Dynamic changes in the brain protein interaction network correlates with progression of Aβ42 pathology in *Drosophila*

Harry M. Scholes[1,*], Adam Cryar[1,*], Fiona Kerr[2,3], David Sutherland[1], Lee A. Gethings[4], Johannes P. C. Vissers[4], Jonathan G. Lees[1,5], Christine A. Orengo[1,#], Linda Partridge[2,6,#], Konstantinos Thalassinos[1,#]

1 Institute of Structural and Molecular Biology, University College London, London, United Kingdom

2 Institute of Healthy Ageing, University College London, London, United Kingdom

3 Department of Biology and Biomedical Sciences, School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, United Kingdom

4 Waters Corporation, Wilmslow, United Kingdom

5 Current address: Faculty of Health and Life Sciences, Oxford Brookes University, United Kingdom

6 Max Planck Institute for Biology of Ageing, Cologne, Germany

* Joint first authors

Corresponding authors

Abstract

Alzheimer's disease (AD), the most prevalent form of dementia, is a progressive and devastating neurodegenerative condition for which there are no effective treatments. Understanding the molecular pathology of AD during disease progression may identify new ways to reduce neuronal damage. Here, we present a longitudinal study tracking dynamic proteomic alterations in the brains of an inducible *Drosophila melanogaster* model of AD expressing the Arctic mutant Aβ42 gene. We identified 3093 proteins from flies that were induced to express Aβ42 and age-matched healthy controls using label-free quantitative ion-mobility data independent analysis mass spectrometry. Of these, 228 proteins were significantly altered by Aβ42 accumulation and were enriched for AD-associated processes. Network analyses further revealed that these proteins have distinct hub and bottleneck properties in the brain protein interaction network, suggesting that several may have significant effects on brain function. Our unbiased analysis provides useful insights into the key processes governing the progression of amyloid toxicity and forms a basis for further functional analyses in model organisms and translation to mammalian systems.

Introduction

Alzheimer's disease (AD) is a progressive and devastating neurodegenerative disease that is the most prevalent form of dementia [1]. Symptoms initially present as episodic memory loss and subsequently develop into widespread cognitive impairment. Two brain lesions are pathological hallmarks of the disease: plaques and neurofibrillary tangles. Plaques are extracellular aggregates of amyloid beta (Aβ) [2], whereas, neurofibrillary tangles are intraneuronal aggregates of hyperphosphorylated tau [3,4]. In addition to these hallmarks, the AD brain experiences many other changes, including metabolic and oxidative dysregulation [5,6], DNA damage [7], cell cycle re-entry [8], axon loss [9] and, eventually, neuronal death [6,10].

Despite a substantial research effort, no cure for AD has been found. Effective treatments are desperately needed to cope with the projected increase in the number of new cases as a result of longer life expectancy and an ageing population. Sporadic onset is the most common form of AD (SAD), for which age is the major risk factor. Familial AD (FAD)—a less common (<1%), but more aggressive, form of the disease—has an early onset of pathology before the age of 65 [11]. FAD is caused by fully penetrant mutations in the Aβ precursor protein (APP) and two subunits—presenilin 1 and presenilin 2—of the ɣ-secretase complex that processes APP in the amyloidogenic pathway to produce Aβ. Whilst the exact disease mechanisms of AD are not yet fully understood, this has provided support for Aβ accumulation as a key player in its cause and progression [1]. Aβ42—a 42 amino acid variant of the peptide—is neurotoxic [12], necessary for plaque deposition [13] and sufficient for tangle formation [14]. The Arctic mutation in Aβ42 (Glu22Gly) [15] causes a particularly aggressive form of familial AD that is associated with an increased rate and volume of plaque deposition [16]. Genetic analyses of SAD, however, suggest a complex molecular pathology, in which alterations in neuro-inflammation, cholesterol metabolism and synaptic recycling pathways may also be required for Aβ42 to initiate the toxic cascade of events leading to tau pathology and neuronal damage in dementia.

Comparison of proteomic analyses of post-mortem human brains have further revealed an increase in metabolic processes and reduction in synaptic function in AD [17]. Oxidised proteins also accumulate at early stages in AD brain, probably as a result of mitochondrial ROS production [18], and redox proteomic approaches suggest that enzymes involved in glucose metabolism are oxidised in mild cognitive impairment and AD [19,20]. Moreover, phospho-proteomic approaches have revealed alterations in phosphorylation of metabolic enzymes and kinases that regulate phosphorylation of chaperones such as HSP27 and crystallin alpha B [21]. Of note, however, there is little proteomic overlap between studies

using post-mortem human brain tissue, which may reflect the low sample numbers available for such studies, differences in comorbidities between patients and confounding post-mortem procedures [17]. Although valuable, post-mortem studies also reflect the end-stage of disease and, therefore, do not facilitate measurement of dynamic alterations in proteins as AD progresses.

Animal models of AD, generated through transgenic over-expression of human APP or tau, provide an opportunity to track proteomic alterations at pre- and post-pathological stages, thus facilitating insight into the molecular mechanisms underlying disease development and revealing new targets for drugs to prevent AD progression. Analyses of transgenic mice models of AD have revealed some overlapping alterations in metabolic enzymes, kinases and chaperones with human AD brain [17]. Only one study, however, has tracked alterations in protein carbonylation over time, showing increases in oxidation of metabolic enzymes (alpha-enolase, ATP synthase α-chain and pyruvate dehydrogenase E1) and regulatory molecules (14-3-3 and Pin1) in correlation with disease progression [22].

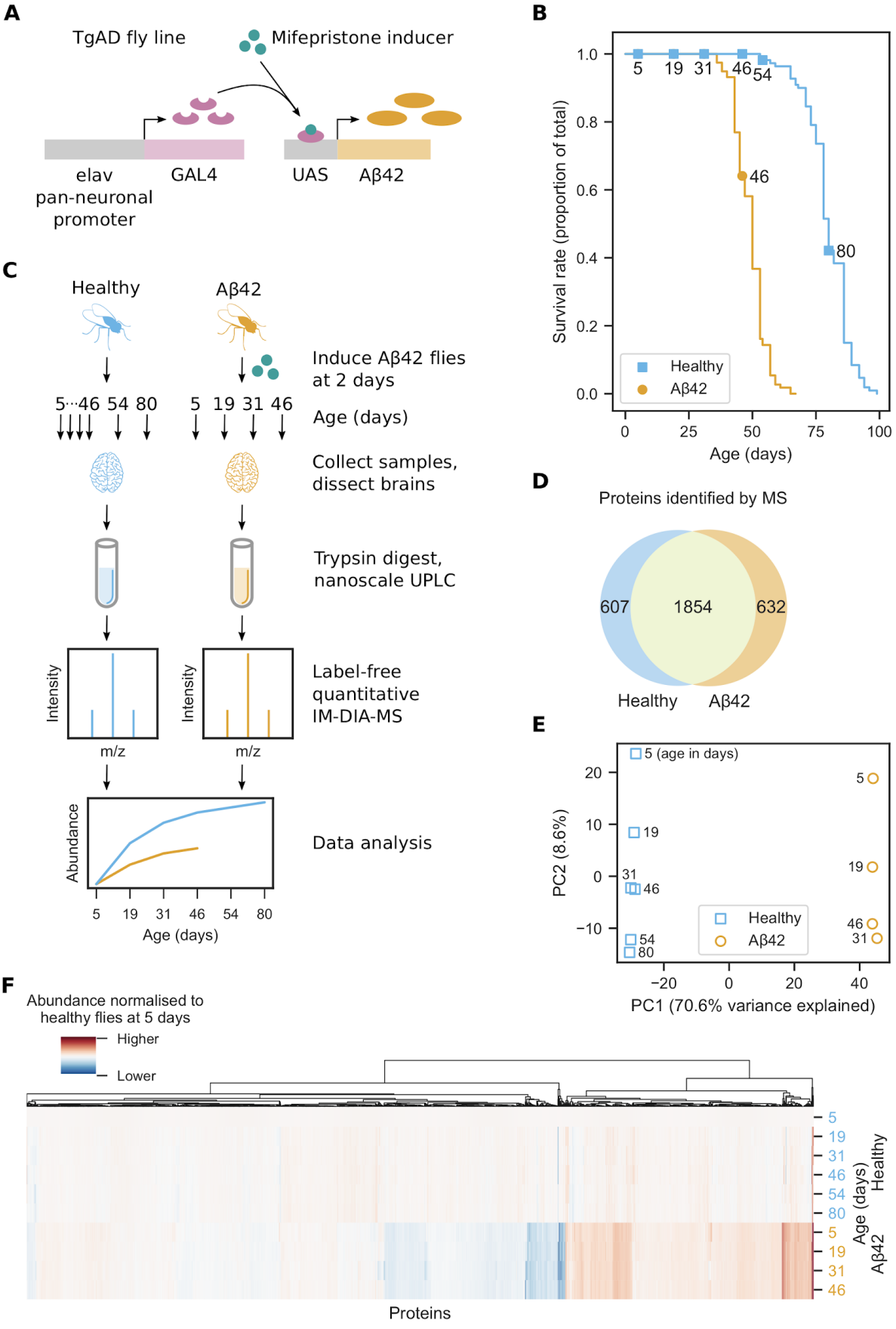
Adult-onset *Drosophila* models of AD have been generated by over-expressing human Aβ42 peptide exclusively in adult fly neurons using inducible expression systems. These models have been shown to develop progressive neurodegenerative phenotypes, such as reduced climbing ability, and shortened lifespan [23]. Taking advantage of the short lifespan of the fly, and the flexible nature of the inducible model, we have performed a longitudinal study of the brain proteome to capture the effects of Aβ42-toxicity in the brain from the point of induction and across life. We identified 3093 proteins using label-free quantitative ion-mobility data independent analysis mass spectrometry (IM-DIA-MS) [24], 1854 of which were common to healthy and Aβ42 flies. Of these, we identified 228 proteins that were significantly altered in AD, some of which overlapped with normal ageing but the majority of which were ageing-independent. Proteins altered in response to Aβ42 were enriched for AD processes and have statistically significant network properties in the brain protein interaction network. We also show that these proteins are likely to be bottlenecks for signalling in the network, suggesting that they comprise important proteins for normal brain function. Our data is a valuable resource to begin to understand the dynamic properties of Aβ42 proteo-toxicity during AD progression. Future functional studies will be required to determine the causal role of these proteins in mediating progression of AD using model organisms and to translate these findings to mammalian systems.

Results

Proteome analysis of healthy and Aβ42-expressing fly brains

Using an inducible transgenic fly line expressing human Arctic mutant Aβ42 (TgAD) [23] (Fig 1A), we confirmed a previously observed [23] reduction in lifespan following Aβ42 induction prior to proteomic analyses (Fig 1B).

To understand how the brain proteome is affected as Aβ42 toxicity progresses, fly brains were dissected from healthy and Aβ42 flies at 5, 19, 31 and 46 days, and at 54 and 80 days for healthy controls, then analysed by label-free quantitative IM-DIA-MS (Fig 1C, Supplementary Data 1). 1854 proteins were identified in both healthy and Aβ42 fly brain from a total of 3093 proteins (Fig 1D), which is typical for recent fly proteomics studies [25,26].



115

116 **Figure 1. Proteome analysis of healthy and AD fly brains.** (**A**) *Drosophila melanogaster*

transgenic model of AD (TgAD) that expresses Arctic mutant Aβ42 in a mifepristone-inducible GAL4/UAS expression system under the pan-neuronal elav promoter. (**B**) Survival curves for healthy and Aβ42 flies. Aβ42 flies were induced to express Aβ42 at 2 days. Markers indicate days that MS samples were collected. (**C**) Experimental design of the brain proteome analysis. Aβ42 flies were induced to express Aβ42 at 2 days. For each of the three biological repeats, 10 healthy and 10 Aβ42 flies were collected at 5, 19, 31 and 46 days, as well as 54 and 80 days for healthy flies. Proteins were extracted from dissected brains and digested with trypsin. The resulting peptides were separated by nanoscale liquid chromatography and analysed by label-free quantitative IM-DIA-MS. (**D**) Proteins identified by IM-DIA-MS. (**E**) Principal component analysis of the IM-DIA-MS data. Axes are annotated with the percentage of variance explained by each principal component. (**F**) Hierarchical biclustering using relative protein abundances normalised to their abundance in healthy flies at 5 days.

