Transcriptome analysis of early-onset familial Alzheimer's disease and Sanfilippo syndrome zebrafish models reveals commonalities in disease mechanisms

Authors:
Karissa Barthelson[1,2*], Kim Hemsley[1] and Michael Lardelli[2]

Affiliations:
[1]Childhood Dementia Research Group. College of Medicine & Public Health. Flinders Health and Medical Research Institute. Flinders University. Sturt Road, Bedford Park SA 5042 Australia.

[2]Alzheimer's Disease Genetics Laboratory. School of Molecular and Biomedical Sciences, Faculty of Sciences, Engineering and Technology, The University of Adelaide. North Terrace Campus. Adelaide SA 5005 Australia

* Corresponding author: Karissa Barthelson: karissa.barthelson@flinders.edu.au

Key words: Alzheimer's disease, Sanfilippo syndrome, childhood dementia, RNA-seq, zebrafish, extracellular matrix, lysosome, ribosome, oxidative phosphorylation

Abstract

Background: Autosomal recessive inheritance of Sanfilippo syndrome (mucopolysaccharidosis type III, MPS III) causes childhood dementia, while Alzheimer's disease is the most common type of adult-onset dementia. There are no approved treatments for Sanfilippo patients, and few options exist for those with Alzheimer's disease. Increasing evidence suggests commonalities in the disease processes. However, a direct comparison of animal models with the two disorders has never been performed.

Method: We used RNA-seq to compare the transcriptome of zebrafish with early-onset familial Alzheimer's disease (EOfAD, *psen1* $^{Q96_K97del/+}$), or MPS IIIB (*naglu* $^{A603fs/A603fs}$) with their wild type siblings at 7 days post fertilisation and at 6 months of age (n = 8 fish per genotype).

Results: Differential gene expression and pathway analysis at each age revealed substantially more differentially expressed genes and pathways in MPS IIIB zebrafish relative to wild type than in the EOfAD-like zebrafish, consistent with MPS IIIB being a more severe, rapidly progressing and earlier onset form of dementia. Similar changes in gene expression were detected between the models in the extracellular matrix receptor interaction pathway in zebrafish larvae, and oxidative phosphorylation, ribosome and lysosome pathways in 6-month-old adult zebrafish brains. Cell type-specific changes in gene expression were detected for MPS IIIB zebrafish brains at 6 months of age, possibly reflecting altered proportions of oligodendrocytes, neural stem cells and inflammatory cells.

Conclusions: Our transcriptome analyses have illuminated possible shared disease mechanisms between EOfAD due to *PSEN1* mutations and MPS IIIB. Future work will investigate the nature of these commonalities.

Background

Alzheimer's disease (AD) is the most common form of dementia (1), and is an increasing economic and social burden as societies age. Since the earliest descriptions of AD in the early 1900's (2), we still do not have effective pharmaceutical treatments which can halt cognitive decline in the long term. This reflects continuing uncertainty regarding the fundamental cause(s) of AD pathology (3, 4).

The majority of AD cases show late-onset (LOAD, after 65 years of age), and occur sporadically with no evidence of familial inheritance. Less commonly, AD can occur in younger individuals and may be linked to dominant mutations i.e. early-onset familial AD (EOfAD). These dominant mutations occur in the genes presenilin 1 (*PSEN1*) (5), presenilin 2 (*PSEN2*) (6, 7), amyloid β precursor protein (*APP*) (8), or sortilin-related receptor 1 (*SORL1*) (9, 10).

It is clear that the endo-lysosomal system plays a role in the pathogenesis of AD. All of the EOfAD genes express proteins with involvement in the cell's endo-lysosomal system (11-14). Additionally, many of the genes associated with increased risk for the development of LOAD show endo-lysosomal system involvement (15). In humans, enlarged endosomes are one of the earliest pathologies observed in AD progression (16). However, it is difficult to dissect the mechanisms which underlie endo-lysosomal dysfunction in AD, particularly *in vivo*. Damage to the brain observed in post-mortem AD brain tissue is extensive (17). Many of the lesions seen may be secondary responses to pathological processes so that the information gained may not be directly relevant to therapeutic interventions. Pathological processes which occur early during disease development are likely the most

73 responsive to therapeutic intervention. However, we cannot easily access early diseased
74 human brain tissue for detailed molecular analyses.

75 Experiments performed in human cell lines can give some insight into the molecular
76 mechanisms underlying AD pathogenesis. However, the relevance of any such
77 observations must be tested *in vivo* (i.e. in animal models), since cultured cells exist in
78 highly artificial environments and some AD-relevant phenomena may be emergent
79 properties of complex multicellular systems such as the brain. A large number of different
80 AD animal models have been developed to investigate AD pathological mechanisms *in*
81 *vivo* (reviewed in (18, 19)), although the quality/relevance of these models is debated (20,
82 21). Much of this uncertainty arises from questions regarding the pathological mode of
83 action of the mutant genes themselves (discussed in (22)). An alternative approach to
84 dissecting the role of the endo-lysosomal system in neurodegeneration is to examine
85 extreme forms of endo-lysosomal system dysfunction, as represented by those lysosomal
86 storage disorders that result in childhood dementia.

87 Sanfilippo syndrome, which is typified by childhood-onset dementia, has an incidence of
88 approximately 0.17–2.35 per 100,000 live births (23) and results from one of four
89 recessively inherited lysosomal storage disorders (mucopolysaccharidosis type III, or MPS
90 IIIA, B, C, D; OMIM: 252900, 252920, 252930, 252940, respectively). Like AD, there are
91 no effective therapies are currently approved to treat children with MPS III. However, the
92 genetics of MPS III and the primary pathological mechanisms underlying the disorder are
93 far better understood. MPS III is caused by recessive mutations in genes encoding
94 lysosomal enzymes required for catabolism of the mucopolysaccharide heparan sulfate:
95 N-sulfoglucosamine sulfohydrolase (*SGSH*) (24), N-acetyl-alpha-glucosaminidase
96 (*NAGLU*) (25), heparan-alpha-glucosaminide N-acetyltransferase (*HGSNAT*) (26) and N-

97 acetylglucosamine-6-sulfatase (*GNS*) (27). Incomplete catabolism of heparan sulfate (HS)
98 leads to its accumulation in lysosomes. Mutations in each of these genes define subtypes
99 of Sanfilippo syndrome, i.e. mutations in *SGSH*, *NAGLU*, *HGSNAT*, & *GNS* define MPS
100 IIIA, IIIB, IIIC, & IIID respectively. (Note that arylsulfatase G deficiency causing MPS IIIE
101 has only been observed in mice (28)).

102 While AD and MPS III are distinct disorders, MPS III is often colloquially referred to as
103 "childhood Alzheimer's", due to the similarity to AD of symptoms experienced by children
104 living with MPS III. At the cellular level in the brain, there are commonalities of phenomena
105 such as oxidative stress (29, 30), inflammation (30-32), and, importantly, dysfunction of
106 the endo-lysosomal system (33). However, these two diseases have never been
107 compared directly within one experimental system.

108 Zebrafish genetic models have emerged as powerful tools for studying cellular/molecular
109 phenomena in neurodegenerative diseases, particularly due to their ~70% genetic
110 similarity to humans (34) and their high fecundity (reviewed in (35)). We have previously
111 generated several zebrafish knock-in models of mutations causing early-onset familial AD
112 (36-40) and MPS III (41) and has analysed their brain transcriptomes. We have noted that
113 all the models of both diseases exhibit differential expression of genes involved in
114 metabolic pathways (41, 42). The most sensitive comparisons of these disease models is
115 best done within single zebrafish families, where genetic and environmental variation is
116 reduced between experimental groups of interest.

117 Here, we use intra-family analyses to compare the gene expression profiles of a zebrafish
118 model of EOfAD; *psen1*^{Q96_K97del} (43), with that of *naglu*^{A603fs}, a model of MPS IIIB (41).
119 We performed our analyses at two ages: 7 days post fertilisation (dpf) which models very

early, pre-symptomatic stages for both diseases, and in young adult brains (6 months of age), which still models pre-symptomatic disease stages for EOfAD but is symptomatic for MPS IIIB. We identified distinct changes in gene expression in each model at both ages. However, important age-related commonalities are also observed in genes implicating the extracellular matrix, energy metabolism, lysosomes and ribosomes, indicating fundamental, underlying similarities in pathological mechanisms.

Methods

Zebrafish husbandry and breeding strategy

The generation of the *naglu*^{A603fs} and *psen1*^{Q96_K97del} lines of zebrafish used in this study was described in (41) and (43) respectively. All fish were maintained and raised in a recirculating water system with a constant temperature of approximately 26°C, salinity levels at approximately 300 ppm, and conductivity approximately 800 µS/cm. These parameters were measured using an ExStik EC500 meter (Extek Instruments, Nashua, New Hampshire, United States). The pH level was measured using pH strips (Sigma Aldrich, St. Louis, Missouri, United States) and was maintained between 6 and 7. All fish were raised under a 14:10 hour light:dark cycle. Fish were fed twice daily with a diet of NRD 5/8 dry flake food (Inve Aquaculture, Dendermonde, Belgium) in the mornings and live *Artemia salina* (Inve Aquaculture) in the afternoon.

To generate the zebrafish samples for RNA-seq, we crossed a *psen1*$^{Q96_K97del/+}$; *naglu*$^{A603fs/+}$ zebrafish with a *psen1*$^{+/+}$; *naglu*$^{A603fs/+}$ zebrafish, to generate a family of 100 sibling larvae with a variety of genotypes (**Fig.1**). These larvae were raised together in a 14 cm diameter plastic petri dish (VWR, Radnor, Pennsylvania, United States) in 150 mL

142 of E3 medium (44) (pH ~7) until 7 days post fertilisation (dpf). The E3 medium was
143 changed every two days.

144 At 7 dpf, all 100 sibling larvae were euthanised by cold shock, then individual larvae were
145 placed in 100 µL of RNA*later*™ Stabilization Solution (Invitrogen™, Waltham,
146 Massachusetts, United States). The larvae in RNA*later* were incubated overnight at 4°C.
147 Then, the tails of the RNA*later*-preserved larvae were removed for genotyping using
148 sterile needles to cut transversely at the level of the cloaca, see **Fig.1A**. The anterior ends
149 of the larvae were stored at -80°C until use.

150 For RNA-seq analysis of 6-month-old adult brains, we raised an additional family of
151 zebrafish spawned by a different parental pair (**Fig.1B**). All fish in the family were
152 euthanised by ice cold water shock, the sex of each fish was confirmed by dissection and
153 observation of testes or ovaries, entire heads were separated from bodies at the level of
154 the gills, and tails were removed for genotyping by PCR. Each entire head was
155 immediately placed in 600 µL of RNA*later* solution and incubated at 4°C overnight.
156 Subsequently, the RNA*later*-preserved brain was dissected from the skull, and then stored
157 in a fresh aliquot of 100 µL of RNA*later* at -80°C until analysis.

158 <u>PCR genotyping</u>

159 Genomic DNA (gDNA) was isolated from the tails of 7 dpf larvae and 6-month-old adults
160 by boiling in sodium hydroxide as described in (44). These gDNA preparations were then
161 used as templates for genotyping by polymerase chain reactions (PCRs). Our method of
162 genotyping involved PCR amplification across mutation sites using the primers described
163 in **Table 1** (purchased from Sigma Aldrich). Then, we subjected the amplified PCR
164 products to electrophoresis using gels consisting of 2-3% agarose (Sigma Aldrich) in 1 x

165 Tris-acetate-EDTA (1 x TAE) buffer for 2 hours at 90V, alongside DNA marker ladders
166 generated from the pHAPE plasmid (45). In this way, we could resolve two PCR products,
167 one each resulting from the wild type allele and/or mutant allele of each gene (see
168 **Fig.S1,2**).

Table 1: Genotyping primer details

Gene	Primer direction	Sequence 5' - 3'
naglu	Forward	AGAGATATTCACAGACTGTTGCC
naglu	Reverse	TTGCTGGAAAGGATGCGG
psen1	Forward	ATGAGGTGGAGCTGAACGG
psen1	Reverse	GGGGTTGTATGGAGTGTGAGG

169

170 RNA extraction and sequencing.

171 Total RNA was isolated from the individual larval zebrafish and 6-month-old adult brain
172 samples using the *mir*Vana™ miRNA Isolation Kit (Invitrogen) following the
173 manufacturer's instructions. Only fish with genotypes *psen1* $^{+/+}$; *naglu* $^{+/+}$ (wild type),
174 *psen1* $^{Q96_K97del/+}$; *naglu* $^{+/+}$ (EOfAD-like) and *psen1* $^{+/+}$; *naglu* $^{A603fs/A603fs}$ (MPS IIIB) were
175 analysed as these represent the genetic state of the respective human diseases. For
176 analysis of adult brain, at least 3 fish of each sex per genotype were analysed. To remove
177 any genomic DNA, we treated the total RNA isolates with DNase I using the *DNA-free*™
178 DNA Removal Kit (Invitrogen) following the manufacturer's instructions for routine DNase
179 digestion. 400 ng of total RNA from each of n = 8 larvae or adult brains per genotype were
180 then delivered to the South Australian Genomics Center (SAGC, Adelaide, South
181 Australia, Australia) for stranded polyA+ library preparation using Nugen Universal Plus
182 mRNA-seq (Tecan, Männedorf, Switzerland), protocol M01442 v2. The libraries were
183 further processed using the MGIEasy Universal Library Conversion Kit, Part No.
184 MGI1000004155 (MGI, Shenzhen, China), then sequenced (2 x 98 paired-end with 8 bp

185 unique molecular identifiers) using the MGI DNBSEQ-G400 platform across four lanes
186 which were subsequently merged.

187 RNA-seq data pre-processing

188 We pre-processed the raw fastq files using a custom pipeline implemented in *snakemake*
189 (46). For details of this pipeline, see **Fig.S3** and https://github.com/karissa-
190 b/2021_MPSIIIBvQ96-RNAseq-7dpfLarve. Briefly, *fastp* (47) was used to trim adaptors
191 and filter the reads by quality and length. Then, the reads were aligned to the zebrafish
192 genome (GRCz11, Ensembl release 104) using *STAR* (48) to generate sorted .bam files.
193 These were subsequently indexed using *samtools* (49). After alignment, PCR duplicates
194 were de-duplicated using the *dedup* function of *umi-tools* (50). The number of reads
195 aligning to gene models of the GRCz11 genome were counted using *featureCounts* (51).
196 *FastQC* (52) and *ngsReports* (53) were used to assess the quality of the raw and trimmed
197 data, as well as the original and deduplicated alignments.

198 Statistical analysis of RNA-seq data

199 Statistical analysis of the RNA-seq data was performed using R (54). We first omitted
200 lowly expressed genes, which are statistically uninformative for differential gene
201 expression analysis. We defined a gene as lowly expressed if it contained a log counts per
202 million (logCPM) of less than 0.5. To account for genes expressed below this limit of
203 detection in some experimental groups, but not in the others, we retained the genes which
204 fell below this threshold if they contained sufficient logCPM values in at least 8 out of the
205 24 total RNA-seq samples (i.e. one of the mutant genotype groups) in each experiment for
206 each age.

207 Differential gene expression analysis was performed with *edgeR* (55), specifying a design
208 matrix with the wild type genotype as the intercept, and each of the mutant genotypes as
209 coefficients. At 7 dpf, zebrafish are yet to become sexually dimorphic and we (36, 38, 39)
210 and others (56, 57) have not detected a substantial effect of sex on zebrafish brain
211 transcriptomes, including in this dataset (**Fig.S4**). Therefore, we did not include the effect
212 of sex in the design matrix. We considered a gene to be differentially expressed (DE) if the
213 p-value after adjusting for the false discovery rate (FDR) was less than 0.05.

