test is based when there are long intervals between two Y-maze arm entries due to low activity. In addition, some zebrafish seemed to move randomly during the free exploration rather than following an alternation search strategy, no matter what genotype they were. Therefore, the assessment of short term spatial working memory using the FMP Y-maze test probably did not work well and requires further optimizations. It might be worth using a training paradigm to detect changes in cognition, such as learning ability or long term memory, although such studies are more technically complex and labour intensive.

5.2 Results

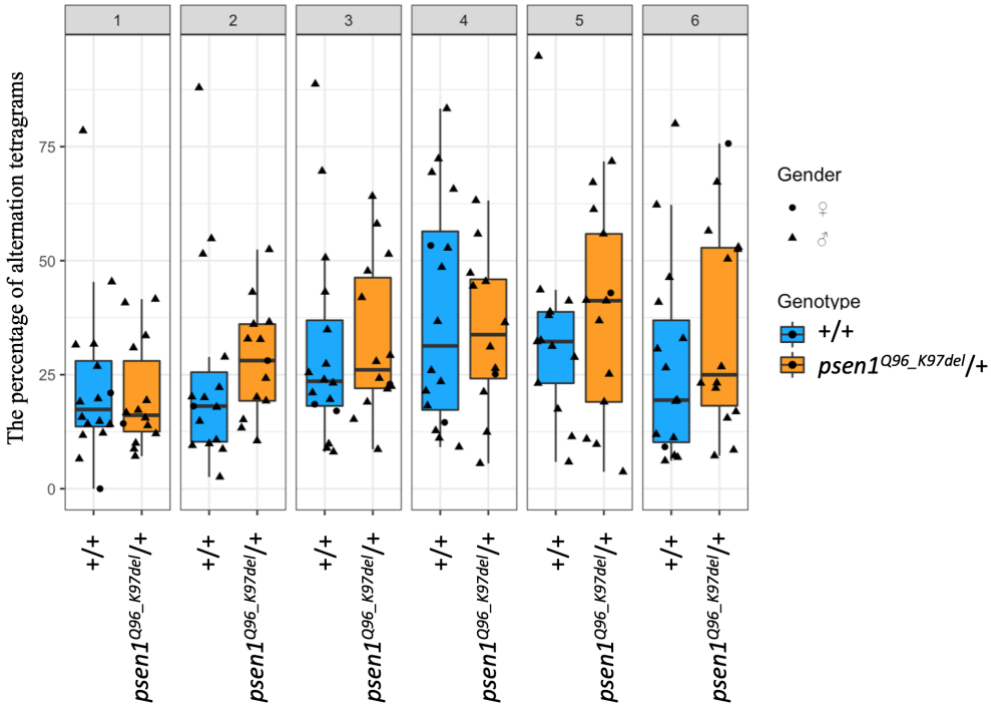
5.2.1 Assessment of spatial working memory in 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish

A FMP Y-maze test, assessing short-term spatial working memory, was firstly performed on the 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish and their wild type siblings. The zebrafish movement patterns were recorded and then analysed as tetragrams, with the tetragrams RLRL and LRLR regarded as alternations, which is the main search strategy used by zebrafish [6]. The total counts of each tetragram in a 1-hour period of free exploration are plotted in **Figure 5.2A**. The alternation tetragrams LRLR and RLRL occurred more frequently for both wild type and mutant zebrafish compared to other tetragrams. The percentage of alternation tetragrams LRLR and RLRL relative to all tetragrams were plotted in six 10-minute bins across the 1-hour period of free exploration (**Figure 5.2B**). None of the bins showed significant differences in the percentage of alternation tetragrams between mutant and wild type zebrafish. Also, the percentage of alternation tetragrams showed a very large spread between zebrafish of the same genotype, indicating there was large variability of individuals in alternation behaviour. The overall probabilities of alternations during the 1-hour test was plotted in **Figure 5.2C**, and no significant differences in alternations could be observed, which is similar to previous results from analyses by Jiayu Ruan (**Figure 5.1A**, [7]). As 6-month-old mutant zebrafish might be too young to show impairments in working memory, the FMP Y-maze test was then performed on older mutant zebrafish.

A) Counts of each tetragram in 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish



B) The percentage of alternation tetragrams in 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish



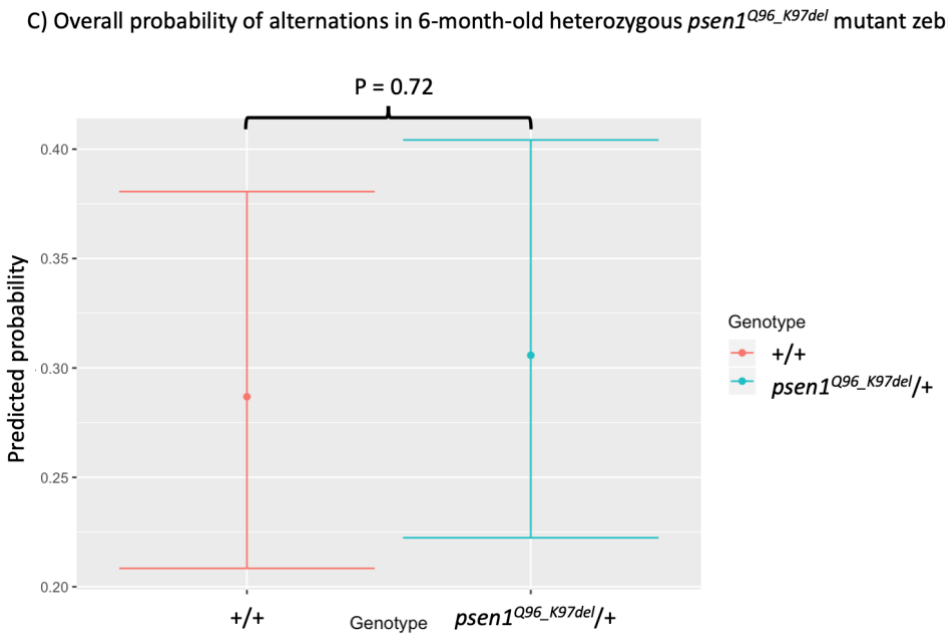
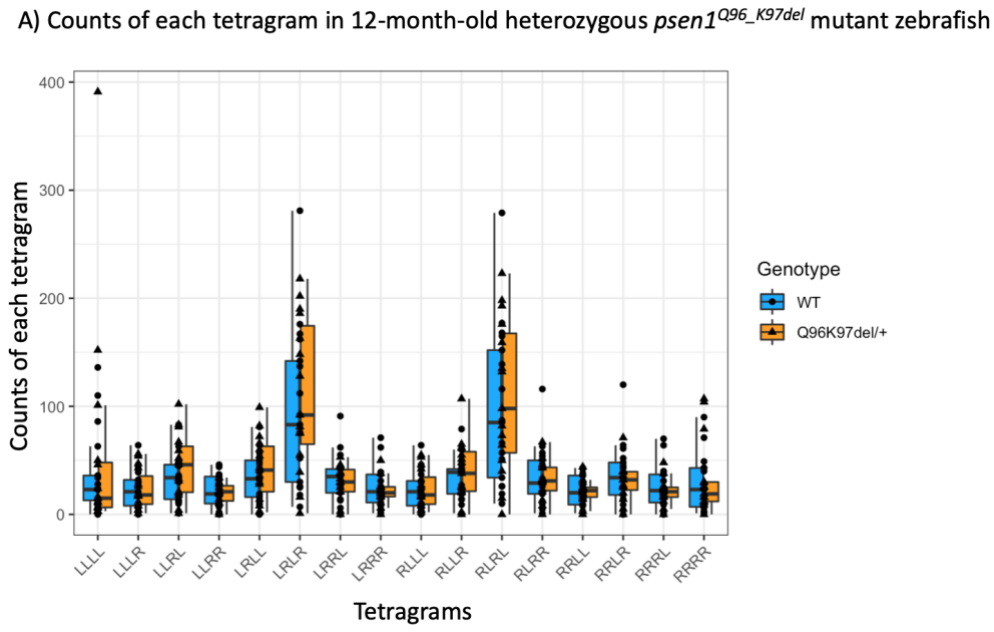


Figure 5.2. Assessment of alternation tetragrams in 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish. A) The counts of each tetragram. B) The percentage of alternation tetragrams LRLR and RLRL relative to all tetragrams, analysed in 10-min bins. C) the overall predicted probability of alternations during the 1-hour test. "●" indicates females, "▲" indicates males.

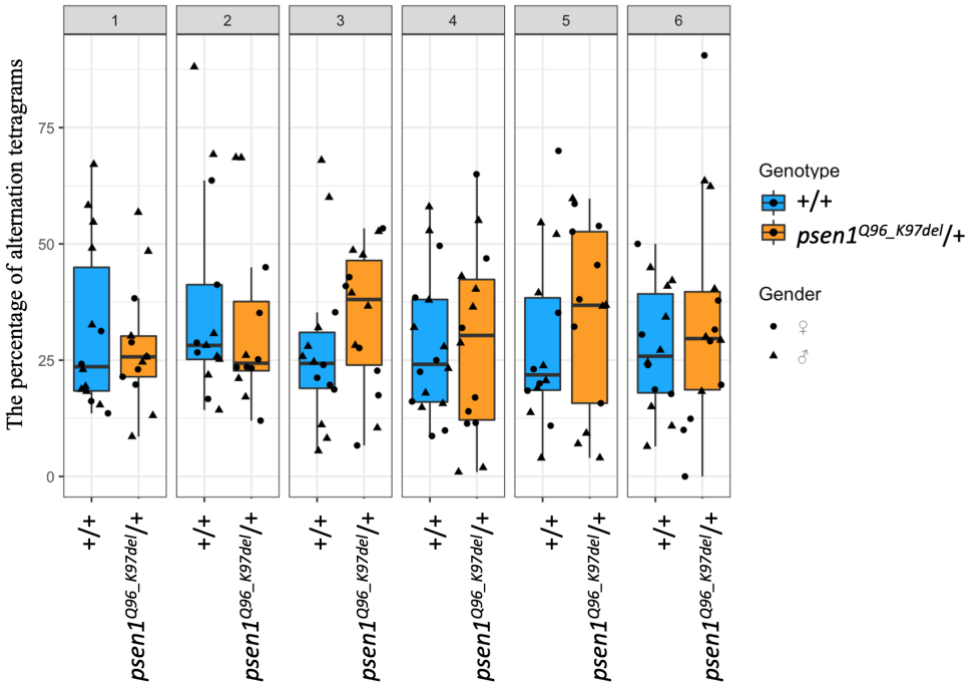
5.2.2 Assessment of spatial working memory in 12-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish

As no differences in alternations could be identified in 6-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish, the FMP Y-maze test was then performed on older 12-month-old mutant zebrafish. In a previous test heterozygous mutant individuals had shown a significant decrease in alternations at this age (**Figure 5.1B**, [7]). Note that, due to a malfunction of the lighting timer system in the room where the zebrafish were maintained, room lighting was continuously on for several months before and during the test period and this may have affected the circadian behaviour of the fish and put them under a degree of stress [9].

The counts of each tetragram from this study are plotted in **Figure 5.3A.** The alternation tetragrams LRLR and RLRL showed a tendency to be more frequent than other tetragrams in both wild type and mutant zebrafish. However, the percentage of alternation tetragrams in each 10-min bin (**Figure 5.3B**) and the predicted overall probabilities of alternations (**Figure 5.3C**) did not show any differences between wild type and mutant zebrafish, which is not consistent with previous results (**Figure 5.1B**, [7]).