For the 1854 proteins identified in both healthy and Aβ42 flies, we assessed the reliability of our data. Proteins were highly correlated between technical and biological repeats (Fig S1). We used principal component analysis of the protein abundances to identify sources of variance (Fig 1E). Healthy and Aβ42 samples are clearly separated in the first principal component, probably due to the effects of Aβ42. In the second principal component, samples are separated by increasing age, due to age-dependent or disease progression changes in the proteome. These results show that whilst ageing does contribute to changes in the brain proteome (8.7% of the total variance), much larger changes are due to expression of Aβ42 (70.6%) and this may reflect either a correlation with the ageing process or progression of AD pathology. We confirmed this result using hierarchical biclustering of protein abundances in Aβ42 versus healthy flies at 5 days (Fig 1F). The results reveal that most proteins do not vary significantly in abundance with age in healthy flies, but many proteins are differentially abundant in Aβ42 flies.

Analysis of brain proteome dysregulation in Aβ42 flies

We next identified the proteins that were significantly altered following Aβ42 expression in the fly brain. To achieve this, we used five methods commonly used to analyse time course RNA-Seq data [27] and classified proteins as significantly altered if at least two methods detected them [28]. We identified 228 significantly altered proteins from 740 proteins that were detected by one or more methods (Fig 2A). A comparison of popular RNA-Seq analysis tools [29] showed that edgeR [30] has a high false positive rate and variable performance on different data sets, whereas, DESeq2 [31] and limma [32] have low false positive rates and perform more consistently. We observed a similar trend in our data set. limma and DESeq2 detected the lowest number of proteins, with 21 proteins in common (Fig S2A). edgeR detected more proteins, of which 38 were also detected by DESeq2 and 16 by limma. EDGE [33] and maSigPro [34] detected vastly more proteins, 464 of which were only detected by one method. Principal component analysis shows that edgeR, DESeq2 and limma detect similar proteins, whereas, EDGE and maSigPro detect very different proteins (Fig S2B).

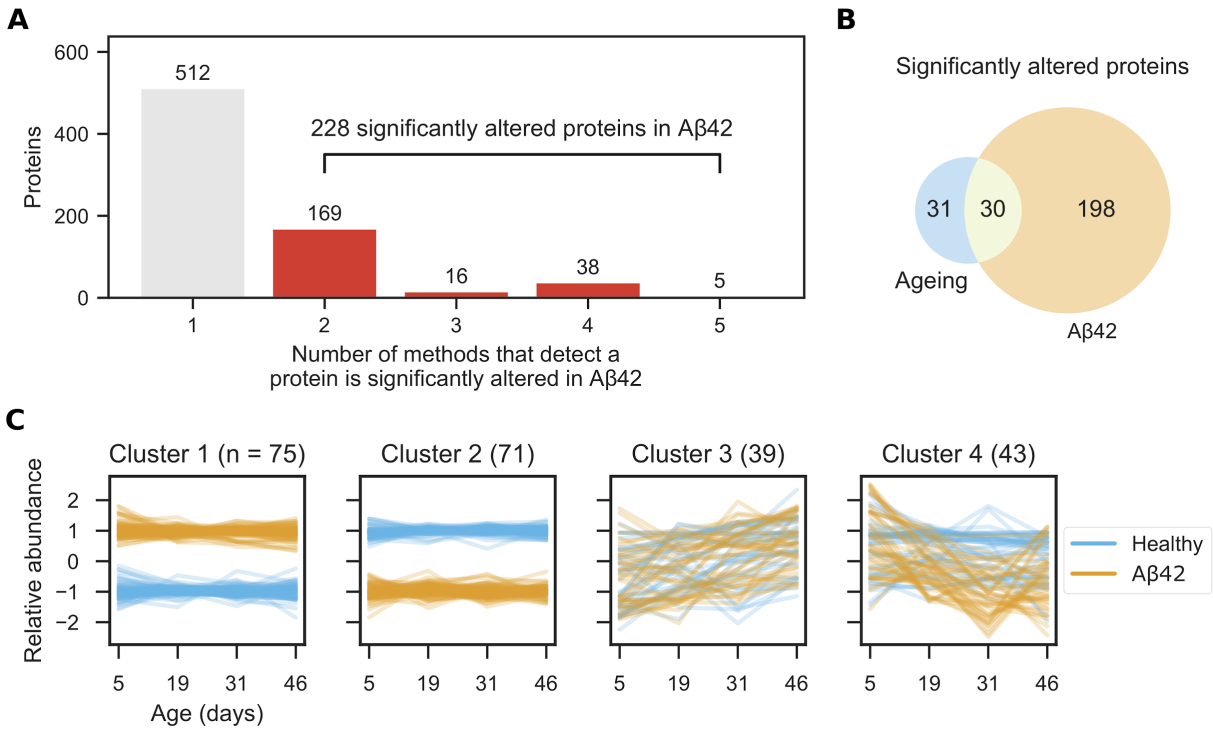


Figure 2. Brain proteome dysregulation in AD. (**A**) Proteins significantly altered in AD were identified using five methods (EDGE, edgeR, DESeq2, limma and maSigPro) and classified as significantly altered if at least two methods detected them. (**B**) Significantly altered proteins in AD (from **A**) and ageing. (**C**) Significantly altered protein abundances were *z* score-transformed and clustered using a Gaussian mixture model.

Although these methods should be able to differentiate between proteins that are altered in Aβ42 flies from those that change during normal ageing, we confirmed this by analysing healthy flies separately. In total, 61 proteins were identified as significantly altered with age (Fig S3), of which 30 were also identified as significantly altered in AD (Fig 2B) and 31 in normal ageing alone. These proteins are not significantly enriched for any pathways or functions. Based on our results, we concluded that the vast majority of proteins that are significantly altered in AD are not altered in normal ageing and that AD causes significant dysregulation of the brain proteome.

To understand the dynamics of protein alterations following Aβ42 induction, we clustered the profiles of proteins significantly altered in Aβ42 flies using a Gaussian mixture model (Fig 2C). The proteins clustered best into four sets (Fig S4). In comparison to healthy flies, cluster 1 contains proteins that have consistently higher abundance in Aβ42 flies. Conversely, cluster 2 contains proteins that have lower abundance in Aβ42 flies. The abundances of proteins from clusters 1 and 2 are affected from the onset of disease at day 5, and remain at similar levels as the disease progresses. Dysregulation of these proteins may initiate AD pathogenesis, be involved in early stages of disease progression, or represent defense mechanisms that could be harnessed for protection. Proteins in cluster 3 follow a similar trend in healthy and Aβ42 flies and increase in abundance with age. However, cluster 4 proteins decrease in abundance as the disease progresses, whilst remaining steady in healthy flies. Further work is required to determine whether reduction of these proteins plays a causal role in disease pathogenesis that could be targeted therapeutically, or whether their decline represents a protective response to damage.

We performed a statistical Gene Ontology enrichment analysis on each cluster, but found no enrichment of terms. Furthermore, we also saw no enrichment when we analysed all 228 proteins together.

Brain proteins significantly altered by Aβ42 have distinct network properties

Following the analyses of brain proteome dysregulation in Aβ42 flies, we analysed the 228 significantly altered proteins in the context of the brain protein interaction network to determine whether their network properties are significantly different to the other brain proteins. Using a subgraph of the STRING [35] network induced on the 3093 proteins identified by IM-DIA-MS, we calculated four graph theoretic network properties (Fig 3A) of the 183 significantly altered proteins contained in this network: *degree*, the number of edges that a node has; *shortest path*, the smallest node set that connect any two nodes; *largest*

198 *connected component*, the largest node set for which all nodes have at least one edge to
199 any of the other nodes; and *betweenness centrality*, the proportion of all the shortest paths
200 in the network that a particular node lies on.

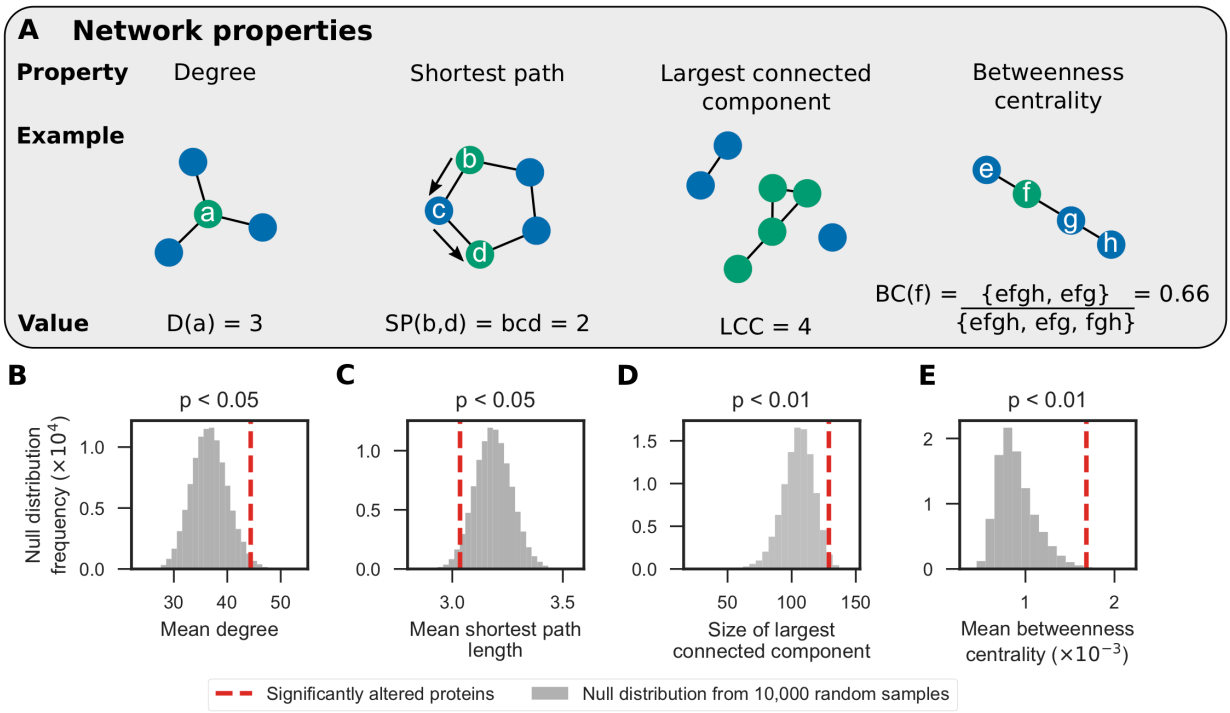


Figure 3: Significantly altered proteins have statistically significant network properties in the brain protein interaction network. (**A**) Network properties that were calculated: *degree*, the number of edges that a node has; *shortest path*, the smallest node set that connect any two nodes; *largest connected component*, the largest node set for which all nodes have at least one edge to any of the other nodes; and *betweenness centrality*, the proportion of all the shortest paths in the network that a particular node lies on. Using a subgraph of the STRING network induced on the 3093 proteins identified by IM-DIA-MS in healthy and Aβ42 flies, the significance of four network characteristics were calculated for the 183 significantly altered proteins contained in this subgraph. (**B**) mean degree; (**C**) mean shortest path length between a node and the remaining 182 nodes; (**D**) the size of the largest connected component in the subgraph induced on these nodes; and (**E**) mean betweenness centrality. Non-parametric p-values were calculated using null distributions of the test statistics, simulated by randomly sampling 183 nodes from the network 10,000 times.