214 We observed biases for differential expression of genes due to their guanine-cytosine
215 content (%GC) and their length (**Fig.S5**). To account for this bias, we used conditional
216 quantile normalisation (*cqn*, (58)). The offset term generated by *cqn* was then included in
217 an additional differential gene expression analysis with *edgeR* to give the final list of DE
218 genes.

219 Over-representation of gene ontology (GO) terms (59, 60) within the DE genes was
220 identified using *goseq* (61), setting a %GC content measure as a co-variate for each gene.
221 This measure was calculated as the average %GC content per transcript, weighted by the
222 transcript length, to account for the fact that different alternative transcripts per gene will
223 have different %GC contents. We considered a GO term to be significantly over-
224 represented in the DE genes if the FDR-adjusted p-value from *goseq* was < 0.05.

225 Due to the highly overlapping nature of GO terms, and the limitation of the hard threshold
226 of considering a gene to be DE, we also performed gene set enrichment analysis of the
227 KEGG (62) pathway gene sets using an ensemble approach. We calculated the statistical
228 significance of the changes to gene expression in the KEGG pathways using three gene
229 set enrichment tools, each having different statistical methodologies: *fry* (63), *camera* (64)

230 and *GSEA* (65, 66) as described previously (38). Then, the individual p-values were
231 combined using the harmonic mean p-value (HMP) method, which has been shown
232 specifically to be robust to dependent tests (67). We considered a gene set to be
233 significantly altered if the FDR-adjusted HMP was < 0.05.

234 To assess the statistical power achieved in these experiments, we used the *ssize_vary*()
235 function of *ssizeRNA* (68) using default parameters. Briefly, we defined the number of
236 genes considered detectable, the average read counts in the wild type group, and the
237 dispersion estimate from each experiment. The proportion of non-DE genes was set to
238 0.9, the pseudo-sample size to 30, the FDR level to 0.05, and the range of absolute logFC
239 values from 0-2.

240 Results

241 What are the molecular pathologies shared by MPS IIIB and *psen1* EOfAD-
242 like zebrafish larvae?

243 We sought to compare directly the changes in global gene expression occurring due to
244 heterozygosity for an EOfAD-like mutation in *psen1* (*psen1*^{Q96_K97del}) or homozygosity for
245 loss-of-function mutations in *naglu* (*naglu*^{A603fs} , i.e. causing MPS IIIB (41)) in 7 dpf
246 zebrafish larvae using RNA-seq. First, to assess the overall similarity between samples,
247 we used principal component analysis (PCA). No distinct clustering of samples by
248 genotype was observed (**Fig. 2A**), supporting that these mutant genotypes do not result in
249 extensive changes to larval zebrafish bulk transcriptomes.

250 We next asked whether any genes were differentially expressed (DE) in both MPS IIIB
251 and EOfAD-like zebrafish larvae, which may support some shared pathological changes.
252 However, we did not identify any genes with an FDR-adjusted p-value < 0.05 in the

253 EOfAD-like larvae (meaning that, at the single gene level, it was not possible to identify
254 any similarity in the changes to gene expression in EOfAD-like and MPS IIIB zebrafish
255 larvae). In contrast, 57 DE genes were identified for the MPS IIIB mutant genotype
256 (**Fig.2B,C, Supplemental Table 1**) consistent with the earlier onset and greater disease
257 severity of MPS III relative to EOfAD. The DE genes in the MPS IIIB larvae were
258 significantly over-represented with 8 GO terms (**Fig.2D, Supplemental Table 2**) and the
259 DE genes within these significant GO terms were typically linked to multiple GO terms.
260 Consequently, the enrichment of DE genes within these GO terms is influenced to some
261 extent by the same set of genes (**Fig.2E**).

262 To detect any subtle changes in gene expression, we then took an ensemble approach to
263 gene set testing, in which the input does not rely on pre-defining DE genes. We applied
264 this method to the gene expression data from both mutants, using the KEGG (62) gene
265 sets. In the MPS IIIB larvae we detected significant changes to gene expression in 13
266 KEGG gene sets (**Fig.3A, Supplemental Table 3**). Not surprisingly, the gene sets for
267 *lysosome*, *glycosaminoglycan degradation* and *other glycan degradation* were found to be
268 significantly altered, consistent with the primary disease mechanisms in MPS III (reviewed
269 in (69)) (**Fig.3A** and **Fig.S6)**. Only 1 KEGG gene set was significantly altered in the
270 EOfAD-like mutant larvae; *ECM receptor interaction*. This gene set is also significantly
271 altered in the MPS IIIB larvae. A positive Pearson correlation was found for the logFC of
272 the genes in this gene set between the EOfAD-like and MPS IIIB mutants (**Fig.3B**), and
273 the genes in the pathway tended to be downregulated (**Fig.S7**). These observations
274 suggest that changes to ECM function could be a common, early disease mechanism
275 between EOfAD and MPS III.

276 We did not detect any differentially expressed genes in the n = 8 individual EOfAD-like
277 mutants (**Fig.2B,C**), while our previous analysis of this EOfAD-like mutation that was
278 performed on n = 6 pools each of 40 larvae at 7 dpf had identified 228 DE genes (70).
279 Therefore, we hypothesised that our experiment was statistically underpowered for the
280 use of individual larvae. Subsequently, we performed a post-hoc power calculation using
281 ssizeRNA (68). Indeed, based on the variation observed within the wild type genotypes,
282 our experimental structure had achieved an estimated 51% power to detect DE genes. A
283 sample of size of > 20 individual larvae would be required to achieve 70% power (**Fig.S8**).
284 This is not feasible to achieve within a single zebrafish family as, in our experience, the
285 maximum size for a single clutch of zebrafish rarely exceeds 150 viable embryos.
286 Therefore, we decided to repeat the analysis using adult zebrafish brains, where we have
287 previously shown that n = 6 generally gives approximately 70% power to detect significant
288 changes in gene expression (38).

289 What do MPS IIIB and *psen1* EOfAD-like zebrafish have in common in young 290 adult brains?

291 PCA performed on brain transcriptome data from 6-month-old adult zebrafish revealed
292 some clustering of samples: the MPS IIIB samples formed a cluster distinct from the wild
293 type and EOfAD-like samples (**Fig.4A**). This suggests that the 6-month-old MPS IIIB brain
294 transcriptome does not represent a wild type state, and that significant changes are
295 apparent. This is consistent with the results of the differential gene expression analysis,
296 where 18 and 259 DE genes were identified in the EOfAD-like and MPS IIIB mutant brains
297 respectively (**Fig.4B, Supplemental Table 1**). The DE genes in the MPS IIIB mutant
298 brains were significantly overrepresented with 53 GO terms and the DE genes in the
299 EOfAD-like mutant brains were significantly over-represented with 9 GO terms. The
300 statistical significance of the GO terms in the EOfAD-like mutants were exclusively driven

301 by three genes: *nenf, mt2* and *slc30a1a*. The significantly enriched GO terms in the MPS
302 IIIB larvae were driven by several shared DE genes (**Fig.S9**). Therefore, the biological
303 interpretation of the significance of these GO terms is challenging. Nevertheless, the 10
304 most significantly over-represented GO terms are shown in **Fig.4C** and the full outputs
305 from goseq can be found in **Supplemental Table 2**.

306 To overcome the complications of the highly overlapping and largely redundant nature of
307 many of the GO terms, and for consistency with our analysis of larvae, we tested the
308 KEGG gene sets using the ensemble approach. By this method, we detected 10 gene
309 sets significantly altered in the EOfAD-like mutant brains and 15 gene sets significantly
310 altered in the MPS IIIB mutant brains **(Fig.5A, Supplemental Table 3)**. The KEGG gene
311 set *Alzheimer's disease* was only significantly altered in the EOfAD-like zebrafish brains,
312 and the gene set for *glycosaminoglycan degradation* was only significantly detected in the
313 MPS IIIB zebrafish, supporting the specificity for each disease of these models.

314 Four KEGG gene sets were found to be affected in the brains of both EOfAD-like and
315 MPS IIIB zebrafish: *lysosome*, *oxidative phosphorylation*, *ribosome*, and *other glycan*
316 *degradation*. Overall, the genes in the *oxidative phosphorylation*, *ribosome* and *lysosome*
317 gene sets show broad upregulation (**Fig.S10-12**). However, a significant positive
318 correlation of the changes to gene expression in these gene sets between the mutants
319 was only observed in the *ribosome* gene set (**Fig.5C-E**). We note that the statistical
320 significance of the *lysosome* and *other glycan degradation* gene sets in EOfAD-like brains
321 appears to be driven partially by the relatively large logFC value of the *hexa* gene
322 (**Fig.5C,E**). Additionally, in the *lysosome* and *oxidative phosphorylation* gene sets,
323 concordance apparently exists in the direction of the logFC values of genes encoding
324 protein subunits of the vacuolar ATPase complex (v-ATPase, **Fig.S10,13**).

325 Are changes in cell type proportions influencing the observed changes to
326 gene expression?

327 In the analysis of bulk tissue using RNA-seq, differences in the proportions of cell types
328 between comparison groups may result in differences in gene transcript representation
329 between those groups that are not related to changes in cellular functions. To assess
330 whether this has occurred for the larval and adult brain tissues we analysed here, we
331 utilised single cell RNA-seq datasets from 5 dpf Tg(olig2:GFP)vu12 zebrafish larvae (71)
332 and from wild type (Tübingen) adult brain (72) to define gene sets consisting of marker
333 genes of the various cell types (we could not find a suitable dataset of 7 dpf zebrafish
334 larvae in the literature). We then tested whether these marker gene sets were significantly
335 enriched with up- or downregulated genes using *fry* (63) with a directional hypothesis. No
336 statistical evidence was found for changes to cell type gene sets in the EOfAD-like
337 mutants at either age (**Fig.6A,B**, **Supplemental Table 3**). However, significant changes to
338 gene expression in marker genes of neural cell types were identified in 6-month-old MPS
339 IIIB brains (**Fig.6B**). Notably, a broad downregulation of genes expressed in
340 oligodendrocytes (**Fig.6C,E**) and neural stem cells (**Fig.6D**), and a broad upregulation of
341 genes expressed in microglia (**Fig.6G**) and macrophages (**Fig.6F**) was observed.
342 Generally, only a small proportion of the leading edge genes (those genes driving the
343 statistical significance of the gene sets) were overlapping between these significantly
344 altered gene sets (**Fig.S13**), although 17 leading edge genes were shared between the
345 *microglia* and *macrophage_grn high* gene sets, representing 71% of all the microglial
346 leading edge genes. Therefore, the significance of the microglial gene set is mostly driven
347 by the same genes that drive the macrophage gene set. This gene expression signal
348 suggests that inflammatory processes are occurring in these brains. Collectively, these
349 results indicate the presence of cell type-specific dysfunction in the brains of 6-month-old

350 MPS IIIB zebrafish, potentially accompanied by alterations in the relative proportions of
351 these cell types.

352 No observed effect on the expression of genes responsive to intracellular iron 353 levels.

354 Our previous analysis of the EOfAD-like mutation *psen1*^{Q96_K97del} in 6-month-old female
355 zebrafish brains (73) revealed significant changes in gene expression in genes regulated
356 by intracellular iron levels. Specifically, the changes in expression occurred for genes
357 containing consensus or non-consensus iron-responsive elements (IREs) in the 3'
358 untranslated regions of their transcripts. However, we did not detect any statistical
359 evidence for changes in expression of genes possessing IREs in either mutant at either
360 age (**Fig. S14, Supplemental Table 3**). Therefore, these datasets do not support that
361 intracellular iron levels are abnormal.

362 Consistency with other changes in gene expression previously observed in 363 EOfAD-like zebrafish.

364 Our failure to observe changes in IRE gene sets in the 6-month-old EOfAD brains
365 described here, led us to examine the consistency between all our analyses of the
366 *psen1*^{Q96_K97del} mutation across independent experiments. We have previously analysed
367 the effect of this mutation in n = 6 pools each of 40 larvae at 7 dpf (70), as well as in
368 female 6-month-old zebrafish brains at n = 4 fish per genotype (43, 73). To assess the
369 overall similarity between samples across the experiments at each age, we performed
370 PCA on the logCPM values of genes considered detectable in each experiment. This
371 revealed that the samples separated based on their experiment of origin and not by
372 genotype. This supports that the overall effect of genotype is very subtle in these samples,
373 relative to the effect of experimental design (**Fig.7A,B**). We then applied our ensemble

374 approach to gene set testing to each of the previous datasets to identify whether similar
375 changes to the KEGG pathways could be identified. In 7 dpf larvae, this detected 85
376 significantly altered KEGG gene sets (**Fig.S15**), with the *ECM receptor interaction* gene
377 set having the highest level of statistical significance (the 10 most significantly altered
378 gene sets from the Dong et al. dataset (70) and their corresponding p-values from the
379 current dataset are shown in **Fig.7C**). In the 6-month-old female adult brain dataset
380 described in (43, 73), we detected 10 significantly altered KEGG gene sets, 6 of which
381 were significantly altered in this experiment (all gene sets found to be significantly altered
382 and their corresponding levels of statistical significance are shown in **Fig.7D**). Notably,
383 there was a lack of correlations of the logFC values of the genes within these gene sets
384 between the independent experiments (**Fig.7E-J**). Together, these results suggest that,
385 while extensive variation is apparent, particularly at the single gene level, when the same
386 mutation is analysed, common biological themes are still identified.

387 Discussion

388 In this study, we exploit the intra-family analysis strategy facilitated by zebrafish to
389 provide, for the first time, a sensitive and direct comparison of genetic models of two
390 inherited forms of early-onset dementia; EOfAD and MPS IIIB. We observed changes to
391 gene expression in the extracellular matrix genes very early in these disease models,
392 followed by changes to mitochondrial (oxidative phosphorylation), lysosomal and
393 ribosomal gene expression at later stages. We also identified changes to marker genes of
394 specific neural cell types in MPS IIIB zebrafish brains, indicating these cell types may be
395 particularly dysfunctional in this disease.

396 **Changes to ECM function could be an early, common disease mechanism in AD and**
397 **MPS III.**

398 The extracellular matrix (ECM) is a complex network of proteins and carbohydrates that
399 surrounds cells in multicellular organisms, providing mechanical support and playing a
400 critical role in cellular functions such as tissue organisation and maintenance of tissue
401 integrity. It consists primarily of proteins such as collagens and elastins, and glycoproteins
402 such as laminins and fibronectin (reviewed in (74, 75)). Glycosaminoglycans incorporating
403 heparan sulfate also form part of the ECM.

404 We observed subtle, yet statistically significant (**Fig.3C**), broad downregulation (**Fig.S7**) of
405 the ECM genes in both MPS IIIB and EOfAD-like zebrafish larvae at 7 dpf. In particular, the
406 collagen and fibronectin genes were downregulated to a similar extent in both disease
407 models (**Fig.S7**). Collagen and fibronectin proteins together with heparan sulfate
408 proteoglycans form part of the basal lamina of the extracellular matrix (76). Given that the
409 core pathology of MPS IIIB hinges on deficiency of NAGLU enzymatic activity, leading to
410 lysosomal accumulation of the heparan sulfate degradation product, the significant
411 alterations in the expression of ECM protein-encoding genes in MPS IIIB larvae is not
412 unexpected. The broad downregulation of these genes implies that the cells making up the
413 bulk larval tissue are detecting excess ECM components. This is consistent with the
414 observation of increased extracellular HS in a *NAGLU* deficient cell line (77). Increased
415 extracellular substrates have been observed in several other lysosomal storage disorders
416 (reviewed in (78)). Collagen-encoding genes were found to be upregulated in MPS IIIB
417 human fibroblasts relative to the other MPS III subtypes and MPS IV (79). However, the
418 collagen genes did not appear to be DE in comparison to a wild type control cell line (79).
419 This may be due to issues in the design of that transcriptome analysis such as low

statistical power (one cell line per MPS disease) and an inappropriate control cell line which may prevent detection of changes to ECM-related gene expression. Additionally, these observations made in human fibroblasts may not capture the processes which occur in neural cells. ECM genes were observed to be downregulated in iPSC-derived neurons generated from fibroblasts of human patients with MPS IIIB (80).

