Journal of Alzheimer's Disease xx (2021) x–xx
DOI 10.3233/JAD-210200
IOS Press

Iron Responsive Element-Mediated Responses to Iron Dyshomeostasis in Alzheimer's Disease

Nhi Hin[a,b,*], Morgan Newman[b], Stephen Pederson[c] and Michael Lardelli[b,*]
[a]*South Australian Genomics Centre, South Australian Health and Medical Research Institute, North Terrace, Adelaide, SA, Australia*
[b]*Alzheimer's Disease Genetics Laboratory, School of Biological Sciences, The University of Adelaide, North Terrace, Adelaide, SA, Australia*
[c]*Dame Roma Mitchell Cancer Research Laboratories, Adelaide Medical School, Faculty of Health & Medical Sciences, University of Adelaide, Adelaide, SA, Australia*

Accepted 17 September 2021

Abstract.
Background: Iron trafficking and accumulation is associated with Alzheimer's disease (AD) pathogenesis. However, the role of iron dyshomeostasis in early disease stages is uncertain. Currently, gene expression changes indicative of iron dyshomeostasis are not well characterized, making it difficult to explore these in existing datasets.
Objective: To identify sets of genes predicted to contain iron responsive elements (IREs) and use these to explore possible iron dyshomeostasis-associated gene expression responses in AD.
Methods: Comprehensive sets of genes containing predicted IRE or IRE-like motifs in their 3′ or 5′ untranslated regions (UTRs) were identified in human, mouse, and zebrafish reference transcriptomes. Further analyses focusing on these genes were applied to a range of cultured cell, human, mouse, and zebrafish gene expression datasets.
Results: IRE gene sets are sufficiently sensitive to distinguish not only between iron overload and deficiency in cultured cells, but also between AD and other pathological brain conditions. Notably, changes in IRE transcript abundance are among the earliest observable changes in zebrafish familial AD (fAD)-like brains, preceding other AD-typical pathologies such as inflammatory changes. Unexpectedly, while some IREs in the 3′ untranslated regions of transcripts show significantly increased stability under iron deficiency in line with current assumptions, many such transcripts instead display decreased stability, indicating that this is not a generalizable paradigm.
Conclusion: Our results reveal IRE gene expression changes as early markers of the pathogenic process in fAD and are consistent with iron dyshomeostasis as an important driver of this disease. Our work demonstrates how differences in the stability of IRE-containing transcripts can be used to explore and compare iron dyshomeostasis-associated gene expression responses across different species, tissues, and conditions.

Keywords: Alzheimer's disease, computational biology, familial Alzheimer's disease, gene expression, iron homeostasis, iron responsive element, transcriptomics

*Correspondence to: Nhi Hin, South Australian Genomics Centre, South Australian Health and Medical Research Institute, North Terrace, Adelaide, SA 5000, Australia. Tel.: +61 8 8128 4621; E-mail: nhi.hin@sahmri.com; Michael Lardelli, Alzheimer's Disease Genetics Laboratory, School of Biological Sciences, The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia. Tel.: +61 8 8313 3212; E-mail: michael.lardelli@adelaide.edu.au.

INTRODUCTION

The pathological processes underlying Alzheimer's disease (AD) commence decades before symptoms become evident [1, 2]. Since the early days

of AD research, iron trafficking and accumulation has been observed to be altered in AD [3–7]. However, it is still unclear whether iron dyshomeostasis represents a late symptom or an early pathological driver of AD. Iron homeostasis is closely linked to many critical biological processes including cellular respiration, hypoxia, and immune responses, all of which are disrupted in AD. Consistent with this, discoveries over the past decade have shown that disruptions to iron homeostasis can drive feedback loops that worsen AD pathology [8,9]. However, evidence for iron dyshomeostasis as an early driver of the disease is only just emerging. While iron accumulation associated with amyloid plaque formation has been observed in various transgenic mouse models of AD [10–13], the age at which iron dyshomeostasis first occurs is uncertain [14].

Currently, gene expression patterns representing responses to iron dyshomeostasis are not well characterized. Cellular responses to iron dyshomeostasis are complex and involve several systems and layers of regulation. The stability of the transcription factor HIF1α (a component of HIF1, a master regulator of responses to hypoxia) is regulated by an iron-dependent mechanism so that transcriptional responses to iron deficiency can resemble hypoxia responses [15]. However, cellular iron homeostasis is also regulated at the post-transcriptional level by the IRP/IRE system [16–18]. In this system, altered levels of available ferrous iron (Fe^{2+}) cause iron regulatory proteins (IRP1 or IRP2) to change conformation or stability respectively [19, 20]. This alters their ability to bind *cis*-regulatory iron responsive element (IRE) stem-loop motifs in the 3′ or 5′ untranslated regions (UTRs) of genes encoding products related to iron metabolism. Only a few IRE-containing genes have been characterized in detail, including transferrin receptor 1 (*TFR1*; 3′ UTR IRE), divalent metal transporter 1 (*DMT1* or *SLC11A2*; 3′ UTR IRE), ferritin heavy chain and ferritin light chain (*FTH1* and *FTL*; both 5′ UTR IRE), and ferroportin (*SLC40A1*; 5′ UTR IRE) [18]. In general, these previously-characterized IRE-containing genes suggest that IRPs binding to 3′ IREs tend to stabilize transcripts to increase protein translation, while binding to 5′ IREs suppresses translation [17]. It is important to note that while the ancestral ferritin IRE is present even in lower metazoans (including sponges), IRE elements in other genes have likely arisen more recently through convergent evolution [21]. Perhaps because of this, the properties of IRE genes can vary in different ways. For example, *SLC11A2* is stabilized by IRP1 binding to a greater extent than by IRP2 binding (as its 3′ IRE has a greater affinity for IRP1) [22]. Similarly, while *CDC42BPA* and *TFRC* both possess 3′ IREs, *CDC42BPA* is stabilized to a greater extent than *TFRC* under iron deficiency [23]. There are also IRE-containing genes that bind only IRP1 or IRP2 exclusively, at least in mice [24]. However, global gene expression changes mediated by genes with IREs have not yet been well-defined, and the overall expression patterns of IRE-containing genes have not been explored in the context of AD. In addition, it is unclear how expression of these genes might differ between AD and other neurodegenerative diseases, or how AD risk factors such as aging and hypoxia might contribute.

In this study, we utilized the SIREs (searching for iron responsive elements) tool [25] to predict and identify sets of IRE-containing genes in human, mouse, and zebrafish. We then applied these gene sets to explore overall IRP/IRE-mediated iron dyshomeostasis-associated responses in datasets involving: 1) a cultured cell line subjected to iron overload and deficiency treatments, 2) a cohort of AD patients, healthy controls, and two other pathological conditions affecting the brain, 3) 5XFAD mice used to model the amyloid and tau pathology seen in AD, and 4) a zebrafish knock-in model possessing a familial AD (fAD)-like mutation.

Our IRE gene sets displayed significant enrichment in AD, the 5XFAD mouse model, and an fAD-like zebrafish model, demonstrating for the first time the early and extensive involvement of IRP/IRE-mediated iron dyshomeostasis-associated gene expression responses in the context of AD. IRE gene sets were sufficiently sensitive to distinguish not only between iron overload and deficiency in a cultured cell line dataset, but also between AD and other pathological conditions affecting the brain (pathological aging and progressive supranuclear palsy), implying that the dysregulation of IRE-containing genes in AD may differ from other conditions. Changes in IRE gene expression were already evident in young adult brains of animal models of AD (3-month-old 5XFAD mice, 6-month-old fAD-like zebrafish). Overall, our observations do not support the current assumption that IRP binding to 3′ IREs generally stabilizes transcripts as, under most conditions, we observed both increases and decreases in the abundances of transcripts with either 3′ or 5′ IREs.

METHODS

Defining IRE gene sets for human, mouse, and zebrafish

We extracted all 3′ and 5′ UTR sequences from the human, mouse, and zebrafish genome assemblies using the Bioconductor packages BSgenome.Hsapiens.UCSC.hg38, BSgenome.Mmusculus.UCSC.mm10, and BSgenome.Drerio.UCSC.danRer11, and gene definitions from the Ensembl 94 release. Note that potential IREs have also been identified within protein coding sequences [24, 26] but were not included in our analyses under the initial assumption that the classic IRE paradigm was valid, as their possible effects on transcript stability are unknown. Each set of UTR sequences was then submitted into the SIREs web server (v.2.0 [25]). The SIREs algorithm assigns quality scores to predicted IREs taking into account whether the sequence is canonical, whether it contains any mismatches and bulges, and the free energy of the secondary structure. Canonical sequences are tagged by SIREs as being high-quality, while IRE-like sequences are tagged as low or medium quality. Given that high-quality IRE predictions miss the majority of true IREs, this enables a more comprehensive sampling of IRE motifs. For human, mouse, and zebrafish, we separately defined the following four gene sets: ***HQ 3′ IREs*** and ***HQ 5′ IREs*** (representing genes with high-quality predicted IREs), along with ***all 3′ IREs*** and ***all 5′ IREs***, which included genes containing any predicted IRE in the respective UTR. Comparisons between gene sets were performed using the UpSetR package (v.1.4.0 [27]) with gene ID mappings between species obtained by BioMart [28].

Over-representation of IRE gene sets in existing MSigDB gene sets

We downloaded the following gene set collections from MSigDB (v.6.0 [29]): Hallmark, C2 (gene sets from online pathway databases and biomedical literature, including KEGG and the REACTOME databases), C3 (motif gene sets based on regulatory targets), and C5 (gene sets defined by Gene Ontology terms). We excluded the following collections from analysis: C1 (specific to human chromosomes and cytogenetic bands, while our analysis involves different species), C4 (computationally-defined cancer-focused gene sets), C6 (oncogenic signatures), and C7 (immunologic signatures). C4, C6, and C7 were not included as the level of detail in the gene sets in these specific collections is more domain-specific rather than broad-level. We used Fisher's exact test to determine whether any IRE gene set was significantly over-represented in each MSigDB gene set. Gene sets were defined as having significant enrichment for IRE gene sets if the FDR-adjusted p-value from Fisher's exact test was below 0.05. UpSet plots were produced using the UpSetR package (v.1.4.0 [27]) while network representations were produced in Gephi (v.0.9.3 [30]). To produce network visualizations, we first exported node and edge tables from R. The nodes table contained the following gene sets: top 15 gene sets (ranked by Fisher's exact test p-value), Hallmark Heme Metabolism gene set, ***all 3′ IREs***, ***all 5′ IREs***, and all genes contained within these gene sets. The edges table contained gene–gene set edges which indicated the gene set(s) that genes belonged to. To create the network plots in Gephi, we used "Force Atlas 2" as the initial layout algorithm, followed by the "Yifan Hu" [31] layout algorithm to improve visual separation between distinct groupings of genes.

Over-representation analysis of transcription factor motifs in IRE gene promoters

Defining promoter regions as being 1500 bp upstream and 200 bp downstream of the transcription start site for each gene, we used the findMotifs.pl script from HOMER (v.4.11) [32, 33] to search for known transcription factor binding site (TFBS) motifs in the promoters of each IRE gene set. The HOMER Motif database contains 363 vertebrate transcription factor binding motifs based on analysis of high-quality public ChIP-seq datasets (http://homer.ucsd.edu/homer/motif/HomerMotifDB/homerResults.html). We considered TFBS motifs as being significantly enriched in a gene set if the FDR-adjusted p-value (q-value as provided in HOMER output) was less than 0.05.

Gene set enrichment testing

We performed all gene set enrichment tests in R v3.6.1 [34] using *fry* [35, 36], *camera* [37], and *fgsea* [38, 39]. For *fry*, and *camera,* we used model fits obtained using *limma* [40, 41], while for *fgsea*, a ranked list was obtained using *moderated t*-statistics taken from *limma*. All genes were used in gene set enrichment tests (i.e., not just DE genes). We combined the raw p-values from *fry, camera*, and *fgsea*

using Wilkinson's method [42] with default parameters, followed by FDR-adjustment. When performing gene set enrichment testing on the MSigDB Hallmark gene sets, we applied FDR-adjustment to combined *p*-values and defined significant enrichment as gene sets having an adjusted *p*-value < 0.05. When performing gene set enrichment on the four IRE gene sets (***all 3′ IREs***, ***all 5′ IREs***, ***HQ 3′ IREs***, ***HQ 5′ IREs***), we applied Bonferroni-adjustment to combined *p*-values to further protect against Type I errors and defined significant enrichment as gene sets having an adjusted *p*-value < 0.05. Depending on the species in the dataset being analyzed, we used the respective IRE gene sets defined for human, mouse, or zebrafish.

Analysis of the Caco-2 cultured cell line dataset

We downloaded processed microarray data from GEO (accession number: GSE3573). This study investigated gene expression responses to iron treatments, including iron deficiency (cells in iron-free medium versus cells cultivated in ferric ammonium nitrate, and iron overload (cells cultivated in DMEM-FBS medium with hemin versus cells in DMEM-FBS medium) [43]. We performed differential gene expression analysis using the "lmFit" and eBayes" functions in *limma* [40]. Genes were defined as differentially expressed when their FDR-adjusted *p*-value < 0.05.

Analysis of the Mayo Clinic RNA-seq dataset

No experiments of human subjects were performed in this study as all human data was accessed from public databases and was anonymous. We downloaded processed CPM count data from Synapse (https://www.synapse.org/#!Synapse:syn5550404). We matched cerebellum and temporal cortex samples by their patient ID, and only retained genes which were present across all samples and patients for which there were both cerebellum and temporal cortex samples ($n = 236$ patients with measurements for cerebellum and temporal cortex, 472 samples in total). We performed analysis using *limma* [40, 41] and determined differentially expressed genes between conditions. In addition, we used the "duplicateCorrelation" function in *limma*, setting the "block" parameter to the patient ID. Genes were considered differentially expressed if their FDR-adjusted *p*-value < 0.05.

Zebrafish husbandry and animal ethics

All zebrafish work was conducted under the auspices of the Animal Ethics Committee (permit numbers S-2017-089 and S-2017-073) and the Institutional Biosafety Committee of the University of Adelaide. Tübingen strain zebrafish were maintained in a recirculated water system.

fAD-like psen1$^{Q96_K97del/+}$ zebrafish

The isolation of the *psen1*$^{Q96\_K97del}$ mutation has previously been described [44]. Zebrafish mutations were only analyzed in the heterozygous state in this study.

Acute hypoxia treatment of zebrafish

psen1$^{Q96\_K97del/+}$ mutants and their wild-type siblings were treated in low oxygen levels by placing zebrafish in oxygen-depleted water for 3 hours (oxygen concentration of 6.6 ± 0.2 mg/L in normoxia and 0.6 ± 0.2 mg/L in hypoxia.

Whole brain removal from adult zebrafish

After normoxia or hypoxia treatment adult fish were euthanized by sudden immersion in an ice water slurry for at least ∼30 s before decapitation and removal of the entire brain for immediate RNA or protein extraction. All fish brains were removed during late morning/noon to minimize any influence of circadian rhythms.

RNA extraction from whole brain

Total RNA was isolated from heterozygous mutant and WT siblings using the *mir*Vana miRNA isolation kit (Thermo Fisher). RNA isolation was performed according to the manufacturer's protocol. First a brain was lysed in a denaturing lysis solution. The lysate was then extracted once with acid-phenol:chloroform leaving a semi-pure RNA sample. The sample was then purified further over a glass-fiber filter to yield total RNA. Total RNA was DNase treated using the DNA-free™ Kit from Ambion, Life Technologies according to the manufacturer's instructions. Total RNA was then sent to the Genomics Facility at the South Australian Health and Medical Research Institute (Adelaide, Australia) to assess RNA quality and for subsequent RNA sequencing (using poly-A enriched RNA-seq technology, and estimated gene

expression from the resulting single-end 75 bp reads using the reference GRCz11 zebrafish assembly transcriptome).

Pre-processing of the fAD-like psen1$^{Q96_K97del/+}$ zebrafish dataset

RNA-seq libraries contained single-end 75 bp Illumina NextSeq reads. We performed quality trimming with *Adapter Removal* using default parameters. Trimmed reads were pseudo-aligned to the reference zebrafish transcriptome using *Kallisto* (v.0.45) [45] and transcript descriptions from Ensembl release 94. The "catchKallisto" function from *edgeR* [46] was used to import and summarize counts from transcript-level to gene-level, with all subsequent analyses performed at the gene-level.

Differential gene expression analysis for fAD-like zebrafish dataset

For differential gene expression analysis, we retained all genes with expression of at least 1 cpm in 4 or more samples, and used voom With Quality Weights to downweight lower quality samples [41]. Contrasts were defined to include all relevant pairwise comparisons between conditions, and genes were considered as differentially expressed using an FDR-adjusted p-value < 0.05. The *limma* model used to create the design matrix for this dataset included biological sex as a variable ($\sim 0 +$ biologicalSex + group, where biologicalSex is either male or female and group refers to the combination of age, hypoxia treatment, and genotype of the samples) to minimize any sex-specific effects in the gene expression analysis. More details regarding the modest effects of biological sex in this dataset are provided in Supplementary Material 1.

Estimation of spliced and unspliced gene expression in fAD-like zebrafish dataset

For spliced transcripts in Ensembl release 94, we additionally defined unspliced genes including intronic regions. Unspliced transcripts were appended to the end of the reference transcriptome and used to build a new Kallisto [45] index. Estimated counts for spliced transcripts and unspliced genes were imported into R using the "catchKallisto" function from *edgeR* [46].

Gene set enrichment tests for other zebrafish datasets

Please refer to [44] for details on RNA-seq data processing and analysis of the non-fAD-like dataset, and ref. [47] (GEO accession: GSE148631) for details on RNA-seq data processing and analysis of the 7-day-old Q96_K97del/+ dataset. In the current work, we used the gene expression counts matrix with *limma*. The voom, design, and contrasts objects produced as part of the *limma* analysis were used for gene set enrichment analysis with the zebrafish IRE gene sets we defined as well as the MSigDB Hallmark gene sets. Significantly enriched Hallmark gene sets had FDR-adjusted p-value < 0.05 while IRE gene sets were considered significantly enriched if the Bonferroni-adjusted p-value < 0.05.

Differential transcript stability analysis

The estimated spliced and unspliced transcript count estimates from *kallisto* [45] were imported into R using the catchKallisto function from *edgeR* [46]. We used *limma* [40] to determine the logFC of spliced transcripts and unspliced transcripts for each comparison. To test for whether there was a significant difference in the logFC of the spliced and unspliced transcripts, we used Welch's t-test with the *s2.prior* values from *limma* as the variances of the spliced and unspliced transcripts. We defined the null (no stabilization of transcript) and alternate (stabilization of transcript) hypotheses for each gene as follows, where s and u refer to the spliced and unspliced versions of a particular gene:

$$H_0 : \log \mathrm{FC_s} = \log \mathrm{FC_u}$$

$$H_a : \log \mathrm{FC_s} \neq \log \mathrm{FC_u}$$

We defined genes with FDR-adjusted p-values < 0.05 as having differential stability.

Checks that observed gene expression differences are not artifacts of differences in proportions of cell types

The Mayo Clinic RNA-seq study, 5XFAD mice, and fAD-like zebrafish datasets are bulk RNA-seq datasets. To confirm that any gene expression changes were likely due to altered transcriptional programs rather than changes in cell type proportions, we compared expression of marker genes for four common neural cell types (astrocytes, neurons,

oligodendrocytes, microglia) in conditions within each dataset. The marker genes for astrocytes, neurons, and oligodendrocytes were obtained from MSigDB gene sets from [48] which were based on studies in mice. The marker genes for microglia were derived from [49] which was based on studies in human and mouse. All gene IDs were converted to human, mouse, or zebrafish Ensembl IDs using BioMart [28] for each dataset. Please refer to Supplementary Material 2 for details.