We performed hypothesis tests and found that these proteins have statistically significant network properties. Firstly, the significantly altered proteins make more interactions than expected (mean degree $p < 0.05$; Fig 3B). Therefore, these proteins may further imbalance the proteome by disrupting the expression or activity of proteins they interact with. Secondly, not only are these proteins close to each other (mean shortest path $p < 0.05$; Fig 3C), but also 129 of them form a connected component (size of largest connected component $p < 0.01$; Fig 3D). These two pieces of evidence suggest that Aβ42 disrupts proteins at the centre of the proteome. Lastly, these proteins lie along shortest paths between many pairs of nodes (mean betweenness centrality $p < 0.01$; Fig 3E) and may control how signals are transmitted in cells. Proteins with high betweenness centrality are also more likely to be essential genes for viability [36]. Taken together, these findings suggest that the proteins significantly altered in AD are important in the protein interaction network, and that dysregulation of these proteins may have significant consequences for the brain proteome and therefore function.

Predicting the severity of Aβ42-induced protein alterations using network properties

We predicted how severely particular Aβ42-associated protein alterations may affect the brain using two network properties—the tendency of a node to be a hub or a bottleneck. In networks, nodes with high degree are hubs for communication, whereas nodes with high betweenness centrality are bottlenecks that regulate how signals propagate through the network. Protein expression tends to be highly correlated to that of its neighbours in the protein interaction network. One exception to this rule, however, are bottleneck proteins, whose expression tends to be poorly correlated with that of its neighbours [36]. This suggests that the proteome is finely balanced and that the expression of bottleneck proteins is tightly regulated to maintain homeostasis. We analysed the hub and bottleneck properties of the significantly altered proteins and identified four hub-bottlenecks and five nonhub-bottlenecks that correlate with Aβ42 expression (Fig 4A) and analysed how their abundances change during normal ageing and as pathology progresses (Fig 4B).

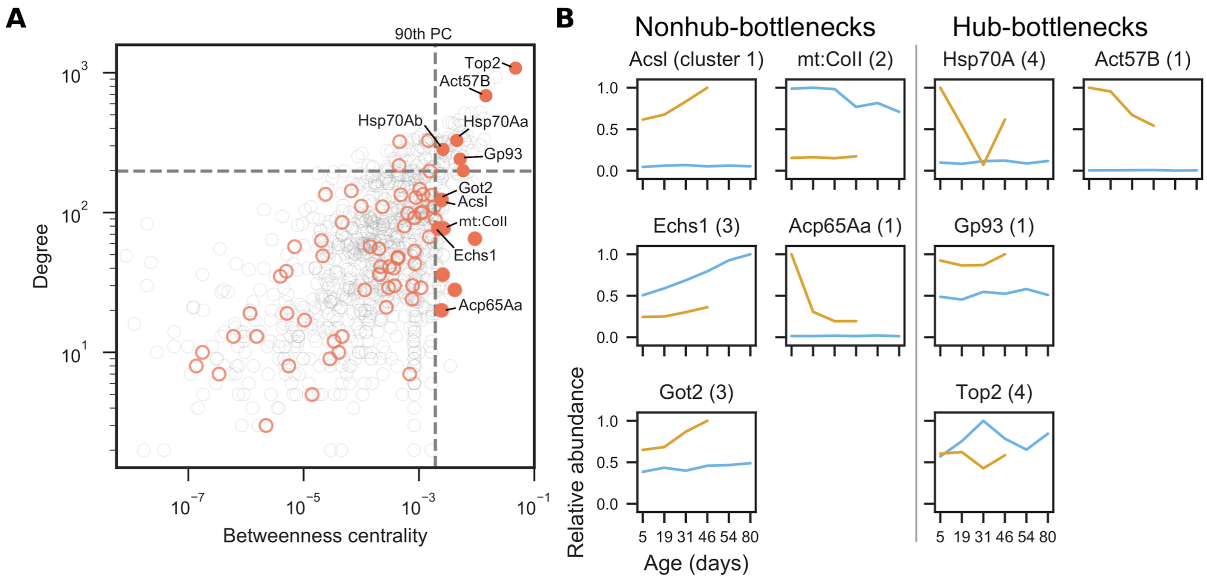


Figure 4. Analysis of hubs and bottlenecks in the brain protein interaction network. In networks, nodes with high degree are hubs and nodes with high betweenness centrality are bottlenecks. (**A**) Degree (hub-ness) is plotted against betweenness centrality (bottleneck-ness) in the brain protein interaction network for all proteins identified by IM-DIA-MS (grey circles). Of the significantly altered proteins (red circles), hub-bottleneck (> 90th percentile (PC) for degree and betweenness centrality) and nonhub-bottleneck proteins (> 90th PC for betweenness centrality) are highlighted (filled red circles). (**B**) Profiles of significantly altered bottleneck proteins implicated in Aβ42 toxicity. Maximum abundances are scaled to 1. Numbers in parentheses denote which cluster from Fig 2C the protein was in.

Nonhub-bottlenecks: Acs1, Echs1, Got2, mt:CoII and Acp65Aa

Three of the nonhub-bottlenecks, Acyl-CoA synthetase long chain (Acs1), Enoyl-CoA hydratase, short chain 1 (Echs1), and Aspartate aminotransferase (Got2), are metabolic enzymes with previous links to neuronal function and damage. Acs1 and Echs1 are involved in the production of acetyl-CoA from fatty acids. Many enzymes involved in acetyl-CoA metabolism associate with AD leading to acetyl-CoA deficits in the brain and loss of cholinergic neurons [6]. Got2 produces the neurotransmitter L-glutamate from aspartate, is involved in assembly of synapses and becomes elevated following brain injury [37]. Brain Acs1 and Got2 levels were stably expressed throughout normal ageing in our healthy flies but increased upon Aβ42 induction and continued to rise with age in Aβ42 flies. This suggests that levels of these proteins increase independently of ageing in AD but correlate closely with disease progression. On the other hand, Echs1 abundance increases in healthy flies during normal ageing, but its levels were reduced upon Aβ42 induction and its ageing-dependent increase was diminished in Aβ42 flies compared to controls. This may reflect a protective response with ageing that is suppressed by Aβ42 toxicity.

Cytochrome c oxidase (COX), complex IV of the mitochondrial electron transport chain, uses energy from reducing molecular oxygen to water to generate a proton gradient across the inner mitochondrial membrane. Levels of mt:CoII (a COX subunit) declined in aged healthy control fly brain. mt:CoII expression was downregulated in Aβ42 flies compared to controls at all time-points and was stably-expressed across age following Aβ42 induction. The link between COX and AD is unclear, although Aβ is known to inhibit COX activity [38]. For example, in AD patients, COX activity—but not abundance—is reduced, resulting in increased levels of ROS [39]. However, in COX-deficient mouse models of AD, plaque deposition and oxidative damage are reduced [40]. Hence, the ageing-dependent decline in mt:CoII may represent either a reduction in COX function which renders the brain vulnerable to damage and is exacerbated by Aβ42 toxicity, or a protective mechanism against both ageing and amyloid toxicity.

The cuticle protein Acp65Aa was also upregulated in Aβ42 flies, but levels fell sharply between 5 and 19 days. However, it is surprising that we identified Acp65Aa in our samples, as it is not expected to be expressed in the brain. One explanation may involve chitin, which has been detected in AD brains and has been suggested to facilitate Aβ nucleation [41]. Amyloid aggregation has previously been shown to plateau around 15 days post-induction [42], which is around the same time that Acp65Aa drops in Aβ42 flies. Our results suggest that Aβ42 causes an increase in Acp65Aa expression early in the disease, but further

experiments are needed to confirm this and to investigate its relationship with nucleation and the aggregation process.

Hub-bottlenecks: Hsp70A, Gp93, Top2 and Act75B

The four hub-bottlenecks are consistent with Aβ42 inducing stress. Hsp70A, a heat shock protein that responds to hypoxia, was significantly upregulated at early time-points (5 days) in Aβ42 flies, compared to healthy controls which exhibited stable expression of this protein throughout life. Although the levels dropped in Aβ42 flies between days 5 and 31 post-induction, at later time-points Hsp70A increased again, possibly suggesting a two-phase response to hypoxia in Aβ42 flies. We found that Gp93—a stress response protein that binds unfolded proteins—to be increased across age in Aβ42 flies compared to controls possibly suggesting an early and sustained protective mechanism against Aβ42-induced damage. DNA topoisomerase 2 (Top2), an essential enzyme for DNA double-strand break repair, was decreased in Aβ42 flies, following a pattern which mirrors changes in its expression with normal ageing. Double-strand breaks occur naturally in the brain as a consequence of neuronal activity—an effect that is aggravated by Aβ[7]. As a consequence of deficient DNA repair machinery, deleterious genetic lesions may accumulate in the brain and exacerbate neuronal loss.

Finally, we found that actin (Act57B) was increased in Aβ42 flies, in agreement with two recent studies on mice brains [43,44]. Kommaddi and colleagues found that Aβ causes depolymerisation of F-actin filaments in a mouse AD model before onset of AD pathology [44]. The authors showed that although the concentration of monomeric G-actin increases, the total concentration of actin remains unchanged. It has long been known that G-, but not F-, actin is susceptible to cleavage by trypsin [45], permitting its detection and quantification by IM-DIA-MS. Hence, the apparent increase of actin in Aβ42 flies may be due to F-actin depolymerisation, which increases the pool of trypsin-digestible G-actin, and is consistent with the findings of Kommaddi *et al*. To confirm whether total actin levels remain the same in the brains of Aβ42 flies, additional experiments would have to be carried out in the future, for example tryptic digestion in the presence of MgADP—which makes F-actin susceptible to cleavage [46]—and transcriptomic analysis of actin mRNA. Furthermore, actin polymerisation is ATP-dependent, so increased levels of G-actin may indicate reduced intracellular ATP. In addition, ATP is important for correct protein folding and therefore reduced levels may lead to increased protein aggregation in AD.

Due to the importance of these hub and bottleneck proteins in the protein interaction network, we predict that AD-associated alterations in their abundance will likely have a significant effect on the cellular dynamics of the brain.

Dysregulated genes are associated with known AD and ageing network modules

Finally, we clustered the protein interaction network into modules and performed a Gene Ontology enrichment analysis on modules that contained any of the 228 significantly altered proteins. We saw no Gene Ontology term enrichment when we tested these proteins clustered according to their abundance profiles (Fig 2C), presumably because the proteins affected in AD are diverse and involved in many different biological processes. However, by testing network modules for functional enrichment, we exploited the principle that interacting proteins are functionally associated. Using a subgraph of the STRING network containing the significantly altered proteins and their directly-interacting neighbours, we used MCODE [47] to find modules of densely interconnected nodes. We chose to include neighbouring proteins to compensate for proteins that may not have been detected in the MS experiments due to the stochastic nature of observing peptides and the wide dynamic range of biological samples [48]. The resulting subgraph contained 4842 proteins, including 183 of the 228 significantly altered proteins, as well as 477 proteins that were only identified in healthy or Aβ42 flies and 3125 proteins that were not identified in our IM-DIA-MS experiments. 12 modules were present in the network (Fig 5A, Supplementary Data 2). The proportion of these modules that were composed of significantly altered proteins ranged from 0–8%. All but one of the modules were enriched for processes implicated in AD and ageing (Fig 5, Supplementary Data 3), including respiration and oxidative phosphorylation, transcription and translation, proteolysis, DNA replication and repair, and cell cycle regulation. These modules contained two proteins that were recently found to be significantly altered in the brain of AD mice [43] and are both upregulated four-fold in AD: adenylate kinase, an adenine nucleotide phosphotransferase, and the armadillo protein Arm, involved in creating long-term memories.

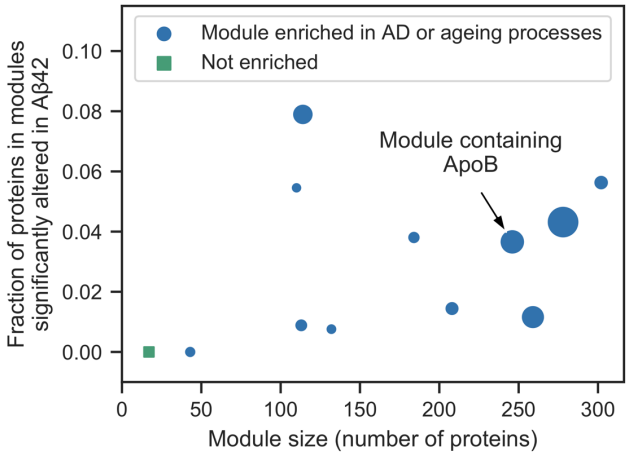


Figure 5. Analysis of network modules enriched for AD or ageing processes. MCODE was used to identify network modules in a subgraph of the STRING network containing the significantly altered proteins and their directly-interacting neighbours. The size of the resulting 12 modules is plotted against the fraction of proteins in these modules that are significantly altered in AD. Module 2 is annotated as containing ApoB. Marker sizes denote the MCODE score for the module.