While the observed changes to ECM gene expression in the EOfAD-like larvae are subtle, we are confident in their reality due to the replication of the result in our previous, independent experiment in pools of larvae at the same age (70). Our observations in EOfAD-like zebrafish larvae are consistent with the transcriptome analyses of several iPSC-derived neuron lines with *PSEN1* mutations (81-83), as well as in directly induced neurons (iNs) from AD patients (84), which more closely retain their epigenetic aging profiles than iPSC-derived neurons. In humans, increased collagen and fibronectin levels have been observed in post-mortem AD brains (85). Changes to extracellular matrix proteins are strongly associated with neuropathological traits and cognition in human AD brains relative to controls (86) and were found to be the earliest detectable biomarker in EOfAD patient cerebrospinal fluid proteomes, occurring 30 years before symptom onset (87). However, the precise role of the ECM very early in AD is poorly understood and requires further investigation.

Common changes to mitochondrial, lysosomal and ribosomal function occur in young adulthood

Zebrafish are recently sexually mature at 6 months of age and we consider this equivalent to early adulthood in humans. For MPS IIIB, this is considered relatively late in disease progression as children typically succumb in their teenage years. This contrasts with

143 EOfAD, where changes in cognition are not expected until the fourth decade of life or
144 later. The dysregulation of genes associated with oxidative phosphorylation observed in
145 the young adult brains of zebrafish models of EOfAD and MPS IIIB (**Fig.5B** and **Fig.S10**)
146 suggests a potential shared molecular pathway underlying the changes which occur in
147 these seemingly disparate conditions.

148 Changes to the expression of the oxidative phosphorylation genes have been observed in
149 young, pre-symptomatic animal models of AD. We previously analysed the effects of
150 seven different AD-related mutations in 6-month-old zebrafish brains and found that these
151 all resulted in changes to oxidative phosphorylation gene expression (42). Mice
152 engineered to express a humanised version of the human ε4 allele of the apolipoprotein E
153 gene (*APOE4*, the strongest genetic risk factor for LOAD) also show changes to gene
154 expression in the oxidative phosphorylation pathway at young ages (42, 88). In humans,
155 changes to the oxidative phosphorylation genes are found early in AD, both in blood (89)
156 and in post-mortem brains (90). Additionally, brain glucose metabolism as measured by
157 fluoro-D-glucose positron emission tomography (FDG-PET) consistently shows changes
158 to glucose uptake during disease progression (reviewed in (91, 92)).

159 The role of mitochondrial dysfunction in MPS III is not extensively characterised. How
160 brain glucose metabolism changes with disease severity in children with MPS III is poorly
161 understood. Mitochondrial abnormalities and/or signs of oxidative stress (that can arise
162 from dysfunctional mitochondrial respiration) have been observed in the brains of mouse
163 models of MPS IIIA (93), IIIB (94), and IIIC (95, 96). Ultrastructural abnormalities (97) and
164 signs of oxidative stress have also been observed in MPS IIIB human brain tissue (30). A
165 mouse model of MPS VII (lysosomal accumulation of several mucopolysaccharides
166 including HS (98, 99)) showed a generally decreased abundance of the oxidative

467 phosphorylation proteins which were not detected by microarray in age matched samples
468 of the same strain (100). While the precise mechanism(s) of how and why mitochondria
469 are affected throughout MPS III disease course is unclear, these may be consequences of
470 defective autophagy/mitophagy (reviewed in (101)) caused by lysosome dysfunction.

471 Our analyses also detected significant changes to genes encoding lysosomal enzymes
472 and proteins (**Fig.5A, Fig.S12**) in both EOfAD-like and MPS IIIB zebrafish brains. While
473 the correlation of the logFC values of these genes was low (**Fig.5C**), they were broadly
474 upregulated in both models (**Fig.S12**). Lysosomes maintain an acidic lumen with pH
475 values generally within a range of 4.5 – 5 (102) via a mechanism involving v-ATPase
476 (reviewed in (103, 104)). The genes encoding the components of this complex are
477 upregulated in both EOfAD-like and MPS IIIB zebrafish brains (**Fig.S10** and **S12**),
478 suggesting that lysosomal acidification may be defective in these models. The
479 upregulation of these genes may reflect how the cells constituting these bulk brain RNA-
480 seq samples are attempting to restore the normal pH of defective lysosomes to maintain
481 cellular homeostasis. Lysosomal acidification defects have been detected in *PSEN1*-
482 deficient and primary *PSEN1*-EOfAD fibroblasts (105), as well as several other models of
483 AD mutations (106-109). Lysosomal pH was not detected as significantly altered in MPS
484 IIIA mouse embryonic fibroblasts (110). However, the existence of defective lysosomal
485 acidification requires exploration in other MPS III subtypes, and in other models.
486 Particularly as it has been observed in several other lysosomal storage disorders (110-
487 113) including MPS I (114).

488 Sufficient lysosomal acidification is required for the release of ferrous iron into the cytosol
489 for absorption by mitochondria (115). We did not observe significant changes to iron
490 homeostasis-related genes in either mutant at either age (**Fig.S14**), despite previous

491 detection of changes to iron transport gene expression in pools of EOfAD-like larvae at 7
492 dpf (70), changes in the *all 3' IRE* gene set (*ire-3all*) in female EOfAD-like brains at 6
493 months of age (73), and in all IRE gene sets in MPS IIIA zebrafish brains at 3 months of
494 age (41). This was unexpected, and the reason(s) for this discrepancy are not fully known
495 (discussed later). However, there is evidence supporting iron dyshomeostasis in mild
496 cognitive impairment (the precursor to AD) (116, 117), AD (118, 119), and in post-mortem
497 AD brains (120, 121). Iron homeostasis in MPS IIIB is far less understood. However,
498 increased iron levels are observed in the cortex of a MPS IIIB mouse model in its
499 symptomatic disease stages (122), as well as in a human MPS IIIB patient (123). We
500 suspect that changes to iron homeostasis genes in young adult zebrafish brains are very
501 subtle, and may be detected more consistently at later ages.

502 We observed broad, yet subtle, upregulation of ribosomal protein genes in EOfAD-like and
503 MPS IIIB zebrafish brains (**Fig.5D**). Upregulation of the ribosomal gene *RPLPA* was
504 observed in a MPS IIIB patient-derived fibroblast line (and downregulation of the
505 ribosomal genes *RPL10* and *RPL23* was observed in fibroblasts from other MPS III
506 subtypes (79)). However, the functionality of ribosomes in MPS III is poorly understood.

507 Early in AD, protein synthesis rates by ribosomes appear to be decreased (124), possibly
508 due to oxidation (by redox-active iron) of ribosomal proteins or RNA (125). The observed,
509 broad upregulation of the ribosomal gene set could reflect a homeostatic response to
510 defects in protein synthesis in both EOfAD and MPS IIIB. Alternatively, we have
511 speculated previously that it may be due to changes in mTOR signalling caused by
512 changes in lysosomal function (discussed in (42)).

513 **Cell type-specific changes to gene expression in MPS IIIB zebrafish brains**

514 In MPS IIIB adult brains, we observed statistical evidence for changes to gene expression
515 in five different cell types. The two most statistically significantly altered marker gene sets
516 were those of oligodendrocytes and neural stem cells. These showed broad
517 downregulation (**Fig.6C**) with minimal overlap in the genes driving the significance
518 (**Fig.S7**). Significant reduction of the expression of the oligodendrocyte gene set was
519 observed in the brains of all three of our zebrafish models at 3 months of age (41), as well
520 as in a mouse model of MPS VII at approximately 5 months of age (126). These gene
521 expression changes are consistent with the observed demyelination in the cerebral cortex
522 of MPS IIIB mice (127) and the corpus collosum of MPS IIIC mice (128) at 6 months of
523 age. Intriguingly, mucopolysaccharides have been shown to inhibit oligodendrocyte
524 maturation (129-131). Together, our results along with previous literature of animal models
525 of MPS diseases support that changes to oligodendrocyte differentiation and/or function
526 occur due to the excess mucopolysaccharides in symptomatic MPS disease stages. In
527 children with MPS III, loss of myelin (132, 133), and reduction of key oligodendrocyte
528 proteins (128) has been observed. However, changes to oligodendrocyte function and
529 myelination are not currently regarded as significant contributors to symptom
530 manifestation.

531 We did not observe any evidence supporting oligodendrocyte dysfunction in EOfAD-like
532 zebrafish at 7 dpf or 6 months of age. However, increasing evidence supports a role for
533 oligodendrocytes in this disease. A brain imaging study of EOfAD patients revealed that
534 white matter hyperintensities may be occurring as early as 22 years prior to symptom
535 onset (134), which likely reflects demyelination and damage to the axons of neurons
536 (135). Intriguingly, expression of *PSEN1* in the human brain is highest in oligodendrocytes

537 (136). It is possible that oligodendrocytes in EOfAD-like zebrafish are affected, just not
538 detectable at the transcriptome level at these ages. Oligodendrocytes in EOfAD-like
539 zebrafish should be investigated further.

540 **Consistent changes to gene expression can be detected at the pathway level across**
541 **independent experiments**

542 The changes to gene expression we observed in the EOfAD-like zebrafish in this
543 experiment were consistent with our previous analyses (42, 43, 70, 73). In each of those
544 datasets, we observed significant changes to gene expression in pathways associated
545 with the ECM in larvae, and in mitochondria, ribosomes and lysosomes in EOfAD-like
546 adult brain. However, comparisons made at the single gene level did not reveal
547 consistency (**Fig.7**). The differences at the single gene level may be attributed to technical
548 factors inherent to conducting experiments independently, such as differences in sample
549 preparation and sequencing platforms. Moreover, the inherent biological variability within
550 the zebrafish models analysed at different times may lead to distinct outcomes when
551 comparisons are not performed within a single experiment. This exemplifies the
552 importance of interpreting results at the pathway or gene set, rather than single gene level
553 when analysing transcriptome data.

554 **Limitations**

555 In the work described in this paper, we have employed zebrafish as a model organism due
556 to its unique advantages of being a commonly used vertebrate model species with high
557 fecundity, genetic tractability, and fairly high genetic similarity to humans (reviewed in
558 (137)). It is important to acknowledge that zebrafish may not fully recapitulate all aspects

559 of human neurodegenerative diseases (particularly for AD), and the relative simplicity and
560 small size of zebrafish brains may limit their equivalence to the human brain diseases we
561 are attempting to model. Additionally, zebrafish are exceptionally regenerative, including
562 the ability to regenerate brain tissue (138). Therefore, zebrafish are not an ideal model to
563 study the mechanisms of neurodegeneration specifically. Consequently, while our
564 zebrafish models offer valuable insights into disease mechanisms preceding
565 neurodegeneration, any findings should be interpreted with consideration of potential
566 species-specific differences. Additionally, our findings should be verified with protein
567 expression and/or functional assays.

568 We observed broad upregulation of microglial/macrophage marker genes and
569 downregulation of oligodendrocyte marker genes in MPS IIIB zebrafish brains. This could
570 be interpreted as indicating that the functions of these cell types are particularly disturbed
571 by the MPS IIIB genotype. Alternatively, it could imply that there are significant changes in
572 the proportions of these cell types constituting the MPS IIIB bulk brain transcriptome
573 samples. In future work, we will dissect the changes to gene expression in MPS IIIB brains
574 at the level of single cells to explore cell-type specific changes to gene expression, and
575 whether there are changes to proportions of these cell types.

576 Conclusion

577 Here, we provide the first direct comparison of zebrafish models of MPS IIIB and EOfAD,
578 which lead to childhood and adult-onset dementia respectively. While the upstream
579 genetic causes of these diseases are independent, our results suggest there may be
580 disease mechanisms in common, particularly with respect to lysosomal dysfunction. This

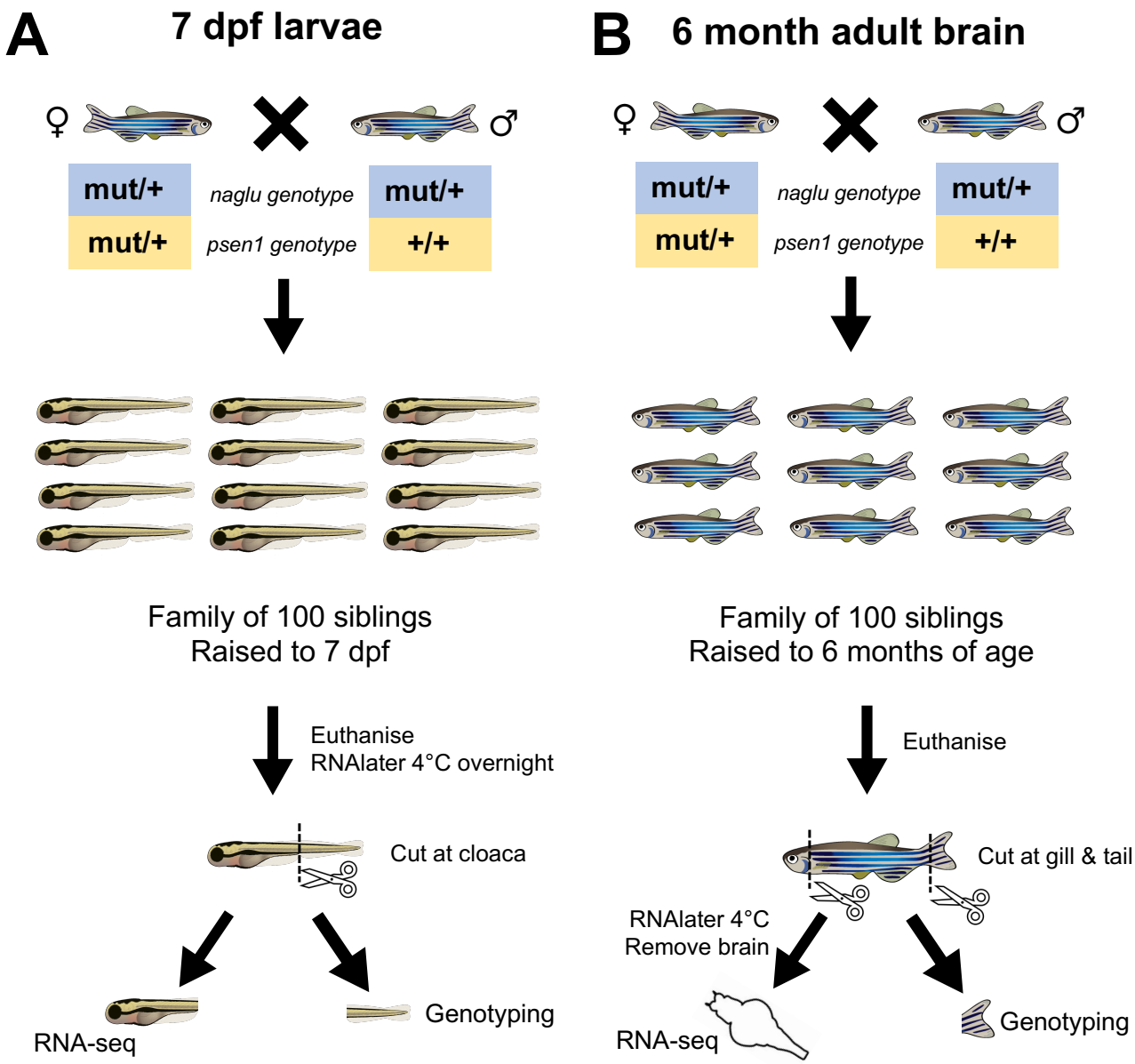
581 poses the question of whether common therapeutic strategies could be applied with

582 benefit for people living with either of these two devastating diseases.