B) The percentage of alternation tetragrams in 12-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish



C) Overall probability of alternations in 12-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish

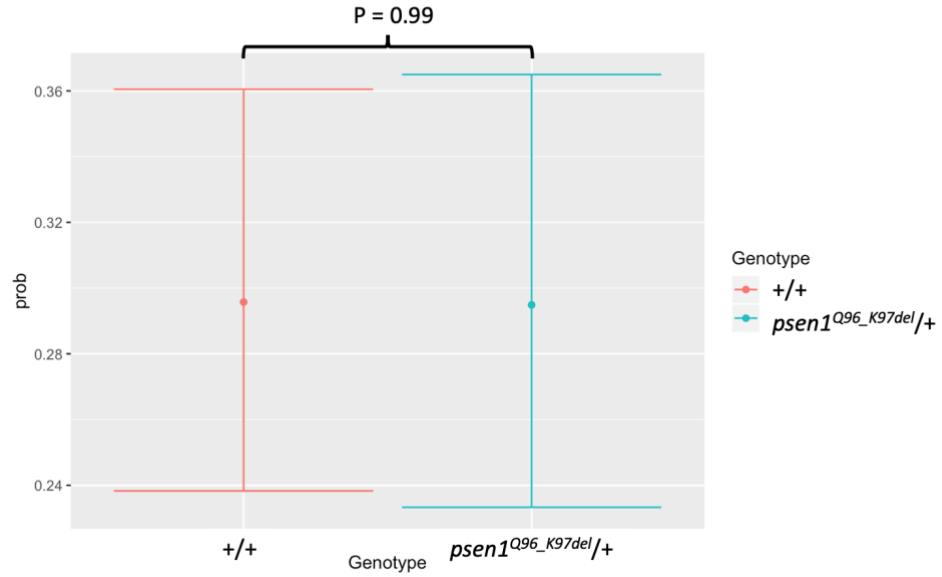
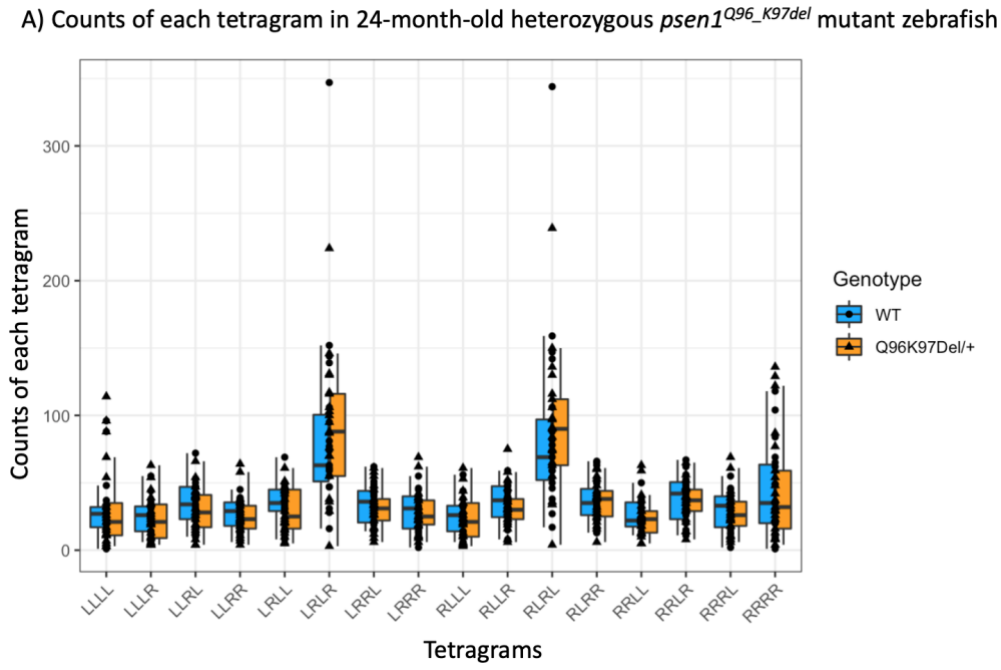


Figure 5.3. Assessment of alternation tetragrams in 12-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish. A) The counts of each tetragram. B) The percentage of alternation tetragrams LRLR and RLRL relative to all tetragrams, analysed in 10-min bins. C) the overall predicted probability of alternations during the 1-hour test. "●" indicates females, "▲" indicates males.

5.2.3 Assessment of spatial working memory in 24-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish

Lastly, the FMP Y-maze test was performed on 24-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish and their wild type siblings. Similar to the results seen in 6-month-old and 12-month-old zebrafish, the alternation tetragrams LRLR and RLRL showed a tendency to be more frequent in both wild type and mutant zebrafish (**Figure 5.4A**), while the percentage of alternation tetragrams in each 10-min bin (**Figure 5.4B**) and the predicted overall probabilities of alternations (**Figure 5.4C**) did not show any differences between wild type and mutant zebrafish.



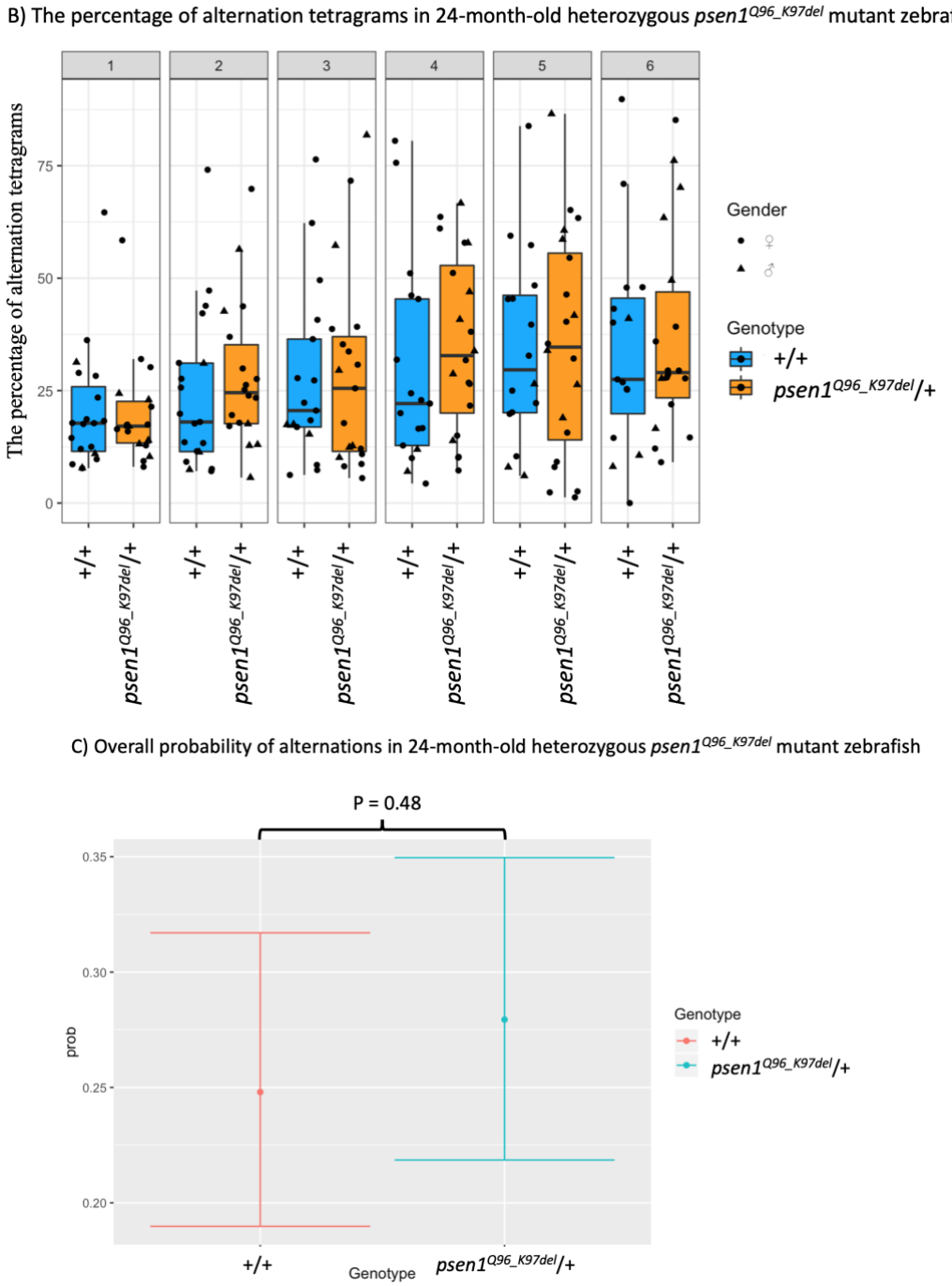


Figure 5.4. Assessment of alternation tetragrams in 24-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish. A) The counts of each tetragram. B) The percentage of alternation tetragrams LRLR and RLRL relative to all tetragrams, analysed in 10-min bins. C) the overall predicted probability of alternations during the 1-hour test. "●" indicates females, "▲" indicates males.

5.3 Discussion

Animals carrying EOfAD-like mutations, as models of human EOfAD, are expected to show similar effects to those seen in human patients. The major clinical symptom of AD is progressive impairments in cognition. As zebrafish larvae represent a powerful model for drug screening, and also show behavioural responses to neuropharmacological treatment similar to adult zebrafish [1], the ADGL initially sought to compare the response of heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish larvae and their wild type siblings to the stimulation of alternating light and dark intervals. No differences in swimming patterns could be identified (unpublished data, study performed by Morgan Newman). As impairments in short-term spatial working memory are observed in AD patients with moderate dementia [5], and the FMP Y-maze provides a simple, rapid and advanced method for the assessment of spatial working memory in zebrafish adults [6], we employed the FMP Y-maze to determine whether impairments in spatial working memory could be observed in our EOfAD-like zebrafish adults. Previous FMP Y-maze tests done by Jiayu Yuan [7] found a slight decrease in FMP Y-maze alternation search strategy, (which are positively correlated with working memory), for 12-month-old heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish compared to their wild type siblings, while no changes in alternation search strategy were observed with 6-month-old and 24-month-old mutant zebrafish. To ensure the reproducibility of these results, the work described in this chapter repeated the FMP Y-maze tests using independent zebrafish families. However, no differences in alternations were observed between heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish and their wild type sibling of all three ages, 6 months, 12 months and 24 months. A power calculation

using effect size [d] =1.2, power=0.8, alpha=0.05 showed that n=14 was required for each sample group [6, 8]. Therefore, the sample sizes used in our behavior tests exceeded that requirement and should have been sufficiently large to detect differences with statistical significance.

The failure to identify differences in alternation search strategy may be due to a number of factors. There was great variability in movement patterns between zebrafish individuals. Although all zebrafish were allowed the same 1-hour period of free exploration in the FMP Y-maze with (as far as we know) no differences in external influences, zebrafish individuals showed great differences in activity. The most active zebrafish could produce approximately 1000 tetragrams, while the most inactive fish could only produce approximately 100. Although the alternation tetragrams were analysed as percentages of all tetragram types, the relatively inactive zebrafish showed long intervals between two consecutive arm entries so that their behaviour might not appropriately represent working memory. Moreover, a number of zebrafish individuals of either mutant or wild type genotype, showed movement behaviour not favouring alternation tetragrams (i.e. all tetragrams had similar frequency). It seemed that not all zebrafish performed the alternation search strategy in the FMP Y-maze tests, and their movement patterns were more likely to be random during the full exploration. As a relatively new method for assessment of animals' spatial working memory, the FMP Y-maze test probably requires more experimental data to improve interpretation of its output and to overcome the problems of low animal activity and outliner samples, to achieve better discrimination between cognitively impaired and normal individuals. Also, the assessment of the changes in cognition in AD should not be limited to working

memory (the FMP Y-maze test). Other behavioural tests involving training can assess long term memory and learning abilities. Although more technically complex, these can be employed in future studies if they detect more easily differences in behaviour between cognitively impaired and normal zebrafish.