In humans, the greatest genetic risk factor for AD is the Ɛ4 allele of ApoE—an apolipoprotein involved in cholesterol transport and repairing brain injuries [49]. A recent study showed that ApoE is only upregulated in regions of the mouse brain that have increased levels of Aβ [43], indicating a direct link between the two proteins. Although flies lack a homolog of ApoE, they do possess a homolog of the related apolipoprotein ApoB (Apolpp) [50], which contributes to AD in mice [51,52] and is correlated with AD in humans [53,54]. Interestingly, whilst it was not identified by IM-DIA-MS, ApoB interacts with 12 significantly altered proteins in the STRING network, so is included in the subgraph induced on the significantly altered proteins and their neighbours. ApoB was found in the second highest scoring module that contains proteins involved in translation and glucose transport (Fig 5) [55].

We analysed the 31 proteins significantly altered in normal ageing, but not AD. Of the 29 proteins that were contained in the STRING network, 24 interact directly with at least one of the AD significantly altered proteins, suggesting an interplay between ageing and AD at the pathway level. Using a subgraph of the STRING network induced on these proteins and their 1603 neighbours, we identified eight network modules that were enriched for ageing processes [56], including respiration, unfolded protein and oxidative damage stress responses, cell cycle regulation, DNA damage repair, and apoptosis.

Discussion

Despite the substantial research effort spent on finding drugs against AD, effective treatments remain elusive. We need to better understand the molecular processes that govern the onset and progression of the complex pathologies observed in AD. This knowledge will help to identify new drug targets to treat and prevent AD.

Analysis of post-mortem human brain tissue is an important way to study dementia, but cannot capture the progression of pathology from the initiation of disease. Due to their short lifespan and ease of genetic manipulation, model organisms such as *Drosophila melanogaster* provide a tractable system in which to examine the progression of AD pathology across life. We performed a longitudinal study of the *Drosophila* brain proteome, using an inducible model of AD, label-free quantitative IM-DIA-MS and network analyses. We were able to track alterations in protein levels from the point of exposure to human Aβ42 and the widespread interaction of Aβ42 with brain signalling networks as pathology progresses.

Our proteomic analyses identified Aβ42-induced alterations in levels of 228 proteins, which clustered into four groups: those which were either elevated (cluster 1) or reduced (cluster 2) in AD relative to controls throughout life, those which were altered in correlation with ageing in healthy and Aβ42 flies (cluster 3), and those which changed in Aβ42 flies across life but independently of ageing-dependent effects in healthy controls (cluster 4). Further computational analysis of these proteins revealed significant network properties within the fly brain proteome. Assessing hub and bottleneck properties, many of the Aβ42-induced proteomic changes represented alterations in bottleneck proteins suggesting that they play key roles in downstream cellular function. Of these, some display non-hub properties indicating that they are important for maintaining cellular homeostasis in a targeted fashion, whereas others also displayed hub properties suggesting that they are central in linking cellular signalling pathways to maintain cell function.

We identified five nonhub-bottleneck proteins and four hub-bottleneck proteins, the expression of which was altered in Aβ42 flies relative to controls across life. Due to the importance of these hub and bottleneck proteins in the protein interaction network, we predict that AD-associated alterations in their abundance will likely have a significant effect on the cellular dynamics of the brain. Indeed, these proteins play key molecular roles in metabolism (AscI, Echs1, Got2), protein homeostasis (Hsp70A, Gp93), and protection against oxidative stress (mt:CoII) and DNA damage (Top2). These processes have been shown to affect neuronal function and protection against proteo-toxicity. Alterations in these

proteins may represent either adaptive responses to the presence of abnormal protein aggregates, such as Aβ42, or mediators of neuronal toxicity. Further functional genomic studies are therefore required to establish the causal role of these processes in governing onset and progression of AD pathology.

Assessing the human orthologs of these genes, identified using DIOPT [57], indicates that several of these bottleneck proteins have been previously implicated in association with AD or other neurological conditions in humans or mammalian models of disease. ACSL4 (Acs1 ortholog) has been shown to associate with synaptic growth cone development and mental retardation [58]. Mutations in ECHS1 (Echs1 ortholog), an enzyme involved in mitochondrial fatty acid oxidation, associate with Leigh Syndrome, a severe developmental neurological disorder [59]. Proteomic studies have revealed that GOT2 (Got2 ortholog) is down-regulated in infarct regions following stroke [60], and in AD patient brain [61]. Integrating data from human post-mortem brain studies, HSPA1A (Hsp70Aa ortholog) upregulates in the protein interaction network of AD patients compared to healthy controls [62], and has recently been suggested to block APP processing and Aβ production in mouse brain [63]. Synthetic, fibrillar, Aβ42 reduces expression of TOP2B (Top2 ortholog) in rat cerebellar granule cells and in a human mesenchymal cell line, suggesting this may contribute to DNA damage in response to amyloid [64]. HSP90B1 (Gp93 ortholog) shows increased expression following TBI in mice [65], and associates with animal models of Huntington's disease [66]. Finally, ACTB (Act57B ortholog) has been implicated as a significant AD risk gene and central hub node using integrated network analyses across GWAS [67].

ACSL4, ECHS1, and HSP90B1 have no reported association with AD or related dementias, however, which suggests that our study has potential to identify new targets in the molecular pathogenesis of this disease. Our study also provides additional information about the homeostasis of these proteins across life from the point of amyloid production. For example, the abundances of Acs1 and Got2 are elevated following Aβ42 induction and continue to increase with age relative to controls. Echs1 is reduced in Aβ42 flies compared to controls but increases across life in parallel with ageing-dependent increases in this protein. Structural proteins Acp65Aa and Act57B are elevated in response to Aβ42 but decline across life whilst remaining stable in control flies. Gp93 and Top2 are either elevated or reduced in response to Aβ42 but mirror ageing-dependent alterations in their expression. mt:CoII is reduced following Aβ42 expression at all time-points, but reduced with ageing in controls. Hsp70A is increased early in Aβ42 flies, reduced to control levels in mid-life then elevated at late pathological stages whilst remaining stable in healthy controls.

Analysing Gene Ontology enrichment using network modules, to capture the diverse

biological processes modified in AD, we identified 12 modules enriched for processes previously implicated in ageing and AD. This validates the use of our *Drosophila* model in identifying progressive molecular changes in response to Aβ42 that are likely to correlate with progression of cognitive decline in human disease. Further work is required to modify the genes identified in our study at different ages, in order to elucidate whether they represent mediators of toxicity as disease progresses, factors which increase neuronal susceptibility to disease with age or compensatory protective mechanisms. Model organisms will be essential in unravelling these complex interactions. Our study therefore forms a basis for future analyses that may identify new targets for disease intervention that are specific to age and/or pathological stage of AD.

Materials and methods

Fly stocks

The TgAD fly line used in this study [23] contains the human transgene encoding the Arctic mutant Aβ42 peptide under the control of an Upstream Activation Sequence (UAS) [68]. Expression of Aβ42 was controlled by GeneSwitch [69]—a mifepristone-inducible GAL4/UAS expression system—under the pan-neuronal elav promoter. All flies were backcrossed for six generations into the w^{1118} genetic background.

Flies were grown in 200 ml bottles on a 12 h/12 h light/dark cycle at constant temperature (25 °C) and humidity. Growth media contained 15 g/l agar, 50 g/l sugar, 100 g/l autolysed yeast, 100 g/l nipagin and 3 ml/l propionic acid. Flies were maintained for two days after eclosion before females were transferred to vials at a density of 25 flies per vial for the lifespan analysis and 10 flies per vial for the IM-DIA-MS analysis. Expression of Aβ42 was induced in TgAD flies by spiking the growth media with mifepristone to a final concentration of 200 μM. Flies were transferred to fresh media three times per week, at which point the number of surviving flies was recorded. For each of the three biological repeats, 10 healthy and 10 Aβ42 flies were collected at 5, 19, 31 and 46 days, as well as 54 and 80 days for healthy flies. Following anesthetisation with CO_2, brains were dissected in ice cold 10 mM phosphate buffered saline snap frozen and stored at -80°C.

Extraction of brain proteins

Brain proteins were extracted by homogenisation on ice into 50 μl of 50 mM ammonium bicarbonate, 10 mM DTT and 0.25% RapiGest detergent. Proteins were solubilised and disulfide bonds were reduced by heating at 80°C for 20 minutes. Free cysteine thiols were alkylated by adding 20 mM IAA and incubating at room temperature for 20 minutes in darkness. Protein concentration was determined and samples were diluted to a final concentration of 0.1% RapiGest using 50 mM ammonium bicarbonate. Proteins were digested with trypsin overnight at 37°C at a 50:1 protein:trypsin ratio. Additional trypsin was added at a 100:1 ratio the following morning and incubated for a further hour. Detergent was removed by incubating at 60°C for 1 hour in 0.1% formic acid. Insoluble debris was removed by centrifugation at 14,000 x g for 30 minutes. Supernatant was collected, lyophilised and stored at -80°C. Prior to lyophilisation peptide concentration was estimated by nanodrop (Thermo Fisher Scientific, Waltham, MA).

Label-free quantitative IM-DIA-MS

Peptides were separated by nanoscale liquid chromatography (LC) by loading 300 ng of protein onto an analytical reversed phase column. IM-DIA-MS analysis was performed using a Synapt G2-Si mass spectrometer (Waters Corporation, Manchester, UK). The time-of-flight analyzer of the instrument was externally calibrated with a NaCsI mixture from m/z 50 to 1990. Spectra were acquired over a range of 50–2000 m/z. Each biological repeat was analysed at least twice to account for technical variation.

LC-MS data were peak detected and aligned by Progenesis QI for proteomics (Waters Corporation). The principles of the embedded search algorithm for DIA data has been described previously [70]. Proteins were identified by searching against the *Drosophila melanogaster* proteome in UniProt, appended with common contaminants, and revered sequence entries to estimate protein identification false discovery rate (FDR) values, using previously specified search criteria [71]. Peptide intensities were normalised to control for variation in protein loading and relative quantification. Abundances were estimated by Hi3-based quantitation [72].

Data analysis

Proteins that were identified in both healthy and Aβ42 flies were considered for further analysis. Missing data were replaced by the minimum abundance measured for any protein in the same repeat [48]. The data were quantile normalised [73], so that different conditions and time points could be compared reliably. Quantile normalisation transforms the abundances so that each repeat has the same distribution.

For PCA analysis, the data were $\log_{10}$-transformed and each protein was standardised to zero mean and unit variance. Hierarchical biclustering was performed using the Euclidean distance metric with the complete linkage method. Prior to clustering, proteins were normalised to their abundance in healthy flies at 5 days.

Proteins that were identified by IM-DIA-MS in either healthy or Aβ42 flies were assessed for overrepresentation of Gene Ontology terms using GOrilla [74], which uses ranked lists of target and background genes. Proteins were ranked in descending order by their mean abundance. The type I error rate was controlled by correcting for multiple testing using the Benjamini-Hochberg method at an FDR of 5%.

Clusters of proteins were assessed for overrepresentation of GO-Slim terms in the Biological Process ontology using Panther (version 13.1) with a custom background of the 3093

proteins identified by IM-DIA-MS in healthy or AD flies.

Identification of significantly altered proteins

Significantly altered proteins were identified using five methods that are frequently used to identify differentially expressed genes in time course RNA-Seq data. DESeq2 [31], EDGE [33], edgeR [30], limma [32] and maSigPro [34] are all available in R through Bioconductor. Dispersions were estimated from the biological and technical repeats. Unless otherwise stated, default parameters were used for all methods under the null hypothesis that a protein does not change in abundance between healthy and AD conditions in normal ageing. The type I error rate was controlled by correcting for multiple testing using the Benjamini-Hochberg method at a FDR of 5%. A protein was classified as significantly altered if two or more methods identified it.