583

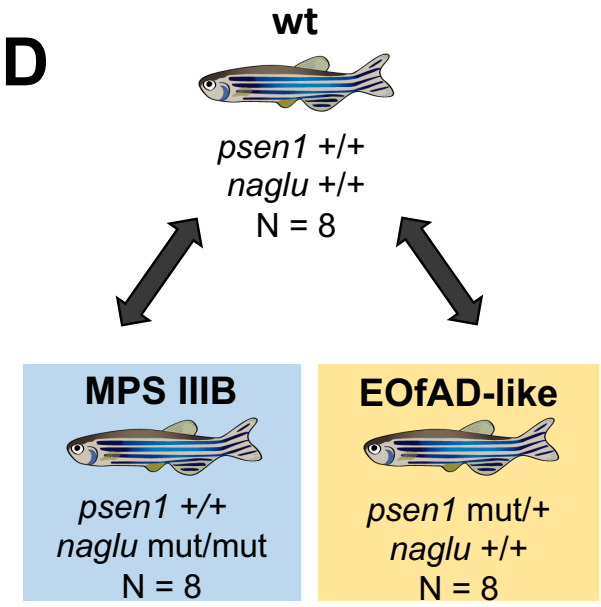
584 # Figures

585 **Fig.1: Intra-family transcriptome analysis strategy. A)** A female zebrafish heterozygous for
586 both the *naglu*^{A603fs} (MPS IIIB) and the *psen1*^{Q96_K97del} (EOfAD-like) mutations was mated with a
587 male zebrafish only heterozygous for *naglu*^{A603fs}. The subsequent family of ~100 sibling zebrafish
588 larvae was raised in a single petri dish at 28.5 °C. At 7 days post fertilisation (dpf), the entire family
589 was euthanised simultaneously and each larva placed individually in RNAlater solution. After 24
590 hours, each larva was bisected transversely at the cloaca. The tail end was used for genomic DNA
591 extraction and genotyping, while the remaining head and trunk was used for RNA-seq. **B)** A
592 different pair of zebrafish with the same *psen1* and *naglu* mutant genotypes was mated to
593 generate the adult brain analysis family. The entire family was raised until 6 months of age. Then
594 all fish in the family were euthanised, and their entire heads removed and immersed in 6 volumes
595 of RNAlater solution each, and stored at 4°C overnight. Brains were subsequently dissected out of
596 the RNAlater-preserved heads while their corresponding tails were used for genotyping by PCR.
597 **C)** Table showing expected genotype proportions in the MPS IIIB/EOfAD-like families. **D)**
598 Schematic showing the pairwise comparisons between MPS IIIB and EOfAD-like zebrafish to wild
599 type zebrafish in each RNA-seq experiment for each age.



C

Genotype		% of progeny
naglu	*psen1*	
+/+	+/+	12.5
mut/+	+/+	25
mut/mut	+/+	12.5
+/+	mut/+	12.5
mut/+	mut/+	25
mut/mut	mut/+	23.5



600

Fig.2: Differential gene expression analysis in zebrafish larvae at 7 days post fertilisation (dpf). A) Principal component analysis (PCA) on the $\log_{10}$ counts per million (logCPM) values after adjusting for %GC and length bias using cqn (58). Each point represents a larval zebrafish transcriptome sample which is coloured by genotype. The percentage variations that principal component 1 (PC1) and 2 (PC2) explain in the dataset are indicated in parentheses in the axis labels. **B)** Numbers of differentially expressed genes for each mutant genotype relative to wild type zebrafish larvae. **C)** Volcano plots showing the log fold change (logFC) and negative $\log_{10}$ of the p-value of the genes expressed in MPS IIIB and EOfAD-like larval zebrafish samples. Each point represents a gene which is coloured red if the differential expression FDR-adjusted p-value was less than 0.05. The x-axis labels are constrained to within -2 and 2 for visualisation purposes. **D)** The 10 most significantly over-represented Gene Ontology (GO) terms among the differentially expressed genes in MPS IIIB zebrafish larvae. **E)** Network visualisation indicating the overlap of DE genes within the significantly over-represented GO terms in MPS IIIB larvae. GO term nodes (black) are connected by unweighted edges to the nodes representing DE genes (blue) to which they are assigned.

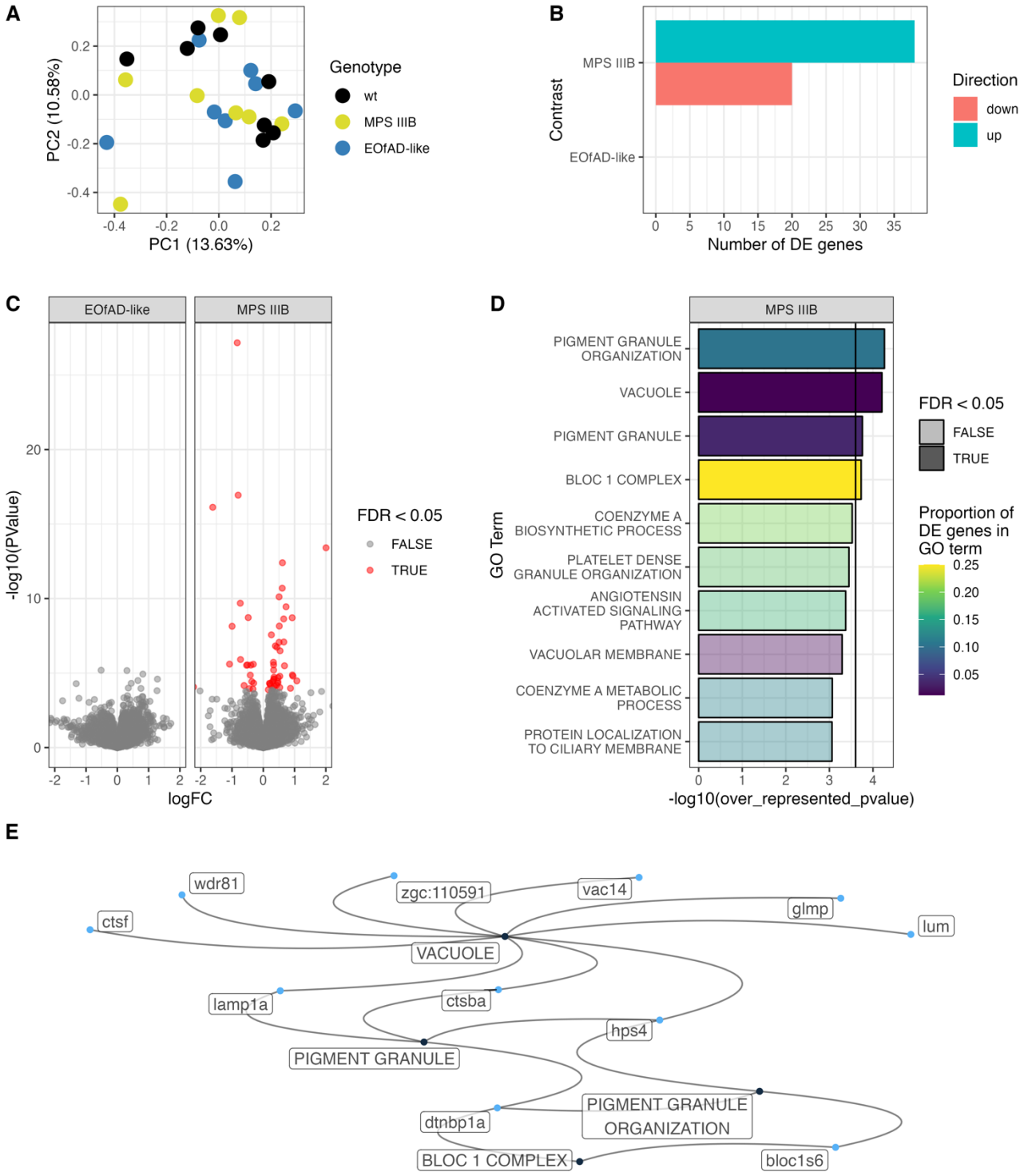

A
PC2 (10.58%)
PC1 (13.63%)
Genotype
wt
MPS IIIB
EOfAD-like
B
Contrast
MPS IIIB
EOfAD-like
Number of DE genes
Direction
down
up
C
EOfAD-like
MPS IIIB
-log10(PValue)
logFC
FDR < 0.05
FALSE
TRUE
D
MPS IIIB
GO Term
PIGMENT GRANULE ORGANIZATION
VACUOLE
PIGMENT GRANULE
BLOC 1 COMPLEX
COENZYME A BIOSYNTHETIC PROCESS
PLATELET DENSE GRANULE ORGANIZATION
ANGIOTENSIN ACTIVATED SIGNALING PATHWAY
VACUOLAR MEMBRANE
COENZYME A METABOLIC PROCESS
PROTEIN LOCALIZATION TO CILIARY MEMBRANE
-log10(over_represented_pvalue)
FDR < 0.05
FALSE
TRUE
Proportion of DE genes in GO term
E
wdr81
zgc:110591
vac14
glmp
ctsf
lum
VACUOLE
lamp1a
ctsba
hps4
PIGMENT GRANULE
PIGMENT GRANULE ORGANIZATION
dtnbp1a
BLOC 1 COMPLEX
bloc1s6

Fig. 3. KEGG pathway analysis in zebrafish larvae at 7 days post fertilisation (dpf). A) Heatmap indicating the KEGG gene sets significantly altered in at least one of the mutant larval zebrafish samples. Cells in the heatmap are coloured by the $\log_{10}$ of the harmonic mean p-value (i.e. the brighter, more yellow cells are more statistically significant). These appear shaded and grey if the p-value did not reach the significance threshold of $p < 0.05$. The numbers inside the heatmap are the FDR-adjusted harmonic mean p-values. **B)** Pearson correlation of the logFC values of genes in the KEGG gene set for extracellular matrix (ECM) receptor interaction. Grey shading indicates confidence intervals.

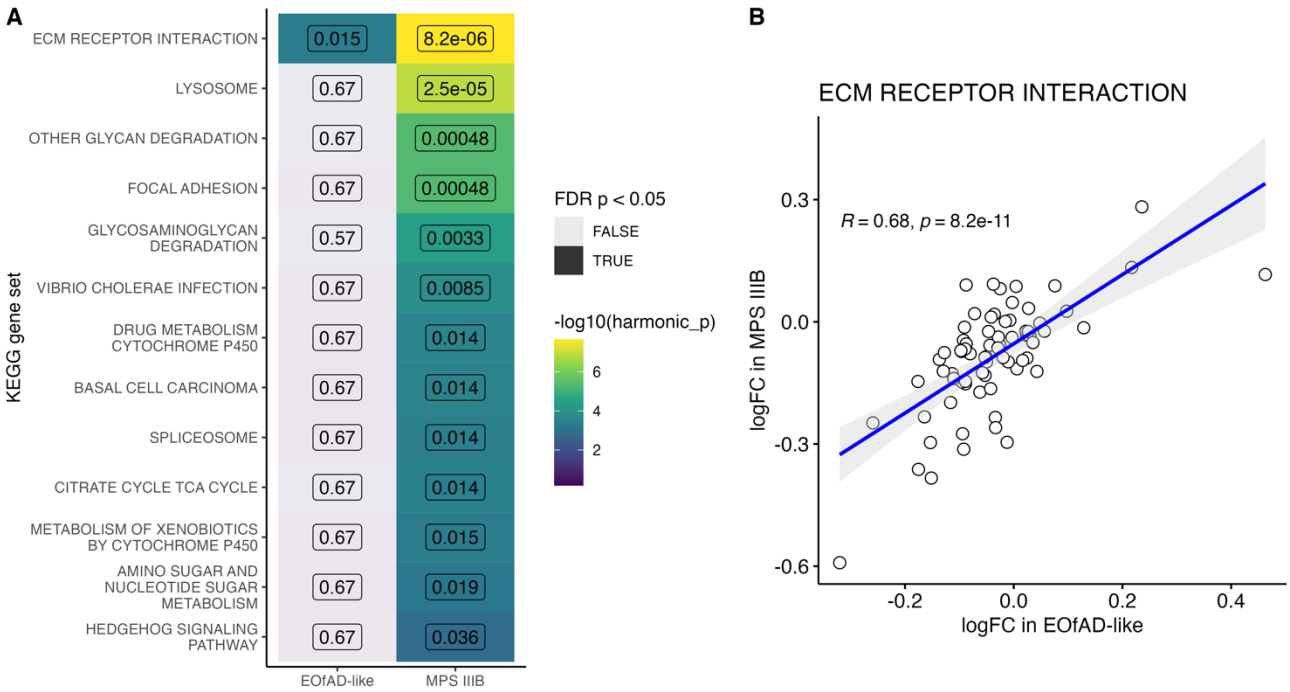
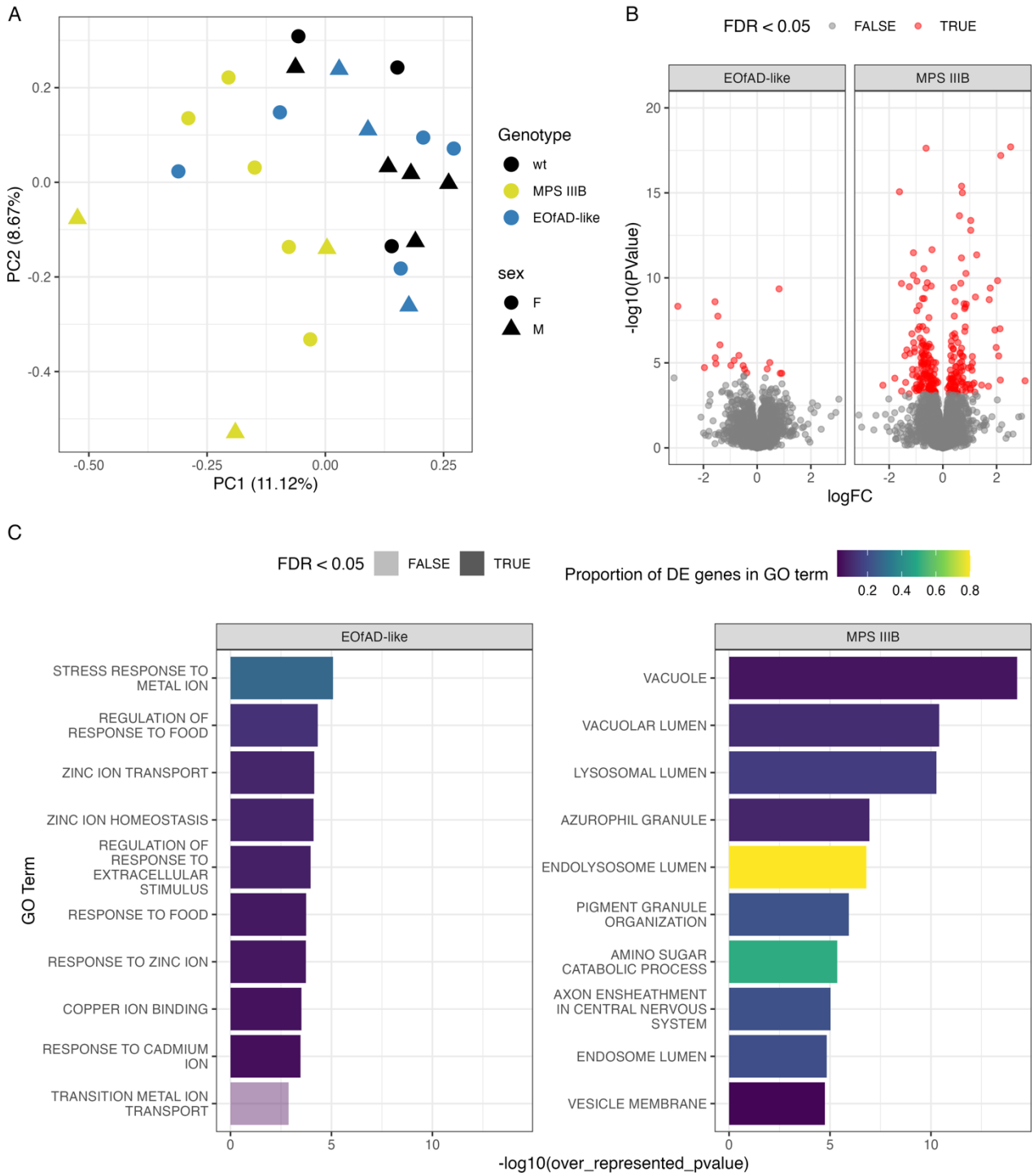


Fig.4: Differential gene expression analysis in zebrafish brains at 6 months of age. A) Principal component analysis (PCA) on the $\log_{10}$ counts per million (logCPM) values after adjusting for %GC and length bias using *cqn* (58). Each point represents an adult zebrafish brain transcriptome sample coloured by genotype and shaped by sex (F: female, M: male, wt: wild type). The percent variations that principal component 1 (PC1) and 2 (PC2) explain in the dataset are indicated in parentheses in the axis labels. **B)** Volcano plots showing the log fold change (logFC) and negative $\log_{10}$ of the p-value of the genes expressed in MPS IIIB and EOfAD-like zebrafish adult brains. Each point represents a gene which is coloured red if the gene was differentially expressed at a FDR-adjusted p-value of less that 0.05. The x-axis labels are constrained to within -2 and 2 for visualisation purposes. **C)** The 10 most significantly over-represented Gene Ontology (GO) terms among the differentially expressed genes in EOfAD-like and MPS IIIB zebrafish brains. The bars are coloured by the proportion of genes within the GO term found to be differentially expressed in each mutant, and appear lighter in colour if the FDR-adjusted p-value was less than 0.05.