5.4 Methods

Zebrafish families for short term spatial working memory tests

Three zebrafish families were used for short term spatial working memory tests at 6 months (31 zebrafish, 16 wild type and 15 mutant), 12 months (32 zebrafish, 17 wild type and 15 mutant) and 24 months (40 zebrafish, 19 wild type and 21 mutant) of age respectively. Each family was generated from a parental heterozygous *psen1*$^{Q96\_K97del}$ mutant zebrafish crossed with a wild type zebrafish, resulting in approximately half wild type and half mutant genotypes among the progeny. Each zebrafish family was raised in the same tank to minimize variable impacts of the environment between individuals.

Short term spatial working memory test using Y-maze

Two zebrafish individuals were collected and isolated for 30 mins before the tests. The tests were performed using the Zantiks AD system (Zantiks Ltd., Cambridge, UK) having a Y-maze with three identical arms oriented at an angle of 120° relative to each other and filled with 1L of aquarium water. A zebrafish to be tested was placed in the Y-maze and its movement was recorded for 1 hour. Testing was performed between 10:00 am and 3:00 pm to reduce circadian influences. Data was outputted as a time series of arm entries and exits, which were then analysed into 16 overlapping tetragrams. Two tetragrams, LRLR and RLRL, represented alternations. Tested zebrafish individuals were then genotyped by PCR with sex recorded. To test whether alternation frequency was affected in *psen1*$^{Q96\_K97del}$

mutants, we fitted the data to a generalised linear mixed model (GLMM) using Template Model Builder (glmmTMB) [10] specifying a betabinomial variance function. The fixed effects were genotype, 10-min interval bins and time (the start time of the FMP Y-maze test), and the random factors were fish identity, day (zebrafish families were tested over a continuous series of days) and interaction of time and day. Tests of significance of the fixed effects were performed by Anova using Type II Wald Chi-square tests [11].

Genomic DNA extraction

Zebrafish were anesthetized using tricaine solution (168μg/mL), and a small tail biopsy (a "tail clip") was then taken from the zebrafish using a surgical blade. 50μL of 1.7mg/μL Proteinase K (Sigma-Aldrich, Missouri, USA) in 1xTE buffer was used to degrade proteins and release genomic DNA from zebrafish tail clips through incubation at 55°C for 3 hours, followed by enzyme inactivation at 95°C for 15 minutes. Genomic DNAs were then stored at -20°C for the following genotyping PCRs.

PCR for genotype determination

PCRs were performed in 25 μL reactions containing genomic DNA, 0.2 mM of each deoxyribonucleotide triphosphate (dNTP, each at a final concentration of 0.2 mM), 0.4 μM of each PCR primer, and 2.5 unit of GoTaq® DNA polymerase

(Promega, Madison, Wisconsin, USA). PCR cycling was performed with 35 cycles of a denaturation temperature of 95°C for 30 s, then an appropriate (primer-dependent) annealing temperature for 30 s and then an extension temperature of 72°C for 45 s. PCR products were electrophoresed through a 1.5% agarose gel in 1×TAE buffer for separation and identification.

5.5 Reference

1. Ahmad, F., et al., *Zebrafish embryos and larvae in behavioural assays.* Behaviour, 2012. **149**(10-12): p. 1241-1281.
2. Burgess, H.A. and M. Granato, *Modulation of locomotor activity in larval zebrafish during light adaptation.* J Exp Biol, 2007. **210**(Pt 14): p. 2526-39.
3. MacPhail, R.C., et al., *Locomotion in larval zebrafish: Influence of time of day, lighting and ethanol.* Neurotoxicology, 2009. **30**(1): p. 52-58.
4. Basnet, R.M., et al., *Zebrafish Larvae as a Behavioral Model in Neuropharmacology.* Biomedicines, 2019. **7**(1).
5. Guariglia, C.C., *Spatial working memory in Alzheimer's disease: A study using the Corsi block-tapping test.* Dement Neuropsychol, 2007. **1**(4): p. 392-395.
6. Cleal, M., et al., *The Free-movement pattern Y-maze: A cross-species measure of working memory and executive function.* Behav Res Methods, 2020.
7. Ruan, J., *The impacts of age, hypoxia, and the psen1Q96_K97del / + mutation on familial Alzheimer's disease,* in *School of Biological Science*. 2019, The university of Adelaide.
8. Cleal, M. and M.O. Parker, *Moderate developmental alcohol exposure reduces repetitive alternation in a zebrafish model of fetal alcohol spectrum disorders.* Neurotoxicol Teratol, 2018. **70**: p. 1-9.
9. Purushothaman, S., et al., *Proteomic and gene expression analysis of zebrafish brain undergoing continuous light/dark stress.* Journal of Sleep Research, 2015. **24**(4): p. 458-465.
10. Brooks, M.E., et al., *glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling.* R Journal, 2017. **9**(2): p. 378-400.
11. Fox, J. and S. Weisberg, *An R Companion to Applied Regression*. 2019, Sage: Thousand Oaks CA.

5.6 Appendix: Primer information for genotyping

Primers specific to psen1 wild type allele and psen1^{Q96_K97del} mutant allele respectively.

Ensembl gene ID	**Target gene**	**Primer sequences**	**Annealing temperature (°C)**
ENSDARG00000004870	*psen1* wild type allele	5'-TCTGTCAGCTTCTACACACAGAAGG-3'	61
		5'-CCATCCCTAAACTGCTCCTACTC-3'	
ENSDARG00000004870	*psen1*$^{Q96\_K97del}$ mutant allele	5'-TGTCAGCTTCTACACAGACGGA-3'	61
		5'-CCATCCCTAAACTGCTCCTACTC-3'	

Chapter 6: Generating aspartate codon-substitution and Acne Inversa-like mutations in zebrafish *psen1* gene

6.1 Introduction

EOfAD is autosomal and dominantly inherited [1], and over 200 EOfAD-causative mutations have been identified in the major locus, *PSEN1* [2]. The PSEN1 protein mainly participates in APP cleavage and amyloid-β production through its γ-secretase activity. However, the involvement of γ-secretase in EOfAD pathology remains controversial, as the effects of EOfAD-causative mutations on γ-secretase activity are variable (see ***Section 1.4.3***). Moreover, EOfAD-causative mutations in the *PSEN1* gene obey a "fAD reading frame preservation rule" such that they do not truncate the open reading frame and maintain production of a mutant transcript that encodes a "full-length" protein [3]. Coding sequence-truncating mutations in *PSEN1* genes do not cause EOfAD but may cause other diseases. A familial form of the skin disease Acne Inversa (Hidradenitis Suppurativa) without EOfAD is caused by a dominant, frameshift mutation in human *PSEN1*, P242LfsX11, which leads to a premature termination codon and produces mutant transcripts encoding a truncated protein [4]. These kinds of non-EOfAD-causative mutations may lead to loss of PSEN1 protein function and probably have some similar functional effects to some EOfAD-causative mutations, such as reduction of γ-secretase activity. However, these effects will not be sufficient for EOfAD pathology. Current transcriptome studies of EOfAD mainly focus on transgenic animal models (see the review by Newman et al., 2017 [5]) and the ADGL has now published or submitted

for publication a number of papers describing knock-in models of EOfAD-like mutations [6-12]. However, many of the changes identified may be irrelevant to AD. Therefore, the generation and analysis of non-EOfAD-like mutations in the EOfAD genes may help define molecular changes that do not directly induce EOfAD pathology (at least by themselves) and can be excluded from set of EOfAD-specific transcriptional effects.

Since γ-secretase activity is an important function of PSEN1, it is important to analyse the effects of its loss on brain transcriptomes in comparison to EOfAD-like mutations but while maintaining the ability to produce a full-length holoprotein. A study of PSEN1 γ-secretase activity conducted by Wolfe et al., 1999 [13] demonstrated that two transmembrane aspartate residues in PSEN1 are critical to the γ-secretase enzyme's catalytic site, and mutating either aspartate to alanine inhibits γ-secretase activity. Therefore, substitution of either aspartate leads to loss of γ-secretase activity but with, presumably, less effect on other protein properties. The generation of a knock-in aspartate codon-substitution mutation model in zebrafish would be useful in transcriptome comparisons with EOfAD-like models to investigate whether or not loss of γ-secretase activity itself induces EOfAD-like transcriptome effects.

In the work described in this chapter, I aimed to generate both an Acne Inversa-like mutation and an aspartate codon-substitution mutation in the zebrafish model, which would help to define the molecular changes caused by truncation of PSEN1 protein and by loss of γ-secretase activity. The CRISPR/Cas system (introduced in ***section 1.6***) was used for the mutation generation with template DNA included to

enable homology-directed repair (HDR). However, due to the low efficiency of HDR in zebrafish (partially due to the rapid development of zebrafish embryos), the desired HDR-induced mutations were not achieved. Instead, two mutations of interest were identified near human the Acne Inversa-equivalent mutation site in the zebrafish *psen1* gene. *Psen1*^{W233fs} substitutes three nucleotides with another four, resulting in at +1 shift in the reading frame. This frameshift mutation is assumed to produce mutant transcripts encoding a truncated protein similar to the human Acne Inversa mutation which can be used in future comparative transcriptome studies. Another mutation, *psen1*$^{W233I\_T234L}$, is a substitution of 6 nucleotides but preserves the reading frame. The mutant transcript of this mutation encodes a full-length protein with a substitution of two amino acids at a membrane transition site, which will likely affect protein structure and hydrophobicity. Therefore, this mutation is similar to many EOfAD-causative mutations in human *PSEN1*, and may be used as another EOfAD-like model in future experiments.

6.2 Results

6.2.1 Attempted generation of an aspartate codon-substitution mutation in *psen1*

The PSEN1 protein possesses two catalytic aspartate residues, both of which are critical for its γ-secretase activity [13]. A mutation substituting one of these aspartate codons with a codon for alanine is desirable for the study of the involvement of γ-secretase activity in AD pathology, as it would cause loss of γ-secretase activity possibly with minimal impact on other properties of the protein and it would preserve the open reading frame similar to EOfAD-causative mutations in human *PSEN1*. Through transcriptome comparison with our EOfAD-like models, this aspartate codon-substitution model may inform us whether a lack of γ-secretase activity itself can induce similar brain transcriptome changes. As the upstream catalytic site aspartate codon is located on an exon-exon boundary, mutating this site may affect transcript splicing. Therefore, we attempted to mutate the second catalytic site aspartate codon to codon for alanine. To facilitate the generation of HDR-induced mutations (the substitution of aspartate with alanine) in zebrafish, template oligo DNA was designed [14] and co-injected with the CRISPR/Cas12a system (that recognises the PAM site 5'-TTTV-3', **Figure 6.1**).

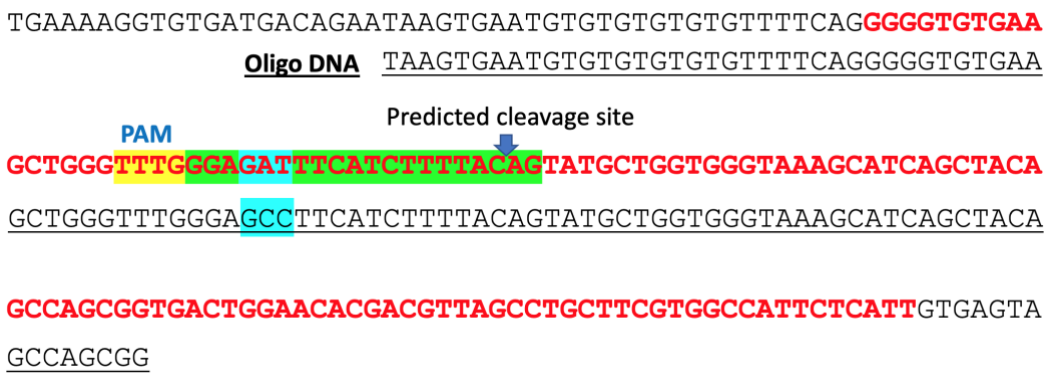


Figure 6.1. The design of sgRNA and oligo DNA template for mutation generation in *psen1* gene. The exon sequence is shown in red, and the oligo DNA sequence is underlined. The CRISPR target sequence is highlighted in green and its PAM site is highlighted in yellow. The downstream catalytic site aspartate codon and the correlated alanine site in the oligo DNA are highlighted in blue.