Expression of previously validated IRE genes in cell line datasets

Gene symbols from Fig. 2 and Supplementary Tables 1 and 2 of Sanchez et al. [50] were matched to corresponding Ensembl gene identifiers using the *biomaRt* R package [28], and filtered to include only genes present in ***HQ 3′ IRE*** and ***HQ 5′ IRE*** gene sets for human and/or mouse. Processed cell line microarray datasets were downloaded from GEO and included a human Caco-2 cell line dataset (GEO accession GSE3573), and mouse splenic B cell dataset (GEO accession GSE77306). The datasets were filtered for validated IRE genes and boxplots were used to visualize expression between groups. *t*-tests were used to test for differences in expression between groups, with p-values ≤ 0.05 considered statistically significant. The Sankey plot was produced using the *networkD3* R package [51].

RESULTS

Defining sets of genes containing iron responsive elements

We utilized the SIREs tool to define species-specific IRE gene sets by searching for IRE and IRE-like motifs in the 3′ and 5′ UTRs of the reference transcriptomes of human (hg38), mouse (mm10), and zebrafish (z11). *SIREs* assigns a quality-score to all detected IREs, with "high-quality" scores corresponding to canonical IREs and "medium-quality" or "low-quality" scores reflecting deviations from the canonical IRE (alternative nucleotide composition in the apical loop, one bulge at the 3′ or one mismatch in the upper stem) that would still produce an IRE-like motif with potential to be functional [25, 52–55]. Figure 1 summarizes the IRP/IRE interaction effects on transcripts and gives examples of canonical and non-canonical IREs.

We defined four gene sets for each species as follows: ***HQ 3′ IREs*** (high-quality predicted 3′ IRE genes), ***HQ 5′ IREs*** (high-quality predicted 5′ IRE genes), **all 3′ *IREs*** (including all low, medium, and high-quality predicted 3′ IRE genes), and ***all 5′ IREs*** (including all low, medium, and high-quality predicted 5′ IRE genes). The size of these gene sets for each species and the overlap present is shown in Fig. 2 and the gene sets are provided in Supplementary Table 1. Overall, searching through human UTRs uncovered the largest number of predicted IRE genes, followed by mouse, and then zebrafish. Overlap between homologous genes from the IRE gene sets of different species was generally poor. The largest sets of genes shared between species were consistently found between human and mouse, which is reflective of their closer evolutionary divergence.

While not many high-quality IRE genes were identified across all three species, the few we identified are consistent with known and well-characterized IREs in the literature [56, 57]. For example, the single ***HQ 3′ IRE*** gene shared between the three species was *TFRC* (transferrin receptor), while the ***HQ 5′ IRE*** genes shared between the three species included *FTH1* (ferritin heavy chain 1) and *ALAS2* (5'-aminolevulinate synthase 2). We note that the scoring of genes for the ***HQ 3′ IRE*** and ***HQ 5′ IRE*** sets is considered stringent (i.e., only canonical IREs). These two gene sets recover known IRE genes in human, including *CDC14A* (3′ IRE; [50]), *CDC42BPA* (3′ IRE; [23, 58]), *SLC40A1* (5′ IRE; [59]), and *FTL* (5′ IRE; [60]). These genes are also recovered in the ***All 3′ IRE*** and ***All 5′ IRE*** sets across mouse and zebrafish too. In addition, the 3′ IRE of *SLC11A2* (also known as *DMT1*) [22] which is only found in mammals, was also recovered in the human ***HQ 5′ IREs*** set and mouse ***All 3′ IREs*** set (but not for zebrafish). The expression of these previously characterized IRE genes across the human, mouse, and zebrafish datasets analyzed in this current work is described in Supplementary Material 3 and discussed further in the analyses we later describe.

IRE gene sets are over-represented within upregulated AD genes, but overall not well-represented in existing gene sets

We explored the biological relevance of the predicted IRE gene sets described above by testing whether genes within them are over-represented in existing gene sets from the Molecular Signatures Database (MSigDB; available at https://www.

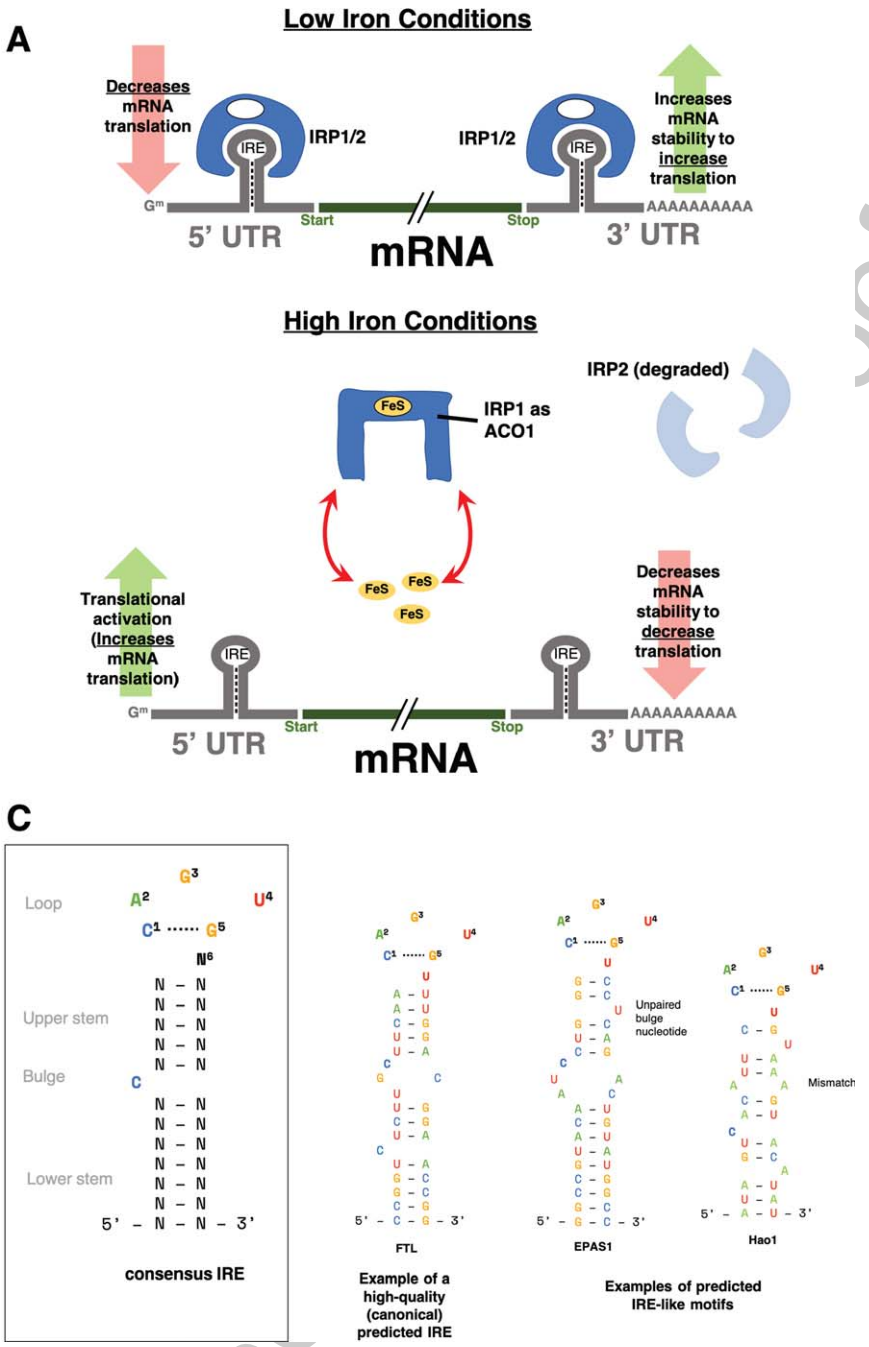

Fig. 1. Iron responsive element (IRE) background. A) The classical paradigm of altered transcript stability due to the IRP/IRE system under iron dyshomeostasis. When cellular iron levels (through FeS) are low, iron regulatory proteins (IRP1 or IRP2) bind IREs in the 5′ or 3′ untranslated regions (UTRs) of genes involved in iron homeostasis. In general, genes with IREs in their 3′ UTRs will be stabilized and hence increased in expression while genes with IREs in their 5′ UTR will have repressed translation from IRP binding. However, when cellular iron levels are high, IRP1 undergoes a conformation change and is now unable to bind IREs. IRP1 now acts as a cytosolic aconitase (ACO1). IRP2 is degraded in high iron conditions and is also unable to bind IREs. Without IRPs binding to IREs, genes with 5′ IREs will have increased translation, while those with 3′ IREs will have decreased mRNA stability, causing their expression to be decreased. B) Consensus IRE secondary structure and examples of high-quality and IRE-like motifs predicted by SIREs. IRE-like motifs with non-canonical structure can be detected by SIREs if they have up to one mismatch pair in the upper stem (e.g., *Hao1*) or 1 unpaired bulge nucleotide on the 3′ strand of the upper stem (e.g., *EPAS1*). For more details on the prediction of non-canonical IRE motifs, please refer to Fig. 1 of Campillos et al. [25].

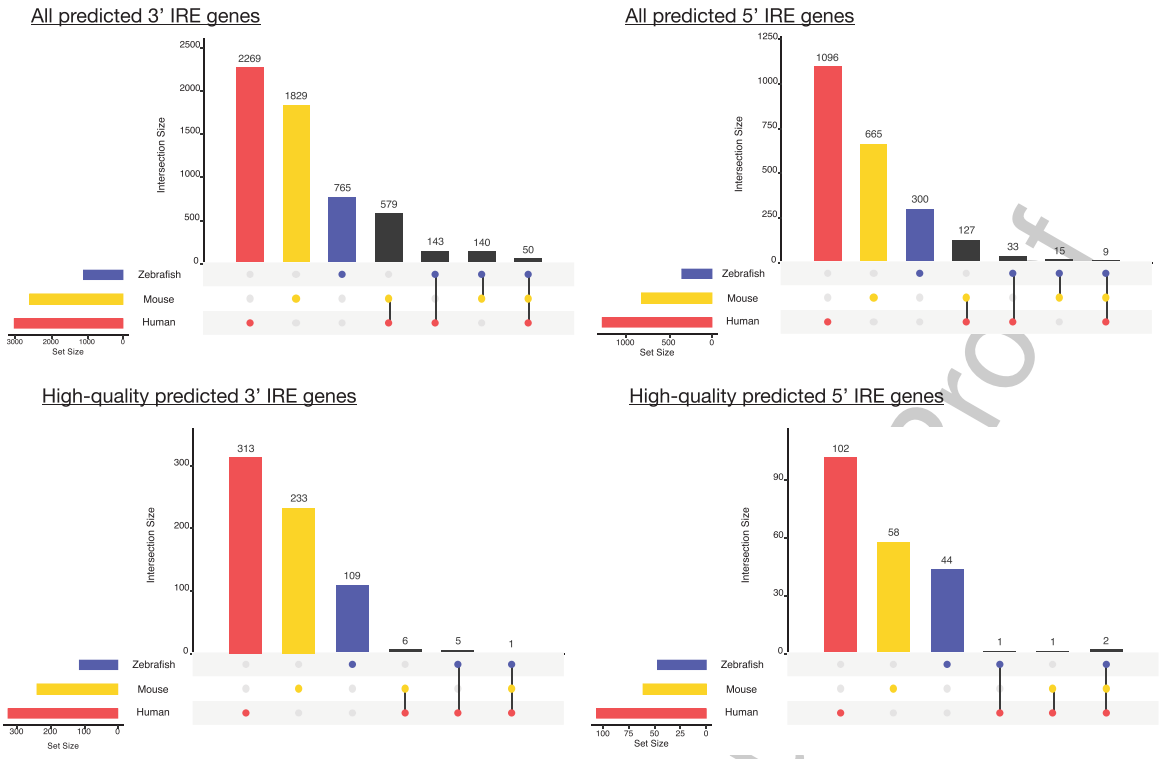

Fig. 2. Overlap between predicted IRE gene sets for human, mouse, and zebrafish. The number of genes in the gene set for each species is shown at the bottom-left bars of each UpSet plot, while genes with shared homologs across species are indicated in the main plot region. IRE genes in mouse and zebrafish gene sets were excluded from this plot if they did not have a human homolog.

gsea-msigdb.org/gsea/msigdb). MSigDB represents one of the largest publicly-available collections of gene sets collated from existing studies and other biological databases [29]. We limited our analysis to gene sets from the following collections: Hallmark (non-redundant sets of ~200 genes each representing various biological activities), C2 (gene sets from databases including KEGG and Reactome and published studies), C3 (gene sets containing genes with motif elements), and C5 (gene sets based on gene ontology terms) (see Methods). We performed over-representation analysis for the predicted IRE gene sets (***all 3′ IREs*** and ***all 5′ IREs***) separately for each species (human, mouse, zebrafish), and used Wilkinson's meta-analytic approach to determine gene sets that were significantly over-represented across the three species' IRE gene sets. Our results indicated that 1,148 of 10,427 tested gene sets displayed over-representation of IRE gene sets across all three species (Bonferroni-adjusted Wilkinson's p-value < 0.05) (Fig. 3; Supplementary Table 2). The bar plots in Fig. 3 show the 15 MSigDB gene sets with the most significant enrichment for human, mouse, and zebrafish. Remarkably, the "Blalock Alzheimer's Disease Up" gene set from the C2 collection was significantly over-represented in both the sets of ***all 3′ IREs*** (Bonferroni-adjusted Wilkinson's p-value = 3.9e–14) and ***all 5′ IREs*** (Bonferroni-adjusted Wilkinson's p-value = 2.1e–26) in the meta-analysis, and was also the gene set with the most significant over-representation of the human ***all 3′ IREs*** set (Bonferroni-adjusted Fisher's exact test p-value = 7.8e–58) (Supplementary Table 2). This supports that disturbance of iron homeostasis particularly distinguishes AD from other disease conditions and pathways represented within the C2 collection. In addition, the top 15 MSigDB gene sets showing the most significant over-representation generally differed between species. However, in all cases, a large proportion of IRE genes from the predicted IRE gene sets were not contained within any of these top-ranked MSigDB gene sets or within the Heme Metabolism gene set belonging to the Hallmark collection (Fig. 3). This is visualized through the network plots of Fig. 3, which show that the predicted IRE genes we defined are not fully captured by existing gene sets. This suggests that IRE sets may be uniquely useful for investigating gene

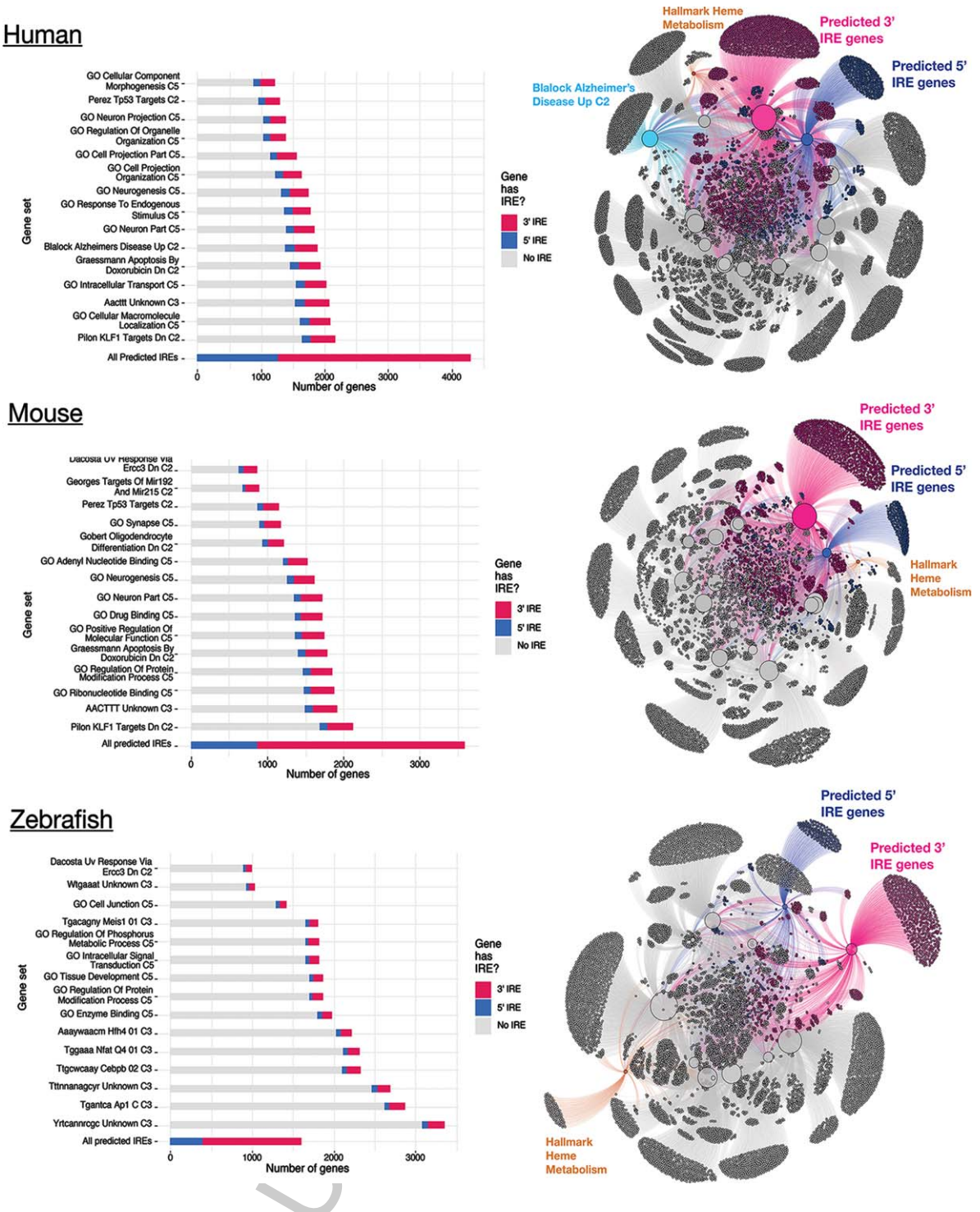

Fig. 3. MSigDB gene sets showing over-representation of predicted 3′ and 5′ IRE genes in human, mouse, and zebrafish. The top 15 MSigDB gene sets ranked by Fisher's exact test *p*-value (testing for over-representation of the ***all 3′ IREs*** and/or ***all 5′ IREs*** sets) are shown for each species. In the network plots, the top 15 MSigDB gene sets are shown as large nodes, with genes represented as small nodes. Edges connecting genes to gene sets indicate the gene set(s) that a gene belongs to. Overall, the ***all 3′ IREs*** and ***all 5′ IREs*** gene sets have a large proportion of genes which are not included in any of the top ranked MSigDB gene sets for each species.