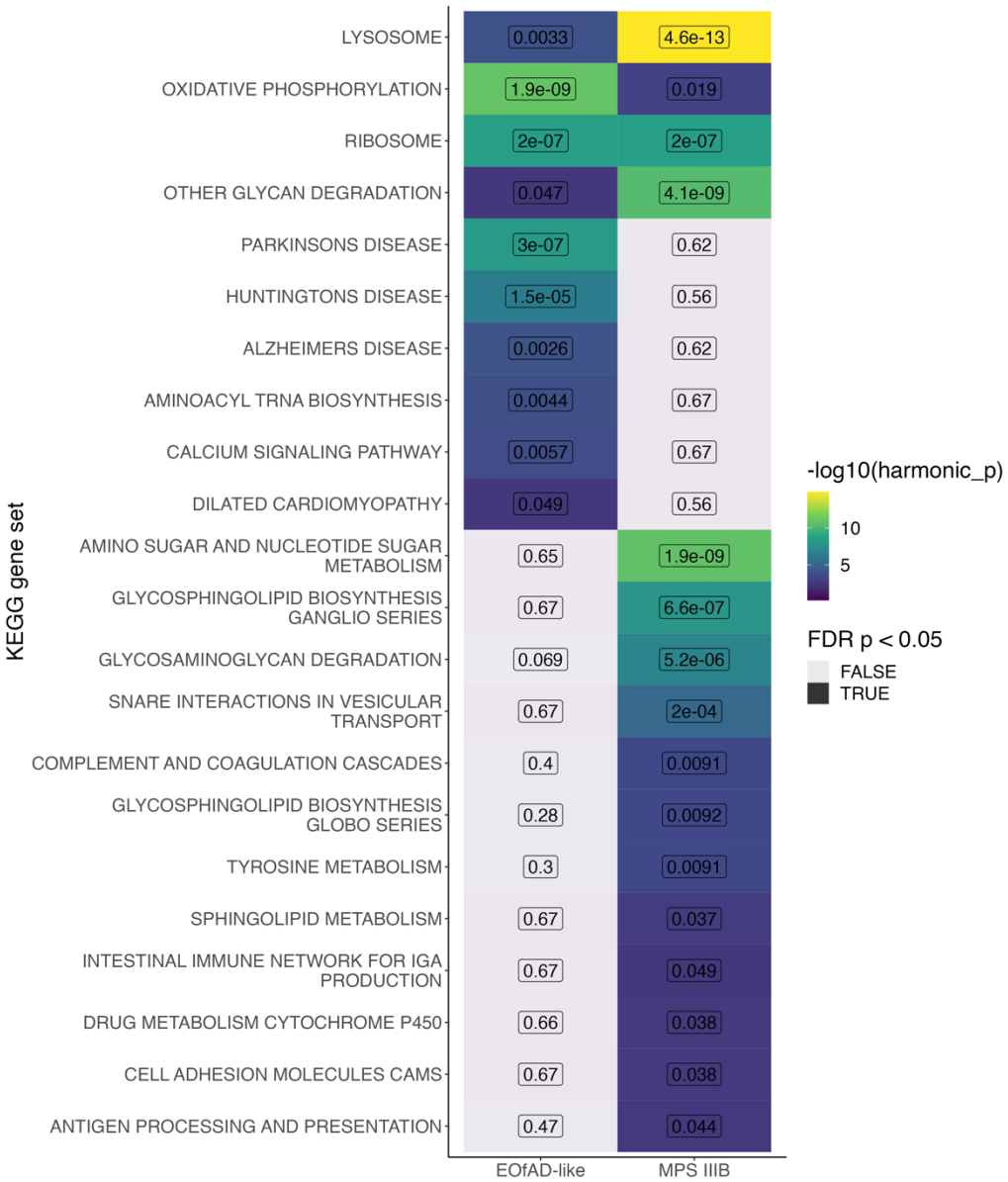

A
PC2 (8.67%)
PC1 (11.12%)
Genotype
wt
MPS IIIB
EOfAD-like
sex
F
M
B
FDR < 0.05
FALSE
TRUE
EOfAD-like
MPS IIIB
-log10(PValue)
logFC
C
FDR < 0.05
FALSE
TRUE
Proportion of DE genes in GO term
EOfAD-like
STRESS RESPONSE TO METAL ION
REGULATION OF RESPONSE TO FOOD
ZINC ION TRANSPORT
ZINC ION HOMEOSTASIS
REGULATION OF RESPONSE TO EXTRACELLULAR STIMULUS
RESPONSE TO FOOD
RESPONSE TO ZINC ION
COPPER ION BINDING
RESPONSE TO CADMIUM ION
TRANSITION METAL ION TRANSPORT
GO Term
MPS IIIB
VACUOLE
VACUOLAR LUMEN
LYSOSOMAL LUMEN
AZUROPHIL GRANULE
ENDOLYSOSOME LUMEN
PIGMENT GRANULE ORGANIZATION
AMINO SUGAR CATABOLIC PROCESS
AXON ENSHEATHMENT IN CENTRAL NERVOUS SYSTEM
ENDOSOME LUMEN
VESICLE MEMBRANE
-log10(over_represented_pvalue)

643

644

645 **Fig.5: KEGG gene sets significantly altered in zebrafish brains at 6 months of age. A)**
646 Heatmap indicating those KEGG gene sets significantly altered in at least one of the mutant
647 zebrafish brains. Cells in the heatmap are coloured by the $\log_{10}$ of the harmonic mean p-value (i.e.
648 the brighter, more yellow cells are more statistically significant). They appear grey if the p-value
649 did not reach the significance threshold of < 0.05. **B)** Scatterplot indicating the Pearson
650 correlations between the log fold change (logFC) of genes in the KEGG gene sets for *oxidative*
651 *phosphorylation*, **C)** *lysosome*, **D)** *ribosome* and **E)** *other glycan degradation* in EOfAD-like and
652 MPS IIIB mutant zebrafish brains.

A


KEGG gene set
LYSOSOME
OXIDATIVE PHOSPHORYLATION
RIBOSOME
OTHER GLYCAN DEGRADATION
PARKINSONS DISEASE
HUNTINGTONS DISEASE
ALZHEIMERS DISEASE
AMINOACYL TRNA BIOSYNTHESIS
CALCIUM SIGNALING PATHWAY
DILATED CARDIOMYOPATHY
AMINO SUGAR AND NUCLEOTIDE SUGAR METABOLISM
GLYCOSPHINGOLIPID BIOSYNTHESIS GANGLIO SERIES
GLYCOSAMINOGLYCAN DEGRADATION
SNARE INTERACTIONS IN VESICULAR TRANSPORT
COMPLEMENT AND COAGULATION CASCADES
GLYCOSPHINGOLIPID BIOSYNTHESIS GLOBO SERIES
TYROSINE METABOLISM
SPHINGOLIPID METABOLISM
INTESTINAL IMMUNE NETWORK FOR IGA PRODUCTION
DRUG METABOLISM CYTOCHROME P450
CELL ADHESION MOLECULES CAMS
ANTIGEN PROCESSING AND PRESENTATION
0.0033
4.6e-13
1.9e-09
0.019
2e-07
2e-07
0.047
4.1e-09
3e-07
0.62
1.5e-05
0.56
0.0026
0.62
0.0044
0.67
0.0057
0.67
0.049
0.56
0.65
1.9e-09
0.67
6.6e-07
0.069
5.2e-06
0.67
2e-04
0.4
0.0091
0.28
0.0092
0.3
0.0091
0.67
0.037
0.67
0.049
0.66
0.038
0.67
0.038
0.47
0.044
EOfAD-like
MPS IIIB
-log10(harmonic_p)
10
5
FDR p < 0.05
FALSE
TRUE

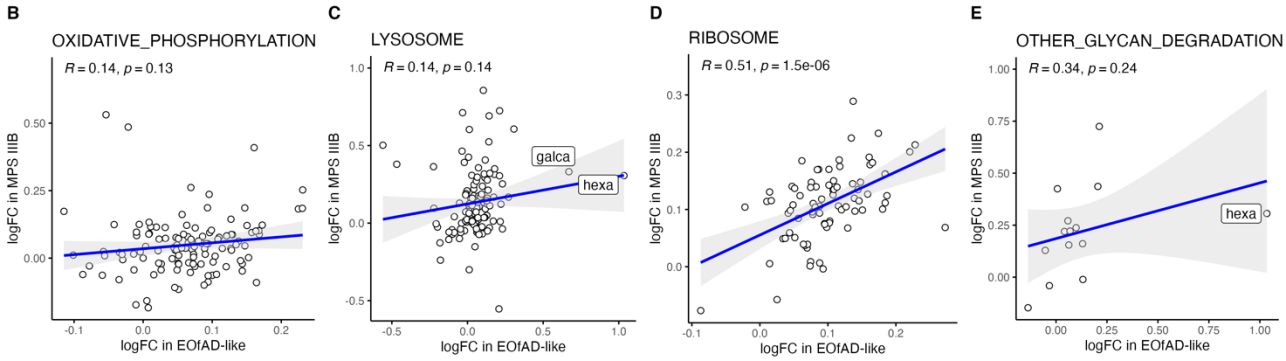
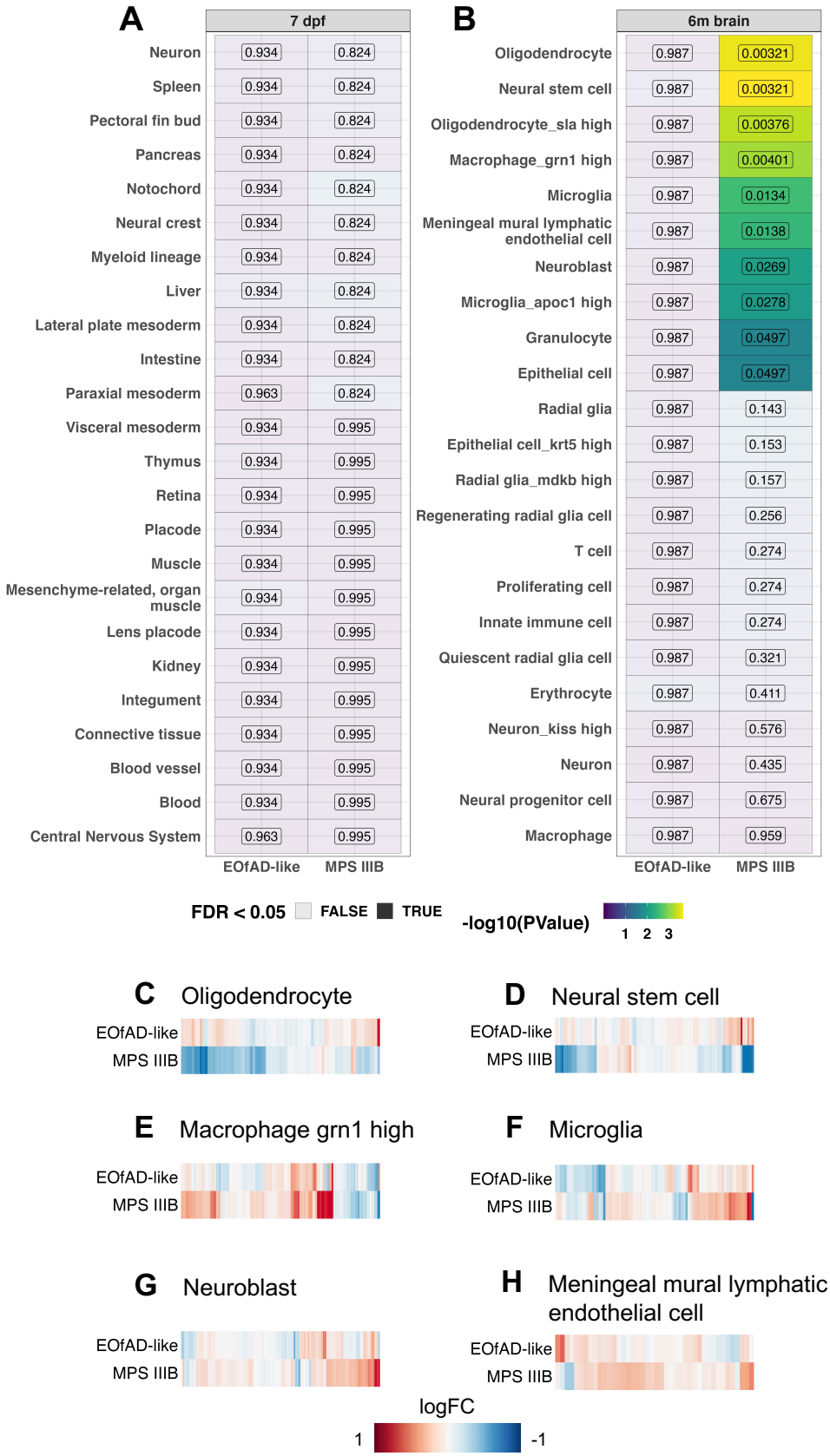

B
OXIDATIVE_PHOSPHORYLATION
R = 0.14, p = 0.13
C
LYSOSOME
R = 0.14, p = 0.14
galca
hexa
D
RIBOSOME
R = 0.51, p = 1.5e-06
E
OTHER_GLYCAN_DEGRADATION
R = 0.34, p = 0.24
hexa
logFC in MPS IIIB
logFC in EOfAD-like