Ten embryos from each injection session (where ~65 embryos were injected in total) were collected for testing using the T7 endonuclease I assay that cleaves double-stranded DNA at mismatch sites, to detect the efficiency of mutagenesis by the CRISPR/Cas12a system alone. The results of T7 endonuclease I assay are shown in **Figure 6.2**. There were no differences between **Lane 2 wild type embryos** and **Lane 3 injected embryos 1**, suggesting that the first injection did not work well. In contrast, two clear cleavage bands of expected sizes were identified in **Lane 4 injected embryos 2**, indicating a good CRISPR/Cas12a cleavage efficiency. The remaining embryos of this injected clutch were raised until 6 months of age for mutation screening.

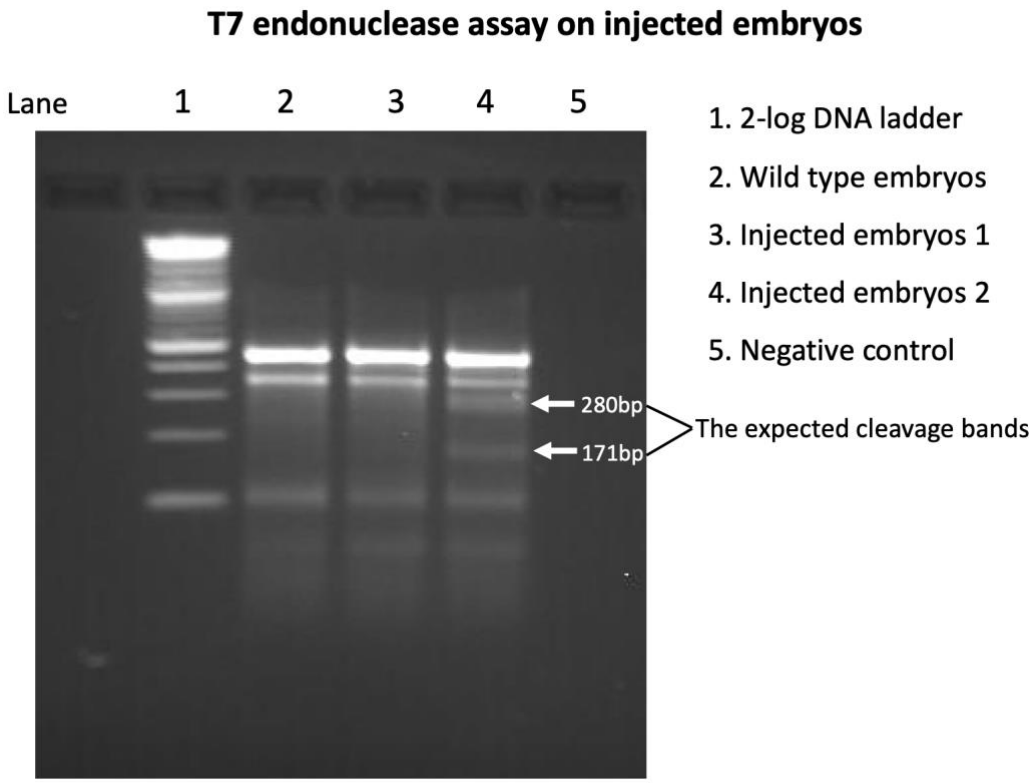


Figure 6.2. T7 endonuclease I assay on injected embryos. The region containing the mutation site was first amplified by PCR and was expected to produce a product of ~451 bp in size. After the T7 endonuclease I assay, two cleavage bands were expected to be produced for DNA from injected embryos having sizes of ~171 bp and ~280 bp. Possibly due to polymorphisms in intron sequence included in the PCR-amplified region, there were irrelevant bands produced from cleavage by T7 endonuclease I.

A small tail biopsy was taken from each 6-month-old injected G0 zebrafish individual for genotyping by PCRs with either wild type (aspartate-specific primer pair) or alanine-specific primer pairs. The genotyping PCR results are shown in **Figure 6.3**. In the tests using the wild type primer pair, all individual samples showed a strong, clear band at the expected product size of 181 bp, while in the tests using the mutant (alanine-specific) primer pair, only fish 2, 7 and 9 showed a

band at approximately 181 bp with different intensities, indicating that these three G0 fish possibly possessed the desired Alanine mutation. Fish 1 showed two strong, clear bands at larger sizes, and thus probably contained a complex DNA rearrangement that included insertion of the DNA template sequence.

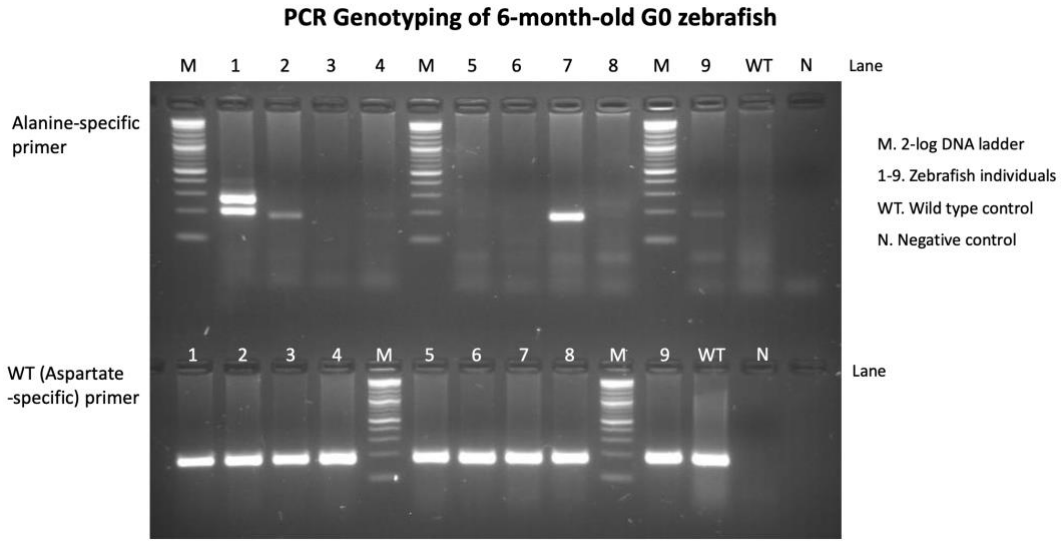


Figure 6.3. Genotyping injected zebrafish individuals at 6 months of age using both wild type (aspartate-specific) and alanine-specific primer pairs. The expected size of the PCR products for both wild type- and mutation-detection (alanine codon detection) primer pairs is 181 bp.

The G0 fish 2, 7 and 9 were subsequently outbred with a wild type zebrafish to generate F1 progeny. The F1 progeny were then genotyped at 6 months of age using both the wild type- and alanine codon-detection primer pairs, but none of them showed a positive signal for alanine codon-detection (results not shown). However, PCRs amplifying a region spanning location of the desired mutation were performed on F1 progeny (**Figure 6.4**), and F1 fish 2 and 7 from G0 fish 2 showed two strong, clear bands. The higher band indicated the normal genomic sequence, which occurred in all samples, while the lower band was novel and so was excised from the gel and purified for sequencing characterisation. Sequencing showed that

this PCR fragment was smaller than that from wild type fish due to a 134 bp deletion spanning approximately half of exon 10 (89 bp) and downstream intronic sequence (45 bp).

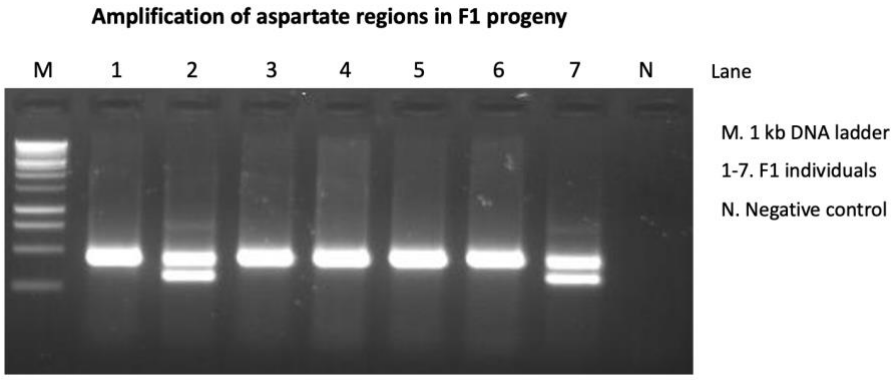


Figure 6.4. PCR amplification spanning the aspartate codon region in the F1 progeny of G0 fish 2. The size of the wild type genomic sequence is 451 bp.

The F1 fish 2 and 7 heterozygous for the 134 bp deletion were mated with each other to generate homozygous F2 progeny, in order to investigate the effects of the deletion on transcript formation. The F2 progeny were genotyped at 6 months of age, and cDNAs for PCRs were derived from one each of a wild type, heterozygous and homozygous mutant brains. PCRs were performed on the cDNAs amplifying a region spanning exon 8 to exon 11 (**Figure 6.5**), and the lower band/smaller fragment amplified from the homozygous mutant sample was purified from the gel and sequenced. The mutant transcripts lack exon 10 (119 bp), resulting in fusion of exon 9 sequence to exon 11 sequence and causing a shift in the reading frame followed downstream by a premature termination codon.

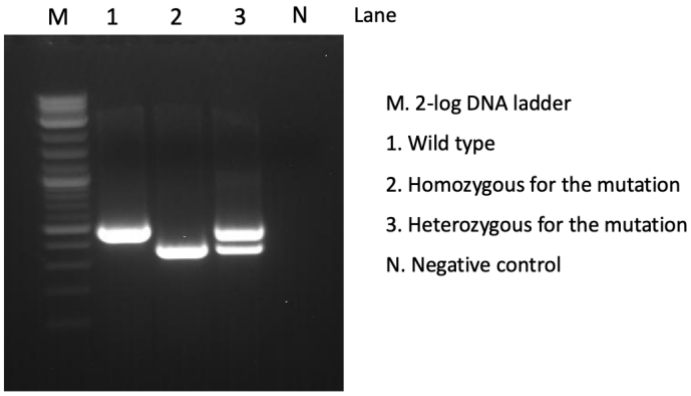


Figure 6.5. PCR amplification of a region of cDNA derived from *psen1* transcripts spanning exon 8 to exon 11 in a wild type, a heterozygous, and a homozygous mutant F2 progeny. The expected size of this region in normal transcripts is 489 bp (higher band), while in mutant transcripts lacking exon 10 it is 370 bp (lower band).

The 134 bp deleted region includes the aspartate codon, which we had aimed to substitute with an alanine codon. Therefore, this mutation almost certainly causes loss of γ-secretase activity. However, the change in the reading frame caused by loss of exon 10 sequence from transcripts mean that, if these transcripts are translated, they will produce a truncated protein. This mutation is, therefore, unlike EOfAD-causative mutations that follow the "fAD reading frame preservation rule" [3]. Nevertheless, the mutation may still find use in identifying any differential effects of truncation of the *psen1* coding sequence at different positions, e.g. in comparison with the *psen1*^{K97fs} mutation that truncates the coding sequence in exon 4 or the *psen1*^{W233fs} mutation described later that truncates the coding sequence in exon 7. Interestingly, we did appear to identify the desired aspartate codon-substitution mutation in DNA from the tail of the injected G0 fish number 7 (see **Figure 6.3**), but none of its progeny carried this mutation. This supports that HDR did occur in this zebrafish but not in cells that were included in the germline. Our

results support that the reported low efficiency of HDR in zebrafish [15]. Therefore, it is possibly inefficient to screen the injected G0 fish for designed HDR-induced mutations as these might not enter the germline. Instead, injected G0 fish could be crossed with wild type fish to produce F1 progeny to be subsequently genotyped by mutation-specific PCR reactions.