DESeq2 models proteins with the negative binomial distribution and performs likelihood ratio tests. A time course experiment was selected in EDGE using the likelihood ratio test and a normal null distribution. edgeR uses the negative binomial distribution and performs quasi-likelihood tests. limma fits linear models to the proteins and performed empirical Bayes F-tests. maSigPro fits generalised linear models to the proteins and performs log-likelihood ratio tests.

Significantly altered proteins were clustered using a Gaussian mixture model. Protein abundances were log10-transformed and *z* scores were calculated. Gaussian mixture models were implemented for 1–228 clusters. The best model was chosen using the Bayesian information criterion (BIC), which penalises complex models:

$$BIC = -2\ln(L) + \ln(n)k$$

where ln(L) is the log-likelihood of the model, n is the number of significantly altered proteins and k is the number of clusters. The model with lowest BIC was chosen.

Networks

All network analysis was performed using the *Drosophila melanogaster* STRING network (version 10) [35]. Low confidence interactions with a ‘combined score’ < 500 were removed in all network analyses.

Network properties of the significantly altered proteins were analysed in the brain protein interaction network. A subgraph of the STRING network was induced on the 3093 proteins identified by IM-DIA-MS in healthy or Aβ42 flies and the largest connected component was

selected (2428 nodes and 44,561 edges). The subgraph contained 183 of the 228 significantly altered proteins. For these proteins, four network properties were calculated as test statistics: mean node degree; mean unweighted shortest path length between a node and the remaining 182 nodes; the size of the largest connected component in the subgraph induced on these nodes; and mean betweenness centrality. Hypothesis testing was performed using the null hypothesis that there is no difference between the nodes in the subgraph. Assuming the null hypothesis is true, null distributions of each test statistic were simulated by randomly sampling 183 nodes from the network 10,000 times. Using the null distributions, non-parametric one-sided p-values were calculated as the probability of observing a test statistic as extreme as the test statistic for the significantly altered proteins.

A subgraph of the STRING network was induced on the proteins significantly altered in AD and their neighbours and the largest connected component was selected (4842 nodes and 182,474 edges). The subgraph contained 198 of the 228 significantly altered proteins and was assessed for enrichment of Gene Ontology terms. Densely connected subgraphs were identified using MCODE [47]. Modules were selected with an MCODE score > 10. As STRING is a functional interaction network, clusters of nodes may correspond to proteins from the same complex, pathway or functional family. Clusters were assessed for overrepresentation of GO-Slim terms in the Biological Process ontology using Panther (version 13.1) [75] with a custom background of the 3093 proteins identified by IM-DIA-MS in healthy or Aβ42 flies. Fisher's exact tests were performed and the type I error rate was controlled by correcting for multiple testing using the Benjamini-Hochberg method at a FDR of 5%.

Open source software

Data analysis was performed in Python 3.6 (Python Software Foundation, http://www.python.org) using SciPy [76], NumPy [77], Pandas [78], scikit-learn [79], NetworkX [80], IPython [81] and Jupyter [82]. Figures were plotted using Matplotlib [83] and seaborn.

Acknowledgements

We thank Dr Damian Crowther (University of Cambridge) for donation of UAS-Aβ42 fly stocks and Dr Hervé Tricoire (CNRS, France) for donation of elavGS fly stocks. Fig 1C: fly graphic by Daan Kauwenberg and brain graphic by Julia Amadeo, both from the Noun

578 Project.

579 Funding

580 H.S. is supported by an ISMB Wellcome PhD studentship [203780/Z/16/A]; A.C. is
581 supported by a BBSRC CASE PhD studentship and Waters Corporation; J.L. is supported
582 by BBSRC [BB/L002817/1]; L.P. and F.K. are supported by a Wellcome Trust Strategic
583 Award to L.P. [098565/Z/12/Z] and an Alzheimer's Research UK Project Grant (ART-2009-
584 4). The work was supported by a Wellcome instrumentation grant [104913/Z/14/Z].

585 Competing interests

586 None

Supplementary Information

Methods

IM-DIA-MS analysis

Nanoscale liquid chromatography (LC) separation of tryptic peptides was performed using a nanoAcquity UPLC system (Waters Corporation) equipped with a UPLC HSS T3 1.7 µm, 75 µm x 250 mm analytical reverse phase column (Waters Corporation). Prior to peptide separation, 300 ng of tryptic peptides were loaded onto a 2G, V/V 5 µm, 180 µm x 20 mm reverse phase trapping column at 5 µl/min for 3 minutes. IM-DIA-MS analysis of tryptic digests was performed using a Synapt GS-Si mass spectrometer equipped with a T-Wave-IMS device. Mass measurements were made in positive-mode ESI with the instrument operated in resolution mode with a typical resolving power of 20,000 full width at half maximum. Prior to analysis the time-of-flight analyzer was externally calibrated with a NaCsI mixture from *m/z* 50 to 1990. The data were post-acquisition lock mass corrected using the double charged monoisotopic ion of [Glu1]-Fibrinopeptide B. To achieve lock mass correction, a 100 fmol/µl solution of [Glu1]-Fibrinopeptide B was infused at a 90° angle to the analytical sprayer. This reference sprayer was sampled every 60 seconds. Accurate IM-DIA-MS data were collected in the DIA mode of analysis, HDMS^E [24,71] IM spectrometry was performed by applying a constant wave height of 40 V whilst a constant wave velocity of 650 m/s was maintained. Wave heights within the trap and transfer were both set at 4 V whilst the wave velocities were 311 and 175 m/s respectively. MS data were acquired over 50-2000 *m/z for* each mode. Spectral acquisition time for each mode was 0.5 s with a 0.015 interscan delay, corresponding to a cycle of low and elevated energy data being acquired every 1.1 s. During the low energy MS mode data was acquired whilst applying a constant collision energy of 4 eV within the transfer. After IMS, MS/MS data was acquired by ramping the collision energy within the transfer region between 15 and 45 eV. To ensure that ions with a *m/z* less than 350 were derived from peptide fragmentation within the transfer region the radio frequency applied to the quadrupole mass analyser was adjusted to optimise transmission within the region of 350 – 2000 Da. Each biological replicate was analysed at least twice.

MS Data Processing

All MS data were processed in Progenesis QI for proteomics. Data were imported into Progenesis to generate a 3D representation of the data (*m/z*, RT and peak intensity). Samples were then time aligned with the software allowed to automatically determine the

best reference run from the dataset. Following alignment, peak picking was performed on MS level data. A peak picking sensitivity of 4 (out of 5) was set. Peptide features were tentatively aligned with their respective fragment ions based primarily on the similarity of their chromatographic and mobility profiles. Requirements for features to be included in post-processing database searching were as follows: 300 counts for low energy ions, 50 counts for high energy ions and 750 counts for deconvoluted precursor intensities. Subsequent data were searched against 20,049 sequences from the UniProt canonical *Drosophila* database (appended with common contaminants). Trypsin was specified as the enzyme of choice and a maximum of two missed cleavages were permitted. Carbamidomethyl (C) was set as a fixed modification whilst oxidation (M) and N-terminal acetylation were set as variable modifications. Peptide identifications were grouped and relative quantification was performed using non-conflicting peptides only.

Data

Supplementary Data 1
supplementary_data_1.xlsx
Proteomics data

Supplementary Data 2
supplementary_data_2.txt
MCODE modules

Supplementary Data 3
supplementary_data_3.xlsx
Gene Ontology enrichment

Figures

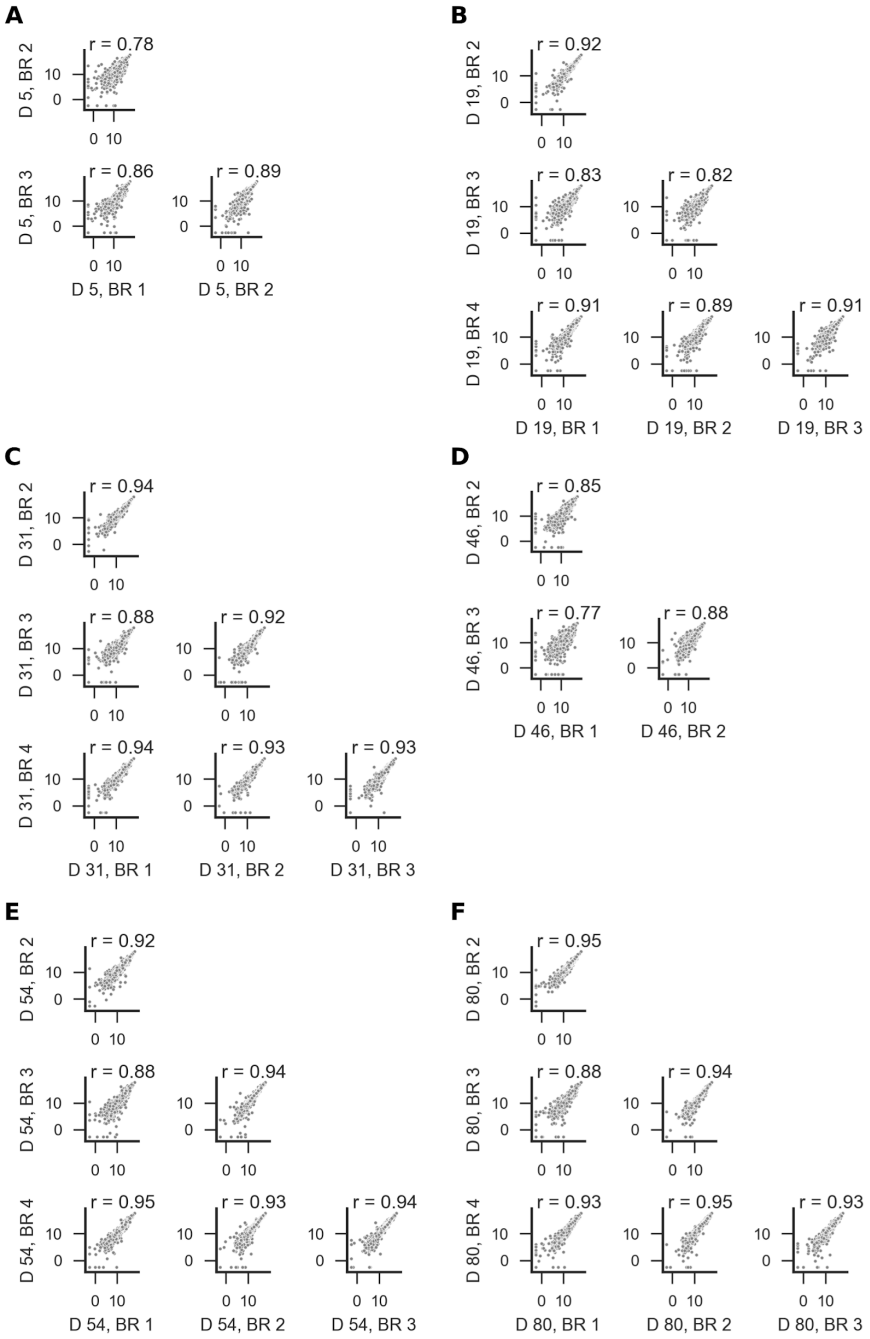


Figure S1: Assessment of experimental reproducibility. Scatter plots comparing protein abundances in different biological repeats (BR) of healthy flies at days (D) (**A**) 5, (**B**) 19, (**C**) 31, (**D**) 46, (**E**) 54 and (**F**) 80. Abundances were log2-transformed before plotting. Pearson correlation coefficients (r) are shown for each pair of biological repeat at each time point.