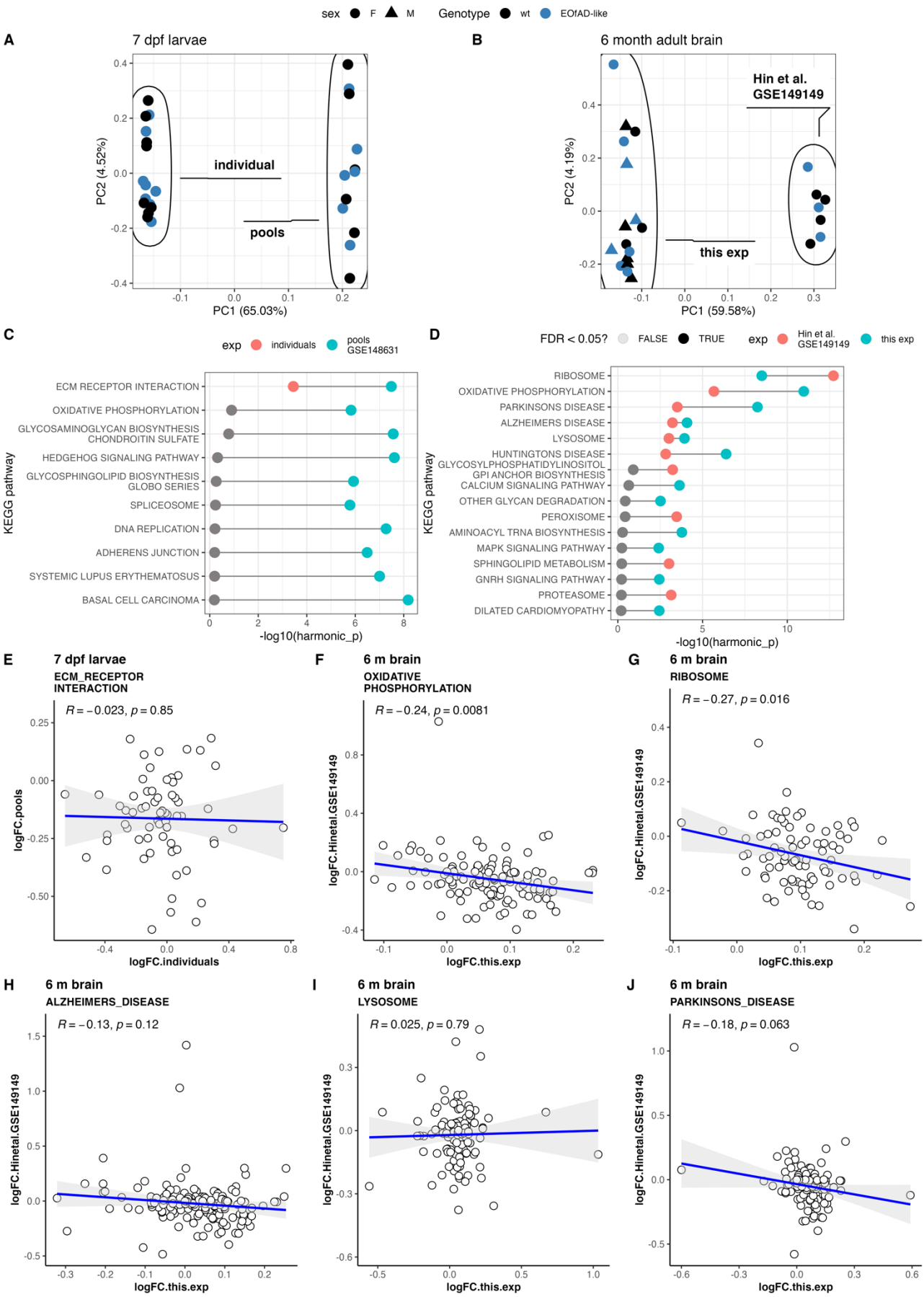
Fig.6 Changes to cell type proportions. A) Heatmap indicating the statistical significance of cell type marker gene sets in 7 days post fertilisation (dpf) larvae and **B)** 6 months of age (6m) adult brains. The cells in the heatmap are coloured by the $\log_{10}$ of the harmonic mean p-value (i.e. the brighter, more yellow cells are more statistically significant). They appear grey if the FDR-adjusted p-value did not reach the threshold of < 0.05. **C)** Heatmap indicating the logFC values of the genes expressed in oligodendrocytes, **D)** neural stem cells, **E)** macrophages with high *grn1* expression, **F)** microglia, **G)** neuroblasts and **H)** meningeal mural lymphatic endothelial cells.


A
7 dpf
Neuron 0.934 0.824
Spleen 0.934 0.824
Pectoral fin bud 0.934 0.824
Pancreas 0.934 0.824
Notochord 0.934 0.824
Neural crest 0.934 0.824
Myeloid lineage 0.934 0.824
Liver 0.934 0.824
Lateral plate mesoderm 0.934 0.824
Intestine 0.934 0.824
Paraxial mesoderm 0.963 0.824
Visceral mesoderm 0.934 0.995
Thymus 0.934 0.995
Retina 0.934 0.995
Placode 0.934 0.995
Muscle 0.934 0.995
Mesenchyme-related, organ muscle 0.934 0.995
Lens placode 0.934 0.995
Kidney 0.934 0.995
Integument 0.934 0.995
Connective tissue 0.934 0.995
Blood vessel 0.934 0.995
Blood 0.934 0.995
Central Nervous System 0.963 0.995
EOfAD-like MPS IIIB
B
6m brain
Oligodendrocyte 0.987 0.00321
Neural stem cell 0.987 0.00321
Oligodendrocyte_sla high 0.987 0.00376
Macrophage_grn1 high 0.987 0.00401
Microglia 0.987 0.0134
Meningeal mural lymphatic endothelial cell 0.987 0.0138
Neuroblast 0.987 0.0269
Microglia_apoc1 high 0.987 0.0278
Granulocyte 0.987 0.0497
Epithelial cell 0.987 0.0497
Radial glia 0.987 0.143
Epithelial cell_krt5 high 0.987 0.153
Radial glia_mdkb high 0.987 0.157
Regenerating radial glia cell 0.987 0.256
T cell 0.987 0.274
Proliferating cell 0.987 0.274
Innate immune cell 0.987 0.274
Quiescent radial glia cell 0.987 0.321
Erythrocyte 0.987 0.411
Neuron_kiss high 0.987 0.576
Neuron 0.987 0.435
Neural progenitor cell 0.987 0.675
Macrophage 0.987 0.959
EOfAD-like MPS IIIB
FDR < 0.05 FALSE TRUE
-log10(PValue) 1 2 3
C Oligodendrocyte
EOfAD-like
MPS IIIB
D Neural stem cell
EOfAD-like
MPS IIIB
E Macrophage grn1 high
EOfAD-like
MPS IIIB
F Microglia
EOfAD-like
MPS IIIB
G Neuroblast
EOfAD-like
MPS IIIB
H Meningeal mural lymphatic endothelial cell
EOfAD-like
MPS IIIB
logFC
1 -1

362

363

Fig.7 Consistent changes to KEGG pathways can be detected across independent experiments analysing EOfAD-like zebrafish. **A)** Principal component (PC) analysis of logCPM values in wild type (wt) and EOfAD-like zebrafish larvae at 7 days post fertilisation when analysing individuals (this experiment) or pools of 40 larvae (70). **B)** PC analysis of logCPM values in 6-month-old wt (wild type) and EOfAD-like zebrafish brains in this experiment and from the data described in Hin et al. (73) **C)** Dumbbell plot showing the $-\log_{10}$ of the harmonic mean p-value of the 10 most significantly altered KEGG gene sets in 7 dpf individuals (red) or pooled (blue) EOfAD-like zebrafish larvae. **D)** Dumbbell plot showing the $-\log_{10}$ of the harmonic mean p-value of the KEGG gene sets significantly altered in the brains of 6-month-old EOfAD-like zebrafish in this experiment (red) or in the experiment described by Hin et al. (73) (blue). Points in **C)** and **D)** appear grey if they did not reach the threshold for being considered statistically significant at FDR < 0.05. **E)** Pearson correlation of the logFC values for genes in the KEGG gene set *ECM receptor interaction* in 7 dpf zebrafish larvae. **F)** Pearson correlation of the logFC values in 6-month-old EOfAD-like zebrafish brains for genes in the KEGG gene sets **F)** *oxidative phosphorylation*, **G)** *ribosome*, **H)** *Alzheimer's disease*, **I)** *lysosome* and **J)** *Parkinson's disease*.


sex F M
Genotype wt EOfAD-like
A
7 dpf larvae
individual
pools
PC2 (4.52%)
PC1 (65.03%)
B
6 month adult brain
Hin et al. GSE149149
this exp
PC2 (4.19%)
PC1 (59.58%)
C
exp individuals pools GSE148631
ECM RECEPTOR INTERACTION
OXIDATIVE PHOSPHORYLATION
GLYCOSAMINOGLYCAN BIOSYNTHESIS CHONDROITIN SULFATE
HEDGEHOG SIGNALING PATHWAY
GLYCOSPHINGOLIPID BIOSYNTHESIS GLOBO SERIES
SPLICEOSOME
DNA REPLICATION
ADHERENS JUNCTION
SYSTEMIC LUPUS ERYTHEMATOSUS
BASAL CELL CARCINOMA
KEGG pathway
-log10(harmonic_p)
D
FDR < 0.05? FALSE TRUE
exp Hin et al. GSE149149 this exp
RIBOSOME
OXIDATIVE PHOSPHORYLATION
PARKINSONS DISEASE
ALZHEIMERS DISEASE
LYSOSOME
HUNTINGTONS DISEASE
GLYCOSYLPHOSPHATIDYLINOSITOL GPI ANCHOR BIOSYNTHESIS
CALCIUM SIGNALING PATHWAY
OTHER GLYCAN DEGRADATION
PEROXISOME
AMINOACYL TRNA BIOSYNTHESIS
MAPK SIGNALING PATHWAY
SPHINGOLIPID METABOLISM
GNRH SIGNALING PATHWAY
PROTEASOME
DILATED CARDIOMYOPATHY
KEGG pathway
-log10(harmonic_p)
E
7 dpf larvae
ECM_RECEPTOR INTERACTION
R = −0.023, p = 0.85
logFC.pools
logFC.individuals
F
6 m brain
OXIDATIVE PHOSPHORYLATION
R = −0.24, p = 0.0081
logFC.Hinetal.GSE149149
logFC.this.exp
G
6 m brain
RIBOSOME
R = −0.27, p = 0.016
logFC.Hinetal.GSE149149
logFC.this.exp
H
6 m brain
ALZHEIMERS_DISEASE
R = −0.13, p = 0.12
logFC.Hinetal.GSE149149
logFC.this.exp
I
6 m brain
LYSOSOME
R = 0.025, p = 0.79
logFC.Hinetal.GSE149149
logFC.this.exp
J
6 m brain
PARKINSONS_DISEASE
R = −0.18, p = 0.063
logFC.Hinetal.GSE149149
logFC.this.exp

Declarations

Ethics approval and consent to participate

All experiments involving zebrafish were conducted under the auspices of the Animal Ethics Committee of the University of Adelaide (approval numbers S-2021–067 and S-2021–041), and Institutional Biosafety Committee (PC2 NLRD 15037).

Availability of data and materials

The raw fastq files from these analyses have been deposited to the GEO database (https://www.ncbi.nlm.nih.gov/geo/) under the accession numbers GSE217196 (7 dpf larvae) and GSE242370 (adult brain). The data arising from our previous analyses of the EOfAD-like mutations can be obtained from the GEO database under accession numbers GSE148631 (pools of 7 dpf larvae) and GSE149149 (6-month-old adult brain). All code to reproduce these analyses can be found at https://github.com/karissa-b/2021_MPSIIIBvQ96-RNAseq-7dpfLarve.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by a Postdoctoral Fellowship awarded to KB from Race Against Dementia and the Dementia Australia Research Foundation, and by funds from The Carthew Family Charity Trust to ML. KB is also supported by funds from Flinders University. ML and KH are academic employees of the University of Adelaide and Flinders University respectively.

Authors' contributions

KB and ML conceived the project. KB performed all experiments and analyses described in this manuscript, prepared the figures, and wrote the first draft under the supervision of KH and ML. All authors read and contributed to the final manuscript.

Acknowledgements

The authors would like to thank Lachlan Baer for his assistance with the snakemake pipeline. The authors acknowledge the South Australian Genomics Centre which provided RNA-seq library preparation and RNA sequencing services. The SAGC is supported by the National Collaborative Research Infrastructure Strategy (NCRIS) via BioPlatforms Australia and by the SAGC partner institutes. This work was supported with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide.

References

1. Rizzi L, Rosset I, Roriz-Cruz M. Global epidemiology of dementia: Alzheimer's and vascular types. Biomed Res Int. 2014;2014:908915. PubMed PMID: 25089278. PMCID: PMC4095986. Epub 20140625. eng.
2. Alzheimer A. Uber eigenartige Erkrankung der Hirnrinde. All Z Psychiatr. 1907;64:146-8.
3. Du X, Wang X, Geng M. Alzheimer's disease hypothesis and related therapies. Transl Neurodegener. 2018;7:2. PubMed PMID: 29423193. PMCID: PMC5789526. Epub 20180130. eng.
4. Kepp KP, Robakis NK, Høilund-Carlsen PF, Sensi SL, Vissel B. The amyloid cascade hypothesis: an updated critical review. Brain. 2023:awad159.
5. Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature. 1995 1995/06/01;375(6534):754-60.
6. Rogaev EI, Sherrington R, Rogaeva EA, Levesque G, Ikeda M, Liang Y, et al. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. Nature. 1995 1995/08/01;376(6543):775-8.

7. Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, et al. Candidate Gene for the Chromosome 1 Familial Alzheimer's Disease Locus. Science. 1995 1995/08/18;269(5226):973-7.

8. Goate A, Chartier-Harlin M-C, Mullan M, Brown J, Crawford F, Fidani L, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature. 1991 1991/02/01;349(6311):704-6.

9. Pottier C, Hannequin D, Coutant S, Rovelet-Lecrux A, Wallon D, Rousseau S, et al. High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. Molecular Psychiatry. 2012 2012/09/01;17(9):875-9.

10. Nicolas G, Charbonnier C, Wallon D, Quenez O, Bellenguez C, Grenier-Boley B, et al. SORL1 rare variants: a major risk factor for familial early-onset Alzheimer's disease. Molecular Psychiatry. 2016 2016/06/01;21(6):831-6.

11. Pasternak SH, Bagshaw RD, Guiral M, Zhang S, Ackerley CA, Pak BJ, et al. Presenilin-1, nicastrin, amyloid precursor protein, and gamma-secretase activity are co-localized in the lysosomal membrane. J Biol Chem. 2003 Jul 18;278(29):26687-94. PubMed PMID: 12736250. Epub 20030507. eng.

12. Kawai M, Cras P, Richey P, Tabaton M, Lowery DE, Gonzalez-DeWhitt PA, et al. Subcellular localization of amyloid precursor protein in senile plaques of Alzheimer's disease. The American journal of pathology. 1992 Apr;140(4):947-58. PubMed PMID: 1562053. PMCID: PMC1886361. eng.

13. Mishra S, Knupp A, Szabo MP, Williams CA, Kinoshita C, Hailey DW, et al. The Alzheimer's gene SORL1 is a regulator of endosomal traffic and recycling in human neurons. Cell Mol Life Sci. 2022 Feb 28;79(3):162. PubMed PMID: 35226190. PMCID: PMC8885486. Epub 20220228. eng.

14. Fedeli C, Filadi R, Rossi A, Mammucari C, Pizzo P. PSEN2 (presenilin 2) mutants linked to familial Alzheimer disease impair autophagy by altering Ca(2+) homeostasis. Autophagy. 2019 Dec;15(12):2044-62. PubMed PMID: 30892128. PMCID: PMC6844518. Epub 20190327. eng.