6.2.2 Generation of a familial Acne Inversa-like mutation in *psen1*

EOfAD-causative mutations in *PSEN1* gene follow a "fAD reading frame preservation rule" [3]. A dominant frameshift mutation in the human *PSEN1* gene, P242LfsX11, does not cause EOfAD but leads to a skin disease Acne Inversa (Hidradenitis Suppurativa) [4]. This mutation causes production of mutant transcripts encoding a truncated protein, which would lack any γ-secretase activity (as it would lack both catalytic aspartate residues). The transcriptome changes caused by this mutation do not induce EOfAD pathology but may be included among changes identified in EOfAD mutation-like models, e.g. if the latter also cause reduction in γ-secretase activity. The identification of these changes would refine our understanding of what transcriptome changes are essential for EOfAD. Therefore, I attempted to generate a model of the Acne Inversa mutation in the zebrafish *psen1* gene through simultaneous injection of CRISPR/Cas9 and a plasmid DNA template to direct HDR. The plasmid DNA template may be more stable than an oligonucleotide template and give a higher efficiency of HDR [16]. The plasmid DNA template was designed to mimic the sequence of human Acne Inversa with a stop codon at the corresponding position (**Figure 6.6**).

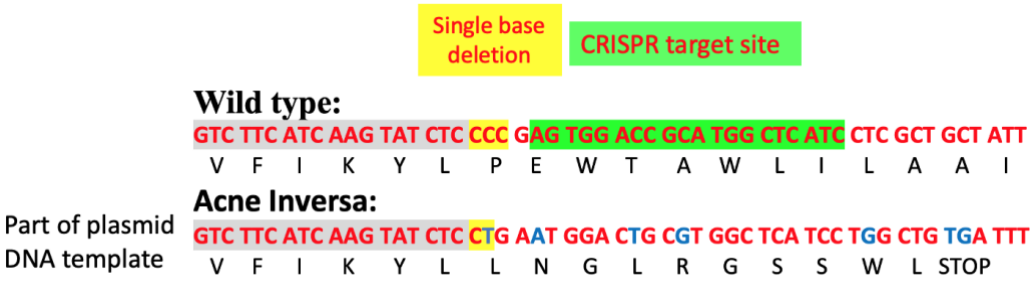


Figure 6.6. A region of the plasmid DNA template (around the desired mutation, the entire sequence is shown in ***section 6.5.1***). The mutation site is highlighted in yellow. The CRISPR target site is highlighted in green. Substituted nucleotides are shown in blue. The nucleotides substituted in addition to those at the P codon were included to allow specific binding of a PCR primer to the mutated region and to give a stop codon at a similar position to that occurring with the P242LfsX11 mutation of the human *PSEN1* gene.

To increase the efficiency of HDR in zebrafish embryos, 20 μM of the NHEJ inhibitor SCR7 was added into the zebrafish embryo culture medium (as suggested by Zhang et al., 2018 [16]). Ten embryos from each batch of injected embryos were pooled and PCR was used to detect any desired HDR events (**Figure 6.7**). Compared to wild type embryos, after gel electrophoresis of the PCR products, a weak band of the expected size occurred in each batch of injections, supporting that HDR-induced mutation occurred at some level in the embryos. However, the low levels of the desired PCR amplification products indicated that the HDR might be at a low frequency.

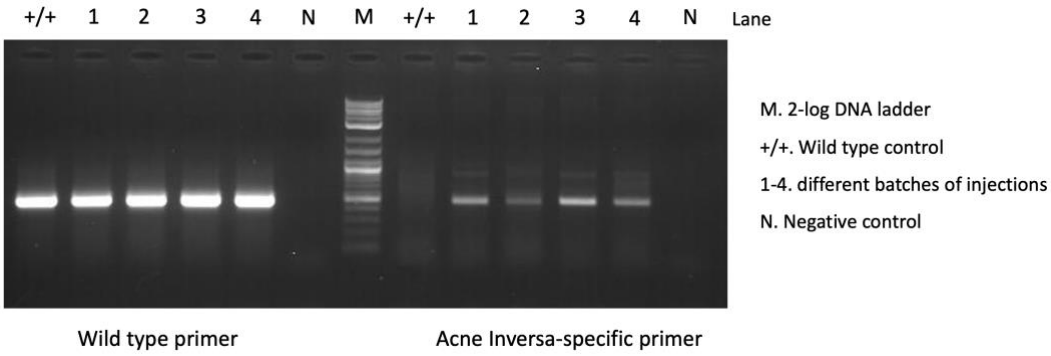


Figure 6.7. Genotyping the injected G0 embryos using PCR primer pairs detecting either the wild type or Acne Inversa mutation-like sequences. The expected size of the region amplified from wild type genomic DNA is 499 bp, and the expected size of the region amplified from the desired mutant sequence is 502 bp.

As discussed above following the attempted aspartate coding mutagenesis, tail biopsies from 6-month-old G0 zebrafish were not genotyped with the Acne Inversa-like mutation-detecting PCR primer pair. Instead, a T7 endonuclease I assay was used to screen for G0 fish carrying any mutation at the targeted mutagenesis site. This identified cleavage products of the expected sizes (**Figure 6.8**). Although electrophoresis revealed unexpected cleavage products in several samples (possibility due to the presence of polymorphisms in the fish DNAs), cleavage products of the expected sizes were identified in G0 fish 1, 2, 3, 4 and 5.

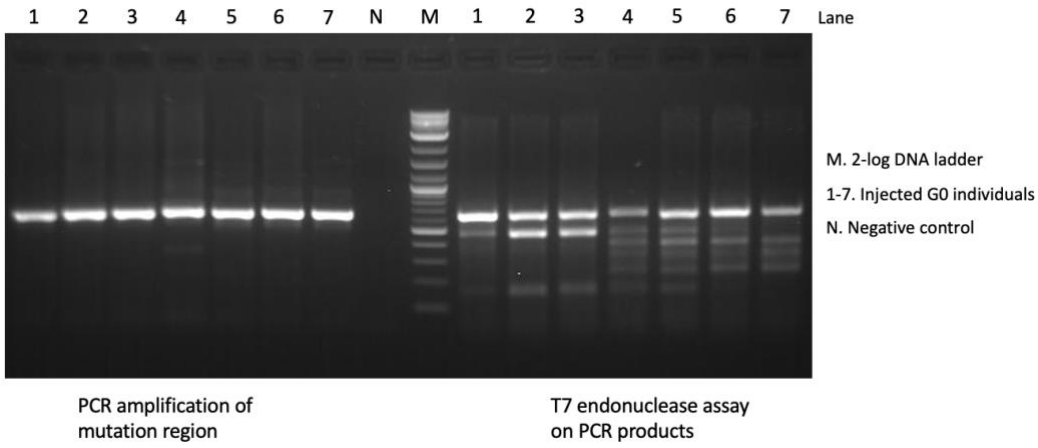


Figure 6.8. T7 endonuclease I assays on DNA amplified from 6-month-old G0 zebrafish individuals. The expected size of the PCR amplification product is 623 bp, while the product sizes after T7 endonuclease I cleavage are expected to be 126 bp and 461 bp.

The five G0 zebrafish carrying mutations at the target site were outcrossed with wild type fish to produce F1 progeny. The F1 progeny were genotyped using the mutation-detecting PCR primer pair, but no diagnostic PCR amplification products were observed, indicating that none of the F1 progeny carried the desired HDR-induced mutation (Results not shown). Therefore, T7 endonuclease I assays were performed on 6-month-old F1 progeny to screen for carriers of mutations differing from that originally desired. The mutation target region were then sequenced from fish identified as possessing such mutations. The following **Table 6.1** lists the mutations subsequently identified, including four in-frame mutations and two frameshift mutations.

Mutations	Wild type sequence	Mutated sequence
In-frame mutations		
21 bp indel	TGG ACC GCA W T A	TTA GCT ACT GTA AAC TGT ATC TCA L A T V N C I S
15 bp indel	GAG TGG ACC E W T	ATC AAA TAT CTC CCC I K Y L P
6 bp indel	TGG ACC W T	ATA CTT I L
9 bp deletion	ACC GCA TGG T A W	
Frameshift mutations		
7 bp deletion	TGG ACC GCA W T A	CA
3 bp indel	TGG ACC W T	GAT GAC D D

Table 6.1. List of mutations identified in F1 progeny. The mutated nucleotides are labeled in red with codons indicated below using the one-letter amino acid code.

The four in-frame mutations maintain the open reading frame and would produce mutant transcripts encoding "full-length" proteins so do EOfAD-causative mutations following the "fAD reading frame preservation rule". These mutations are located at membrane transition sites (as the third lumenal loop of the Psen1 protein is very short), and thus should cause the changes in protein structure and hydrophobicity. Among these four in-frame mutations, the 6 bp indel mutation, *psen1*$^{W233I\_T234L}$, is the simplest structurally as it only changes two codons, one of which is equivalent to a codon in human *PSEN1* affected by an EOfAD mutation. Therefore, this *psen1*$^{W233I\_T234L}$ mutation has potential use as an additional EOfAD mutation-like model for the investigation of EOfAD-specific effects using zebrafish.

The two frame-shifting mutations would result in the production of mutant transcripts encoding truncated proteins. The mutant transcript of the indel mutation, *psen1*^{W233fs}, would have an earlier premature termination codon than the 7 bp deletion mutation and so would encode a smaller protein more similar to that from the transcripts of the human familial Acne Inversa mutant allele P242LfsX11. The data from sequencing of the *psen1*^{W233fs} mutation is shown in **Figure 6.9**.

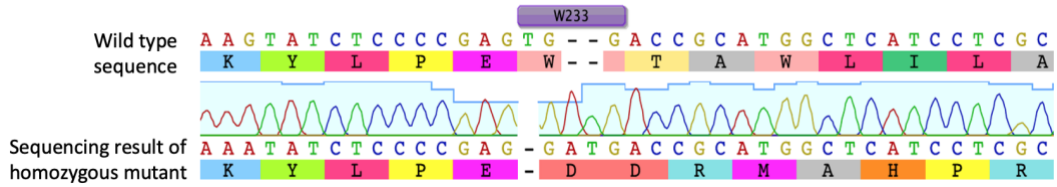


Figure 6.9. Sequence spanning the *psen1*^{W233fs} indel mutation. Two nucleotides TG in the wild type sequence are substituted with three nucleotides GAT, causing a frame-shift.

Future work will exploit the *psen1*^{W233fs} mutation as a model of P242LfsX11 in transcriptome comparisons with wild type and EOfAD-like mutations. As patients with the P242LfsX11 mutation do not suffer EOfAD, these transcriptome comparisons can identify molecular changes due to mutation of psen1 that are not by themselves, sufficient to drive EOfAD.

6.3 Discussion

6.3.1 The difficulties in generating HDR-induced mutations in zebrafish models

The zebrafish mutation models intended to be generated in this work were to be used to study human EOfAD. Ideally, they would be as similar as possible to the original mutations in human PSEN1. Therefore, I attempted to generate them through template-guided HDR. However, HDR can be difficult to achieve in zebrafish. Normally the frequency of HDR-induced mutations in zebrafish is far less than the frequency of non-homology end joining (NHEJ)-induced mutations. To increase the likelihood of achieving HDR to create a model of the familial Acne Inversa mutation, I utilised a strategy suggested by Zhang et al. in 2018 [16] in which the DNA template is carried within a plasmid vector and the NHEJ inhibitor SCR7 is added to the embryo culture medium. However, this gave little improvement in HDR efficiency as the desired mutation, although detectable by PCR in injected G0 embryos, could not be found in F1 progeny. Instead, all the subsequently isolated mutations proved to be products of NHEJ.