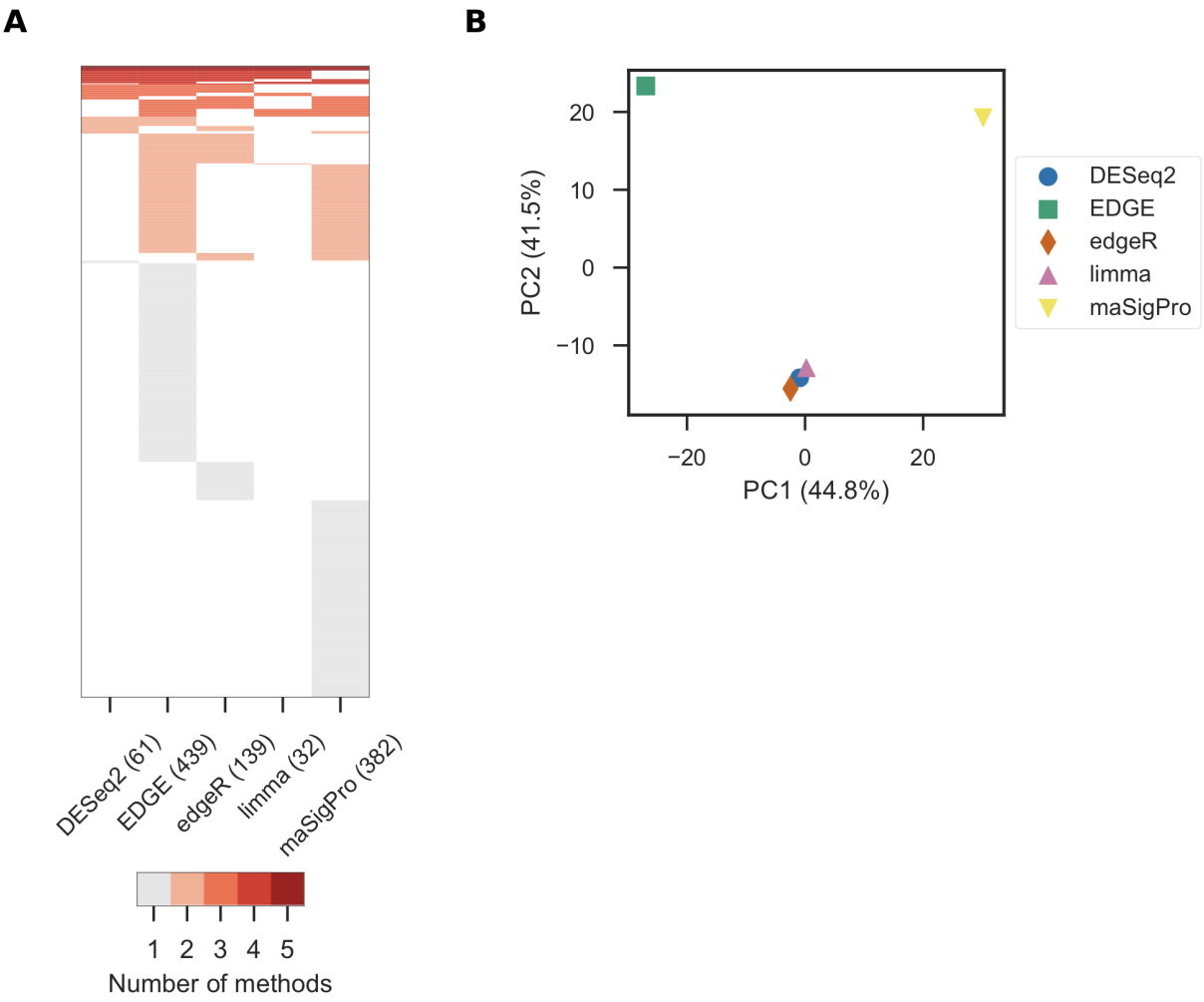


Figure S2: Analysis of the five statistical methods used to identify significantly altered proteins.

(**A**) Heat map of the proteins detected by each method. (**B**) Principal component analysis of these results. Axes are annotated with the percentage of variance explained by each principal component.

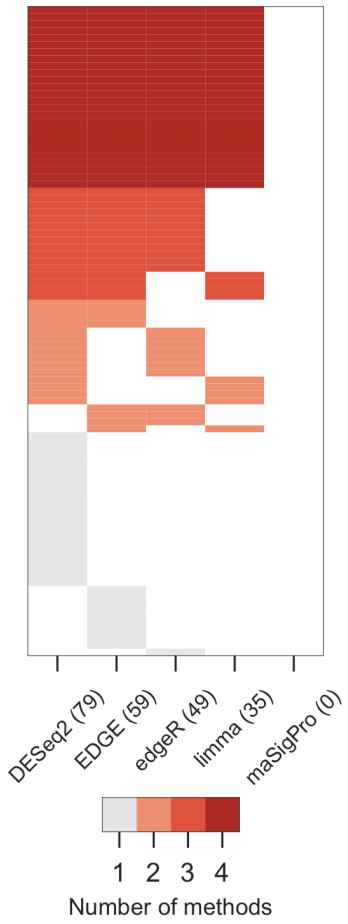


Figure S3: Identification of significantly altered proteins during normal ageing. Heat map of the proteins detected by each method.

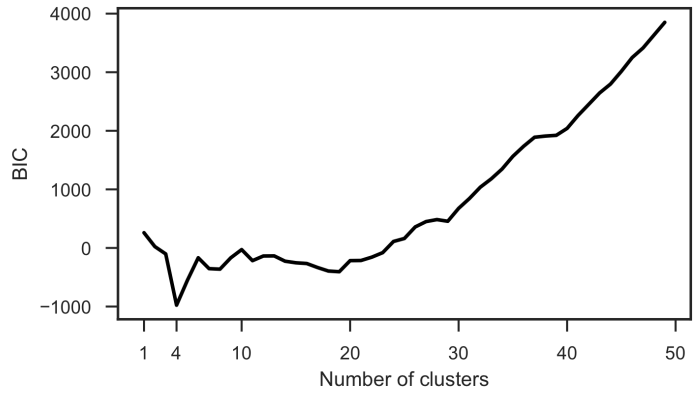


Figure S4: Model selection for clustering of the significantly altered proteins using a Gaussian mixture model. The best model was chosen using the Bayesian information criterion (BIC), which penalises complex models.

References

1. Lane CA, Hardy J, Schott JM. Alzheimer's disease. Eur J Neurol. 2018;25: 59–70. doi:10.1111/ene.13439

2. Glenner GG, Wong CW. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. 1984. Biochem Biophys Res Commun. 2012;425: 534–539. doi:10.1016/j.bbrc.2012.08.020

3. Grundke-Iqbal I, Iqbal K, Tung YC, Quinlan M, Wisniewski HM, Binder LI. Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. Proc Natl Acad Sci U S A. 1986;83: 4913–4917. Available: https://www.ncbi.nlm.nih.gov/pubmed/3088567

4. Goedert M, Wischik CM, Crowther RA, Walker JE, Klug A. Cloning and sequencing of the cDNA encoding a core protein of the paired helical filament of Alzheimer disease: identification as the microtubule-associated protein tau. Proc Natl Acad Sci U S A. 1988;85: 4051–4055. Available: https://www.ncbi.nlm.nih.gov/pubmed/3131773

5. Cai H, Cong W-N, Ji S, Rothman S, Maudsley S, Martin B. Metabolic Dysfunction in Alzheimer's Disease and Related Neurodegenerative Disorders. Curr Alzheimer Res. 2011;9: 5–17. doi:10.2174/156720512799015064

6. Szutowicz A, Bielarczyk H, Jankowska-Kulawy A, Pawełczyk T, Ronowska A. Acetyl-CoA the key factor for survival or death of cholinergic neurons in course of neurodegenerative diseases. Neurochem Res. 2013;38: 1523–1542. doi:10.1007/s11064-013-1060-x

7. Suberbielle E, Sanchez PE, Kravitz AV, Wang X, Ho K, Eilertson K, et al. Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-β. Nat Neurosci. 2013;16: 613–621. doi:10.1038/nn.3356

8. Raina AK, Monteiro MJ, McShea A, Smith MA. The role of cell cycle-mediated events in Alzheimer's disease. Int J Exp Pathol. 1999;80: 71–76. doi:10.1046/j.1365-2613.1999.00106.x

9. Kanaan NM, Pigino GF, Brady ST, Lazarov O, Binder LI, Morfini GA. Axonal degeneration in Alzheimer's disease: when signaling abnormalities meet the axonal transport system. Exp Neurol. 2013;246: 44–53. doi:10.1016/j.expneurol.2012.06.003

10. Donev R, Kolev M, Millet B, Thome J. Neuronal death in Alzheimer's disease and therapeutic opportunities. J Cell Mol Med. 2009;13: 4329–4348. doi:10.1111/j.1582-4934.2009.00889.x

11. Van Cauwenberghe C, Van Broeckhoven C, Sleegers K. The genetic landscape of Alzheimer disease: clinical implications and perspectives. Genet Med. The Author(s); 2016;18: 421–430. doi:10.1038/gim.2015.117

12. Zhang Y, McLaughlin R, Goodyer C, LeBlanc A. Selective cytotoxicity of intracellular amyloid beta peptide1-42 through p53 and Bax in cultured primary human neurons. J Cell Biol. 2002;156: 519–529. doi:10.1083/jcb.200110119

13. McGowan E, Pickford F, Kim J, Onstead L, Eriksen J, Yu C, et al. Abeta42 is essential for parenchymal and vascular amyloid deposition in mice. Neuron. 2005;47: 191–199. doi:10.1016/j.neuron.2005.06.030

14. Götz J, Chen F, van Dorpe J, Nitsch RM. Formation of neurofibrillary tangles in P301l tau transgenic mice induced by Abeta 42 fibrils. Science. 2001;293: 1491–1495. doi:10.1126/science.1062097

15. Mullan M, Crawford F, Axelman K, Houlden H, Lilius L, Winblad B, et al. A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. Nat Genet. 1992;1: 345–347. doi:10.1038/ng0892-345

16. Nilsberth C, Westlind-Danielsson A, Eckman CB, Condron MM, Axelman K, Forsell C, et al. The "Arctic" APP mutation (E693G) causes Alzheimer's disease by enhanced Abeta protofibril formation. Nat Neurosci. 2001;4: 887–893. doi:10.1038/nn0901-887

17. Moya-Alvarado G, Gershoni-Emek N, Perlson E, Bronfman FC. Neurodegeneration and Alzheimer's disease (AD). What Can Proteomics Tell Us About the Alzheimer's Brain? Mol Cell Proteomics. 2016;15: 409–425. doi:10.1074/mcp.R115.053330

18. Lynn BC, Wang J, Markesbery WR, Lovell MA. Quantitative changes in the mitochondrial proteome from subjects with mild cognitive impairment, early stage, and late stage Alzheimer's disease. J Alzheimers Dis. 2010;19: 325–339. doi:10.3233/JAD-2010-1254

19. Butterfield DA, Di Domenico F, Swomley AM, Head E, Perluigi M. Redox proteomics analysis to decipher the neurobiology of Alzheimer-like neurodegeneration: overlaps in Down's syndrome and Alzheimer's disease brain. Biochem J. 2014;463: 177–189. doi:10.1042/BJ20140772

20. Aluise CD, Robinson RAS, Cai J, Pierce WM, Markesbery WR, Butterfield DA. Redox proteomics analysis of brains from subjects with amnestic mild cognitive impairment compared to brains from subjects with preclinical Alzheimer's disease: insights into memory loss in MCI. J Alzheimers Dis. 2011;23: 257–269. doi:10.3233/JAD-2010-101083

21. Dammer EB, Lee AK, Duong DM, Gearing M, Lah JJ, Levey AI, et al. Quantitative phosphoproteomics of Alzheimer's disease reveals cross-talk between kinases and small heat shock proteins. Proteomics. 2015;15: 508–519. doi:10.1002/pmic.201400189

22. Sultana R, Robinson RAS, Di Domenico F, Abdul HM, St Clair DK, Markesbery WR, et al. Proteomic identification of specifically carbonylated brain proteins in APP(NLh)/APP(NLh) × PS-1(P264L)/PS-1(P264L) human double mutant knock-in mice model of Alzheimer disease as a function of age. J Proteomics. 2011;74: 2430–2440. doi:10.1016/j.jprot.2011.06.015