15. Gao S, Casey AE, Sargeant TJ, Mäkinen V-P. Genetic variation within endolysosomal system is associated with late-onset Alzheimer's disease. Brain. 2018;141(9):2711-20.

16. Cataldo AM, Peterhoff CM, Troncoso JC, Gomez-Isla T, Hyman BT, Nixon RA. Endocytic pathway abnormalities precede amyloid beta deposition in sporadic Alzheimer's disease and Down syndrome: differential effects of APOE genotype and presenilin mutations. The American journal of pathology. 2000 Jul;157(1):277-86. PubMed PMID: 10880397. PMCID: PMC1850219. eng.

17. Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT. Neuropathological alterations in Alzheimer disease. Cold Spring Harb Perspect Med. 2011 Sep;1(1):a006189. PubMed PMID: 22229116. PMCID: PMC3234452. eng.

18. McKean NE, Handley RR, Snell RG. A Review of the Current Mammalian Models of Alzheimer's Disease and Challenges That Need to Be Overcome. Int J Mol Sci. 2021 Dec 6;22(23). PubMed PMID: 34884970. PMCID: PMC8658123. Epub 20211206. eng.

19. Drummond E, Wisniewski T. Alzheimer's disease: experimental models and reality. Acta Neuropathol. 2017 Feb;133(2):155-75. PubMed PMID: 28025715. PMCID: PMC5253109. Epub 20161226. eng.

20. Hargis KE, Blalock EM. Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us? Behavioural Brain Research. 2017 2017/03/30/;322:311-28.

21. Mullane K, Williams M. Preclinical Models of Alzheimer's Disease: Relevance and Translational Validity. Current Protocols in Pharmacology. 2019 2019/03/01;84(1):e57.

22. Lardelli M. An Alternative View of Familial Alzheimer's Disease Genetics. Journal of Alzheimer's Disease. 2023;Preprint:1-27.

23. Zelei T, Csetneki K, Vokó Z, Siffel C. Epidemiology of Sanfilippo syndrome: results of a systematic literature review. Orphanet journal of rare diseases. 2018 Apr 10;13(1):53. PubMed PMID: 29631636. PMCID: PMC5891921. Epub 2018/04/11. eng.

24. Scott HS, Blanch L, Guo X-H, Freeman C, Orsborn A, Baker E, et al. Cloning of the sulphamidase gene and identification of mutations in Sanfilippo A syndrome. Nature genetics. 1995 1995/12/01;11(4):465-7.

25. Zhao HG, Li HH, Bach G, Schmidtchen A, Neufeld EF. The molecular basis of Sanfilippo syndrome type B. Proceedings of the National Academy of Sciences. 1996 1996/06/11;93(12):6101-5.

26. Klein U, Kresse H, von Figura K. Sanfilippo syndrome type C: deficiency of acetyl-CoA:alpha-glucosaminide N-acetyltransferase in skin fibroblasts. Proceedings of the National Academy of Sciences. 1978 1978/10/01;75(10):5185-9.

27. Kresse H, Paschke E, von Figura K, Gilberg W, Fuchs W. Sanfilippo disease type D: deficiency of N-acetylglucosamine-6-sulfate sulfatase required for heparan sulfate degradation. Proceedings of the National Academy of Sciences. 1980 1980/11/01;77(11):6822-6.

28. Kowalewski B, Lamanna WC, Lawrence R, Damme M, Stroobants S, Padva M, et al. Arylsulfatase G inactivation causes loss of heparan sulfate 3-O-sulfatase activity and mucopolysaccharidosis in mice. Proceedings of the National Academy of Sciences. 2012 2012/06/26;109(26):10310-5.

29. Villani GRD, Domenico CD, Musella A, Cecere F, Napoli DD, Natale PD. Mucopolysaccharidosis IIIB: Oxidative damage and cytotoxic cell involvement in the neuronal pathogenesis. Brain Research. 2009 2009/07/07/;1279:99-108.

30. Hamano K, Hayashi M, Shioda K, Fukatsu R, Mizutani S. Mechanisms of neurodegeneration in mucopolysaccharidoses II and IIIB: analysis of human brain tissue. Acta Neuropathol. 2008 2008/05/01;115(5):547-59.

31. Jones MZ, Alroy J, Rutledge JC, Taylor JW, Alvord EC, Jr., Toone J, et al. Human Mucopolysaccharidosis IIID: Clinical, Biochemical, Morphological and Immunohistochemical Characteristics. Journal of Neuropathology & Experimental Neurology. 1997;56(10):1158-67.

32. Valle DAd, Santos MLSF, Telles BA, Cordeiro ML. Neurological, neurobehavioral, and radiological alterations in patients with mucopolysaccharidosis III (Sanfilippo's syndrome) in Brazil. Frontiers in Neurology. 2022;13.

33. Viana GM, Priestman DA, Platt FM, Khan S, Tomatsu S, Pshezhetsky AV. Brain Pathology in Mucopolysaccharidoses (MPS) Patients with Neurological Forms. J Clin

Med. 2020 Feb 1;9(2). PubMed PMID: 32024172. PMCID: PMC7073982. Epub 20200201. eng.

34. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, et al. The zebrafish reference genome sequence and its relationship to the human genome. Nature. 2013 2013/04/01;496(7446):498-503.

35. Liu Y. Zebrafish as a Model Organism for Studying Pathologic Mechanisms of Neurodegenerative Diseases and other Neural Disorders. Cellular and Molecular Neurobiology. 2023 2023/04/01.

36. Barthelson K, Dong Y, Newman M, Lardelli M. PRESENILIN 1 mutations causing early-onset familial Alzheimer's disease or familial acne inversa differ in their effects on genes facilitating energy metabolism and signal transduction. Journal of Alzheimer's Disease. 2021;82(1):327-47.

37. Barthelson K, Pederson SM, Newman M, Jiang H, Lardelli M. In-frame and frameshift mutations in zebrafish Presenilin 2 affect different cellular functions in young adult brains. Journal of Alzheimer's Disease Reports. 2021;5(1):395-404.

38. Barthelson K, Pederson SM, Newman M, Lardelli M. Brain transcriptome analysis reveals subtle effects on mitochondrial function and iron homeostasis of mutations in the SORL1 gene implicated in early onset familial Alzheimer's disease. Molecular Brain. 2020 2020/10/19;13(1):142.

39. Barthelson K, Pederson SM, Newman M, Lardelli M. Brain transcriptome analysis of a protein-truncating mutation in sortilin-related receptor 1 associated with early-onset familial Alzheimer's disease indicates early effects on mitochondrial and ribosome function. Journal of Alzheimer's Disease. 2021;79(3):1105-19.

40. Jiang H, Pederson SM, Newman M, Dong Y, Barthelson K, Lardelli M. Transcriptome analysis indicates dominant effects on ribosome and mitochondrial function of a premature termination codon mutation in the zebrafish gene psen2. Plos one. 2020;15(7):e0232559.

41. Gerken E, Ahmad S, Barthelson K, Lardelli M. Zebrafish models of Mucopolysaccharidosis types IIIA, B, and C show hyperactivity and changes in oligodendrocyte state. bioRxiv. 2023:2023.08.02.550904.

42. Barthelson K, Newman M, Lardelli M. Brain transcriptomes of zebrafish and mouse Alzheimer's disease knock-in models imply early disrupted energy metabolism. Disease Models & Mechanisms. 2022;15(1):dmm049187.

43. Newman M, Hin N, Pederson S, Lardelli M. Brain transcriptome analysis of a familial Alzheimer's disease-like mutation in the zebrafish presenilin 1 gene implies effects on energy production. Mol Brain. 2019 May 3;12(1):43. PubMed PMID: 31053140. PMCID: PMC6500017. Epub 2019/05/06. eng.

44. Westerfield M. The zebrafish book: a guide for the laboratory use of zebrafish. http://zfin org/zf_info/zfbook/zfbk html. 2000.

45. Allen AG, Barthelson K, Lardelli M. pHAPE: a plasmid for production of DNA size marker ladders for gel electrophoresis. Biology Methods and Protocols. 2023;8(1):bpad015.

46. Mölder F, Jablonski KP, Letcher B. Sustainable data analysis with Snakemake.[version 2; peer review: 2 approved]. F1000Research 10: 33. Crossref, PubMed. 2021.

366 47. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor.
367 Bioinformatics. 2018;34(17):i884-i90.

368 48. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR:
369 ultrafast universal RNA-seq aligner. Bioinformatics. 2013 Jan 1;29(1):15-21. PubMed
370 PMID: 23104886. PMCID: PMC3530905. Epub 2012/10/30. eng.

371 49. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence
372 Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.

373 50. Smith TS, Heger A, Sudbery I. UMI-tools: Modelling sequencing errors in Unique
374 Molecular Identifiers to improve quantification accuracy. Genome Research. 2017.

375 51. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for
376 assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30.

377 52. Andrews S. FastQC: a quality control tool for high throughput sequence data.
378 Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom; 2010.

379 53. Ward CM, To TH, Pederson SM. ngsReports: a Bioconductor package for managing
380 FastQC reports and other NGS related log files. Bioinformatics. 2020 Apr
381 15;36(8):2587-8. PubMed PMID: 31841127. Epub 2019/12/17. eng.

382 54. Team RC. R: A Language and Environment for Statistical Computing. In: Computing
383 RFfS, editor. Vienna, Austria2020.

384 55. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for
385 differential expression analysis of digital gene expression data. Bioinformatics. 2010
386 Jan 1;26(1):139-40. PubMed PMID: 19910308. PMCID: PMC2796818. Epub
387 2009/11/17. eng.

388 56. Lee SLJ, Horsfield JA, Black MA, Rutherford K, Gemmell NJ. Identification of sex
389 differences in zebrafish (Danio rerio) brains during early sexual differentiation and
390 masculinization using 17α-methyltestoterone†. Biology of Reproduction.
391 2018;99(2):446-60.

392 57. Wong RY, McLeod MM, Godwin J. Limited sex-biased neural gene expression
393 patterns across strains in Zebrafish (Danio rerio). BMC Genomics. 2014
394 2014/10/17;15(1):905.

395 58. Hansen KD, Irizarry RA, Wu Z. Removing technical variability in RNA-seq data using
396 conditional quantile normalization. Biostatistics (Oxford, England). 2012
397 Apr;13(2):204-16. PubMed PMID: 22285995. PMCID: PMC3297825. Epub
398 2012/01/31. eng.

399 59. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene
900 ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature
901 genetics. 2000 May;25(1):25-9. PubMed PMID: 10802651. PMCID: PMC3037419.
902 Epub 2000/05/10. eng.

903 60. Consortium GO. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids
904 Res. 2021 Jan 8;49(D1):D325-d34. PubMed PMID: 33290552. PMCID:
905 PMC7779012. Epub 2020/12/09. eng.

906 61. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-
907 seq: accounting for selection bias. Genome Biology. 2010 2010/02/04;11(2):R14.

908 62. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic
909 Acids Research. 2000;28(1):27-30.

910 63. Wu D, Lim E, Vaillant F, Asselin-Labat M-L, Visvader JE, Smyth GK. ROAST:
911 rotation gene set tests for complex microarray experiments. Bioinformatics.
912 2010;26(17):2176-82.

913 64. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene
914 correlation. Nucleic Acids Research. 2012;40(17):e133-e.

915 65. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast
916 gene set enrichment analysis. bioRxiv. 2021:060012.

917 66. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al.
918 Gene set enrichment analysis: A knowledge-based approach for interpreting
919 genome-wide expression profiles. Proceedings of the National Academy of Sciences.
920 2005 2005/10/25;102(43):15545-50.

921 67. Wilson DJ. The harmonic mean p-value for combining dependent tests. Proceedings
922 of the National Academy of Sciences. 2019 2019/01/22;116(4):1195-200.

923 68. Bi R, Liu P. Sample size calculation while controlling false discovery rate for
924 differential expression analysis with RNA-sequencing experiments. BMC
925 bioinformatics. 2016 Mar 31;17:146. PubMed PMID: 27029470. PMCID:
926 PMC4815167. Epub 2016/04/01. eng.

927 69. Fedele AO. Sanfilippo syndrome: causes, consequences, and treatments. The
928 application of clinical genetics. 2015;8:269-81. PubMed PMID: 26648750. PMCID:
929 PMC4664539. Epub 2015/12/10. eng.

930 70. Dong Y, Newman M, Pederson SM, Barthelson K, Hin N, Lardelli M. Transcriptome
931 analyses of 7-day-old zebrafish larvae possessing a familial Alzheimer's disease-like
932 mutation in psen1 indicate effects on oxidagtive phosphorylation, ECM and MCM
933 functions, and iron homeostasis. BMC Genomics. 2021 2021/03/24;22(1):211.

934 71. Farnsworth DR, Saunders LM, Miller AC. A single-cell transcriptome atlas for
935 zebrafish development. Developmental Biology. 2020 2020/03/15/;459(2):100-8.

936 72. Jiang M, Xiao Y, E W, Ma L, Wang J, Chen H, et al. Characterization of the Zebrafish
937 Cell Landscape at Single-Cell Resolution. Frontiers in Cell and Developmental
938 Biology. 2021;9.

939 73. Hin N, Newman M, Pederson S, Lardelli M. Iron responsive element-mediated
940 responses to iron dyshomeostasis in Alzheimer's disease. Journal of Alzheimer's
941 Disease. 2021;84(4):1597-630.

942 74. Karamanos NK, Theocharis AD, Piperigkou Z, Manou D, Passi A, Skandalis SS, et
943 al. A guide to the composition and functions of the extracellular matrix. The FEBS
944 Journal. 2021 2021/12/01;288(24):6850-912.

945 75. Theocharis AD, Skandalis SS, Gialeli C, Karamanos NK. Extracellular matrix
946 structure. Advanced Drug Delivery Reviews. 2016 2016/02/01/;97:4-27.

947 76. Laurie GW, Leblond CP, Martin GR. Localization of type IV collagen, laminin,
948 heparan sulfate proteoglycan, and fibronectin to the basal lamina of basement
949 membranes. Journal of Cell Biology. 1982;95(1):340-4.

950 77. Roy E, Bruyère J, Flamant P, Bigou S, Ausseil J, Vitry S, et al. GM130 gain-of-
951 function induces cell pathology in a model of lysosomal storage disease. Human
952 Molecular Genetics. 2012;21(7):1481-95.

953 78. Batzios SP, Zafeiriou DI, Papakonstantinou E. Extracellular matrix components: An
954 intricate network of possible biomarkers for lysosomal storage disorders? FEBS
955 Letters. 2013 2013/04/17/;587(8):1258-67.

956 79. Wiśniewska K, Gaffke L, Krzelowska K, Węgrzyn G, Pierzynowska K. Differences in
957 gene expression patterns, revealed by RNA-seq analysis, between various
958 Sanfilippo and Morquio disease subtypes. Gene. 2022 2022/02/20/;812:146090.

959 80. Lemonnier T, Blanchard S, Toli D, Roy E, Bigou S, Froissart R, et al. Modeling
960 neuronal defects associated with a lysosomal disorder using patient-derived induced
961 pluripotent stem cells. Human Molecular Genetics. 2011;20(18):3653-66.

962 81. Kwart D, Gregg A, Scheckel C, Murphy EA, Paquet D, Duffield M, et al. A Large
963 Panel of Isogenic APP and PSEN1 Mutant Human iPSC Neurons Reveals Shared
964 Endosomal Abnormalities Mediated by APP β-CTFs, Not Aβ. Neuron.
965 2019;104(2):256-70.e5.

966 82. Corsi GI, Gadekar VP, Haukedal H, Doncheva NT, Anthon C, Ambardar S, et al. The
967 transcriptomic landscape of neurons carrying PSEN1 mutations reveals changes in
968 extracellular matrix components and non-coding gene expression. Neurobiology of
969 Disease. 2023 2023/03/01/;178:105980.