It remains quite difficult to generate mutations by HDR in zebrafish, and improvements to the efficiency of this process should be investigated. The generation of mutations by HDR is essential for the investigation of molecular changes due to particular genetic states, and greater success in HDR-induced mutagenesis in zebrafish would promote the application of this model in the study of human disease.

6.3.2 The generation of EOfAD transcriptome profiles using multiple EOfAD-like models

Due to the low success rate of generating mutations by HDR in zebrafish, it can be difficult to introduce exact equivalent human EOfAD-causative mutations into the corresponding zebrafish genes. Fortunately, for the EOfAD in the *PRESENILIN* genes, mutations generated through NHEJ which follow "fAD reading frame preservation rule" and cause slight changes in protein structure and hydrophilicity like the many human EOfAD-causative mutations, can be used as EOfAD-like models. The *psen1*$^{Q96\_K97del}$ mutation is an example. This mutation maintains the open reading frame but causes structural and hydrophilicity changes in the first lumenal loop of the translated protein similar to EOfAD mutations such as *PSEN1*$^{L113\_I114insT}$ [17] and *PSEN1*^{P117L} [18], which distort the first luminal loop of human PSEN1. However, not all the molecular changes induced by the *psen1*$^{Q96\_K97del}$ mutation are necessarily involved in driving development of AD pathology. For example, the effects of different EOfAD-causative mutations on γ-secretase activity appear to be variable [19, 20]. Within one EOfAD-like model, it is hard to define whether or not a particular transcriptome signature element is involved in AD pathology. Therefore, more EOfAD-like mutation models must be generated to identify transcriptome signature elements which are specifically universal to EOfAD. Here, I have generated a 6 bp mutation, *psen1*$^{W233I\_T234L}$, substituting two amino acids in the encoded protein. While this mutation is not precisely equivalent to any known human EOfAD mutation in PSEN1, it obeys the "fAD reading preservation rule" of EOfAD mutations in *PSEN1* and *PSEN2* (and

APP). The position of this mutation at the lumenal loop 3, transmembrane domain 6 transition site likely changes the structure and hydrophobicity of the encoded protein, similar to the human EOfAD-causative mutations, *PSEN1*^{T245P} [21], *PSEN1*^{A246E} [22] and *PSEN1*^{A246P} [23], near the *psen1*$^{W233I\_T234L}$-equivalent position in the human PSEN1 protein. Therefore, the *psen1*$^{W233I\_T234L}$ mutation has the potential to be used as another EOfAD-like model to aid in refining EOfAD transcriptome signatures.

6.3.3 Use of the Acne Inversa-like mutation model in refinement of the EOfAD transcriptome signature

Human familial Acne Inversa is an autosomal, dominantly inherited skin disease caused by a frameshift mutation in human *PSEN1*, P242LfsX11, which does not cause EOfAD [4]. This mutation leads to the production of mutant transcripts encoding truncated proteins and so does not obey the "fAD reading frame preservation rule". Analysis of this mutation, or mutations like it, can, by exclusion, help us to refine our understanding of which elements of the EOfAD signature are critical for EOfAD. (Of course, it is possible that two elements operating in combination are critical for EOfAD so we cannot automatically exclude from the signature any element which is present in both the familial Acne Inversa and all EOfAD mutation-like brain transcriptomes.)

Although a human Acne Inversa-equivalent mutation in zebrafish was not achieved through HDR, two frameshift mutations were identified near the Acne Inversa

mutation-equivalent site. Both of these mutations are predicted to cause premature termination of the coding sequence. However, the 3 bp indel mutation, *psen1*^{W233fs}, has an earlier premature termination codon more similar to that seen in human Acne Inversa mutant transcripts. Therefore, the *psen1*^{W233fs} mutation will be used as an Acne Inversa mutation-like model in further studies.

The *psen1*^{W233fs} mutation is currently being exploited in a "3-way" comparative transcriptome analysis performed by Karissa Barthelson. A zebrafish heterozygous for *psen1*^{W233fs} was crossed with a zebrafish heterozygous for the EOfAD-like mutation *psen1*^{T428del} (that is an exact equivalent of the human EOfAD mutation *PSEN1*^{T440del} [24]) to produce a family/tank of siblings with either wild type, or heterozygous, or transheterozygous genotypes (**Figure 6.10**). Raising zebrafish siblings in the same tank reduces environmental noise in the subsequent transcriptome data obtained and a comparative brain transcriptome analysis will be performed on the wild type, heterozygous *psen1*^{W233fs} mutant, and heterozygous *psen1*^{T428del} mutant genotypes. Hopefully, this analysis will reveal cellular functions affected uniquely by the EOfAD mutation-equivalent *psen1*^{T428del} allele and not the familial Acne Inversa mutation-like *psen1*^{W233fs} allele. The brain transcriptome of the heterozygous *psen1*^{T428del} mutants can also be compared to that already defined for the EOfAD mutation-like heterozygous *psen1*$^{Q96\_K97del}$ mutants to identify transcriptome changes in common. Also of interest, the transcriptome effects of the Acne Inversa mutation-like *psen1*^{W233fs} allele will be compared with the known brain transcriptome changes caused by heterozygosity for the *psen*^{K97fs} frameshift allele. It is curious that only one familial Acne Inversa mutation has been identified

so far in human *PSEN1*, raising the possibility that the position of the frameshift in the coding sequence is critical to the disease.

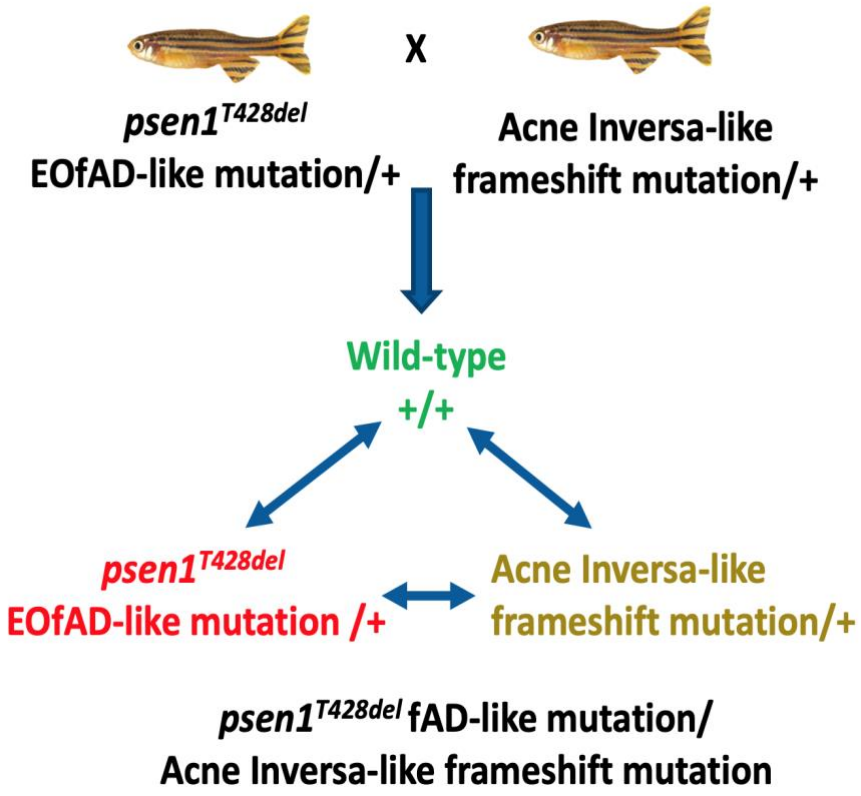


Figure 6.10. The mating strategy of a "3-way" comparative transcriptome analysis. The family of siblings consists of 4 genotypes but only wild type, heterozygous *psen1*^{W233fs} mutants and heterozygous *psen1*^{T428del} mutants are used in transcriptome analysis since the transheterozygote has no known relevance to human disease genetics.

6.4 Methods

Micro-injection

The template DNA oligonucleotide to generate an Aspartate codon-substitution mutation by HDR was synthesized by Sigma-Aldrich (Missouri, USA), while the template DNA plasmid to generate the familial Acne Inversa-like mutation was synthesized by Biomatik (Cambridge, Ontario, Canada). The designed sgRNA (CRISPR RNA, crRNA), trans-activating crRNA (tracrRNA) and Cas proteins were purchased from Integrated DNA Technologies (IDT, Coralville, USA).

The injection mixture for attempted induction of the Aspartate codon-substitution mutation consisted of 21 μM crRNA, 21 μM Cas12a protein and 1.6 μM oligo DNA template (no tracrRNA was required), while the injection mixture for attempted induction of the Acne Inversa-like mutation contained 20 μM crRNA, 20 μM tracrRNA, 12.8 μM Cas9 protein and 107 ng plasmid DNA template. The injection mixtures were incubated at 37°C for 10 minutes before injection. 5-10 nL of injection mixture was injected into zebrafish embryos at the one cell stage. Embryos injected with the CRISPR/Cas12a system were incubated at 34°C, while embryos injected with the CRISPR/Cas9 system were incubated at 28°C. Ten embryos from each injected clutch were collected after 24 hours to test whether the CRISPR/Cas system had worked, and the remaining embryos were raised to 6 months of age, before outcrossing with wild type fish to produce F1 progeny.

Zebrafish breeding

The injected G0 embryos were raised to reach sexual maturity (approximately 6 months of age). Each G0 zebrafish with mutations detected in a tail biopsy (using either the T7 endonuclease I assay or specific PCRs) was outcrossed with a wild type zebrafish to produce F1 progeny.

Mutations in F1 progeny are normally stable and heritable. The F1 progeny were then raised to 6 months of age for mutation detection and sequencing. Any F1 individual potentially bearing the mutation of interest was outcrossed with a wild type fish to produce F2 progeny. The F2 individuals heterozygous for the mutation were used for further experiments (such as "3-way" transcriptome analysis), and were also incrossed to generate homozygous individuals that were then used to establish a line of fish for propagation of the mutation.

Genomic DNA extraction

Zebrafish were anesthetized using tricaine solution (168µg/mL), and a small tail biopsy (a "tail clip") was then taken from the zebrafish using a surgical blade.

50µL of 1.7mg/µL Proteinase K (Sigma-Aldrich, Missouri, USA) in 1xTE buffer was used to degrade proteins and release genomic DNA from pools of ten zebrafish embryos or from tail clips of individual adult fish after incubation at 55°C for 3 hours, followed by enzyme inactivation at 95°C for 15 minutes. This genomic DNA was then stored at -20°C for the following genotyping PCRs.

PCR for genotype determination

PCRs were performed in 25 μL reactions containing genomic DNA, 0.2 mM of each deoxyribonucleotide triphosphate (dNTP, each at a final concentration of 0.2 mM), 0.4 μM of each PCR primer, and 2.5 unit of GoTaq® DNA polymerase (Promega, Madison, Wisconsin, USA). PCR cycling was performed with 35 cycles of a denaturation temperature of 95°C for 30 s, then an appropriate (primer-dependent) annealing temperature for 30 s and then an extension temperature of 72°C for 45 s. PCR products were electrophoresed through a 1.5% agarose gel in 1×TAE buffer for separation and identification.

T7 endonuclease I assay

The T7 endonuclease I assay [25] was used to detect mutant DNA strands among wild type PCR products. T7 endonuclease I recognises sites of base mismatches in double-stranded DNA and cleaves at these sites. Samples with the expected cleavage sizes in T7 endonuclease I assays can be regarded as having mutations at the targeted site, although irrelevant cleavage bands may occur e.g. due to the existence of sequence polymorphisms. PCR products found to include mutation by the T7 endonuclease I assay were subsequently sent to Australian Genome Research Facility (AGRF, Melbourne, VIC, Australia) for sequencing.