23. Sofola O, Kerr F, Rogers I, Killick R, Augustin H, Gandy C, et al. Inhibition of GSK-3 Ameliorates Aβ Pathology in an Adult-Onset Drosophila Model of Alzheimer's Disease. Lu B, editor. PLoS Genet. Public Library of Science; 2010;6: e1001087. doi:10.1371/journal.pgen.1001087

24. Rodriguez-Suarez E, Hughes C, Gethings L, Giles K, Wildgoose J, Stapels M, et al. An Ion Mobility Assisted Data Independent LC-MS Strategy for the Analysis of Complex Biological Samples. Curr Anal Chem. 2013;9: 199–211. Available: https://www.ingentaconnect.com/content/ben/cac/2013/00000009/00000002/art00006

25. Brown CJ, Kaufman T, Trinidad JC, Clemmer DE. Proteome changes in the aging Drosophila melanogaster head. Int J Mass Spectrom. 2018;425: 36–46. doi:10.1016/j.ijms.2018.01.003

748 26. Tain LS, Sehlke R, Jain C, Chokkalingam M, Nagaraj N, Essers P, et al. A proteomic
749 atlas of insulin signalling reveals tissue-specific mechanisms of longevity assurance.
750 Mol Syst Biol. 2017;13: 939. doi:10.15252/msb.20177663

751 27. Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, Huber W, et al. Count-
752 based differential expression analysis of RNA sequencing data using R and
753 Bioconductor. Nat Protoc. 2013;8: 1765–1786. doi:10.1038/nprot.2013.099

754 28. Zhang ZH, Jhaveri DJ, Marshall VM, Bauer DC, Edson J, Narayanan RK, et al. A
755 comparative study of techniques for differential expression analysis on RNA-Seq data.
756 PLoS One. 2014;9: e103207. doi:10.1371/journal.pone.0103207

757 29. Seyednasrollah F, Laiho A, Elo LL. Comparison of software packages for detecting
758 differential expression in RNA-seq studies. Brief Bioinform. 2015;16: 59–70.
759 doi:10.1093/bib/bbt086

760 30. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential
761 expression analysis of digital gene expression data. Bioinformatics. 2010;26: 139–140.
762 doi:10.1093/bioinformatics/btp616

763 31. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for
764 RNA-seq data with DESeq2. Genome Biol. 2014;15: 550. doi:10.1186/s13059-014-
765 0550-8

766 32. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential
767 expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res.
768 2015;43: e47. doi:10.1093/nar/gkv007

769 33. Woo S, Leek JT, Storey JD. A computationally efficient modular optimal discovery
770 procedure. Bioinformatics. 2011;27: 509–515. doi:10.1093/bioinformatics/btq701

771 34. Nueda MJ, Tarazona S, Conesa A. Next maSigPro: updating maSigPro bioconductor
772 package for RNA-seq time series. Bioinformatics. 2014;30: 2598–2602.
773 doi:10.1093/bioinformatics/btu333

774 35. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING
775 database in 2017: quality-controlled protein-protein association networks, made broadly
776 accessible. Nucleic Acids Res. 2017;45: D362–D368. doi:10.1093/nar/gkw937

777 36. Yu H, Kim PM, Sprecher E, Trifonov V, Gerstein M. The importance of bottlenecks in
778 protein networks: correlation with gene essentiality and expression dynamics. PLoS
779 Comput Biol. 2007;3: e59. doi:10.1371/journal.pcbi.0030059

780 37. Maas AI. Cerebrospinal fluid enzymes in acute brain injury. 2. Relation of CSF enzyme
781 activity to extent of brain injury. J Neurol Neurosurg Psychiatry. 1977;40: 666–674.
782 doi:10.1136/jnnp.40.7.666

783 38. Casley CS, Canevari L, Land JM, Clark JB, Sharpe MA. Beta-amyloid inhibits integrated
784 mitochondrial respiration and key enzyme activities. J Neurochem. 2002;80: 91–100.
785 doi:10.1046/j.0022-3042.2001.00681.x

786 39. Cardoso SM, Proença MT, Santos S, Santana I, Oliveira CR. Cytochrome c oxidase is
787 decreased in Alzheimer's disease platelets. Neurobiol Aging. 2004;25: 105–110.
788 doi:10.1016/S0197-4580(03)00033-2

789 40. Fukui H, Diaz F, Garcia S, Moraes CT. Cytochrome c oxidase deficiency in neurons
790 decreases both oxidative stress and amyloid formation in a mouse model of Alzheimer's

disease. Proc Natl Acad Sci U S A. 2007;104: 14163–14168. doi:10.1073/pnas.0705738104

41. Castellani RJ, Siedlak SL, Fortino AE, Perry G, Ghetti B, Smith MA. Chitin-like polysaccharides in Alzheimer's disease brains. Curr Alzheimer Res. 2005;2: 419–423. Available: https://www.ncbi.nlm.nih.gov/pubmed/16248847

42. Rogers I, Kerr F, Martinez P, Hardy J, Lovestone S, Partridge L. Ageing Increases Vulnerability to Aβ42 Toxicity in Drosophila. Iijima KM, editor. PLoS One. 2012;7: e40569. doi:10.1371/journal.pone.0040569

43. Savas JN, Wang Y-Z, DeNardo LA, Martinez-Bartolome S, McClatchy DB, Hark TJ, et al. Amyloid Accumulation Drives Proteome-wide Alterations in Mouse Models of Alzheimer's Disease-like Pathology. Cell Rep. 2017;21: 2614–2627. doi:10.1016/j.celrep.2017.11.009

44. Kommaddi RP, Das D, Karunakaran S, Nanguneri S, Bapat D, Ray A, et al. Aβ mediates F-actin disassembly in dendritic spines leading to cognitive deficits in Alzheimer's disease. J Neurosci. 2018;38: 1085–1099. doi:10.1523/JNEUROSCI.2127-17.2017

45. Jacobson GR, Rosenbusch JP. ATP binding to a protease-resistant core of actin. Proc Natl Acad Sci U S A. 1976;73: 2742–2746. Available: https://www.ncbi.nlm.nih.gov/pubmed/134374

46. Hozumi T. Structural aspects of skeletal muscle F-actin as studied by tryptic digestion: evidence for a second nucleotide interacting site. J Biochem. 1988;104: 285–288. Available: https://www.ncbi.nlm.nih.gov/pubmed/2972700

47. Bader GD, Hogue CWV. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003;4: 2. doi:10.1186/1471-2105-4-2

48. Lazar C, Gatto L, Ferro M, Bruley C, Burger T. Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies. J Proteome Res. 2016;15: 1116–1125. doi:10.1021/acs.jproteome.5b00981

49. Liu C-C, Liu C-C, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat Rev Neurol. 2013;9: 106–118. doi:10.1038/nrneurol.2012.263

50. Palm W, Sampaio JL, Brankatschk M, Carvalho M, Mahmoud A, Shevchenko A, et al. Lipoproteins in Drosophila melanogaster--assembly, function, and influence on tissue lipid composition. PLoS Genet. 2012;8: e1002828. doi:10.1371/journal.pgen.1002828

51. Bereczki E, Bernát G, Csont T, Ferdinandy P, Scheich H, Sántha M. Overexpression of human apolipoprotein B-100 induces severe neurodegeneration in transgenic mice. J Proteome Res. 2008;7: 2246–2252. doi:10.1021/pr7006329

52. Löffler T, Flunkert S, Havas D, Sántha M, Hutter-Paier B, Steyrer E, et al. Impact of ApoB-100 expression on cognition and brain pathology in wild-type and hAPPsl mice. Neurobiol Aging. 2013;34: 2379–2388. doi:10.1016/j.neurobiolaging.2013.04.008

53. Caramelli P, Nitrini R, Maranhão R, Lourenço AC, Damasceno MC, Vinagre C, et al. Increased apolipoprotein B serum concentration in Alzheimer's disease. Acta Neurol Scand. 1999;100: 61–63. doi:10.1111/j.1600-0404.1999.tb00724.x

54. Zhang R, Barker L, Pinchev D, Marshall J, Rasamoelisolo M, Smith C, et al. Mining

biomarkers in human sera using proteomic tools. Proteomics. 2004;4: 244–256. doi:10.1002/pmic.200300495

55. Niccoli T, Cabecinha M, Tillmann A, Kerr F, Wong CT, Cardenes D, et al. Increased Glucose Transport into Neurons Rescues Aβ Toxicity in Drosophila. Curr Biol. 2016;26: 2291–2300. doi:10.1016/j.cub.2016.07.017

56. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell. 2013;153: 1194–1217. doi:10.1016/j.cell.2013.05.039

57. Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N, et al. An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinformatics. 2011;12: 357. doi:10.1186/1471-2105-12-357

58. Meloni I, Muscettola M, Raynaud M, Longo I, Bruttini M, Moizard M-P, et al. *FACL4*, encoding fatty acid-CoA ligase 4, is mutated in nonspecific X-linked mental retardation. Nat Genet. 2002;30: 436–440. doi:10.1038/ng857

59. Peters H, Buck N, Wanders R, Ruiter J, Waterham H, Koster J, et al. ECHS1 mutations in Leigh disease: a new inborn error of metabolism affecting valine metabolism. Brain. 2014;137: 2903–2908. doi:10.1093/brain/awu216

60. Datta A, Akatsu H, Heese K, Sze SK. Quantitative clinical proteomic study of autopsied human infarcted brain specimens to elucidate the deregulated pathways in ischemic stroke pathology. J Proteomics. 2013;91: 556–568. doi:10.1016/j.jprot.2013.08.017

61. McKenzie AT, Moyon S, Wang M, Katsyv I, Song W-M, Zhou X, et al. Multiscale network modeling of oligodendrocytes reveals molecular components of myelin dysregulation in Alzheimer's disease. Mol Neurodegener. 2017;12: 82. doi:10.1186/s13024-017-0219-3

62. Chi L-M, Wang X, Nan G-X. In silico analyses for molecular genetic mechanism and candidate genes in patients with Alzheimer's disease. Acta Neurol Belg. 2016;116: 543–547. doi:10.1007/s13760-016-0613-6

63. Gerber H, Mosser S, Boury-Jamot B, Stumpe M, Piersigilli A, Goepfert C, et al. The APMAP interactome reveals new modulators of APP processing and beta-amyloid production that are altered in Alzheimer's disease. Acta Neuropathol Commun. 2019;7: 13. doi:10.1186/s40478-019-0660-3

64. Terzioglu-Usak S, Negis Y, Karabulut DS, Zaim M, Isik S. Cellular Model of Alzheimer's Disease: Aβ1-42 Peptide Induces Amyloid Deposition and a Decrease in Topo Isomerase IIβ and Nurr1 Expression. Curr Alzheimer Res. 2017;14: 636–644. doi:10.2174/1567205014666170117103217

65. Tzekov R, Dawson C, Orlando M, Mouzon B, Reed J, Evans J, et al. Sub-Chronic Neuropathological and Biochemical Changes in Mouse Visual System after Repetitive Mild Traumatic Brain Injury. PLoS One. 2016;11: e0153608. doi:10.1371/journal.pone.0153608

66. Kalathur RKR, Giner-Lamia J, Machado S, Barata T, Ayasolla KRS, Futschik ME. The unfolded protein response and its potential role in Huntington's disease elucidated by a systems biology approach. F1000Res. 2015;4: 103. doi:10.12688/f1000research.6358.2

67. Talwar P, Silla Y, Grover S, Gupta M, Agarwal R, Kushwaha S, et al. Genomic convergence and network analysis approach to identify candidate genes in Alzheimer's

877 disease. BMC Genomics. 2014;15: 199. doi:10.1186/1471-2164-15-199

878 68. Crowther DC, Kinghorn KJ, Miranda E, Page R, Curry JA, Duthie FAI, et al.
879 Intraneuronal Abeta, non-amyloid aggregates and neurodegeneration in a Drosophila
880 model of Alzheimer's disease. Neuroscience. 2005;132: 123–135.
881 doi:10.1016/j.neuroscience.2004.12.025

882 69. Osterwalder T, Yoon KS, White BH, Keshishian H. A conditional tissue-specific
883 transgene expression system using inducible GAL4. Proc Natl Acad Sci U S A. 2001;98:
884 12596–12601. doi:10.1073/pnas.221303298

885 70. Li G-Z, Vissers JPC, Silva JC, Golick D, Gorenstein MV, Geromanos SJ. Database
886 searching and accounting of multiplexed precursor and product ion spectra from the
887 data independent analysis of simple and complex peptide mixtures. Proteomics. 2009;9:
888 1696–1719. doi:10.1002/pmic.200800564

889 71. Distler U, Kuharev J, Navarro P, Levin Y, Schild H, Tenzer S. Drift time-specific collision
890 energies enable deep-coverage data-independent acquisition proteomics. Nat Methods.
891 2014;11: 167–170. doi:10.1038/nmeth.2767

892 72. Silva JC, Gorenstein MV, Li G-Z, Vissers JPC, Geromanos SJ. Absolute quantification
893 of proteins by LCMSE: a virtue of parallel MS acquisition. Mol Cell Proteomics. 2006;5:
894 144–156. doi:10.1074/mcp.M500230-MCP200

895 73. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods
896 for high density oligonucleotide array data based on variance and bias. Bioinformatics.
897 2003;19: 185–193. doi:10.1093/bioinformatics/19.2.185

898 74. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and
899 visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:
900 48. doi:10.1186/1471-2105-10-48

901 75. Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis
902 with the PANTHER classification system. Nat Protoc. 2013;8: 1551–1566.
903 doi:10.1038/nprot.2013.092

904 76. Oliphant TE. SciPy: Open source scientific tools for Python. Computing in Science and
905 Engineering. 2007;9: 10–20.