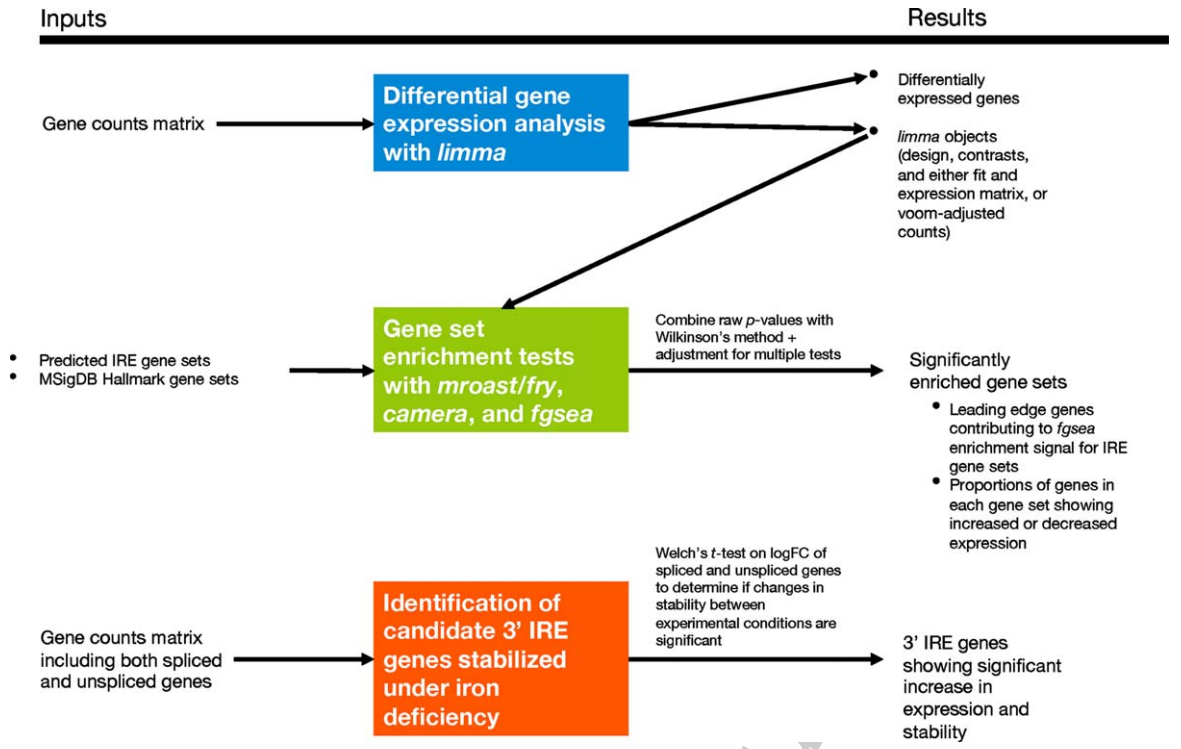

Fig. 4. IRE-containing gene expression analysis workflow. The section including identification of candidate 3′ IRE genes stabilized under iron deficiency was only applied to the fAD-like zebrafish dataset due to unavailable raw RNA-seq reads for the other datasets needed to identify expression of unspliced genes.

expression changes during the IRE-IRP response to iron dyshomeostasis.

Involvement of transcription factor regulation within IRE gene sets

To be confident that IRE gene sets accurately capture information about IRP/IRE-mediated responses, we needed to investigate possible co-ordinate regulation by other factors. As a starting point, we examined whether known binding motifs for transcription factors were significantly over-represented in the promoter regions of genes within each IRE gene set (***all 3′ IREs***, ***all 5′ IREs***, ***HQ 3′ IREs***, and ***HQ 5′ IREs***) for each species. We detected significant over-representation of several transcription factors including the Klf14 motif in the zebrafish ***all 3′ IREs*** set (FDR-adjusted p-value = 0.049), the E2F motif in the human ***all 5′ IREs*** set (FDR-adjusted p-value = 0.049) and the Bach1 (FDR-adjusted p-value = 0.012), Nrf2 (FDR-adjusted p-value = 0.013), and NF-E2 (FDR-adjusted p-value = 0.019) motifs in the zebrafish ***HQ 5′ IREs*** set (Supplementary Table 3). This suggests that the following results should be interpreted with the knowledge that expression of subsets of genes in some of the IRE gene sets may be influenced by factors other than iron ions.

Gene set enrichment testing approach

Our predicted IRE gene sets can be used in any standard gene set enrichment testing approach to detect potential changes in iron homeostasis between conditions. Our workflow, which we later apply to human, mouse, and zebrafish datasets, is shown in Fig. 4. Due to variability in the results produced by different gene set enrichment testing methods, we were inspired by the EGSEA framework [61] to combine the results from different methods. Based on an initial analysis using EGSEA, we chose *fry/mroast* [35, 36], *camera* [37], and *fgsea* [38, 39] as the representative methods to use. (See Supplementary Figure 2 for a principal component analysis of results from different gene set enrichment analysis approaches.) A summary of the different characteristics of the three methods is shown in Table 1. The raw enrichment p-values from these approaches can be combined to obtain an overall enrichment p-value

Table 1
Summary of gene set testing approaches used in our analyses

Method	Hypothesis	Focused / battery approach	Accounts for inter-gene correlation	R implementation
Fry / Roast	Self-contained null hypothesis (genes within a set have no association with experimental condition)	Focused (each gene set is tested on its own terms, and multiple testing adjustment is applied afterwards)	Yes	"fry" and "mroast" functions in the *limma* Bioconductor package [35, 36]
Camera	Competitive null hypothesis (genes within a set do not have a stronger association with experimental condition compared to a random gene set)	Focused (each gene set is tested on its own terms, and multiple testing adjustment is applied afterwards)	Yes	"camera" function in the *limma* Bioconductor package [37]
fgsea (fast implementation of GSEA in R)	Self-contained null hypothesis (genes within a set have no association with experimental condition)	Battery (gene sets are pitted against each other to determine which ones are more significantly enriched)	No	"fgsea" function in the *fgsea* Bioconductor package [38, 39]

for each gene set. In accordance with EGSEA default parameters, we used Wilkinson's method to combine raw *p*-values, and then applied adjustment for multiple testing on all combined *p*-values. Along with performing this gene set enrichment testing on our IRE gene sets, we also recommend using the same method to test for enrichment for the MSigDB Hallmark gene sets as the diverse biological activities they represent help to provide context for interpreting the IRE enrichment results. To explore further the results of IRE gene set enrichment analysis, we use UpSet plots to display the overlap between sets of "leading-edge" genes in the ***all 3′ IREs*** and ***all 5′ IREs*** gene sets. The leading-edge genes can be interpreted as the core (often biologically important) genes of a gene set that account for the significant enrichment as calculated by GSEA [39].

Differences in IRE gene set enrichment during iron deficiency and iron overload in a cultured cell line

We first tested how our enrichment approach illuminated the effects of iron overload and deficiency in a cultured cell line microarray dataset from Caco-2 cells (GEO accession: GSE3573). As only one cell type contributed to this dataset, interpretation of the IRE enrichment results is simplified by not having to consider the differing iron requirements of different cell types. This allowed us to focus on whether the iron dyshomeostasis treatments (iron overload and iron deficiency) could be detected and distinguished in terms of their IRP/IRE system-driven transcript abundance response. Because the dataset was from a microarray experiment, we performed only the differential gene expression and gene set enrichment testing portions of our workflow. The results of principal component analysis and differential gene expression analysis are provided in Supplementary Figure 3. We note that this particular dataset was chosen despite not being derived from a central nervous system cell type, as maintaining iron homeostasis is fundamental to all cell types (regardless of their tissue of origin). In addition, this particular dataset included iron deficiency and iron overload treatments in the same batch and with sufficient sample number, and no other cultured cell line gene expression dataset was publicly available at the time that fulfilled these criteria.

In general, we found iron deficiency and iron overload treatments resulted in different gene expression responses. In terms of classic IRE genes such as *TFRC, FTL,* and *FTH1*, iron deficiency is associated with increased expression of *TFRC* (consistent with increased transcript stabilization of this 3′ IRE gene) and decreased expression of the 5′ IRE genes *FTL* and *FTH1,* and these changes are in the opposite direction during iron overload (see Supplementary Material 3 and Supplementary Figure 1A). Differential gene expression across all detected genes in the dataset additionally revealed that iron deficiency was associated with 96 differentially expressed genes of which 10 possessed predicted IREs while iron overload was associated with 212 differentially expressed genes (FDR-adjusted *p*-value from *limma* < 0.05) of which 33 possessed predicted IREs (Supplementary Figure 3B). There were 17 differentially expressed genes in common between iron deficiency and iron overload treatments, and all moved in opposite directions according to the treatment (i.e., increased abundance

under iron deficiency and decreased abundance under iron overload, or vice versa). These clear differences between iron deficiency and iron overload were reflected in gene set enrichment analyses using the MSigDB Hallmark gene sets, where the gene sets involved and the proportions of gene transcripts with increased or decreased abundance differed (Fig. 5A). As expected, IRE gene sets also showed significant enrichment under iron deficiency and overload conditions (Fig. 5B) (Bonferroni adjusted p-value < 0.05).

We expected that iron deficiency would result in increased expression of genes with 3′ IREs under the IRP/IRE paradigm similar to the *TFRC* gene as seen in Supplementary Material 3. However, both the 3′ and 5′ IRE gene sets displayed mixed patterns of increased and decreased expression under both the iron deficiency and iron overload treatments (Fig. 5B). This indicates it would be difficult to distinguish between these conditions based purely on overall increases or decreases in the expression of IRE gene sets. Despite this, we see that the iron deficiency and iron overload treatments can be distinguished by their "leading-edge" genes (those genes contributing most to the enrichment signal for the predicted IRE gene sets (***all 3′ IREs*** and ***all 5′ IREs***) (Fig. 5C). This supports that gene set enrichment using predicted IRE gene sets may be sufficiently sensitive to detect whether iron dyshomeostasis is present and potentially more informative than individual IRE genes in distinguishing between different IRP-IRE system-mediated gene expression responses in iron deficiency and iron overload treatments in this cell line.

A distinct iron homeostasis response in human AD patients compared to other neuropathologies

Given that IRE gene sets could distinguish between iron overload and deficiency in a cultured cell line, we next tested our gene sets on a more complex dataset including cerebellum and temporal cortex tissue samples from postmortem human brains. The brains originated from either healthy controls or patients afflicted with one of three conditions: AD; pathological aging (PA), a condition involving amyloid pathology but no significant dementia symptoms; or progressive supranuclear palsy (PSP), a tauopathy without amyloid pathology [62]. An important characteristic of the dataset is that both cerebellum and temporal cortex tissue samples were available from each patient. A summary of the 236 patients whose data we included in this analysis is shown in Table 2 and the results of differential gene expression analysis and IRE gene set enrichment analysis are shown in Table 3 and Fig. 6A and summarized in comparison to the mouse and zebrafish datasets later analyzed in Table 5. The expression of classic IRE genes including *TFRC*, *FTL,* and *FTH1* in this dataset is shown in Supplementary Material 3 and Supplementary Figure 1B and 1C. In our analyses, we focus mainly on comparing conditions within each tissue rather than between tissues. This is because we found significant differences in the AD versus control comparison in the temporal cortex compared to the cerebellum (see Supplementary Figure 4).

Overall, our IRE enrichment analyses indicate significant enrichment of IRE gene sets in all pathological conditions (AD, PA, or PSP) compared to healthy controls within the cerebellum and temporal cortex (Fig. 6A). In all pathological conditions, 3′ and 5′ IRE gene transcripts show overall mixed patterns of abundance (e.g., increased and decreased). Further examination of the leading-edge genes from these IRE gene sets gives more insight regarding potential differences and similarities between AD, PA, and PSP (Fig. 6B). Overall, AD, PSP, and PA appear to involve distinct yet partially overlapping IRE gene expression responses in a tissue-specific manner. Within the temporal cortex, there are 435 3′ IRE leading-edge genes and 178 5′ IRE leading-edge genes exclusively present in the "AD versus control" comparison. These are greater numbers than for the other two conditions. Interestingly, AD and PA share only relatively small numbers of leading-edge genes despite the fact that many regard PA as a prodrome of AD [63]. Also interesting is that, in the cerebellum, while AD and PSP share many 3′ and 5′ IRE leading-edge genes, PA is associated with a large number of unique leading-edge 3′ and 5′ IRE genes, further emphasizing its difference from AD and also PSP. These observations are supported when considering expression of previously characterized IRE genes individually (see Supplementary Material 3 and Supplementary Figure 1B, C). The results in this Supplementary Material suggest that, in particular, some 3′ IRE genes (e.g., *TFRC*, *CAV3*, *CDC14A*, *CDC42BPA, SERTAD2, TRAM1*) show increased expression in AD (relative to controls) whereas PA and PSP do not. Overall, our results suggest that IRP-IRE system-mediated gene expression changes may be sufficiently sensitive to discern and identify potential differences in the way iron homeostasis is regulated in AD compared to PA and PSP.

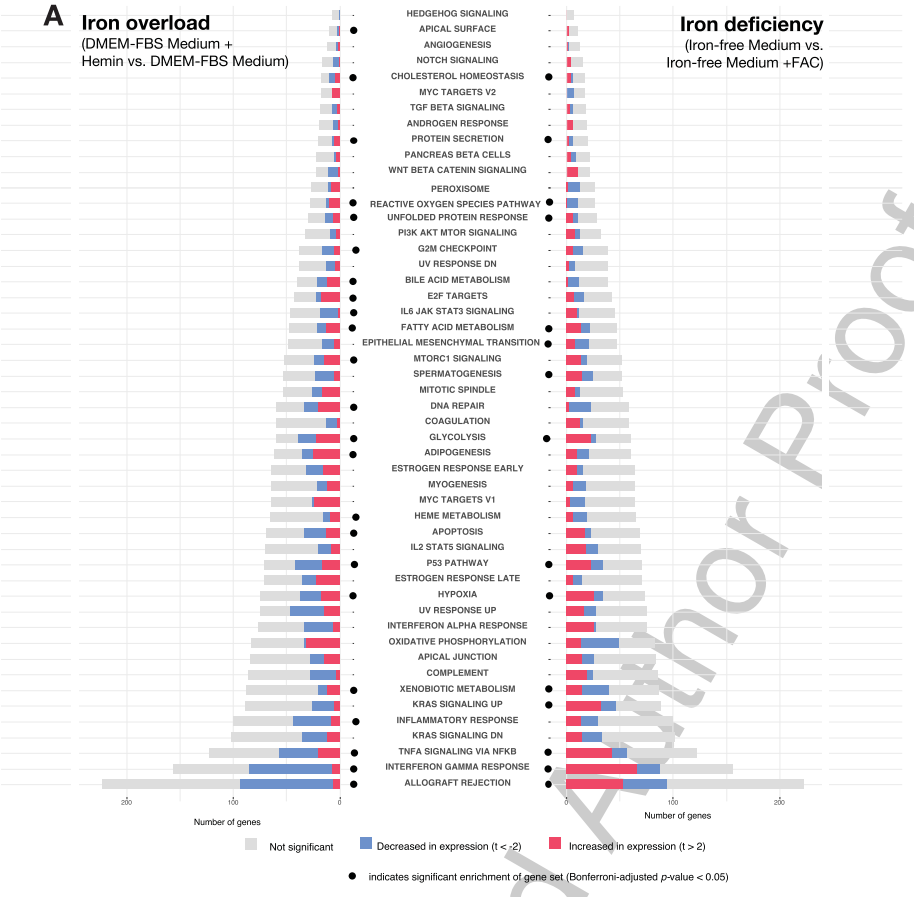

Fig. 5. *(Continued)*

Fig. 5. Analysis of Caco-2 cultured cell line microarray dataset. A) Gene set enrichment testing results for human predicted IRE gene sets for the iron overload and iron deficiency treatments. Dots indicate if a gene set was considered significantly enriched (Bonferroni-adjusted p-value <0.05) in the iron overload (on left) or iron deficiency (on right) treatment. B) Gene set enrichment testing results for MSigDB Hallmark gene sets in the iron overload and iron deficiency treatments. Dots indicate if a gene set was considered significantly enriched (FDR-adjusted p-value <0.05) in the iron overload (on left) or iron deficiency (on right) treatments. C) UpSet plots showing overlap between iron overload and iron deficiency treatments in GSEA leading-edge genes for the *All 3′ IREs* and *All 5′ IREs* gene sets. The bars to the lower-left indicate the number of the leading-edge genes for iron overload and iron deficiency treatments, while the bars in the main plot region indicate the number of leading-edge genes which are unique or shared between the treatments.

Table 2
Samples analyzed from Mayo Clinic RNA-seq study

Group	Total number	% Female	Mean age ± s.d. (y)	Diagnosis
Alzheimer's disease (AD)	134	56.7	82.6 ± 7.3	Braak Stage ≥ 4, diagnosis according to NINCDS-ADRDA criteria
Control	130	46.2	82.9 ± 8.3	Braak Stage ≤ 3, No or sparse CERAD neuritic and cortical plaque density, no diagnoses for any neurodegenerative disease
Pathological aging (PA)	44	54.5	84.5 ± 4.3	Braak Stage ≤ 3, CERAD neuritic and cortical plaque density of 2 or more, no diagnoses for any neurodegenerative disease or mild cognitive impairment
Progressive supranuclear palsy (PSP)	164	40.2	73.8 ± 6.5	Braak Stage ≤ 3, no or sparse CERAD neuritic and cortical plaque density

Table 3
Differential gene expression and IRE gene set enrichment results from Mayo Clinic RNA-seq study.

	Comparison	No. of DE genes (FDR $p<0.05$ and abs(logFC) >0.5)	*All 3′ IREs* enrichment *p*	*All 5′ IREs* enrichment *p*	*HQ 3′ IREs* enrichment *p*	*HQ 5′ IREs* enrichment *p*
Effect of AD	AD versus control (cerebellum)	Down: 1,422 Up: 1,308	0.0109*	0.128	0.770	0.770
	AD versus control (temporal cortex)	Down: 1,650, Up: 1,929	9.03E-15*	2.44E-14*	4.50E-15*	8.95E-08*
Effect of pathological aging (amyloid pathology, no dementia symptoms)	PA versus control (cerebellum)	Down: 254, Up: 463	0.000116*	0.00139*	0.0201*	0.770
	PA versus control (temporal cortex)	Down: 466, Up: 512	0.000521*	0.0461*	0.000065*	0.0000682*
Effect of PSP	PSP versus control (cerebellum)	Down: 2,550, Up: 1,080	1.81E-22*	3.45E-22*	1.43E-12*	0.0399*
	PSP versus control (temporal cortex)	Down: 1,271, Up: 622	3.58E-11*	5.17E-13*	1.39E-06*	0.0399*
Differences between AD and pathological aging	AD versus PA (cerebellum)	Down: 1,728, Up: 1,006	0.000116*	0.00139*	0.0201*	0.770
	AD versus PA (temporal cortex)	Down: 2,633, Up: 2,174	0.000521*	0.0461*	0.000065*	0.0000682*
Differences between AD and PSP	AD versus PSP (cerebellum)	Down: 108, Up: 659	0.0019*	0.00184*	0.158	0.158
	AD versus PSP (temporal cortex)	Down: 2,052, Up: 2,354	9.12E-22*	2.10E-22*	5.35E-16*	0.040*
Differences between cerebellum and temporal cortex tissue	Cerebellum versus temporal cortex in controls	Down: 10,925, Up: 9,302	0.00E+00*	0.00E+00*	2.82E-269*	1.60E-81*
	Cerebellum versus temporal cortex in AD	Down: 11,966, Up: 9,950	0.00E+00*	0.00E+00*	7.71E-302*	9.35E-92*
	Cerebellum versus temporal cortex in PA	Down: 10,200, Up: 8,737	1.87E-153*	5.88E-160*	1.22E-89*	1.31E-88*
	Cerebellum versus temporal cortex in PSP	Down: 11,552, Up: 9,576	0.00E+00*	0.00E+00*	0.00E+00*	4.74E-104*

Asterisks (*) indicate significant enrichment (Bonferroni-adjusted Wilkinson's p-value from *fry*, *camera*, and *fgsea* <0.05). AD, Alzheimer's disease; PSP, progressive supranuclear palsy; PA, Pathological aging