6.5 References

1. Masters, C.L., et al., *Alzheimer's disease.* Nature Reviews Disease Primers, 2015. **1**.
2. Lanoiselee, H.M., et al., *APP, PSEN1, and PSEN2 mutations in early-onset Alzheimer disease: A genetic screening study of familial and sporadic cases.* PLoS Med, 2017. **14**(3): p. e1002270.
3. Jayne, T., et al., *Evidence For and Against a Pathogenic Role of Reduced gamma-Secretase Activity in Familial Alzheimer's Disease.* J Alzheimers Dis, 2016. **52**(3): p. 781-99.
4. Newman, M., 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): p. 602-17.
5. Newman, M., et al., *Animal Models of Alzheimer's Disease*, in *Animal Models for the Study of Human Disease* C. P.M., Editor. 2017, Elsevier Science & Technology: Amsterdam, Netherlands. p. 1031-1085.
6. Jiang, H., M. Newman, and M. Lardelli, *The zebrafish orthologue of familial Alzheimer's disease gene PRESENILIN 2 is required for normal adult melanotic skin pigmentation.* PLoS One, 2018. **13**(10): p. e0206155.
7. Newman, M., et al., *Accelerated loss of hypoxia response and biased allele expression in zebrafish with Alzheimer's disease-like mutations.* bioRxiv, 2019: p. 526277.
8. Hin, N., et al., *Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2.* PLoS One, 2020. **15**(1): p. e0227258.
9. Newman, M., et al., *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): p. 43.
10. Dong, Y., et al., *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: p. 2020.05.03.075424.

11. Barthelson, K., et al., *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.
12. Newman, M., et al., *Accelerated loss of hypoxia response in zebrafish with familial Alzheimer's disease-like mutation of Presenilin 1*. bioRxiv, 2020.
13. Wolfe, M.S., et al., *Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity.* Nature, 1999. **398**(6727): p. 513-7.
14. Jacobi, S.A., et al., *Efficient homology-directed repair using single-stranded DNA templates*.
15. Boel, A., et al., *CRISPR/Cas9-mediated homology-directed repair by ssODNs in zebrafish induces complex mutational patterns resulting from genomic integration of repair-template fragments*. Disease Models & Mechanisms, 2018. **11**(10).
16. Zhang, Y.B., Z.W. Zhang, and W. Ge, *An efficient platform for generating somatic point mutations with germline transmission in the zebrafish by CRISPR/Cas9-mediated gene editing.* Journal of Biological Chemistry, 2018. **293**(17): p. 6611-6622.
17. De Jonghe, C., et al., *Aberrant splicing in the presenilin-1 intron 4 mutation causes presenile Alzheimer's disease by increased Abeta42 secretion.* Hum Mol Genet, 1999. **8**(8): p. 1529-40.
18. Wisniewski, T., et al., *A novel Polish presenilin-1 mutation (P117L) is associated with familial Alzheimer's disease and leads to death as early as the age of 28 years.* Neuroreport, 1998. **9**(2): p. 217-21.
19. Bai, X.C., et al., *An atomic structure of human gamma-secretase.* Nature, 2015. **525**(7568): p. 212-7.
20. Sun, L., et al., *Analysis of 138 pathogenic mutations in presenilin-1 on the in vitro production of Abeta42 and Abeta40 peptides by gamma-secretase.* Proc Natl Acad Sci U S A, 2017. **114**(4): p. E476-E485.

21. Edwards-Lee, T., et al., *A presenilin-1 mutation (T245P) in transmembrane domain 6 causes early onset Alzheimer's disease.* Neurosci Lett, 2006. **398**(3): p. 251-2.
22. Sherrington, R., et al., *Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease.* Nature, 1995. **375**(6534): p. 754-60.
23. Roeber, S., et al., *Three novel presenilin 1 mutations marking the wide spectrum of age at onset and clinical patterns in familial Alzheimer's disease.* J Neural Transm (Vienna), 2015. **122**(12): p. 1715-9.
24. Ishikawa, A., et al., *A mutant PSEN1 causes dementia with Lewy bodies and variant Alzheimer's disease.* Ann Neurol, 2005. **57**(3): p. 429-34.
25. Babon, J.J., M. McKenzie, and R.G.H. Cotton, *The use of resolvases T4 endonuclease VII and T7 endonuclease I in mutation detection.* Molecular Biotechnology, 2003. **23**(1): p. 73-81.

6.6 Appendix

6.6.1 SgRNAs and DNA templates for CRISPR/Cas systems

Aspartate codon-substitution mutation: CRISPR/Cas12a system

SgRNA:

GGAGATTTCATCTTTTACAG

Template DNA oligo:

TAAGTGAATGTGTGTGTGTGTTTTCAGGGGGTGTGAAGCTGGGTTTGG

GAGCCTTCATCTTTTACAGTATGCTGGTGGGTAAAGCATCAGCTACAG

CCAGCGG

Acne Inversa-like mutation: CRISPR/Cas9 system

SgRNA:

GATCAGCGCTCTCATGGCTC

Template DNA plasmid:

GATGAGCCATGCGGTCCACTCGGTATACATTATTAAAAAAAGCACACA
ATACTGTCTGTTTCTTTTGCATAGACCTGAATAATGAATACACTTTTAC
AGTGCTGTAAAATACAAGCCTACACACAACCTTTTTTTTTTTGCTATTT
TGCTGAAAACTAAAAATTTTTAGGGCAAGTTTGACATTTGCATGGAAT
TGCTCAGTATTAACGTGATTTCCTGCCTCTTAACGATGTTTCTGTTCCC
TCCAGGGAAGTGTTCAAGACGTATAACGTGGCGATGGATTACTTCACG
CTGGCGTTGATCATCTGGAACTTCGGTGTGGTGGGAATGATCTGCATC
CACTGGAAGGGGCCGCTGCGGCTCCAGCAGGCCTATCTGATCATGATC
AGCGCTCTCATGGCTCTGGTCTTCATCAAGTATCTCCTGAATGGACTGC
GTGGCTCATCCTGGCTGTGATTTCAGTCTACGGTCAGTCAGTCTGCAA
ACAGAGCAACACATTCACTTCTGTGTACTGATAGGTTTATTGTTCTGTA
ATGCTGCTTTGGAACAATAATTAGCCTGAAATGAATTGAATAATTTCA
AAAACACAGCCCCTTAAATGTCTTTCCTGCTCCTGTGTAGATCTTCTGG
CAGTGTTGTGTCCGAAAGGCCCTCTGCGAATCCTTGTGGAAACAGCTC
AAGAGAGGAATGAGGCCATTTTCCCAGCGCTCATCTACTCCTGTAAGA
AAAACCCCTCAGCACTTCCATCTTCTCAGACTGTGAATGTCCATCTGTT
TAATTGTGTTATCGTCACGTTTCTATGAAATTTCTCCTTCGATTTGTCA
GCTACGATGGTGTGGCTCTTCAATATGGCGGACCGAGTGGACCGCATG
GCTCATC

SgRNA sequences are indicated in red, PAM sites are highlighted in blue, and mutation sites are highlighted in yellow.

6.6.2 Primer information

Target gene	Primer sequences	Annealing temperature (°C)	Description
Primer for aspartate codon-substitution mutation			
psen1	5'-TGACGGGACGTGAAAATAAA-3'	60	Aspartate site region
	5'- TCCACCTTTTCCAAATGACC-3'		
psen1	5'-TGACGGGACGTGAAAATAAA-3'	60	Common forward primer
	5'-CCAGCATACTGTAAAAGATGAAAT-3'		Wild type (aspartate) reverse primer
	5'-CCAGCATACTGTAAAAGATGAAGG-3'		Alanine-specific reverse primer
Primer for Acne Inversa-like mutation			
psen1	5'-CGTGGCGATGGATTACTTCA-3'	60	Acne Inversa region
	5'-CACCCTGGTTCTCCTGCTGA-3'		
psen1	5'-TATCTCCCCGAGTGGACCGC-3'	60	Wild type forward primer
	5'-AAGTATCTCCTGAATGGACTGCGT-3'		Acne Inversa-specific forward primer
	5'- CACCCTGGTTCTCCTGCTGA-3'		Common reverse primer

Discussion

Alzheimer's disease (AD) is a major health issues worldwide, and its incidence shows an increasing trend [1]. The study of AD pathology and the investigation of possible treatments have been ongoing for decades [2]. Multiple hypotheses have been proposed based on different features seen in disease progression. Each hypothesis has its limitations and cannot comprehensively explain the cause(s) and pathology of the disease. As the early-onset familial form AD (EOfAD) shows dominant inheritance [3], EOfAD-causative mutations can be introduced into animal models to explore the molecular basis of the disease *in vivo*. Rodents, as genetically manipulable mammals, are the most common animal models used in AD studies. However, the failure in clinical trials of AD therapies which had succeeded in preclinical trials using rodent models [4], and the inconsistency in brain transcriptome profiles between different transgenic mouse models of AD and between transgenic mouse models and human patients [5] has illuminated the limitations of rodent models in AD research, indicating that they do not reflect comprehensively the disease state of human patients. On the other hand, rodent models have a relatively short lifespan, and do not naturally develop NFTs in their brains [6], suggesting rodent models might not replicate the pathology seen in human AD patients. Their short lifespan is another limitation of rodent models, as it has been argued that humans' long post-reproductive life span and longevity might contribute to the onset of AD and its pathological signs, such as the development of NFTs in brains [6]. Therefore, it is desirable to use other animal models that may reveal different aspects of the molecular interactions and cellular

functions related to the human EOfAD genes, *PSEN1*, *PSEN2* and *APP*. The application of diverse animal models in AD studies may achieve a more thorough understanding of AD pathology in human patients. The Alzheimer's Disease Genetic Laboratory (ADGL) uses zebrafish to model the molecular effects of AD mutations. Although zebrafish models also have the same limitation of a relatively short lifespan as rodent models, the unique characteristics of EOfAD genes identified in zebrafish (reviewed in Newman et al., 2014 [7]) reveal different aspects of the functions of these genes. Also, zebrafish embryos are easily accessible for genetic manipulation, and zebrafish have the capability to produce large families of siblings. These advantages are ideal for experimental proposes.

Two EOfAD-like mutations have been generated in the zebrafish *psen1* gene, $psen1^{K97fs}$ and $psen1^{Q96\_K97del}$. Previous work in the ADGL investigated molecular changes induced by EOfAD-like mutations through brain transcriptome analysis. Brain transcriptome analysis of young adult (6 months of age) and aged adult (24 months of age) zebrafish heterozygous for the $psen1^{K97fs}$ mutation showed a transcription "inversion" with age (the differential expression of a number of genes compared to wild type was in opposite directions at two ages) [8]. This is reminiscent of the transcriptional inversion seen in humans in a postmortem transcriptome comparison of MCI, AD and age-matched normal brains [9]. However, the changes occurring in the regulation of genes during AD progression remain unknown. This means that we do not know whether this apparent gene expression inversion is due to a slow and irreversible accumulation of cellular damage or to a rapid shift into a pathological state (which might be reversible). Therefore, a month-by-month study was designed to monitor changes in

transcriptome state during aging from 6 months to 24 months, taking advantage of the zebrafish's ability to produce large families of siblings to provide sufficient experimental samples. However, the early qPCR results from this study failed to validate the changes in gene expression observed in the transcriptome analyses upon which it was based, so that DE genes identified in the transcriptome analyses could not be used to monitor age-dependent changes in brain transcriptome state. Although this inconsistency ultimately led to the study being abandoned, it showed us the huge variability in brain gene expression that exists even between the fish of the same genotype, and also that an appropriate sample size is critical for a transcriptome analysis to reveal changes in cellular function. A power calculation subsequently performed by Stephen Pederson using the strategy from Zhao et al., 2018 [10] found that n=6 could provide a power of ~70% for detection of fold-change > 2 at a false discovery rate (FDR) of 0.05 across the vast majority of expressed transcripts in zebrafish brain transcriptomes. (The previous brain transcriptome analyses from *psen1*^{K97fs}/+ and *psen1*$^{Q96\_K97del}$/+ fish relative to their wild type siblings had used n=3 and n=4 respectively.) The n=6 sample size was subsequently used in the ADGL's transcriptome analyses including my analysis of 7 day post fertilization (dpf) heterozygous *psen1*$^{Q96\_K97del}$ mutant larvae.