970 83. Caldwell AB, Liu Q, Zhang C, Schroth GP, Galasko DR, Rynearson KD, et al.
971 Endotype reversal as a novel strategy for screening drugs targeting familial
972 Alzheimer's disease. Alzheimer's & Dementia. 2022 2022/11/01;18(11):2117-30.

973 84. Mertens J, Herdy JR, Traxler L, Schafer ST, Schlachetzki JCM, Böhnke L, et al. Age-
974 dependent instability of mature neuronal fate in induced neurons from
975 Alzheimer’s patients. Cell Stem Cell. 2021;28(9):1533-48.e6.

976 85. Lepelletier FX, Mann DMA, Robinson AC, Pinteaux E, Boutin H. Early changes in
977 extracellular matrix in Alzheimer's disease. Neuropathology and Applied
978 Neurobiology. 2017 2017/02/01;43(2):167-82.

979 86. Johnson ECB, Carter EK, Dammer EB, Duong DM, Gerasimov ES, Liu Y, et al.
980 Large-scale deep multi-layer analysis of Alzheimer's disease brain reveals strong
981 proteomic disease-related changes not observed at the RNA level. Nat Neurosci.
982 2022 Feb;25(2):213-25. PubMed PMID: 35115731. PMCID: PMC8825285. Epub
983 20220203. eng.

984 87. Johnson ECB, Bian S, Haque RU, Carter EK, Watson CM, Gordon BA, et al.
985 Cerebrospinal fluid proteomics define the natural history of autosomal dominant
986 Alzheimer’s disease. Nature Medicine. 2023 2023/08/01;29(8):1979-88.

987 88. Lee S, Devanney NA, Golden LR, Smith CT, Schwartz JL, Walsh AE, et al.
988 APOE modulates microglial immunometabolism in response to age,
989 amyloid pathology, and inflammatory challenge. Cell reports. 2023;42(3).

990 89. Lunnon K, Keohane A, Pidsley R, Newhouse S, Riddoch-Contreras J, Thubron EB,
991 et al. Mitochondrial genes are altered in blood early in Alzheimer's disease.
992 Neurobiology of Aging. 2017 2017/05/01/;53:36-47.

993 90. Manczak M, Park BS, Jung Y, Reddy PH. Differential expression of oxidative
994 phosphorylation genes in patients with Alzheimer’s disease. NeuroMolecular
995 Medicine. 2004 2004/04/01;5(2):147-62.

996 91. Demetrius LA, Eckert A, Grimm A. Sex differences in Alzheimer’s disease:
997 metabolic reprogramming and therapeutic intervention. Trends in Endocrinology &
998 Metabolism. 2021;32(12):963-79.

999 92. Dave A, Hansen N, Downey R, Johnson C. FDG-PET Imaging of Dementia and
000 Neurodegenerative Disease. Seminars in Ultrasound, CT and MRI. 2020
001 2020/12/01/;41(6):562-71.

002 93. Settembre C, Fraldi A, Jahreiss L, Spampanato C, Venturi C, Medina D, et al. A
003 block of autophagy in lysosomal storage disorders. Human Molecular Genetics.
004 2008;17(1):119-29.

005 94. Villani GRD, Gargiulo N, Faraonio R, Castaldo S, Gonzalez y Reyero E, Di Natale P.
006 Cytokines, neurotrophins, and oxidative stress in brain disease from
007 mucopolysaccharidosis IIIB. Journal of Neuroscience Research. 2007
008 2007/02/15;85(3):612-22.

009 95. Martins C, Hůlková H, Dridi L, Dormoy-Raclet V, Grigoryeva L, Choi Y, et al.
010 Neuroinflammation, mitochondrial defects and neurodegeneration in
011 mucopolysaccharidosis III type C mouse model. Brain. 2015 Feb;138(Pt 2):336-55.
012 PubMed PMID: 25567323. PMCID: PMC4306821. Epub 20150106. eng.

013 96. Pará C, Bose P, Bruno L, Freemantle E, Taherzadeh M, Pan X, et al. Early defects in
014 mucopolysaccharidosis type IIIC disrupt excitatory synaptic transmission. JCI Insight.
015 2021 08/09/;6(15).

016 97. Haust MD. Mitochondrial budding and morphogenesis of cytoplasmic vacuoles in
017 hepatocytes of children with the hurler syndrome and sanfilippo disease.
018 Experimental and Molecular Pathology. 1968 1968/10/01/;9(2):242-57.

019 98. Sly WS, Quinton BA, McAlister WH, Rimoin DL. Beta glucuronidase deficiency:
020 Report of clinical, radiologic, and biochemical features of a new
021 mucopolysaccharidosis. The Journal of Pediatrics. 1973 1973/02/01/;82(2):249-57.

022 99. Tomatsu S, Gutierrez MA, Ishimaru T, Peña OM, Montaño AM, Maeda H, et al.
023 Heparan sulfate levels in mucopolysaccharidoses and mucolipidoses. Journal of
024 Inherited Metabolic Disease. 2005 2005/09/01;28(5):743-57.

025 100. Parente MK, Rozen R, Seeholzer SH, Wolfe JH. Integrated analysis of proteome and
026 transcriptome changes in the mucopolysaccharidosis type VII mouse hippocampus.
027 Molecular Genetics and Metabolism. 2016 2016/05/01/;118(1):41-54.

028 101. Heon-Roberts R, Nguyen ALA, Pshezhetsky AV. Molecular Bases of
029 Neurodegeneration and Cognitive Decline, the Major Burden of Sanfilippo Disease. J
030 Clin Med. 2020;9(2):344. PubMed PMID: 32012694. eng.

031 102. Johnson DE, Ostrowski P, Jaumouillé V, Grinstein S. The position of lysosomes
032 within the cell determines their luminal pH. Journal of Cell Biology. 2016;212(6):677-
033 92.

034 103. Mindell JA. Lysosomal Acidification Mechanisms. Annual Review of Physiology.
035 2012 2012/03/17;74(1):69-86.

036 104. Colacurcio DJ, Nixon RA. Disorders of lysosomal acidification-The emerging role of
037 v-ATPase in aging and neurodegenerative disease. Ageing Res Rev. 2016
038 2016/12//;32:75-88. PubMed PMID: 27197071. eng.

039 105. Lee JH, McBrayer MK, Wolfe DM, Haslett LJ, Kumar A, Sato Y, et al. Presenilin 1
040 Maintains Lysosomal Ca(2+) Homeostasis via TRPML1 by Regulating vATPase-

Mediated Lysosome Acidification. Cell reports. 2015 Sep 1;12(9):1430-44. PubMed PMID: 26299959. PMCID: PMC4558203. Epub 2015/08/25. eng.

106. Jiang Y, Sato Y, Im E, Berg M, Bordi M, Darji S, et al. Lysosomal Dysfunction in Down Syndrome Is APP-Dependent and Mediated by APP-βCTF (C99). J Neurosci. 2019 Jul 3;39(27):5255-68. PubMed PMID: 31043483. PMCID: PMC6607756. Epub 20190501. eng.

107. Lee J-H, Yang D-S, Goulbourne CN, Im E, Stavrides P, Pensalfini A, et al. Faulty autolysosome acidification in Alzheimer's disease mouse models induces autophagic build-up of Aβ in neurons, yielding senile plaques. Nature Neuroscience. 2022 2022/06/01;25(6):688-701.

108. Prasad H, Rao R. Amyloid clearance defect in ApoE4 astrocytes is reversed by epigenetic correction of endosomal pH. Proc Natl Acad Sci U S A. 2018 Jul 10;115(28):E6640-e9. PubMed PMID: 29946028. PMCID: PMC6048470. Epub 20180626. eng.

109. Im E, Jiang Y, Stavrides PH, Darji S, Erdjument-Bromage H, Neubert TA, et al. Lysosomal dysfunction in Down syndrome and Alzheimer mouse models is caused by v-ATPase inhibition by Tyr(682)-phosphorylated APP βCTF. Sci Adv. 2023 Jul 28;9(30):eadg1925. PubMed PMID: 37494443. PMCID: PMC10371027. Epub 20230726. eng.

110. Fraldi A, Annunziata F, Lombardi A, Kaiser H-J, Medina DL, Spampanato C, et al. Lysosomal fusion and SNARE function are impaired by cholesterol accumulation in lysosomal storage disorders. The EMBO Journal. 2010 2010/11/03;29(21):3607-20.

111. Bach G, Chen C-S, Pagano RE. Elevated lysosomal pH in Mucolipidosis type IV cells. Clinica Chimica Acta. 1999;280(1-2):173-9.

112. Bourdenx M, Daniel J, Genin E, Soria FN, Blanchard-Desce M, Bezard E, et al. Nanoparticles restore lysosomal acidification defects: Implications for Parkinson and other lysosomal-related diseases. Autophagy. 2016;12(3):472-83.

113. Holopainen JM, Saarikoski J, Kinnunen PK, Järvelä I. Elevated lysosomal pH in neuronal ceroid lipofuscinoses (NCLs). European Journal of Biochemistry. 2001;268(22):5851-6.

114. Pereira VG, Gazarini ML, Rodrigues LC, Da Silva FH, Han SW, Martins AM, et al. Evidence of lysosomal membrane permeabilization in mucopolysaccharidosis type I: rupture of calcium and proton homeostasis. Journal of cellular physiology. 2010;223(2):335-42.

115. Yambire KF, Rostosky C, Watanabe T, Pacheu-Grau D, Torres-Odio S, Sanchez-Guerrero A, et al. Impaired lysosomal acidification triggers iron deficiency and inflammation in vivo. eLife. 2019 2019/12/03;8:e51031.

116. Sternberg Z, Hu Z, Sternberg D, Waseh S, Quinn JF, Wild K, et al. Serum Hepcidin Levels, Iron Dyshomeostasis and Cognitive Loss in Alzheimer's Disease. Aging and disease. 2017;8(2):215-27.

117. Smith MA, Zhu X, Tabaton M, Liu G, McKeel Jr DW, Cohen ML, et al. Increased Iron and Free Radical Generation in Preclinical Alzheimer Disease and Mild Cognitive Impairment. Journal of Alzheimer's Disease. 2010;19:363-72.

118. Crespo ÂC, Silva B, Marques L, Marcelino E, Maruta C, Costa S, et al. Genetic and biochemical markers in patients with Alzheimer's disease support a concerted

systemic iron homeostasis dysregulation. Neurobiology of Aging. 2014 2014/04/01/;35(4):777-85.

119. Damulina A, Pirpamer L, Soellradl M, Sackl M, Tinauer C, Hofer E, et al. Cross-sectional and Longitudinal Assessment of Brain Iron Level in Alzheimer Disease Using 3-T MRI. Radiology. 2020 2020/09/01;296(3):619-26.

120. Kenkhuis B, Somarakis A, de Haan L, Dzyubachyk O, Ijsselsteijn ME, de Miranda NFCC, et al. Iron loading is a prominent feature of activated microglia in Alzheimer's disease patients. Acta Neuropathologica Communications. 2021 2021/02/17;9(1):27.

121. Ayton S, Wang Y, Diouf I, Schneider JA, Brockman J, Morris MC, et al. Brain iron is associated with accelerated cognitive decline in people with Alzheimer pathology. Mol Psychiatry. 2020 Nov;25(11):2932-41. PubMed PMID: 30778133. PMCID: PMC6698435. Epub 20190218. eng.

122. Puy V, Darwiche W, Trudel S, Gomila C, Lony C, Puy L, et al. Predominant role of microglia in brain iron retention in Sanfilippo syndrome, a pediatric neurodegenerative disease. Glia. 2018 2018/08/01;66(8):1709-23.

123. Brady J, Trehan A, Landis D, Toro C. Mucopolysaccharidosis type IIIB (MPS IIIB) masquerading as a behavioural disorder. BMJ case reports. 2013 May 8;2013. PubMed PMID: 23661660. PMCID: PMC3669838. Epub 2013/05/11. eng.

124. Ding Q, Markesbery WR, Chen Q, Li F, Keller JN. Ribosome dysfunction is an early event in Alzheimer's disease. J Neurosci. 2005 Oct 5;25(40):9171-5. PubMed PMID: 16207876. PMCID: PMC6725754. eng.

125. Honda K, Smith MA, Zhu X, Baus D, Merrick WC, Tartakoff AM, et al. Ribosomal RNA in Alzheimer Disease Is Oxidized by Bound Redox-active Iron*. Journal of Biological Chemistry. 2005 2005/06/03/;280(22):20978-86.

126. Parente MK, Rozen R, Cearley CN, Wolfe JH. Dysregulation of Gene Expression in a Lysosomal Storage Disease Varies between Brain Regions Implicating Unexpected Mechanisms of Neuropathology. PLOS ONE. 2012;7(3):e32419.

127. DiRosario J, Divers E, Wang C, Etter J, Charrier A, Jukkola P, et al. Innate and adaptive immune activation in the brain of MPS IIIB mouse model. Journal of Neuroscience Research. 2009 2009/03/01;87(4):978-90.

128. Mahsa T, Erjun Z, Irene L, Sheng-Kwei S, Sophie W, Jonathan DC, et al. Severe Central Nervous System Demyelination in Sanfilippo Disease. bioRxiv. 2023:2023.04.12.536631.

129. Yellajoshyula D, Pappas SS, Rogers AE, Choudhury B, Reed X, Ding J, et al. THAP1 modulates oligodendrocyte maturation by regulating ECM degradation in lysosomes. Proceedings of the National Academy of Sciences. 2021 2021/08/03;118(31):e2100862118.

130. Sloane JA, Batt C, Ma Y, Harris ZM, Trapp B, Vartanian T. Hyaluronan blocks oligodendrocyte progenitor maturation and remyelination through TLR2. Proceedings of the National Academy of Sciences. 2010 2010/06/22;107(25):11555-60.

131. Siebert JR, Osterhout DJ. The inhibitory effects of chondroitin sulfate proteoglycans on oligodendrocytes. Journal of Neurochemistry. 2011 2011/10/01;119(1):176-88.

132. Tamagawa K, Morimatsu Y, Fujisawa K, Hara A, Taketomi T. Neuropathological study and chemico-pathoiogical correlation in sibling cases of Sanfilippo syndrome type B. Brain and Development. 1985 1985/01/01/;7(6):599-609.

133. Barone R, Nigro F, Triulzi F, Musumeci S, Fiumara A, Pavone L. Clinical and neuroradiological follow-up in mucopolysaccharidosis type III (Sanfilippo syndrome). Neuropediatrics. 1999 Oct;30(5):270-4. PubMed PMID: 10598840. eng.

134. Lee S, Viqar F, Zimmerman ME, Narkhede A, Tosto G, Benzinger TLS, et al. White matter hyperintensities are a core feature of Alzheimer's disease: Evidence from the dominantly inherited Alzheimer network. Annals of Neurology. 2016 2016/06/01;79(6):929-39.

135. Prins ND, Scheltens P. White matter hyperintensities, cognitive impairment and dementia: an update. Nature Reviews Neurology. 2015 2015/03/01;11(3):157-65.

136. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. Science. 2015 2015/01/23;347(6220):1260419.

137. Lardelli M, Baer L, Hin N, Allen A, Pederson SM, Barthelson K. The Use of Zebrafish in Transcriptome Analysis of the Early Effects of Mutations Causing Early Onset Familial Alzheimer’s Disease and Other Inherited Neurodegenerative Conditions. Journal of Alzheimer's Disease. 2023;Preprint:1-15.

138. Kizil C, Kaslin J, Kroehne V, Brand M. Adult neurogenesis and brain regeneration in zebrafish. Dev Neurobiol. 2012 Mar;72(3):429-61. PubMed PMID: 21595047. eng.