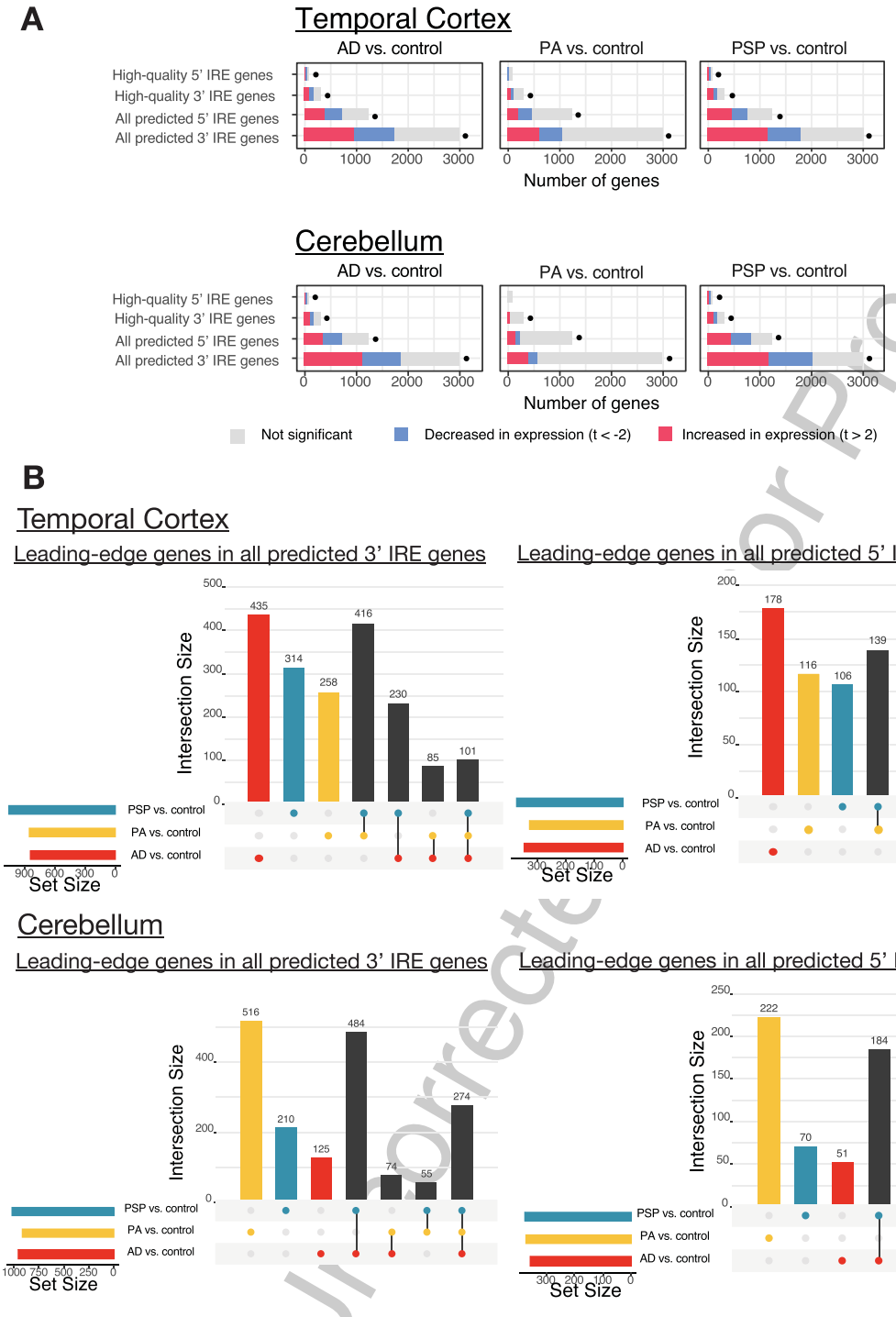

Fig. 6. Analysis of Mayo Clinic RNA-seq dataset with human IRE gene sets. A) Enrichment analysis results of IRE gene sets in comparisons of AD (Alzheimer's disease), PA (pathological aging), and PSP (progressive supranuclear palsy) versus control. The proportions of IRE genes within each IRE gene set with increased, decreased, or unchanged expression are indicated as colored bars. The dots indicate that the IRE gene set was significantly enriched (Bonferroni-adjusted enrichment p-value < 0.05). B) UpSet plots showing overlap between leading-edge IRE genes in comparisons of AD, PA, and PSP versus control. The bars to the lower-left region of each UpSet plot indicate the number of leading-edge IRE genes for each comparison, while bars in the main plot region indicate the number of leading-edge IRE genes which are unique or shared between comparisons.

Age-dependent disruption of IRE-driven iron homeostasis in the 5XFAD mouse model

AD is thought to develop over decades [64–66]. However, detailed molecular studies of AD brains can only be based on postmortem tissues. To reveal the early molecular changes that initiate the progression to AD, we must study the brains of animal disease models. Given that the IRE gene sets appear to work well in human datasets, we then tested our mouse IRE gene sets on an RNA-seq dataset derived from brain cortex tissue of the 5XFAD transgenic mouse AD model (GEO: GSE140286). The 5XFAD mouse is one of the most common systems used to model the amyloid-β histopathology of AD brains and, in a concordance analysis, brain gene expression patterns in this model were previously shown to resemble those of human AD brains to a greater degree than for four other mouse models (3xTg, CK-p25, J20, and Tg2576) [67]. 5XFAD mice possess two transgenes that include a total of five different mutations, each of which individually causes fAD in humans. In this dataset, the mice analyzed were either 3, 6, or 12 months of age. We acknowledge that the small sample size of this particular dataset means that results should be interpreted tentatively (see Discussion).

Using gene set enrichment testing methods as before, we observed significant enrichment of the ***all 3′ IREs*** and ***all 5′ IREs*** gene sets in several comparisons. These included 5XFAD versus wild type comparisons and wild type aging comparisons (Bonferroni-adjusted enrichment $p < 0.05$) (Fig. 7). Notably, even the youngest age group of 5XFAD mutant mice (3 months) displayed significant enrichment of genes containing 3′ or 5′ IREs compared to age-matched wild types. This is consistent with an enrichment of immune-related Hallmark gene sets that we observed in this age group (see Fig. 7B) and with a previous transcriptome analysis suggesting immune activation in 5XFAD mice as early as 2–4 months [68]. UpSet plots of overlapping leading-edge genes suggest that the IRE responses due to aging and due to the 5XFAD transgenes may involve partially overlapping gene expression changes (Fig. 8). In addition, the UpSet plots reveal subsets containing large numbers of IRE-containing genes which uniquely contribute to enrichment of IRE gene sets in only one comparison (e.g., 256 3′ IRE genes only contribute to enrichment in the "*5xFAD versus WT at 6 months*" comparison). Notably, there are 126 shared 3′ IRE genes contributing to enrichment in the "*5xFAD versus WT at 6 months*" and "*5xFAD versus WT at 3 months*" comparisons, but no genes shared between these comparisons and the "*5xFAD versus WT at 12 months*" comparison. These 126 genes were found to show over-representation of Gene Ontology terms including "Hsp70 protein binding" and "nuclear import signal receptor" (FDR-adj. $p < 0.05$) (Supplementary Table 4), suggesting that subsets of leading-edge genes may have biological relevance, and that these may give insight into the biologically distinct age-dependent iron dyshomeostasis responses caused by the 5XFAD transgenes. Although beyond the scope of our current analysis, further investigation into these subsets of genes may contribute to definition of differences in iron dyshomeostasis responses under different biological conditions of interest.

Similarities in IRE-driven iron homeostasis responses during hypoxia and from a familial AD-like mutation in a zebrafish model

Concerns have been raised over the relevance of transgenic mouse models in modelling the early stages of AD (reviewed by [69]). Meanwhile, knock-in models of fAD (that possess a single, heterozygous mutation in an endogenous gene like the human version of the disease) only exhibit subtle pathological changes with little to no visible amyloid or tau pathology present [70]. Despite the comparatively more subtle pathology seen in heterozygous, single fAD-like mutation knock-in models, since they closely mimic the genetic state of fAD, analysis of their brain gene expression may reveal molecular changes relevant to early fAD states in human brains. We had access to a whole-brain RNA-seq dataset from a knock-in zebrafish model of fAD possessing a heterozygous, single fAD-like mutation in its endogenous *psen1* gene (*psen1*$^{Q96\_K97del}$/+) (henceforth referred to as the "fAD-mutation-like" model in this paper). Previous analysis of a subset of this dataset involving young adult (6-month-old) brains revealed gene expression changes related to energy metabolism and mitochondrial function, especially ATP synthesis and other ATP-dependent functions including lysosomal acidification [44]. (In contrast, the 3-month-old young adult 5XFAD mouse brain dataset is dominated by immune/inflammation responses, Fig. 7B). Considering the critical role of iron homeostasis in energy metabolism, we decided to revisit this zebrafish dataset. We performed IRE gene set enrichment on the entire dataset in order

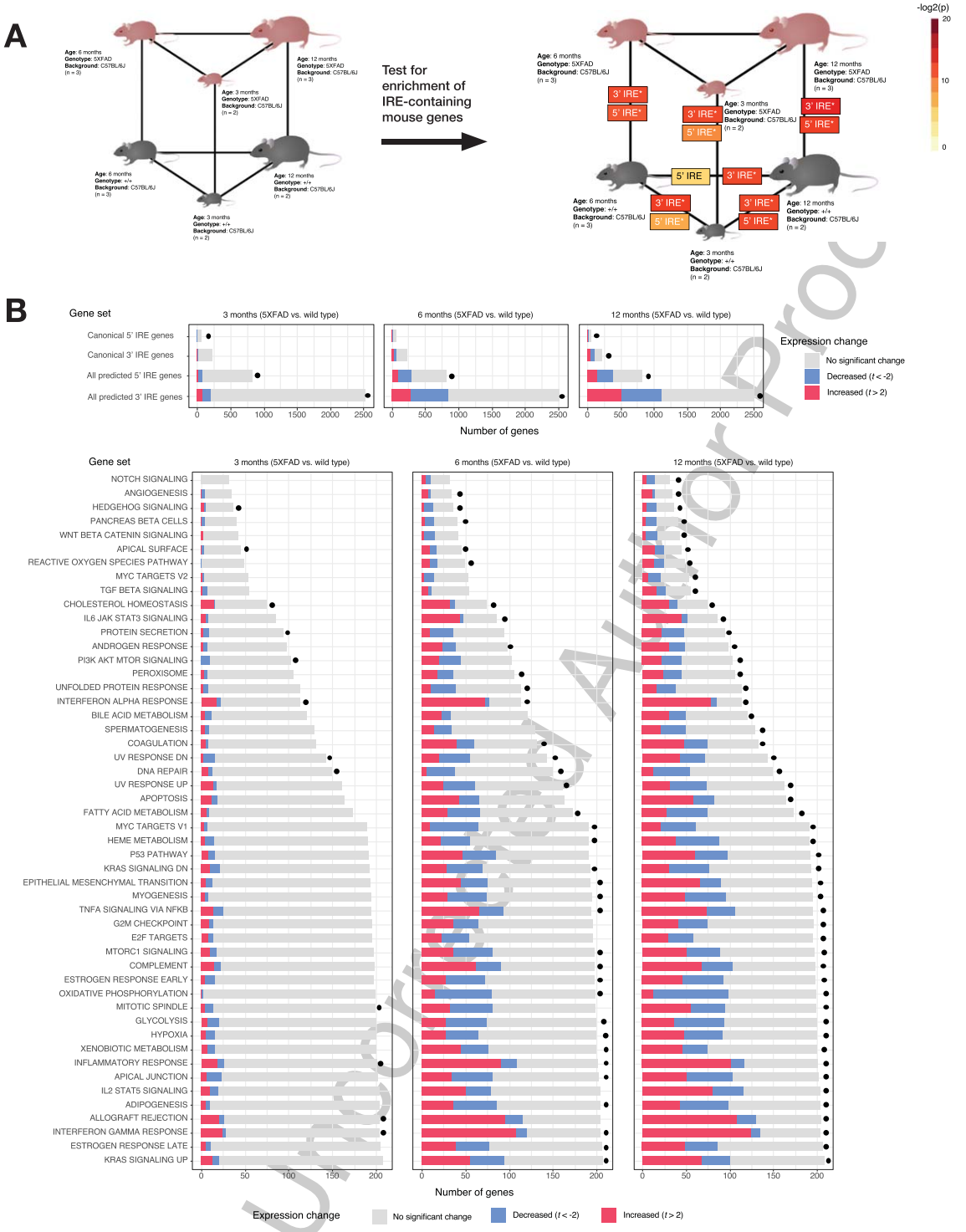

Fig. 7. Analysis of 5XFAD mouse dataset. A) Experimental design and results of IRE gene set enrichment analysis. The gene sets ***all 3′ IREs*** and ***all 5′ IREs*** derived from searching for IRE sequences in the UTRs of genes in the reference mouse genome mm10 are represented here as "3′ IRE" and "5′ IRE" respectively. Asterisks (*) indicate that the gene set was significantly enriched in a particular comparison (Bonferroni-adjusted p-value < 0.05). B) Proportions of genes in IRE and MSigDB Hallmark gene sets which are increased ($t > 2$) or decreased ($t < -2$) in expression in all "5XFAD versus wild type" comparisons. A dot next to a bar indicates that the gene set was significantly enriched (FDR-adjusted p-value < 0.05).

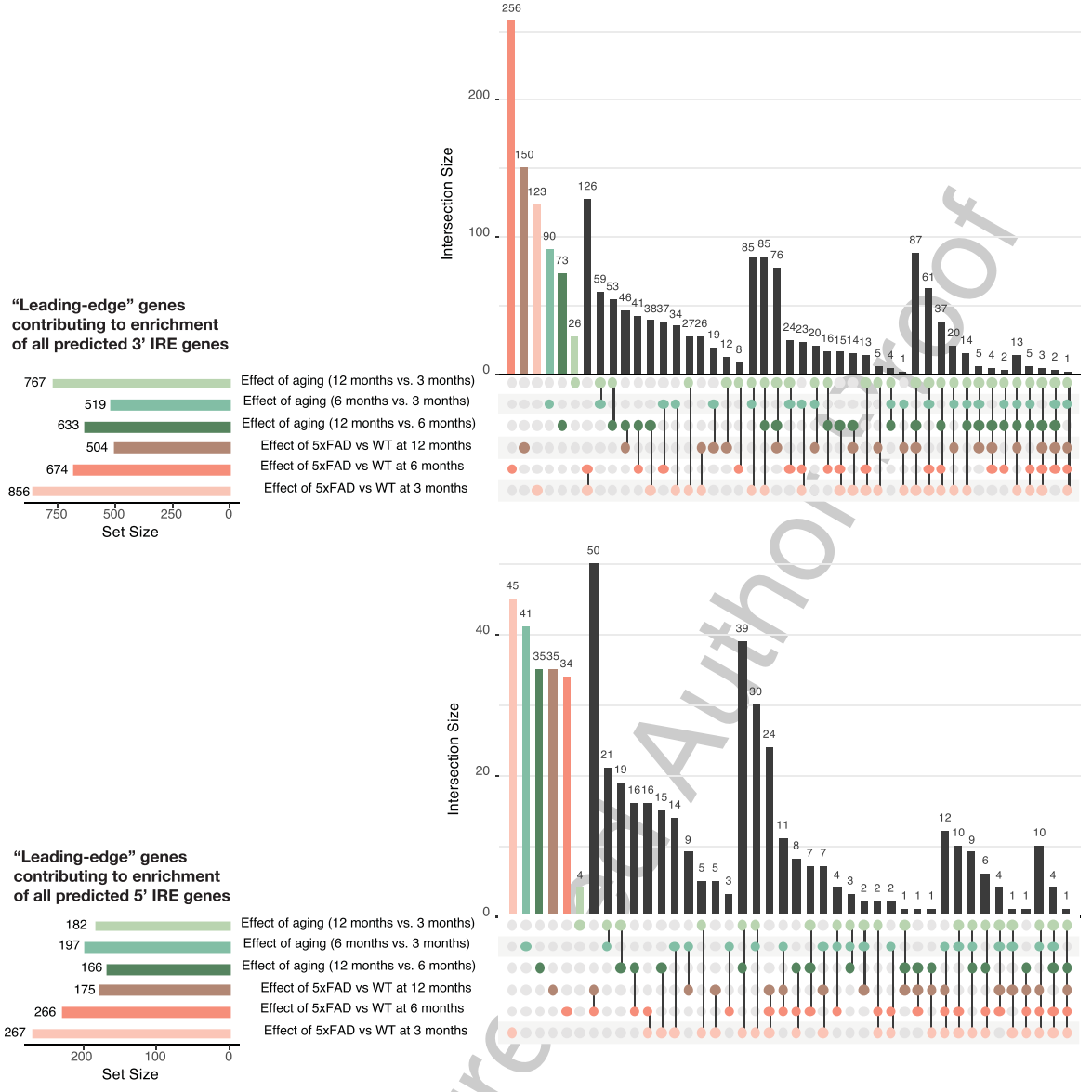

Fig. 8. UpSet plots showing the overlap in GSEA leading-edge genes between all comparisons in the 5XFAD mouse datasets for the gene sets ***all 3′ IREs and*** *all 5′* ***IREs***. Numbers of genes for each intersection are shown above intersection bars.

to include exploration of the effects of aging and acute hypoxia (two important risk factors for sporadic late onset AD) and to analyze how these effects interact with the fAD-like *psen1*$^{Q96\_K97del}$/+mutant genotype. The experimental design and the results of the differential gene expression analyses and gene set enrichment tests are summarized in Fig. 9.