Our EOfAD-like mutation, *psen1*^{K97fs}, mimics the human *PSEN2*^{K115fs} mutation which was thought to be the only AD-causative frameshift mutation in the *PRESENILIN* genes. However, the identification of an alternative transcript splicing isoform from human *PSEN2* that restores the reading frame in *PSEN2*^{K115fs} transcripts [11] indicates that the *PSEN2*^{K115fs} mutation, like other EOfAD mutations, obeys the "fAD reading frame preservation rule". This challenged the

validity of the *psen1*^{K97fs} mutation as a realistic model of human *PSEN2*^{K115fs}. Subsequent attempts were made to detect such alternatively spliced transcripts in heterogynous *psen1*^{K97fs} mutant brains using RT-PCR to amplify regions spanning the mutation site and neighbouring exons, but no alternatively spliced isoforms were identified. Based on this, we redefined the *psen1*^{K97fs} mutation, and the brain transcriptome changes it causes, as non-EOfAD-like. However, this mutation can still prove valuable for modelling the function of the PS2V isoform of PSEN2 that shows increased expression in late onset AD brains.

Zebrafish larvae represent a powerful tool for high-throughput screening of chemicals and drug candidates. Wagner et al., 2015 [1] showed that the most effective drugs in an animal model of dyslipidemia were those able to revert disease transcriptome signatures back to normal. To translate this idea to our EOfAD-like model *psen1*$^{Q96\_K97del}$ to find AD-preventative drugs would require identification in larvae of a number of differentially expressed (DE) genes, which could be monitored to represent the presumably pathological transcriptome state [12]. Therefore, we performed transcriptome analysis on clutches of 7 dpf heterozygous mutant larvae. This led to the prediction of affected cellular pathways similar to those seen in mutant brains, but there was negligible concordance between the genes identified as DE in the larvae and in the mutant brains. Also, the huge variability in gene expression between clutches of larvae of the same genotype made it difficult to define a set of reproducibly DE genes to represent a pathological transcriptome state and that could be used in high-throughput chemical library screening. Morgan Newman subsequently performed behavioural testing of heterozygous *psen1*$^{Q96\_K97del}$ mutant larvae through assessment of their response to

simulation by repeating intervals of 3-min light and 1-min dark. However, she did not observe significant differences in activity patterns. Therefore, it does not appear viable to use larvae with EOfAD-like mutations for drug screening through assessment of either gene expression states or behaviour. This is understandable as EOfAD is an adult onset neurodegenerative disease of brain, and so transcriptome analysis of whole larvae might fail to observe changes due to insufficient age or the dilution of transcriptome changes in CNS with those occurring in non-neural tissues. Moreover, AD is an age-dependent disease, and so, at early ages, any pathological transcriptome or behavioural changes are expected to be subtle.

The transcriptome analysis of 7 dpf heterozygous mutant larvae found these unsuitable for chemical library screening but did identify several larva-specific effects of the *psen1*$^{Q96\_K97del}$ mutation, particularly an effect on regulation of the genes encoding the minichromosome maintenance (MCM) proteins. The eukaryotic MCM complex functions as a DNA helicase essential for DNA replication and cell division. An involvement of EOfAD mutant *PSEN1* in cell sensitivity to DNA damage has been reported in several studies [13, 14]. This observation supports a DNA replication stress hypothesis of AD by Yurov et al. [15] proposing that DNA replication stress due to incomplete DNA replication or improper repair results in DNA damage and genomic instability. However, the effects on DNA replication and cell cycle were only observed in the whole 7 dpf growing mutant larvae and not in transcriptome analyses of 6-month-old and 24-month-old mutant brains (where relatively very little cell replication occurs). Therefore, the relevance for the mature CNS of *PSEN1* mutation effects on DNA replication and the cell cycle require further investigation. Nevertheless, zebrafish

carrying the *psen1*$^{Q96\_K97del}$ mutation may be an advantageous model in which to investigate such effects further.

In addition to brain molecular changes, changes in behaviour are another essential focus when using animal models in the study of human neurological disease, as relevant animal models are expected to show similar symptoms to those occurring in humans. Progressive cognitive impairment is the major clinical symptom in human AD patients. Impairment of short-term spatial working memory have been observed in AD patients with moderate dementia [16] and the short-term spatial working memory of animals can be easily assessed using the free movement pattern (FMP) Y-maze. Therefore, the FMP Y-maze test was employed as our first attempt to assess changes in the behaviour of mutant zebrafish at different ages. However, no statistically significant differences were observed between heterozygous *psen1*$^{Q96\_K97del}$ mutant fish and their wild type siblings at any of three ages, 6 months, 12 months and 24 months. Not all zebrafish of either mutant or wild type genotype showed an obvious search strategy in their behaviour, and some fish showed extremely low activity during the tests. Therefore, our studies found either that the current FMP Y-maze test might not be ideal for assessment of short-term spatial working memory in zebrafish, or that it requires further optimization. Future experimentation should investigate other behavioural tests involving training, which can assess long-term memory and learning abilities and may achieve a better discrimination between cognitively impaired and normal individuals. For example, the learning abilities of zebrafish can be assessed using the method described by Aoki et al., 2017 [17]. Zebrafish adults were treated with a 2-hour training session in a Y-maze using a specific colour paired with electric shock at one arm, and the

fish's efficiency of choosing the correct arm after training was used to represent their learning ability. As electric shock is a strong stimulation, zebrafish subsequently do not move randomly as observed in FMP Y-maze tests, and this may allow better discrimination between cognitively impaired and normal individuals.

In addition to studies using our existing EOfAD-like mutants, the generation of other EOfAD-like or EOfAD-relevant mutations is required for the refinement of a brain transcriptome signature of EOfAD. We failed to generate the desired aspartate codon-substitution and Acne Inversa-like mutations using HDR due to the low efficiency of HDR in zebrafish. Nevertheless, two probably informative mutations were identified near the human Acne Inversa mutation-equivalent site. The 6 bp indel in-frame mutation, *psen1*$^{W233I\_T234L}$, can represent another EOfAD-like mutation and should help to refine the transcriptome signature that is specifically common to EOfAD mutations. Another 3 bp indel mutation, *psen1*^{W233fs}, causes a frameshift close to the site equivalent to the human Acne Inversa mutation and can be used, by exclusion, to refine our understanding of which transcriptome elements are critical for EOfAD. This frameshift mutation is currently being exploited by Karissa Barthelson in a "3-way" comparative transcriptome analysis expected to reveal brain cellular functions affected uniquely by the EOfAD mutation-equivalent *psen1*^{T428del} allele and not the familial Acne Inversa mutation-like *psen1*^{W233fs} allele. Also, the *psen1*^{W233fs} frameshift mutation can be used in comparison with our other frameshift mutations, such as *psen1*^{K97fs}, to investigate differential effects of truncating the *psen1* coding sequence at different positions. This is an important issue since previous research by the ADGL has demonstrated differential dominant

effects of forced expression of truncated forms of Psen1 protein including truncations similar to those produced by the human *PSEN2*^{K115Efs} [18] and *PSEN1*^{P242Lfs} [19] alleles.

Overall, the research presented in this thesis has focused on identifying the effects of EOfAD-like mutations in the zebrafish model. The studies span analysis of both transcriptome and behaviour for a heterozygous EOfAD-like mutation model *psen1*$^{Q96\_K97del}$, as well as the generation of additional mutations in the *psen1* gene. The research has been valuable in illuminating the great variability in gene expression that occurs in individual bran transcriptomes and in tests of spatial working memory using the FMP Y-maze. Future work should investigate behavioural tests involving training to assess impairments in cognition. Also, transcriptome analysis of the EOfAD-like mutation *psen1*$^{W233I\_T234L}$, and the Acne Inversa-like mutation *psen1*^{W233fs}, may provide greater insight into the critical cellular changes driving EOfAD.

Reference

1. Gaugler, J., et al., *2019 Alzheimer's disease facts and figures.* Alzheimers & Dementia, 2019. **15**(3): p. 321-387.
2. Teri, L., E.B. Larson, and B.V. Reifler, *Behavioral disturbance in dementia of the Alzheimer's type.* J Am Geriatr Soc, 1988. **36**(1): p. 1-6.
3. Masters, C.L., et al., *Alzheimer's disease.* Nat Rev Dis Primers, 2015. **1**: p. 15056.
4. Drummond, E. and T. Wisniewski, *Alzheimer's disease: experimental models and reality.* Acta Neuropathol, 2017. **133**(2): p. 155-175.
5. Hargis, K.E. and E.M. Blalock, *Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us?* Behav Brain Res, 2017. **322**(Pt B): p. 311-328.
6. Gunn-Moore, D., et al., *Alzheimer's disease in humans and other animals: A consequence of postreproductive life span and longevity rather than aging.* Alzheimers & Dementia, 2018. **14**(2): p. 195-204.
7. Newman, M., E. Ebrahimie, and M. Lardelli, *Using the zebrafish model for Alzheimer's disease research.* Front Genet, 2014. **5**: p. 189.
8. Hin, N., et al., *Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2.* PLoS One, 2020. **15**(1): p. e0227258.
9. Berchtold, N.C., et al., *Brain gene expression patterns differentiate mild cognitive impairment from normal aged and Alzheimer's disease.* Neurobiol Aging, 2014. **35**(9): p. 1961-72.
10. Zhao, S., et al., *RnaSeqSampleSize: real data based sample size estimation for RNA sequencing.* BMC Bioinformatics, 2018. **19**(1): p. 191.
11. Braggin, J.E., et al., *Alternative splicing in a presenilin 2 variant associated with Alzheimer disease.* Ann Clin Transl Neurol, 2019. **6**(4): p. 762-777.
12. Love, D.R., et al., *Technology for high-throughput screens: the present and future using zebrafish.* Current Opinion in Biotechnology, 2004. **15**(6): p. 564-571.
13. Chan, S.L., et al., *Presenilin-1 mutations sensitize neurons to DNA damage-induced death by a mechanism involving perturbed calcium homeostasis*

and activation of calpains and caspase-12. Neurobiol Dis, 2002. **11**(1): p. 2-19.

14. Michelsen, K.A., et al., *Presenilin 1-related alterations in DNA integrity in a transgenic mouse model of Alzheimer's disease.* Brain Research, 2010. **1316**: p. 139-144.
15. Yurov, Y.B., S.G. Vorsanova, and I.Y. Iourov, *The DNA replication stress hypothesis of Alzheimer's disease.* ScientificWorldJournal, 2011. **11**: p. 2602-12.
16. Guariglia, C.C., *Spatial working memory in Alzheimer's disease: A study using the Corsi block-tapping test.* Dement Neuropsychol, 2007. **1**(4): p. 392-395.
17. Aoki, R., T. Tsuboi, and H. Okamoto, *Y-maze avoidance: an automated and rapid associative learning paradigm in zebrafish.* Neurosci Res, 2015. **91**: p. 69-72.
18. Moussavi Nik, S.H., et al., *Alzheimer's disease-related peptide PS2V plays ancient, conserved roles in suppression of the unfolded protein response under hypoxia and stimulation of gamma-secretase activity.* Hum Mol Genet, 2015. **24**(13): p. 3662-78.
19. Newman, M., 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): p. 602-17.