We first turned our attention to the comparison between young adult (6-month-old) *psen1*$^{Q96\_K97del}$/+mutant zebrafish and their wild-type siblings. At this age, gene expression changes in the mutant fish likely represent early stresses driving the development of fAD in humans. Gene set enrichment tests in this comparison identify alteration of energy metabolism-related Hallmark gene sets (e.g., OXIDATIVE PHOSPHORYLATION, GLYCOLYSIS) (Fig. 9B) which is consistent with a previous analysis of gene ontology terms with this dataset [44]. In addition, we see enrichment of other gene sets including FATTY ACID METABOLISM, PI3K AKT MTOR SIGNALLING, MTORC1 SIGNALLING, and HEME METABOLISM. All of

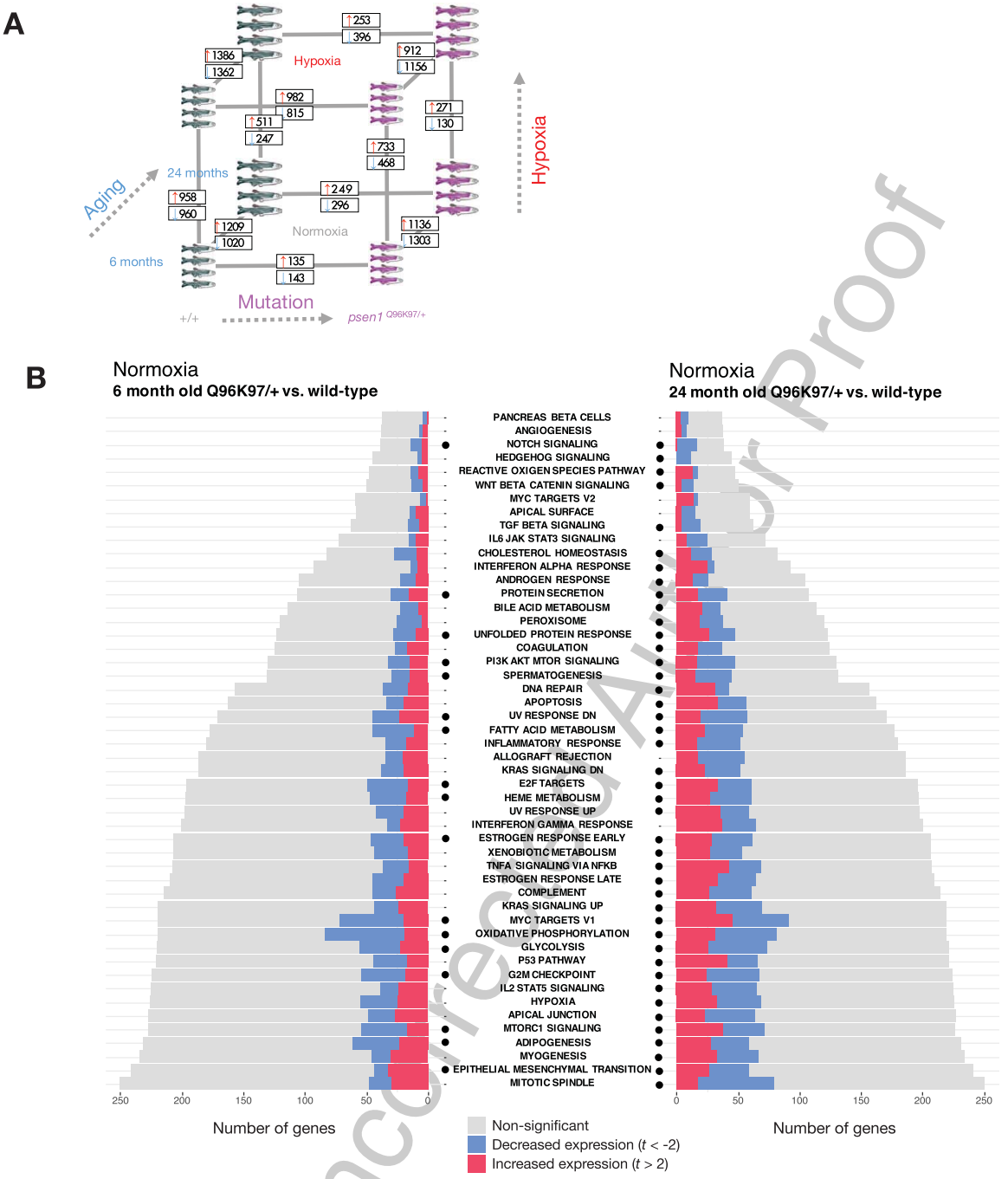

Fig. 9. Differential gene expression and gene set enrichment analysis in the fAD-like zebrafish dataset. A) Results of differential gene expression analysis. Genes which were significantly increased or decreased in expression are indicated in boxes. These differentially expressed genes have FDR-adjusted p-value < 0.05. B) Gene set enrichment with MSigDB Hallmark gene sets. The comparisons between Q96_ K97del/+ fAD-like mutants and their wild-type siblings are shown for the 6-month-old (young adult) and 24-month-old (infertile adult) age groups. Dots indicate gene sets which are significantly enriched (FDR-adjusted p-value < 0.05).

these gene sets showing enrichment in 6-month-old *psen1*$^{Q96\_K97del}$/+mutants relative to wild-type siblings are also enriched in 24-month-old ("middle-aged") *psen1*$^{Q96\_K97del}$/+mutants relative to their wild-type siblings (Fig. 9B). This supports that the biological activities represented in these gene sets

are among the earliest altered in this particular fAD mutation model.

In Lumsden et al. [14], we predicted that fAD mutations in the major locus *PSEN1* might cause an early cellular deficiency of ferrous iron due to the observation of insufficient acidification of the endolysosomal pathway in *in-vitro PSEN1* mutation studies [71,72]. In Newman et al. [44], we saw that Gene Ontology enrichment analysis of 6-month-old *psen1*$^{Q96\_K97del}$/+zebrafish brains supported that lysosomal acidification was affected. Therefore, we applied our IRE enrichment analysis to test for evidence of iron dyshomeostasis in these fish. The enrichment of the ***all 3′ IREs*** gene set in 6-month-old *psen1*$^{Q96\_K97del}$/+zebrafish brains supports that iron dyshomeostasis is an important stress in the early stages of fAD (Fig. 10A).

While almost all pairwise comparisons in the zebrafish dataset show significant enrichment of at least one IRE gene set (Fig. 10A), the expression of IRE genes appears to differ in terms of the proportions of IRE-containing transcripts which show increased versus decreased abundance (Fig. 10B). In addition, the principal component analysis plot of expression of the ***all 3′ IREs*** gene set over all samples shown in Fig. 10C suggests that different conditions appear to drive distinct patterns of expression of these genes. Across the first principal component, different age groups (6- and 24-month-old brains) differ in their expression of predicted 3′ IRE genes, while the second principal component appears to separate *psen1*$^{Q96\_K97del}$/+mutants from their wild-type siblings. This separation between *psen1*$^{Q96\_K97del}$/+mutants and wild type siblings is even more pronounced when both are exposed to hypoxia.

To gain more insight into the similarities and differences between the IRE responses in the *psen1*$^{Q96\_K97del}$/+versus wild type comparison, we plotted UpSet plots of overlapping leading-edge genes (Fig. 10D). These plots suggest that the IRE responses during hypoxia, during aging, and due to this fAD-like mutation, are mostly distinct from each other with unique leading-edge genes. However, similarities in the IRE responses between different conditions are suggested by the existence of some shared leading-edge genes. For example, for the set of ***all 3′ IREs***, the "*6-month-old psen1*$^{Q96\_K97del}$*/+versus wild type*" comparison shares 19 leading-edge genes with the "*24-month-old psen1*$^{Q96\_K97del}$*/+versus wild type*" comparison. As an initial exploration of the biological relevance of these genes, we submitted them to the STRINGR tool. This indicated that the proteins products of these genes were significantly associated with each other (e.g., in terms of text-mining from PubMed articles, experimentally determined interactions, and co-expression). These proteins were significantly over-represented in the gene sets "MAPK pathway", "AP-1 transcription factor", "Jun-like transcription factor", and "signaling by TGF-beta family members" (FDR-adjusted over-representation $p < 0.05$; Supplementary Figure 5). The AP-1 and MAPK pathways have previously been shown to be stimulated by iron depletion [73, 74]. Therefore, mechanistically, it is possible that the IRE-IRP response to iron dyshomeostasis in the fAD-like zebrafish mutant involves these pathways.

The fact that aging, hypoxia, and a fAD-like mutation all cause changes in the abundance of IRE-containing transcripts in zebrafish raised concerns regarding the specificity of such changes. Therefore, as a negative control, we examined changes in IRE transcript abundance in a brain transcriptome dataset derived from zebrafish heterozygous for a *psen1* mutation now thought not to be fAD-like, *psen1*^{K97fs} [75]. In this dataset, *psen1*^{K97fs}/+zebrafish are compared to their wild type siblings at 6 and 24 months. We tested this dataset for enrichment using our zebrafish 3′ and 5′ IRE gene sets but found no significant enrichment of any of our predicted IRE gene sets in *psen1*^{K97fs}/+versus wild type comparisons at any age (Table 4; Supplementary Figure 6). Reassuringly, we still observed significant enrichment of both 3′ and 5′ IRE gene sets during wild type aging (24-month-old wild types versus 6-month-old wild types), consistent with the equivalent comparison in the *psen1*$^{Q96\_K97del}$/+dataset. These results support that IRE-containing transcript abundance changes are sufficiently sensitive to reflect differences in iron homeostasis between different mutation models.

Simultaneous stabilization of some 3′ IRE transcripts and destabilization of others

In the cultured cell line dataset analyzed above, we noticed that even a straightforward iron deficiency treatment resulted in the simultaneous increase and decrease in expression of 3′ IRE-containing genes. These findings are difficult to reconcile with the current paradigm that stabilization of 3′ IRE-containing genes occurs under iron deficiency and suggest that the current model of the IRP/IRE system may be incomplete or insufficient for describing the regulation of a broader collection of IRE genes. Given

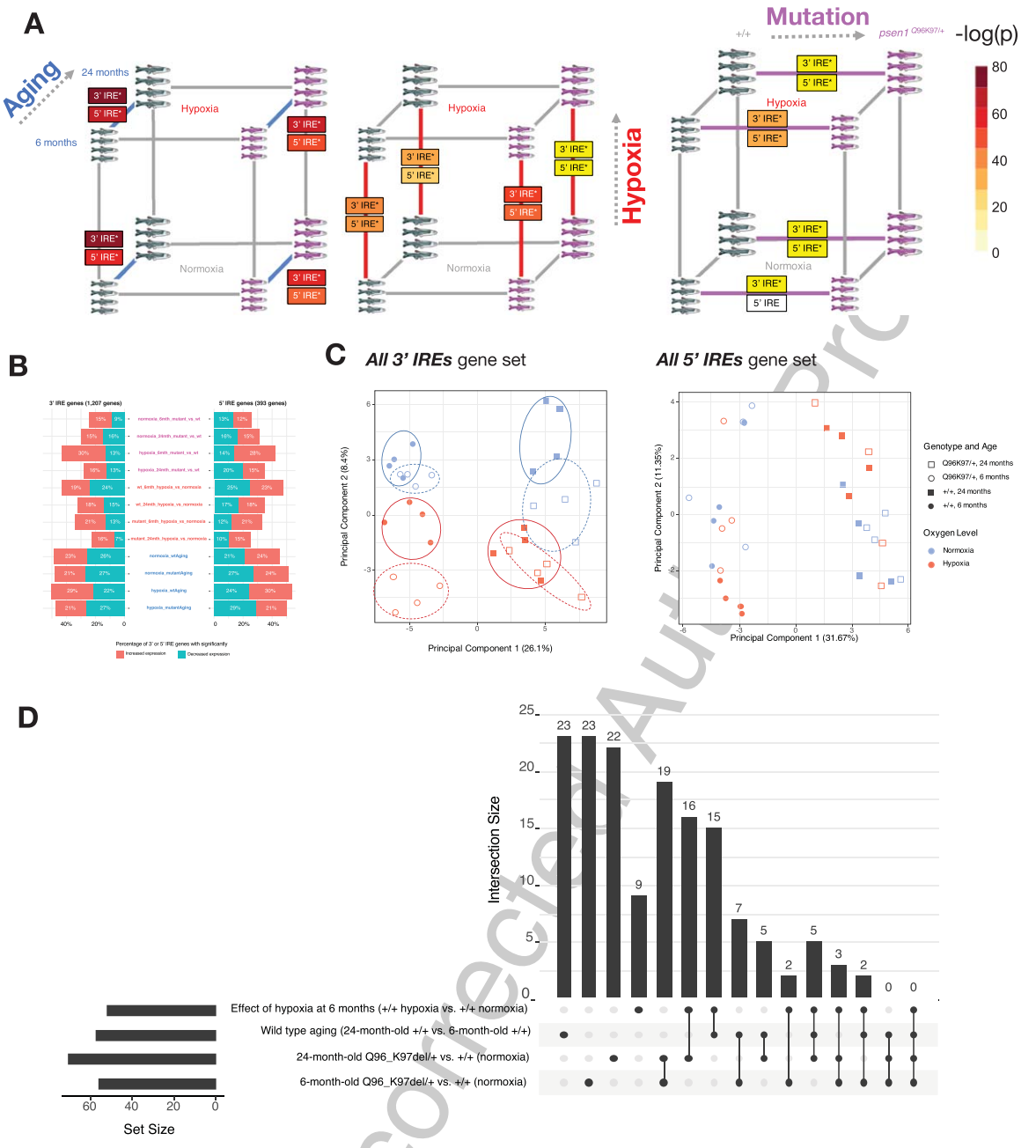

Fig. 10. Iron responsive element (IRE)-containing gene expression in the fAD-like zebrafish dataset. A) Results of gene set enrichment testing using predicted IRE gene sets. We represent the gene sets ***all 3′ IREs*** and ***all 5′ IREs*** derived from searching for IRE and IRE-like sequences from z11 reference zebrafish gene UTRs as "3′ IRE" and "5′ IRE" in the panel. B) Proportions of predicted IRE genes which are increased ($t > 2$) or decreased ($t < -2$) in expression for each pairwise comparison in the dataset. C) Principal component analysis of all genes in the sets ***all 3′ IREs*** and ***all 5′ IREs*** for all samples. Circles on the 3′ IRE plot show that different conditions generally have distinct expression of genes in the ***all 3′ IREs set*** but not in the ***all 5′ IREs*** set. D) UpSet plot showing overlap in leading-edge genes for the ***all 3′ IREs*** gene set for select comparisons.

that many predicted 3′ IRE genes with non-canonical IREs (e.g., in the ***all 3′ IREs*** set) displayed enrichment and different gene expression patterns in the different genotype/age/oxygenation groups of the fAD-like zebrafish dataset, we decided to explore further the stability changes of these genes by comparing the expression of spliced and unspliced transcripts for each gene (Supplementary Material 4). We found

Table 4
Enrichment of Iron Responsive Element (IRE) gene sets in fAD-like zebrafish dataset. Raw *p*-values from *fry*, *camera* and *fgsea* were combined with Wilkinson's method, with combined *p*-values then Bonferroni-adjusted for multiple testing. The same process was repeated for the K97fs/+dataset, which involves an independent family of fish (†)

	Comparison	*All 3′ IREs* enrichment *p*	*All 5′ IREs* enrichment *p*	*HQ 3′ IREs* enrichment *p*	*HQ 5′ IREs* enrichment *p*
Effect of *psen1* mutation	6-month-old Q96_K97del/+ versus 6-month-old +/+ all under normoxia	0.0109	0.128	0.770	0.770
	6-month-old Q96_K97del/+ versus 6-month-old +/+ all under hypoxia	9.03E-15	2.44E-14	4.50E-15	8.95E-08
	24-month-old Q96_K97del/+ versus 24-month-old +/+ all under normoxia	0.000116	0.00139	0.0201	0.770
	24-month-old Q96_K97del/+ versus 24-month-old +/+ all under hypoxia	0.000521	0.0461	0.000065	0.0000682
	6-month-old K97fs/+ versus 6-month-old +/+ all under normoxia†	0.199	0.289	0.0197	0.289
	24-month-old K97fs/+ versus 24-month-old +/+ all under normoxia†	0.116	0.115	0.00000118	0.979
Effect of hypoxia	6-month-old +/+ under hypoxia versus 6-month-old +/+ under normoxia	1.10E-18	1.10E-18	8.19E-13	1.27E-08
	24-month-old +/+ under hypoxia versus 24-month-old +/+ under normoxia	3.67E-15	7.98E-11	2.94E-07	3.31E-05
	6-month-old Q96_K97del/+ under hypoxia versus 6-month-old Q96_K97del/+ under normoxia	1.13E-24	9.88E-22	2.28E-13	2.86e- 3
	24-month-old Q96_K97del/+ under hypoxia versus 24-month-old Q96_K97del/+ under normoxia	0.000521	0.015	0.128	0.770
Effect of aging	24-month-old +/+ versus 6-month-old +/+ all under normoxia	6.80E-35	4.82E-27	3.41E-18	1.26E-21
	24-month-old +/+ versus 6-month-old +/+ all under hypoxia	1.55E-35	6.84E-31	2.36E-20	7.40E-23
	24-month-old Q96_K97del/+ versus 6-month-old Q96_K97del/+ all under normoxia	1.85E-27	1.03E-21	8.09E-17	1.07E-17
	24-month-old Q96_K97del/+ versus 6-month-old Q96_K97del/+ all under hypoxia	3.77E-30	1.59E-23	2.11E-19	2.46E-16
	24-month-old +/+ versus 6-month-old +/+ all under normoxia (K97fs family)†	1.69E-10	3.96E-08	5.95E-10	4.59E-06
	24-month-old K97fs/+ versus 6-month-old K97fs/+ all under normoxia†	0.199	0.289	0.0197	0.289

that transcripts of some predicted 3′ IRE genes were significantly increased in stability while others were significantly decreased in stability (Supplementary Figure 7; Supplementary Table 5).

Some validated IRE genes demonstrate expression changes inconsistent with the classic IRE paradigm

The work described above discovered both increased and decreased stability of transcripts predicted to possess IREs in fAD-like zebrafish brains relative to wild-type brains. However, it did not address directly whether such changes in stability occur in transcripts confirmed as targets of IRP binding and in systems subjected to disturbance of iron homeostasis. To support that our sets of 5′- and 3′-IRE-containing genes include both transcripts that increase in abundance and transcripts that decrease in abundance as iron levels change, we examined several IRE genes previously demonstrated by Sanchez et al. [24] to produce transcripts bound by IRPs (see below). We examined these transcripts in transcriptome datasets produced from cell lines subjected to changes in iron levels (Fig. 11).

In Sanchez et al. [24], a total of 258 mRNAs were identified that immunoprecipitated with mouse IRP1 and/or IRP2 (44 immunoprecipitating with both IRP1 and IRP2, 113 immunoprecipitating exclusively with IRP1, and 101 immunoprecipitating exclusively with IRP2). We filtered these mRNAs for those with high-quality IREs (as predicted using SIREs) in the UTR of either the human or the mouse Ensembl gene model (Fig. 11A). We found 30 mRNAs with high-quality IREs in human Ensembl gene models, and 50 mRNAs with high-quality IREs in mouse Ensembl gene models. These were then separated into those with IREs in their 3′ UTRs and those with IREs in their 5′ UTRs

(see Fig. 11A for more details), since transcripts with 3′ IREs are expected to show altered abundance under the classic IRE paradigm while the abundance of transcripts containing 5′ IREs is not expected to change (see Fig. 1). Lists of experimentally-validated IRE genes filtered using these steps are available in Supplementary Table 6.

We then examined the expression of these experimentally-validated IRE-containing genes in transcriptome datasets from human and mouse cells subjected to iron level-altering treatments. We sought to observe whether transcript abundance differences were consistent with those predicted under the classic IRE paradigm. The datasets included the

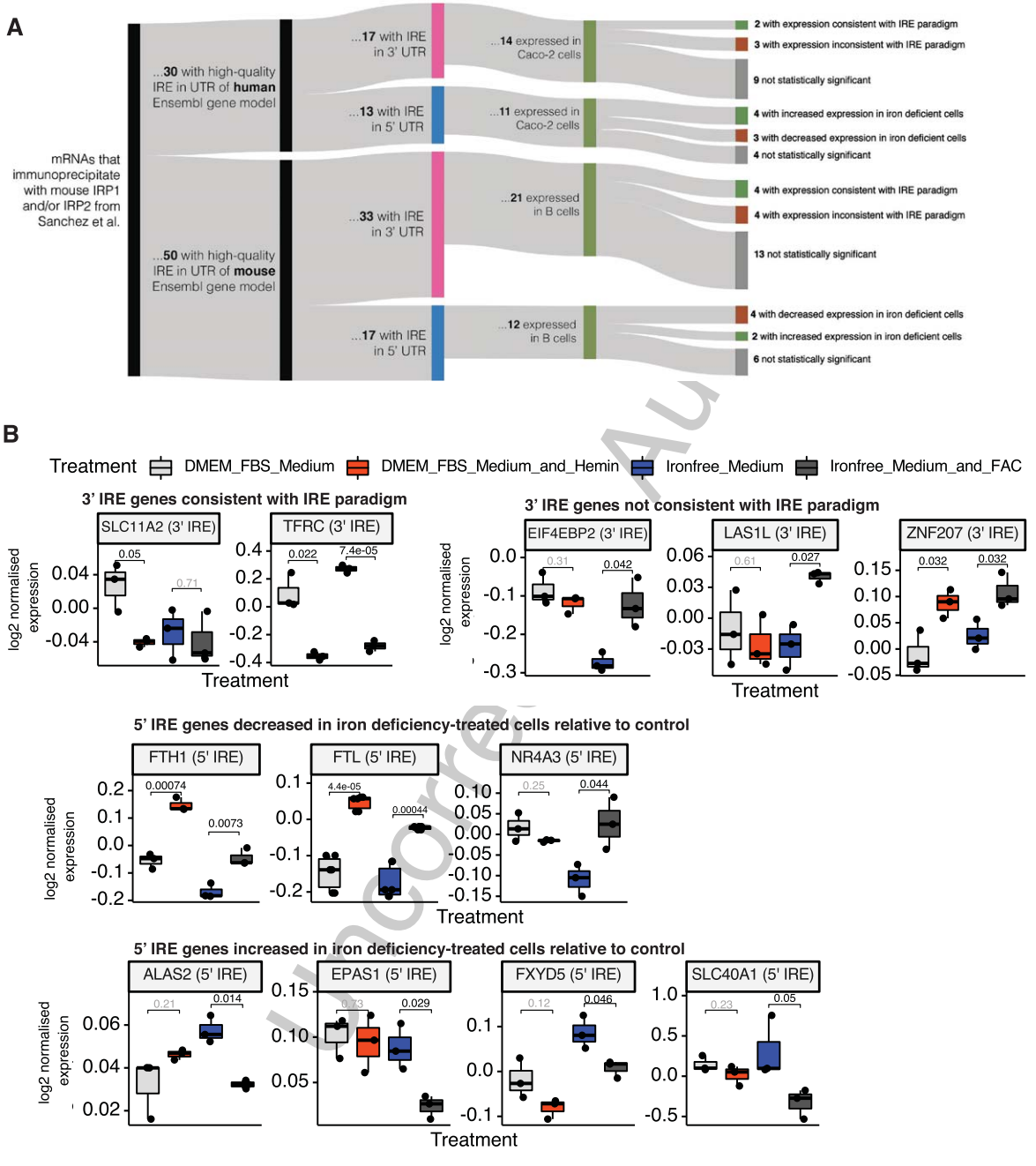

Fig. 11. *(Continued)*

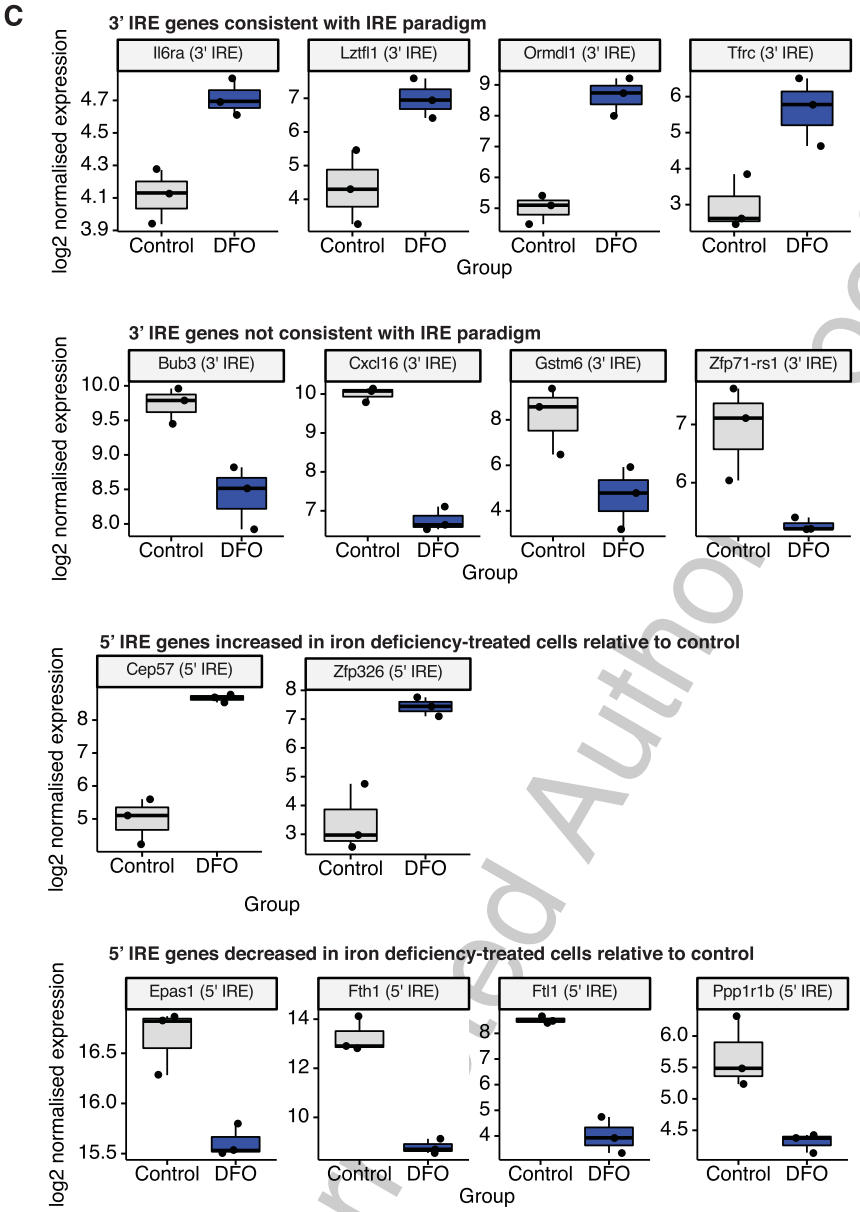

Fig. 11. Expression of validated IRE genes in two cell line datasets. A) Summary of workflow. mRNAs found to bind mouse IRP1 and/or IRP2 via immunoprecipitation in Sanchez et al. [24] were filtered for those which contained a high-quality IRE in their 3′ or 5′ UTR in human and mouse Ensembl gene models. Expression of these validated IRE genes was examined in a mouse splenic B cell dataset and a human Caco-2 cell line dataset. Within each dataset, *t*-tests were used to test for significant expression differences of validated IRE genes between iron deficiency-treated cells and control cells. *t*-test *p*-values ≤ 0.05 were considered to indicate statistically significant differences. Statistically significant 3′ IRE genes were categorized as either "consistent with the IRE paradigm" (increased expression in iron deficiency-treated cells) or "inconsistent with the IRE paradigm" (decreased expression in iron deficiency-treated cells). 5′ IRE genes were not expected to show significant differences according to the IRE paradigm, but 5′ IRE genes showing significant differences in either direction in iron deficiency-treated cells were also recorded and plotted. B) Expression of validated IRE genes showing statistically significant differences in iron deficiency-treated human Caco-2 cells. Iron deficiency-treated cells refers to the "Iron-free Medium" cells relative to "Iron-free Medium + FAC" cells. Numbers over boxplots are *p*-values from *t*-tests between groups. C) Expression of validated IRE genes showing statistically significant differences in iron deficiency-treated mouse splenic B cells. Iron deficiency-treated cells refers to the deferoxamine (DFO)-treated cells relative to control cells. *p*-values are from *t*-tests between DFO-treated and control cells.

human Caco-2 cultured cell line dataset (previously described above), and a mouse splenic B cell dataset [76] (GEO Accession: GSE77306).

The Caco-2 dataset includes both iron overload (DMEM-FBS Medium + Hemin versus DMEM-FBS Medium) and iron deficiency (Iron-free Medium versus Iron-free Medium + FAC) treatments, while the mouse B cell dataset includes only a treatment with deferoxamine (DFO) to create iron deficiency (DFO-treated cells versus control). We note that the DFO and control samples in the mouse B cell dataset are not directly comparable within each group, due to the unstimulated or stimulated anti-IgM or LPS treatments also used in this study. However, a principal component analysis plot of the samples shows that the DFO-treated samples apparently separate from the control samples across Principal Component 2, indicating that DFO-treatment is likely to be one of the main causes of differential gene expression in this dataset (Supplementary Figure 8).

Overall, in the mouse B cells and human Caco-2 cells affected by changes in iron levels, approximately half of the genes identified with validated IREs show significant differential expression. Of those, similar proportions show increased or decreased transcript abundance in either the 3′ IRE or 5′ IRE groups and, of those, there are numerous genes showing changes in expression inconsistent with the IRE paradigm. In particular, in the Caco-2 cells, three 3′ IRE genes (*EIF4EB2*, *LAS1L*, and *ZNF207*) show decreased expression in cells treated to create iron deficiency (Fig. 11B; *t*-test *p*-values ≤ 0.05). In the mouse B cells, DFO-treatment to cause iron deficiency increases the abundance of transcripts of four 3′ IRE-containing genes, including *Tfrc,* in accordance with the classic paradigm (Fig. 11C; *t*-test *p*-values ≤ 0.05). In contrast, transcripts of four such 3′ IRE genes show decreased abundance (*Bub3, Cxcl16*, *Gstm6*, *Zfp71-rs1*) (Fig. 11C; *t*-test *p*-values ≤ 0.05).

Although these sample sizes are small, they clearly indicate that the presence of an IRE in either the 5′ or 3′ UTR of a transcript cannot, by itself, be used to predict how the abundance of the transcript will change in response to iron availability.

DISCUSSION

In this study we identified comprehensive sets of genes predicted to contain IREs in humans, mice, and zebrafish. The idea of analyzing gene expression datasets using gene sets is not novel, with many well-established techniques for gene set testing [35–39] that offer the opportunity to analyze the behavior of large groups of genes to detect statistically significant changes in state, even though these may be quite subtle for individual genes. Table 5 summarizes representative enriched Hallmark gene sets (representing common biological pathways/activities) across the human, mouse, and zebrafish datasets analyzed in this work. Surprisingly, we found that IRE gene sets are generally not well-represented in existing gene sets. The existing gene set in which IRE genes are most highly represented is one previously shown to be upregulated in AD brains (the "Blalock Alzheimer's Disease Up" gene set from MSigDB). This supports a unique importance of IRE genes in AD. Consequently, the IRE gene sets identified here represent a novel resource for researchers interested in exploring iron dyshomeostasis gene expression responses in detail in different species, conditions, and treatments. In our results, we demonstrate that IRE gene sets are sufficiently sensitive to reveal broad differences across different conditions (both in human AD and animal models), while also offering the opportunity to study individual genes/transcripts contributing to enrichment in more detail. IRE gene sets displayed significant enrichment in postmortem brains from a human AD cohort, the 5XFAD mouse model, and a zebrafish model of a fAD-like mutation. Taken together, our results demonstrate, for the first time, the involvement of coordinated IRP/IRE-mediated gene expression responses associated with AD.

Overall, IRE gene sets were sufficiently sensitive to reveal several interesting phenomena. The relatively tightly controlled conditions in the mouse and zebrafish model datasets revealed strongly age-dependent effects on the transcript abundances of these predicted IRE-containing genes. Notably, 3′ IRE gene expression changes were among the earliest changes observable in the zebrafish heterozygous single fAD-like mutation model alongside changes in energy metabolism. These 3′ IRE gene expression changes preceded other signals of pathological change in the transcriptome (such as altered expression of inflammatory response pathways) commonly associated with AD. Our findings raise the question of how early IRE gene expression changes, and iron dyshomeostasis more broadly, may begin in fAD brains. In a recent analysis of transcriptome data from pools of 7-day-old zebrafish larvae heterozygous for the fAD-like mutation *psen1*$^{Q96\_K97del}$ (compared to wild type half-siblings), we observed

Table 5
Summary of enriched gene sets from the MSigDB Hallmark Gene Set Collection in Datasets Analyzed.

Dataset	Comparison	Hallmark Gene Sets Enriched*
Caco-2 cell line (see Fig. 5)	Iron Overload	↓ Interferon Gamma Response
		↑ Reactive Oxygen Species Pathway
	Iron Deficiency	↑ Interferon Gamma Response
		↓ Reactive Oxygen Species Pathway
Human Mayo-Clinic RNA-seq Study (see Supplementary Table 7)	AD versus control (Temporal Cortex)	↓ Oxidative Phosphorylation
		↑ TNFA Signaling via NFKB
		↑ Interferon Alpha Response
		↑ Inflammatory Response
		↑ Interferon Gamma Response
	AD versus control (Cerebellum)	↑ G2M Checkpoint
		↑ Mitotic Spindle
		↑ PI3K AKT MTOR Signaling
		↑ Hypoxia
	PA versus control (Temporal Cortex)	Estrogen Response
		Heme Metabolism
		Complement
		Adipogenesis
		Inflammation
	PA versus control (Cerebellum)	Estrogen Response
		Glycolysis
		Hypoxia
		Unfolded Protein Response
	PSP versus control (Temporal Cortex)	↓ Interferon Alpha Response
		↑ Reactive Oxygen Species Signaling Pathway
		↑ Glycolysis
		↑ Hypoxia
		↑ Estrogen Response
	PSP versus control (Cerebellum)	↑ Myc Targets
		↑ G2M Checkpoint
		↑ Mitotic Spindle
		↑ Heme Metabolism
		↑ Hypoxia
		↓ Fatty Acid Metabolism
Mouse 5XFAD model (see Fig. 7)	5XFAD versus WT (3 months old)	↑ Interferon Gamma Response
		↑ Interferon Alpha Response
		↑ Cholesterol Homeostasis
	5XFAD versus WT (6 months old)	Above also enriched + many other gene sets, see Fig. 7 for more details)
	5XFAD versus WT (12 months old)	Above also enriched + many other gene sets, see Fig. 7 for more details)
Zebrafish *psen1*Q96K97del/+ model (see Fig. 9)	Q96K97del/+ versus WT (6 months)	↓ Oxidative Phosphorylation
		↓ Fatty Acid Metabolism
		MTORC1 signaling
		PI3K AKT MTOR signaling
	Q96K97del/+ versus WT (24 months)	Above also enriched + many other gene sets, see Fig. 9 for more details)

Gene sets shown here are representative enriched gene sets from the MSigDB Hallmark Collection (Bonferroni-adjusted Wilkinson's p-value from *fry*, *camera*, and *fgsea* < 0.05). For full lists of enriched gene sets, see relevant figures or Supplementary Tables). For gene sets which had a significant direction, an up (↑) or down (↓) arrow is used to indicate whether most of the genes in the gene set are increased or decreased in expression. No arrow indicates gene sets involving both increased and decreased expression from different subsets of genes.

over-representation of the GO term "iron ion transport" without significant changes in 3′ IRE or 5′ IRE transcript abundance [47]. This supports that the 3′ IRE transcript abundance signal in 6-month-old brains of this mutant line reflects changes in ferrous iron abundance that develop with age in response to an underlying defect in cellular iron regulation.

In relation to our findings associated with the 5XFAD mouse model, we acknowledge that low sample number is a limitation in this analysis and constrains interpretation of our results. Unfortunately, to our knowledge, no other publicly-available 5XFAD mouse dataset was available containing 1) both young and aged adults in the same batch (where young mice

are defined as ∼2–4 months old as this is when the earliest signs of pathology are observed for this model [77, 78]), 2) 5XFAD and WT genotype mice in the same batch, 3) information about biological sex for each sample (so that sex-specific effects could be accounted for in the statistical model for differential gene expression), and 4) separation of groups by biological factors (e.g., age, genotype) rather than by unidentifiable batch effects when inspected through principal component analysis. Our original rationale for analyzing the 5XFAD mouse model came from results from an independent concordance analysis of various mouse models and human AD [67]. In that study, it was found that 5XFAD mouse brain transcriptomes showed greatest (although still fairly poor) concordance to human AD compared to other mouse models including the 3xTg model. However, we also acknowledge the limitations associated with only analyzing the 5XFAD mouse model here. It would be useful to analyze other mouse models of AD in the future to compare the involvement of IRE genes across different models. Likewise, we anticipate that further analysis of fAD-like mutation models in general will help to extend our knowledge of the early molecular changes contributing to the development of AD.

Notably, both classic IRE genes (e.g., *TFRC*, *FTL*, and *FTH1*) and the IRE gene sets were sufficiently sensitive to distinguish not only between iron overload and deficiency in a cultured cell line dataset, but also between AD and other neuropathologies (i.e., PA and PSP). This suggests that the dysregulation of IRE-containing genes, (and iron homeostasis in general), in AD may differ from other neuropathological conditions. The differences in both classic IRE genes and IRE gene set expression between AD and PA are of particular interest. Considering that presence of amyloid pathology is required to diagnose individuals with AD [79,80] and contributes to diagnosis of preclinical AD [1, 2], it will be interesting to explore further the biological relevance and impact of these IRE gene expression differences using other techniques to gauge whether PA is a prodrome of, or a separate pathological condition to, AD.

One of our original intentions for IRE gene set analysis of the human AD and animal model datasets was to gain insight into whether these brains were in a state of iron deficiency or overload using gene expression-level evidence. Unfortunately, it became clear that interpretation of IRE gene expression changes was more complex than expected, both because of the use of bulk brain tissue samples in the human and animal model datasets (which obscures differences between cell type and tissue types) and because the regulation of both 3′ and 5′ IRE-containing genes we observed did not follow the expected paradigm. To address the former issue, future work should expand the use of IRE gene sets to single-cell RNA-seq (scRNA-seq) datasets so that the IRE gene expression patterns of different cell types can be deconvoluted. Recently, a large-scale scRNA-seq study of AD brains indicated highly cell-type specific changes related to inflammation during early stages of disease progression within the prefrontal cortex [81], highlighting the importance of studying iron homeostasis responses at the single-cell level in the context of AD. The failure of 3′ IRE genes to follow the expected expression paradigm is a limitation we did not anticipate when we began this work. In particular, our observations indicating that 1) the stability of some 3′ IRE transcripts (rather than their "expression"/"levels" as addressed in preceding analyses) is significantly different under different conditions in zebrafish, and that 2) validated 3′ IRE genes show both increased and decreased expression under iron deficiency in two cell line datasets (suggesting they are not stabilized as predicted by the IRE paradigm), challenge the simplistic generalization that 3′ IREs stabilize transcripts under ferrous iron deficiency.

Previous research found that, under iron deficiency, the transcripts of 3′ IRE-containing genes such as *TFRC* are stabilized causing increased expression [20] to increase ferrous iron availability. However, our findings reveal that this principle (the "classic IRE paradigm") does not necessarily apply to other 3′ IRE-containing genes, including those with canonical IREs (e.g., *SLC11A2* in the Caco-2 cultured cell line), and genes with IRE-like elements deviating from the canonical IRE sequence. Interestingly, Tybl et al. [82] found that the 3′ IRE in *SLC11A2* (*DMT1)* appears to influence transcript abundance in a development-specific manner in mice, by promoting transcript expression during postnatal growth and suppressing expression in adulthood. Together with our findings, this demonstrates that the regulation of IRE genes is more complex than commonly appreciated.

Aside from 3′ IRE genes, 5′ IRE genes also showed unexpected behavior in the datasets we analyzed. In general, the classic IRE paradigm holds that binding of IRPs to 5′ IREs blocks translation but is not expected to change transcript abundance. We initially sought to investigate 5′ IRE transcripts as a negative control group to which changes in the stability of the 3′ IRE transcripts could be compared.

Unexpectedly, our results indicated that 5′ IRE gene sets are significantly enriched under various conditions. Upon reflection, this is not unreasonable, considering that binding of any protein to a transcript must, inevitably, changes the transcript's stability to some degree. However, the direction of that stability change is difficult to predict as there is no generalizable cellular "rule" for the outcome of protein binding. This is consistent with the reality that each gene represents a largely independently evolving system constrained only by cells' requirements for its functionality. For example, inactivation of RNA-binding proteins has been demonstrated previously to cause both up- and down-regulation of different genes simultaneously [83]. In addition, the 5′ canonical IRE and surrounding acute box sequences of the H-ferritin (*FTH*) transcript have been shown to stabilize it, which might appear contrary to the expected function of a 5′ IRE (of decreasing protein translation under iron deficient conditions) [84]. H-ferritin is one of the better-characterized IRE genes, and it is possible that other, yet-to-be characterized regulatory mechanisms involving 5′ IREs could contribute to stability changes in other transcripts. Aside from 3′ IREs, other 3′ UTR regulatory motifs are known that bind to proteins to regulate multiple biological activities (e.g., AU-rich 3′ UTR motifs as described in [85]). Therefore, mechanistically, it is reasonable to assume that IRE-based regulation of transcript stability may be co-opted for other biological activities beyond iron homeostasis (and see below). Alternative splicing of IRE genes may be one mechanism for this, where a non-IRE-containing isoform is expressed in certain tissues or conditions, as is the case for mammalian *DMT-1* (*SLC11A2*) [22]. Differential transcript isoform inclusion of IREs is beyond the scope of our current work but will be important to explore in future. We expect our findings to provide a starting point for further research involving IRE genes and their complex regulation beyond the classic paradigm.

It is important to acknowledge that some of the complexity involved with IRE gene expression, and the large number of genes identified with putative IREs, may reflect the involvement of IREs in functions other than iron homeostasis. These functions include regulation of oxidative stress [86–88], oxidative phosphorylation [89], and cell cycle control [90, 91]. In 2019, it was proposed that the tricarboxylic acid cycle (TCA) may have evolved around the reactivity of iron [92], and TCA intermediates themselves also perform important cellular regulatory functions [93]. In addition, our analyses identified several transcription factor motifs potentially associated with other biological pathways as over-represented in IRE gene sets. Similarly, although beyond the scope of our current work, a regulatory role for micro RNAs (miRNAs) in binding IRE structures is emerging [94–97]. This may be particularly relevant to AD, as a miRNA has been described that can bind the 5′ IRE-like structure in AβPP transcripts to upregulate AβPP its expression [98]. Co-regulation of IRE-containing genes within other biological pathways may explain some of the IRE gene set enrichment observed in the datasets we analyzed, and some of the unexpected expression changes (e.g., decreased expression of 3′ IRE genes in conditions where stabilization of their transcripts might be expected, and increased expression of 5′ IRE genes even when no change in stability would be expected). Hence, the changes in expression of predicted IRE genes we observed must be interpreted carefully as reflecting not only iron dyshomeostasis responses, but also other co-occurring changes in linked processes. This was particularly evident in our analysis of the Caco-2 cultured cell line dataset where even straightforward iron overload and iron deficiency treatments resulted in significant enrichment of gene sets related to reactive oxygen species and hypoxia pathways. We note that the relationship between iron homeostasis and hypoxia would be interesting to explore in future work, particularly their similarities at the level of gene expression. Previous research has demonstrated that desferrioxamine (that induces iron deficiency) can also induce HIF-1 activity, which activates erythropoietin (important for red blood cell formation to carry oxygen) [99]. This supports that iron deficiency and hypoxia involve similar molecular changes in cells (i.e., the view that iron deficiency generates a form of "pseudohypoxia"). We expect that future work deconvoluting the expression of IRE genes at the cellular level will be required to determine to what extent such expression changes are influenced by cell/tissue type compared to other linked biological activities.

Most previous work has assumed that accumulation of iron in the brain with age (a phenomenon observed broadly across animal phyla [100, 101]) is indicative of cellular iron overload driving oxidative stress [102]. Indeed, iron accumulation and oxidative stress apparently contribute to the insolubility of amyloid plaques and neurofibrillary tangles in AD brains [103]. However, a recent publication by Yambire et al. [104] showed disturbed

lysosomal function leading to increased lysosomal pH and causing a deficiency of functional ferrous iron (Fe^{2+}) while non-functional ferric iron (Fe^{3+}) accumulated in lysosomes. The deleterious effects of this on the brains of mice (defective mitochondrial biogenesis and function and stimulation of inflammatory responses) could be alleviated by dietary iron supplementation. The observations of Nixon and colleagues that acidification of the endolysosomal pathway is affected both by fAD mutations in *PSEN1* [71] and excessive dosage of the amyloid beta A4 precursor protein (*AβPP*) gene [105], together with our observations from our fAD-mutation-like *psen1*$^{Q96\_K97del}$/+mutant zebrafish in both this study and in [44] (where gene expression changes associated with lysosome and mitochondria for this model were first described), support the possibility that AD brains may suffer a ferrous iron deficiency in a background of ferric iron overload. This is discussed further below.

Although the evidence from human AD brains is more tentative at this stage, our findings of increased expression of *TFRC* expression and other well-characterized 3′ IRE genes in the temporal cortex tissue of human AD brains is consistent with the existence of a cellular ferrous iron deficiency, although more research is required to confirm this at the single-cell level. Intriguingly, the greatest genetic risk factor for late onset AD, the ε4 allele of the gene *APOE*, appears to increase lysosomal pH [106] but ε4′s increased risk of AD is alleviated in individuals who possess the *HFE* 282Y allele that predisposes to the iron overload disease hemochromatosis [107]. Given that many other risk loci for sporadic late onset AD also affect endolysosomal pathway function (reviewed in [108]), it is reasonable to suggest that disturbed iron homeostasis may afflict brains with this disease.

Our work provides evidence consistent with Yambire et al. [104] and Lee et al. [71] to suggest that fAD mutations may drive a ferrous iron deficiency that contributes to AD pathology. However, iron export is a complex process, and it remains unclear whether iron deficiency or overload is more relevant to the pathology seen in AD. Much of the existing research has focused on the role of AβPP and its contributions to elevating iron levels within cells and to increasing iron export to the cerebral interstitium in AD. For example, Venkataramani et al. [109] demonstrated that decreased AβPP levels resulted in stabilization of ferroportin, resulting in increased iron export from neuroblastoma cells in two mouse models. Later work by McCarthy et al. [110] proposed a mechanism for this phenomenon, involving a motif in sAβPPa (cleavage product of AβPP) binding to ferroportin and increasing iron export from cells through increasing the number of ferroportin molecules in the plasma membrane but not their activity. A complicating factor is that ferroportin-mediated iron export from cells requires both (1) stabilization of ferroportin in the membrane, and (2) ferroxidase activity (provided by either hephaestin or ceruloplasmin) [110]. The role of hephestin as a ferroxidase in this process of iron export from cells has sometimes been overlooked, despite that it is well-established as complexing with, and stabilizing, ferroportin to export iron [111–113]). For example, recently Tsatsanis et al. [114] described the effects on intracellular iron and iron export of changes in amyloidogenic cleavage (via β-secretase) and non-amyloidogenic cleavage (via γ-secretase) of AβPP. It was observed that amyloidogenic cleavage of AβPP led to destabilization of ferroportin on the cell surface and increased iron retention in cells. In addition, inhibiting β-secretase mitigated this effect, increasing ferroportin stabilization and decreasing iron retention in cells. However, an alternative explanation not considered by those authors is that hephestin may also be cleaved by β-secretase (see Fig. 7 in [115]), so it is possible that reduced β-secretase activity could increase the hephestin available and also contribute to the effects on iron export observed. Overall, the situation around iron export from cells is a complex area of ongoing research involving many cellular factors, and it will be increasingly important in future work to understand these mechanisms more completely in the context of AD.

We have previously noted that all the fAD mutations of AβPP (including gene duplications) are expected to have the singular quality of increasing the abundance of the amyloidogenic β-CTF/C99 fragment of AβPP (see Fig. 2 in Lumsden et al. [14]). Recently Jiang et al. [105] showed that increased β-CTF reduces lysosomal acidification. According to the results of Yambire et al. cited above, this would be expected to reduce cellular ferrous iron levels (while causing accumulation of ferric iron in lysosomes). This idea is consistent with the existence of an IRE in the 5′ UTR of AβPP transcripts that inhibits AβPP translation (and hence β-CTF production) when cytosolic ferrous iron levels are low [116]. Thus, AβPP may influence cellular ferrous iron levels by modulating lysosomal pH to regulate importation of ferrous iron into the cytosol, rather

than by modulating ferroportin stability to regulate iron export from the cytosol.

Overall, our results demonstrate for the first time the involvement of a coordinated IRP/IRE-mediated gene expression response in the context of AD. By searching entire human, mouse, and zebrafish transcriptomes for all genes containing potential IREs, we formed comprehensive gene sets applicable to the analysis of any human, mouse, or zebrafish gene expression dataset. For the first time, our work suggests the existence of gene expression changes associated with the IRP/IRE system in the young, pre-pathology brains of *PSEN1* fAD mutation carriers, and is consistent with the idea of iron dyshomeostasis as an early contributor to the disease process. It also reinforces the critical importance of IRE genes in general in the development of the late onset, sporadic form of AD. More broadly, our approach highlights how changes in the stability and abundance of IRE-containing transcripts can be used to give insight into iron dyshomeostasis-associated gene expression responses in different species, tissues, and conditions.

ACKNOWLEDGMENTS

The authors thank the Carthew Family and Prof. David Adelson for their encouragement and helpful discussions. This work was financially supported by grants GNT1061006 and GNT1126422 from the National Health and Medical Research Council of Australia (NHMRC) and by funds from the Carthew Family Charity Trust. Lead author NH was supported by an Australian Government Research Training Program Scholarship and by the South Australian Genomics Centre (SAGC) post-graduation. Co-author MN was supported by funds from the grants listed above. Joint senior co-authors SP and ML were/are academic employees of the University of Adelaide respectively. The authors also thank the University of Adelaide's Phoenix HPC team for providing supercomputing resources for this work.

Authors' disclosures available online (https://www.j-alz.com/manuscript-disclosures/21-0200r2).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/JAD210200.

DATASETS AND DATA ARTICLES

The fAD-like zebrafish dataset supporting the conclusions of this article is available in the GEO repository with accession number GSE149149.

REFERENCES

[1] Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Holtzman DM, Jagust W, Jessen F, Karlawish J, Liu E, Molinuevo JL, Montine T, Phelps C, Rankin KP, Rowe CC, Scheltens P, Siemers E, Snyder HM, Sperling R, Elliott C, Masliah E, Ryan L, Silverberg N (2018) NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* **14**, 535–562.

[2] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH (2011) Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 280–292.

[3] Bartzokis G, Sultzer D, Mintz J, Holt LE, Marx P, Kelly Phelan C, Marder SR (1994) *In vivo* evaluation of brain iron in Alzheimer's disease and normal subjects using MRI. *Biol Psychiatry* **35**, 480–487.

[4] Bartzokis G, Tishler TA (2000) MRI evaluation of basal ganglia ferritin iron and neurotoxicity in Alzheimer's and Huntington's disease. *Cell Mol Biol* **46**, 821–833.

[5] Ding B, Chen KM, Ling HW, Sun F, Li X, Wan T, Chai WM, Zhang H, Zhan Y, Guan YJ (2009) Correlation of iron in the hippocampus with MMSE in patients with Alzheimer's disease. *J Magn Reson Imaging* **29**, 793–798.

[6] Langkammer C, Ropele S, Pirpamer L, Fazekas F, Schmidt R (2014) MRI for iron mapping in Alzheimer's disease. *Neurodegener Dis* **13**, 189–191.

[7] Van Duijn S, Bulk M, Van Duinen SG, Nabuurs RJA, Van Buchem MA, Van Der Weerd L, Natté R (2017) Cortical iron reflects severity of Alzheimer's disease. *J Alzheimers Dis* **60**, 1533–1545.

[8] Mills E, Dong XP, Wang F, Xu H (2010) Mechanisms of brain iron transport: Insight into neurodegeneration and CNS disorders. *Future Med Chem* **2**, 51–64.

[9] Yamamoto A, Shin R-W, Hasegawa K, Naiki H, Sato H, Yoshimasu F, Kitamoto T (2004) Iron (III) induces aggregation of hyperphosphorylated τ and its reduction to iron (II) reverses the aggregation: Implications in the formation of neurofibrillary tangles of Alzheimer's disease. *J Neurochem* **82**, 1137–1147.

[10] Telling ND, Everett J, Collingwood JF, Dobson J, van der Laan G, Gallagher JJ, Wang J, Hitchcock AP (2017) Iron biochemistry is correlated with amyloid plaque morphology in an established mouse model of Alzheimer's disease. *Cell Chem Biol* **24**, 1205-1215.e3.

[11] Leskovjan AC, Kretlow A, Lanzirotti A, Barrea R, Vogt S, Miller LM (2011) Increased brain iron coincides with early plaque formation in a mouse model of Alzheimer's disease. *Neuroimage* **55**, 32–38.

[12] Gong NJ, Dibb R, Bulk M, van der Weerd L, Liu C (2019) Imaging beta amyloid aggregation and iron accumulation

in Alzheimer's disease using quantitative susceptibility mapping MRI. *Neuroimage* **191**, 176–185.

[13] El Tannir El Tayara N, Delatour B, Le Cudennec C, Guégan M, Volk A, Dhenain M (2006) Age-related evolution of amyloid burden, iron load, and MR relaxation times in a transgenic mouse model of Alzheimer's disease. *Neurobiol Dis* **22**, 199–208.

[14] Lumsden AL, Rogers JT, Majd S, Newman M, Sutherland GT, Verdile G, Lardelli M (2018) Dysregulation of neuronal iron homeostasis as an alternative unifying effect of mutations causing familial Alzheimer's disease. *Front Neurosci* **12**, 533.

[15] Lane DJR, Merlot AM, Huang MLH, Bae DH, Jansson PJ, Sahni S, Kalinowski DS, Richardson DR (2015) Cellular iron uptake, trafficking and metabolism: Key molecules and mechanisms and their roles in disease. *Biochim Biophys Acta* **1853**, 1130–1144.

[16] Gray NK, Hentze MW (1994) Iron regulatory protein prevents binding of the 43S translation pre-initiation complex to ferritin and eALAS mRNAs. *EMBO J* **13**, 3882–3891.

[17] Hentze MW, Muckenthaler MU, Andrews NC (2004) Balancing acts: Molecular control of mammalian iron metabolism. *Cell* **117**, 285–297.

[18] Muckenthaler MU, Galy B, Hentze MW (2008) Systemic iron homeostasis and the iron-responsive element/iron-regulatory protein (IRE/IRP) regulatory network. *Annu Rev Nutr* **28**, 197–213.

[19] Basilion JP, Rouault TA, Massinople CM, Klausner RD, Burgesst WH (1994) The iron-responsive element-binding protein: Localization of the RNA-binding site to the aconitase active-site cleft (UV cross-linking/base hydrolysis). *Biochemistry* **91**, 574–578.

[20] Pantopoulos K (2004) Iron metabolism and the IRE/IRP regulatory system: An update. *Ann N Y Acad Sci* **1012**, 1–13.

[21] Piccinelli P, Samuelsson T (2007) Evolution of the iron-responsive element. *RNA* **13**, 952–966.

[22] Gunshin H, Allerson CR, Polycarpou-Schwarz M, Rofts A, Rogers JT, Kishi F, Hentze MW, Rouault TA, Andrews NC, Hediger MA (2001) Iron-dependent regulation of the divalent metal ion transporter. *FEBS Lett* **509**, 309–316.

[23] Cmejla R, Petrak J, Cmejlova J (2006) A novel iron responsive element in the 3′UTR of human MRCKα. *Biochem Biophys Res Commun* **341**, 158–166.

[24] Sanchez M, Galy B, Schwanhaeusser B, Blake JJ, Bähr-Ivacevic T, Benes V, Selbach M, Muckenthaler MU, Hentze MW (2011) Iron regulatory protein-1 and -2: Transcriptome-wide definition of binding mRNAs and shaping of the cellular proteome by iron regulatory proteins. *e-Blood* **118**, e168–e179.

[25] Campillos M, Cases I, Hentze MW, Sanchez M (2010) SIREs: Searching for iron-responsive elements. *Nucleic Acids Res* **38**, W360-367.

[26] Tanzi RE, Hyman BT (1991) Alzheimer's mutation. *Nature* **350**, 564–564.

[27] Conway JR, Lex A, Gehlenborg N (2017) UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938-2940.

[28] Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, Kasprzyk A (2009) BioMart - Biological queries made easy. *BMC Genomics* **10**, 22.

[29] Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP, Bateman A (2011) Databases and ontologies Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740.

[30] Bastian M, Heymann S, Jacomy M (2009) Gephi: An Open Source Software for Exploring and Manipulating Networks Visualization and Exploration of Large Graphs. *Third Int AAAI Conf Weblogs Soc Media*.

[31] Hu Y (2011) *Algorithms for Visualizing Large Networks*.

[32] Benner C. HOMER (Hypergeometric Optimization of Motif EnRichment).

[33] Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576–589.

[34] Team RC (2018) R: A language and environment for statistical computing.

[35] Giner G, Smyth GK (2016) FRY: A fast approximation to ROAST gene set test with mean aggregated set statistics. *F1000Research* **5**, https://doi.org/10.7490/f1000research.1113351.1.

[36] Wu D, Lim E, Vaillant F, Asselin-Labat M-L, Visvader JE, Smyth GK (2010) ROAST: Rotation gene set tests for complex microarray experiments. *Bioinformatics* **26**, 2176–2182.

[37] Wu D, Smyth GK (2012) Camera: A competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res* **40**, e133.

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

[39] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545–15550.

[40] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47.

[41] Law CW, Chen Y, Shi W, Smyth GK (2014) Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29.

[42] Wilkinson B (1951) A statistical consideration in psychological research. *Psychol Bull* **48**, 156–158.

[43] Chicault C, Toutain B, Monnier A, Aubry M, Fergelot P, Le Treut A, Galibert MD, Mosser J (2006) Iron-related transcriptomic variations in CaCo-2 cells, an *in vitro* model of intestinal absorptive cells. *Physiol Genomics* **26**, 55–67.

[44] Newman M, Hin N, Pederson S, Lardelli M (2019) Brain transcriptome analysis of a familial Alzheimer's disease-like mutation in the zebrafish presenilin 1 gene implies effects on energy production. *Mol Brain* **12**, 43.

[45] Bray NL, Pimentel H, Melsted P, Pachter L (2016) Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**, 525–527.

[46] Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140.

[47] Dong Y, Newman M, Pederson S, Hin N, Lardelli M (2020) Transcriptome analyses of 7-day-old zebrafish larvae possessing a familial Alzheimer's disease-like mutation in psen1 indicate effects on oxidative phosphorylation, mcm functions, and iron homeostasis. *bioRxiv*, 2020.05.03.075424.

[48] Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, Chen L, Chen L, Chen TM, Chin MC, Chong J, Crook BE, Czaplinska A, Dang CN, Datta S, Dee NR, Desaki AL, Desta T, Diep E, Dolbeare TA, Donelan MJ, Dong HW, Dougherty JG, Duncan BJ, Ebbert AJ, Eichele G, Estin LK, Faber C, Facer BA, Fields R, Fischer SR, Fliss TP, Frensley C, Gates SN, Glattfelder KJ, Halverson KR, Hart MR, Hohmann JG, Howell MP, Jeung DP, Johnson RA, Karr PT, Kawal R, Kidney JM, Knapik RH, Kuan CL, Lake JH, Laramee AR, Larsen KD, Lau C, Lemon TA, Liang AJ, Liu Y, Luong LT, Michaels J, Morgan JJ, Morgan RJ, Mortrud MT, Mosqueda NF, Ng LL, Ng R, Orta GJ, Overly CC, Pak TH, Parry SE, Pathak SD, Pearson OC, Puchalski RB, Riley ZL, Rockett HR, Rowland SA, Royall JJ, Ruiz MJ, Sarno NR, Schaffnit K, Shapovalova N V., Sivisay T, Slaughterbeck CR, Smith SC, Smith KA, Smith BI, Sodt AJ, Stewart NN, Stumpf KR, Sunkin SM, Sutram M, Tam A, Teemer CD, Thaller C, Thompson CL, Varnam LR, Visel A, Whitlock RM, Wohnoutka PE, Wolkey CK, Wong VY, Wood M, Yaylaoglu MB, Young RC, Youngstrom BL, Yuan XF, Zhang B, Zwingman TA, Jones AR (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176.

[49] Bonham LW, Sirkis DW, Yokoyama JS (2019) The transcriptional landscape of microglial genes in aging and neurodegenerative disease. *Front Immunol* **10**, 1170.

[50] Sanchez M, Galy B, Dandekar T, Bengert P, Vainshtein Y, Stolte J, Muckenthaler MU, Hentze MW (2006) Iron regulation and the cell cycle: Identification of an iron-responsive element in the 3′-untranslated region of human cell division cycle 14A mRNA by a refined microarray-based screening strategy. *J Biol Chem* **281**, 22865–22874.

[51] Allaire J, Christopher G, Russell K, Yetman C (2017) networkD3: D3 JavaScript Network Graphs from R.

[52] Goforth JB, Anderson SA, Nizzi CP, Eisenstein RS (2010) Multiple determinants within iron-responsive elements dictate iron regulatory protein binding and regulatory hierarchy. *RNA* **16**, 154–169.

[53] Henderson BR, Menotti E, Kühn LC (1996) Iron regulatory proteins 1 and 2 bind distinct sets of RNA target sequences. *J Biol Chem* **271**, 4900–4908.

[54] Henderson BR, Menotti E, Bonnard C, Kuhn LC (1994) Optimal sequence and structure of iron-responsive elements. Selection of RNA stem-loops with high affinity for iron regulatory factor. *J Biol Chem* **269**, 17481–17489.

[55] Butt J, Kim HY, Basilion JP, Cohen S, Iwai K, Philpott CC, Altschul S, Klausner RD, Rouault TA (1996) Differences in the RNA binding sites of iron regulatory proteins and potential target diversity. *Proc Natl Acad Sci U S A* **93**, 4345–4349.

[56] Surinya KH, Cox TC, May BK (1998) Identification and characterization of a conserved erythroid-specific enhancer located in intron 8 of the human 5-aminolevulinate synthase 2 gene. *J Biol Chem* **273**, 16798–16809.

[57] Theil EC (1990) Regulation of ferritin and transferrin receptor mRNAs. *J Biol Chem* **265**, 4771–4774.

[58] Cmejla R, Ptackova P, Petrak J, Savvulidi F, Cerny J, Sebesta O, Vyoral D (2010) Human MRCKα is regulated by cellular iron levels and interferes with transferrin iron uptake. *Biochem Biophys Res Commun* **395**, 163–167.

[59] Lymboussaki A, Pignatti E, Montosi G, Garuti C, Haile DJ, Pietrangelo A (2003) The role of the iron responsive element in the control of ferroportin1/IREG1/MTP1 gene expression. *J Hepatol* **39**, 710–715.

[60] Aziz N, Munro HN (1987) Iron regulates ferritin mRNA translation through a segment of its 5′ untranslated region. *Proc Natl Acad Sci U S A* **84**, 8478–8482.

[61] Alhamdoosh M, Law CW, Tian L, Sheridan JM, Ng M, Ritchie ME (2017) Easy and efficient ensemble gene set testing with EGSEA. *F1000Research* **6**, 2010.

[62] Allen M, Carrasquillo MM, Funk C, Heavner BD, Zou F, Younkin CS, Burgess JD, Chai HS, Crook J, Eddy JA, Li H, Logsdon B, Peters MA, Dang KK, Wang X, Serie D, Wang C, Nguyen T, Lincoln S, Malphrus K, Bisceglio G, Li M, Golde TE, Mangravite LM, Asmann Y, Price ND, Petersen RC, Graff-Radford NR, Dickson DW, Younkin SG, Ertekin-Taner N (2016) Human whole genome genotype and transcriptome data for Alzheimer's and other neurodegenerative diseases. *Sci Data* **3**, 160089.

[63] Ossenkoppele R, Van Berckel BN, Prins ND (2011) Amyloid imaging in prodromal Alzheimer's disease. *Alzheimers Res Ther* **3**, 26.

[64] Protas HD, Chen K, Langbaum JBS, Fleisher AS, Alexander GE, Lee W, Bandy D, De Leon MJ, Mosconi L, Buckley S, Truran-Sacrey D, Schuff N, Weiner MW, Caselli RJ, Reiman EM (2013) Posterior cingulate glucose metabolism, hippocampal glucose metabolism, and hippocampal volume in cognitively normal, late-middle-aged persons at 3 levels of genetic risk for Alzheimer disease. *JAMA Neurol* **70**, 320–325.

[65] Caldwell CC, Yao J, Brinton RD (2015) Targeting the prodromal stage of Alzheimer's disease: Bioenergetic and mitochondrial opportunities. *Neurotherapeutics* **12**, 66–80.

[66] Vermunt L, Sikkes SAM, van den Hout A, Handels R, Bos I, van der Flier WM, Kern S, Ousset PJ, Maruff P, Skoog I, Verhey FRJ, Freund-Levi Y, Tsolaki M, Wallin ÅK, Olde Rikkert M, Soininen H, Spiru L, Zetterberg H, Blennow K, Scheltens P, Muniz-Terrera G, Visser PJ, Vellas B, Reynish E, Ousset PJ, Andrieu S, Burns A, Pasquier F, Frisoni G, Salmon E, Michel JP, Zekry DS, Boada M, Dartigues JF, Olde-Rikkert MGM, Rigaud AS, Winblad B, Malick A, Sinclair A, Frölich L, Scheltens P, Ribera C, Touchon J, Robert P, Salva A, Waldemar G, Bullock R, Tsolaki M, Rodriguez G, Spiru L, Jones RW, Stiens G, Stoppe G, Eriksdotter Jönhagen M, Cherubini A, Lage PM, Gomez-Isla T, Camus V, Agüera-Morales E, Lopez F, Savy S, Cantet C, Coley N (2019) Duration of preclinical, prodromal, and dementia stages of Alzheimer's disease in relation to age, sex, and APOE genotype. *Alzheimers Dement* **15**, 888–898.

[67] Hargis KE, Blalock EM (2017) Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us? *Behav Brain Res* **322**, 311–328.

[68] Bundy JL, Vied C, Badger C, Nowakowski RS (2019) Sex-biased hippocampal pathology in the 5XFAD mouse model of Alzheimer's disease: A multi-omic analysis. *J Comp Neurol* **527**, 462–475.

[69] Drummond E, Wisniewski T (2017) Alzheimer's disease: Experimental models and reality. *Acta Neuropathol* **133**, 155–175.

[70] Xia D, Watanabe H, Wu B, Lee SH, Li Y, Tsvetkov E, Bolshakov VY, Shen J, Kelleher RJ (2015) Presenilin-1 knockin mice reveal loss-of-function mechanism for familial Alzheimer's disease. *Neuron* **85**, 967–981.

[71] Lee JH, Yu WH, Kumar A, Lee S, Mohan PS, Peterhoff CM, Wolfe DM, Martinez-Vicente M, Massey AC, Sovak G, Uchiyama Y, Westaway D, Cuervo AM, Nixon RA (2010) Lysosomal proteolysis and autophagy require

presenilin 1 and are disrupted by Alzheimer-related PS1 mutations. *Cell* **141**, 1146–1158.

[72] Lee JH, McBrayer MK, Wolfe DM, Haslett LJ, Kumar A, Sato Y, Lie PPY, Mohan P, Coffey EE, Kompella U, Mitchell CH, Lloyd-Evans E, Nixon RA (2015) Presenilin 1 maintains lysosomal Ca2+homeostasis via TRPML1 by regulating vATPase-mediated lysosome acidification. *Cell Rep* **12**, 1430–1444.

[73] Yu Y, Richardson DR (2011) Cellular iron depletion stimulates the JNK and p38 MAPK signaling transduction pathways, dissociation of ASK1-thioredoxin, and activation of ASK1. *J Biol Chem* **286**, 15413–15427.

[74] Riera H, Afonso V, Collin P, Lomri A (2015) A central role for JNK/AP-1 pathway in the pro-oxidant effect of pyrrolidine dithiocarbamate through superoxide dismutase 1 gene repression and reactive oxygen species generation in hematopoietic human cancer cell line U937. *PLoS One* **10**, e0127571.

[75] Hin N, Newman M, Kaslin J, Douek AM, Lumsden A, Nik SHM, Dong Y, Zhou XF, Manucat-Tan NB, Ludington A, Adelson DL, Pederson S, Lardelli M (2020) Accelerated brain aging towards transcriptional inversion in a zebrafish model of the K115fs mutation of human PSEN2. *PLoS One* **15**, e0227258.

[76] Jiang Y, Li C, Wu Q, An P, Huang L, Wang J, Chen C, Chen X, Zhang F, Ma L, Liu S, He H, Xie S, Sun Y, Liu H, Zhan Y, Tao Y, Liu Z, Sun X, Hu Y, Wang Q, Ye D, Zhang J, Zou S, Wang Y, Wei G, Liu Y, Shi Y, Eugene Chin Y, Hao Y, Wang F, Zhang X (2019) Iron-dependent histone 3 lysine 9 demethylation controls B cell proliferation and humoral immune responses. *Nat Commun* **10**, 2935.

[77] Oakley H, Cole SL, Logan S, Maus E, Shao P, Craft J, Guillozet-Bongaarts A, Ohno M, Disterhoft J, Van Eldik L, Berry R, Vassar R (2006) Intraneuronal beta-amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: Potential factors in amyloid plaque formation. *J Neurosci* **26**, 10129–10140.

[78] Landel V, Baranger K, Virard I, Loriod B, Khrestchatisky M, Rivera S, Benech P, Féron F (2014) Temporal gene profiling of the 5XFAD transgenic mouse model highlights the importance of microglial activation in Alzheimer's disease. *Mol Neurodegener* **9**, 33.

[79] Deture MA, Dickson DW (2019) The neuropathological diagnosis of Alzheimer's disease. *Mol Neurodegener* **14**, 32.

[80] Jack CR, Albert MS, Knopman DS, McKhann GM, Sperling RA, Carrillo MC, Thies B, Phelps CH (2011) Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 257–262.

[81] Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, Menon M, He L, Abdurrob F, Jiang X, Martorell AJ, Ransohoff RM, Hafler BP, Bennett DA, Kellis M, Tsai LH (2019) Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337.

[82] Tybl E, Gunshin H, Gupta S, Barrientos T, Bonadonna M, Celma NF, Palais G, Karim Z, Sanchez M, Andrews NC, Galy B, Celma Nos F, Palais G, Karim Z, Sanchez M, Andrews NC, Galy B (2020) Control of systemic iron homeostasis by the 3′ iron-responsive element of divalent metal transporter 1 in mice. *HemaSphere* **4**, e459.

[83] Hasan A, Cotobal C, Duncan CDS, Mata J (2014) Systematic analysis of the role of RNA-binding proteins in the regulation of RNA stability. *PLoS Genet* **10**, e1004684.

[84] Thomson AM, Cahill CM, Cho HH, Kassachau KD, Epis MR, Bridges KR, Leedman PJ, Rogers JT (2005) The acute box cis-element in human heavy ferritin mRNA 5′- untranslated region is a unique translation enhancer that binds poly(C)-binding proteins. *J Biol Chem* **280**, 30032–30045.

[85] Ostareck-Lederer A, Ostareck DH, Standart N, Thiele BJ (1994) Translation of 15-lipoxygenase mRNA is inhibited by a protein that binds to a repeated sequence in the 3′ untranslated region. *EMBO J* **13**, 1476.

[86] Martins EAL, Robalinho RL, Meneghini R (1995) Oxidative stress induces activation of a cytosolic protein responsible for control of iron uptake. *Arch Biochem Biophys* **316**, 128–134.

[87] Pantopoulos K, Hentze MW (1995) Rapid responses to oxidative stress mediated by iron regulatory protein. *EMBO J* **14**, 2917–2924.

[88] Yang L, Wang D, Wang XT, Lu YP, Zhu L (2018) The roles of hypoxia-inducible Factor-1 and iron regulatory protein 1 in iron uptake induced by acute hypoxia. *Biochem Biophys Res Commun* **507**, 128–135.

[89] Popovic Z, Templeton DM (2007) Inhibition of an iron-responsive element/iron regulatory protein-1 complex by ATP binding and hydrolysis. *FEBS J* **274**, 3108–3119.

[90] Yu Y, Kovacevic Z, Richardson DR (2007) Tuning cell cycle regulation with an iron key. *Cell Cycle* **6**, 1982–1994.

[91] Fu D, Richardson DR (2007) Iron chelation and regulation of the cell cycle: 2 Mechanisms of posttranscriptional regulation of the universal cyclin-dependent kinase inhibitor p21CIP1/WAF1 by iron depletion. *Blood* **110**, 752–761.

[92] Muchowska KB, Varma SJ, Moran J (2019) Synthesis and breakdown of universal metabolic precursors promoted by iron. *Nature* **569**, 104–107.

[93] Martínez-Reyes I, Chandel NS (2020) Mitochondrial TCA cycle metabolites control physiology and disease. *Nat Commun* **11**, 102.

[94] Andolfo I, Falco L De, Asci R, Russo R, Colucci S, Gorrese M, Zollo M, Iolascon A (2010) Regulation of divalent metal transporter 1 (DMT1) non-IRE isoform by the microRNA Let-7d in erythroid cells. *Haematologica* **95**, 1244.

[95] Davis M, Clarke S (2013) Influence of microRNA on the maintenance of human iron metabolism. *Nutrients* **5**, 2611–2628.

[96] Freiesleben S, Hecker M, Zettl UK, Fuellen G, Taher L (2016) Analysis of microRNA and gene expression profiles in multiple sclerosis: Integrating interaction data to uncover regulatory mechanisms. *Sci Rep* **6**, 34512.

[97] Sangokoya C, Doss JF, Chi J-T (2013) Iron-responsive miR-485-3p regulates cellular iron homeostasis by targeting ferroportin. *PLoS Genet* **9**, e1003408.

[98] Long JM, Maloney B, Rogers JT, Lahiri DK (2018) Novel upregulation of amyloid-β precursor protein (APP) by microRNA-346 via targeting of APP mRNA 5′-untranslated region: Implications in Alzheimer's disease. *Mol Psychiatry* **24**, 345–363.

[99] Guo M, Song L-P, Jiang Y, Liu W, Yu Y, Chen G-Q (2006) Hypoxia-mimetic agents desferrioxamine and cobalt chloride induce leukemic cell apoptosis through different hypoxia-inducible factor-1α independent mechanisms. *Apoptosis* **11**, 67–77.

[100] Zecca L, Youdim MBH, Riederer P, Connor JR, Crichton RR, Stankowski JN, Dawson VL, Dawson TM, Kawamoto

M, Horibe T, Kohno M, Kawakami K, Duce JA, Tsatsanis A, Cater MA, James SA, Robb E, Wikhe K, Leong SL, Perez K, Johanssen T, Mark A, Cho H-H, Galatis D, Moir RD, Masters CL, McLean C, Tanzi RE, Cappai R, Barnham KJ, Ciccotosto D, Rogers JT, Bush AI, Greenough M a, Cho H-H, Galatis D, Moir RD, Masters CL, McLean C, Tanzi RE, Cappai R, Barnham KJ, Ciccotosto GD, Rogers JT, Bush AI, De Domenico I, McVey Ward D, Kaplan J (2011) Iron-export ferroxidase activity of β-amyloid precursor protein is inhibited by zinc in Alzheimer's disease. *Cell* **142**, 359.

[101] Jacomin A-C, Geraki K, Brooks J, Tjendana-Tjhin V, Collingwood JF, Nezis IP (2019) Impact of autophagy and aging on iron load and ferritin in Drosophila brain. *Front Cell Dev Biol* **7**, 142.

[102] Ayton S, Lei P, Bush AI (2013) Metallostasis in Alzheimer's disease. *Free Radic Biol Med* **62**, 76–89.

[103] Smith MA, Harris PLR, Sayre LM, Perry G (1997) Iron accumulation in Alzheimer disease is a source of redox-generated free radicals. *Proc Natl Acad Sci U S A* **94**, 9866–9868.

[104] Yambire KF, Rostosky C, Watanabe T, Pacheu-Grau D, Torres-Odio S, Sanchez-Guerrero A, Senderovich O, Meyron-Holtz EG, Milosevic I, Frahm J, Phillip West A, Raimundo N (2019) Impaired lysosomal acidification triggers iron deficiency and inflammation *in vivo*. *Elife* **8**, e51031.

[105] Jiang Y, Sato Y, Im E, Berg M, Bordi M, Darji S, Kumar A, Mohan PS, Bandyopadhyay U, Diaz A, Cuervo AM, Nixon RA (2019) Lysosomal dysfunction in down syndrome is app-dependent and mediated by APP-βCTF (c99). *J Neurosci* **39**, 5255–5268.

[106] Prasad H, Rao R (2018) Amyloid clearance defect in ApoE4 astrocytes is reversed by epigenetic correction of endosomal pH. *Proc Natl Acad Sci U S A* **115**, E6640–E6649.

[107] Tisato V, Zuliani G, Vigliano M, Longo G, Franchini E, Secchiero P, Zauli G, Paraboschi EM, Singh AV, Serino ML, Ortolani B, Zurlo A, Bosi C, Greco A, Seripa D, Asselta R, Gemmati D, Vikram Singh A, Serino ML, Ortolani B, Zurlo A, Bosi C, Greco A, Seripa D, Asselta R, Gemmati D (2018) Gene-gene interactions among coding genes of iron-homeostasis proteins and APOE-alleles in cognitive impairment diseases. *PLoS One* **13**, e0193867.

[108] Van Acker ZP, Bretou M, Annaert W (2019) Endo-lysosomal dysregulations and late-onset Alzheimer's disease: Impact of genetic risk factors. *Mol Neurodegener* **14**, 1–20.

[109] Venkataramani V, Doeppner TR, Willkommen D, Cahill CM, Xin Y, Ye G, Liu Y, Southon A, Aron A, Au-Yeung HY, Huang X, Lahiri DK, Wang F, Bush AI, Wulf GG, Ströbel P, Michalke B, Rogers JT (2018) Manganese causes neurotoxic iron accumulation via translational repression of amyloid precursor protein and H-Ferritin. *J Neurochem* **147**, 831–848.

[110] McCarthy RC, Park Y, Kosman DJ (2014) sAPP modulates iron efflux from brain microvascular endothelial cells by stabilizing the ferrous iron exporter ferroportin. *EMBO Rep* **15**, 809–815.

[111] Yeh K, Yeh M, Glass J (2011) Interactions between ferroportin and hephaestin in rat enterocytes are reduced after iron ingestion. *Gastroenterology* **141**, 292-299.e1.

[112] Yeh K, Yeh M, Mims L, Glass J (2009) Iron feeding induces ferroportin 1 and hephaestin migration and interaction in rat duodenal epithelium. *Am J Physiol Liver Physiol* **296**, G55–G65.

[113] Han O, Kim E-Y (2007) Colocalization of ferroportin-1 with hephaestin on the basolateral membrane of human intestinal absorptive cells. *J Cell Biochem* **101**, 1000–1010.

[114] Tsatsanis A, Wong BX, Gunn AP, Ayton S, Bush AI, Devos D, Duce JA (2020) Amyloidogenic processing of Alzheimer's disease β-amyloid precursor protein induces cellular iron retention. *Mol Psychiatry* **25**, 1958-1966.

[115] Pigoni M, Wanngren J, Kuhn PH, Munro KM, Gunnersen JM, Takeshima H, Feederle R, Voytyuk I, De Strooper B, Levasseur MD, Hrupka BJ, Müller SA, Lichtenthaler SF (2016) Seizure protein 6 and its homolog seizure 6-like protein are physiological substrates of BACE1 in neurons. *Mol Neurodegener* **11**, 67.

[116] Rogers JT, Randall JD, Cahill CM, Eder PS, Huang X, Gunshin H, Leiter L, McPhee J, Sarang SS, Utsuki T, Greig NH, Lahiri DK, Tanzi RE, Bush AI, Giordano T, Gullans SR (2002) An iron-responsive element type II in the 5′-untranslated region of the Alzheimer's amyloid precursor protein transcript. *J Biol Chem* **277**, 45518–45528.