906 77. Oliphant TE. A guide to NumPy [Internet]. Trelgol Publishing USA; 2006. Available:
907 http://ns.ael.ru/ports/distfiles/numpybook.pdf

908 78. McKinney W, Others. Data structures for statistical computing in python. Proceedings of
909 the 9th Python in Science Conference. Austin, TX; 2010. pp. 51–56. Available:
910 https://pdfs.semanticscholar.org/f6da/c1c52d3b07c993fe52513b8964f86e8fe381.pdf

911 79. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-
912 learn: Machine Learning in Python. J Mach Learn Res. 2011;12: 2825–2830. Available:
913 http://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

914 80. Hagberg A, Swart P, S Chult D. Exploring network structure, dynamics, and function
915 using NetworkX [Internet]. Los Alamos National Lab.(LANL), Los Alamos, NM (United
916 States); 2008. Available: https://www.osti.gov/biblio/960616

917 81. Perez F, Granger BE. IPython: A System for Interactive Scientific Computing.
918 Computing in Science Engineering. 2007;9: 21–29. doi:10.1109/MCSE.2007.53

82. Kluyver T, Ragan-Kelley B, Pérez F, Granger BE, Bussonnier M, Frederic J, et al. Jupyter Notebooks-a publishing format for reproducible computational workflows. ELPUB. 2016. pp. 87–90. Available: https://books.google.com/books?hl=en&lr=&id=Lgy3DAAAQBAJ&oi=fnd&pg=PA87&dq=jupyter&ots=N1AZ8RuCbo&sig=s-Q__Qm1hHR7bJmgi6x_Y8NGCzE

83. Hunter JD. Matplotlib: A 2D Graphics Environment. Comput Sci Eng. IEEE Computer Society; 2007;9: 90–95. doi:10.1109/MCSE.2007.55

84. Müller UC, Deller T, Korte M. Not just amyloid: physiological functions of the amyloid precursor protein family. Nat Rev Neurosci. 2017;18: 281–298. doi:10.1038/nrn.2017.29

85. Venegas C, Kumar S, Franklin BS, Dierkes T, Brinkschulte R, Tejera D, et al. Microglia-derived ASC specks cross-seed amyloid-β in Alzheimer's disease. Nature. 2017;552: 355–361. doi:10.1038/nature25158

86. Yagi Y, Tomita S, Nakamura M, Suzuki T. Overexpression of human amyloid precursor protein in Drosophila. Mol Cell Biol Res Commun. 2000;4: 43–49. doi:10.1006/mcbr.2000.0248

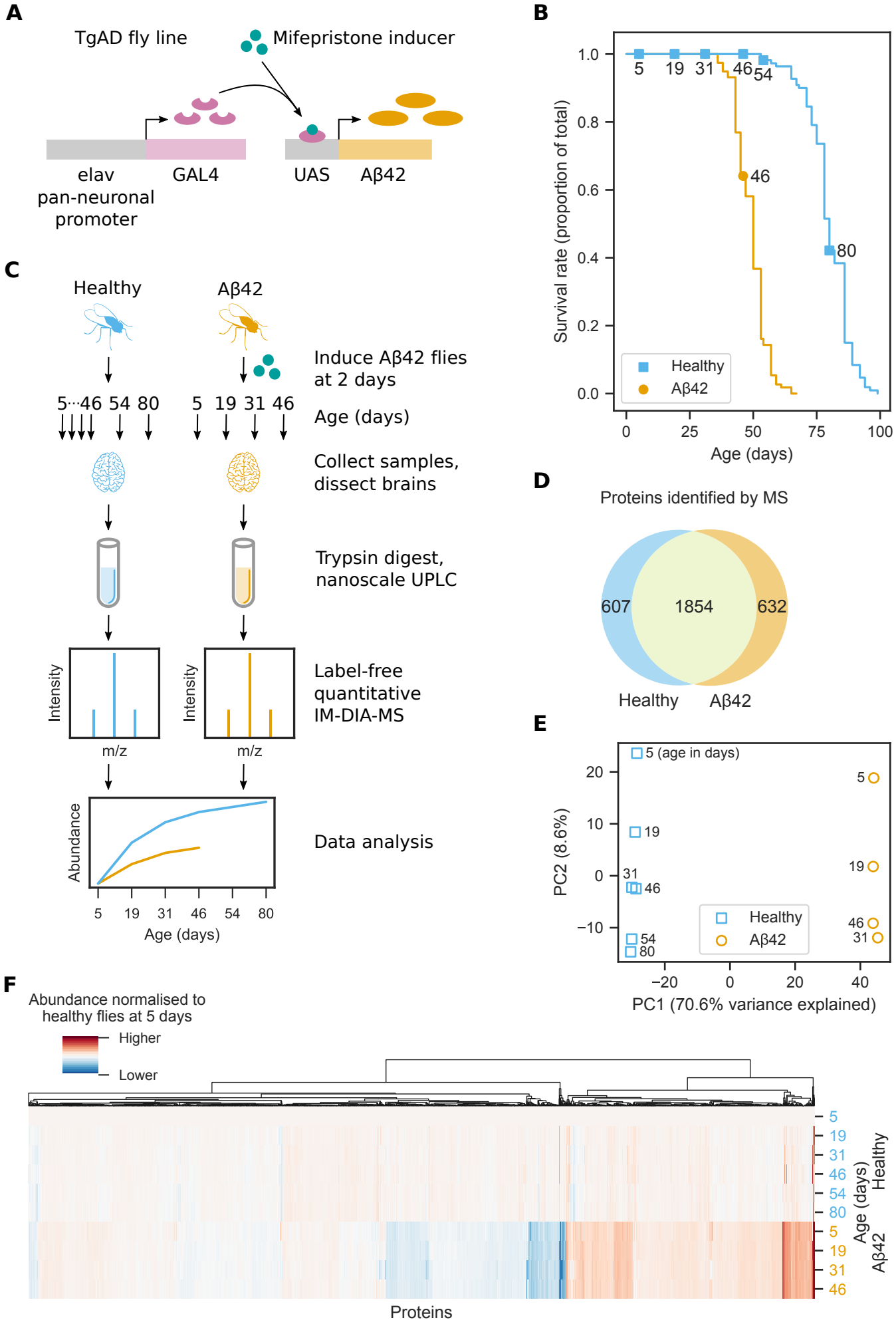
87. Li Y, Liu T, Peng Y, Yuan C, Guo A. Specific functions of Drosophila amyloid precursor-like protein in the development of nervous system and nonneural tissues. J Neurobiol. 2004;61: 343–358. doi:10.1002/neu.20048

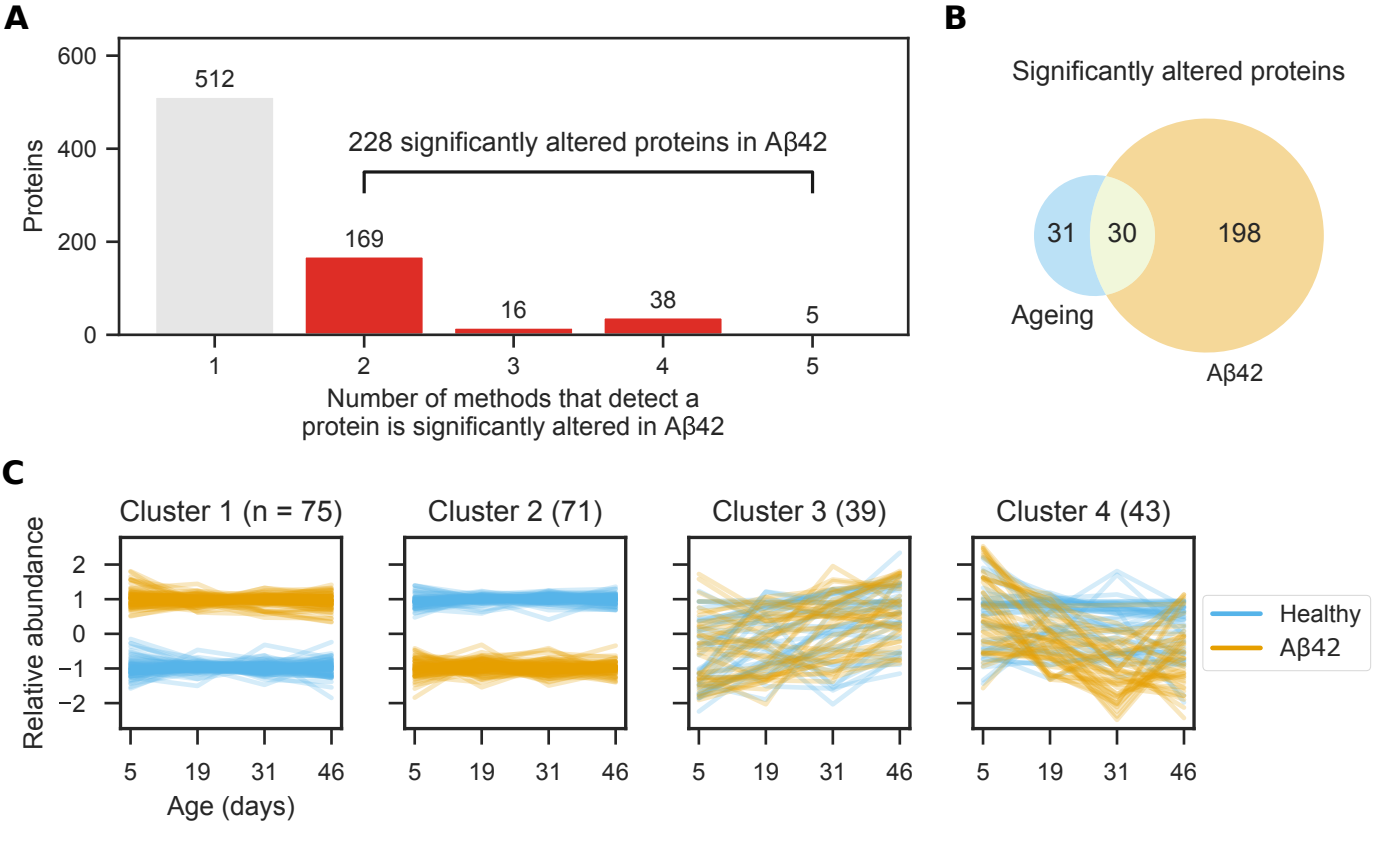
88. Torroja L, Chu H, Kotovsky I, White K. Neuronal overexpression of APPL, the Drosophila homologue of the amyloid precursor protein (APP), disrupts axonal transport. Curr Biol. 1999;9: 489–492. doi:10.1016/S0960-9822(99)80215-2

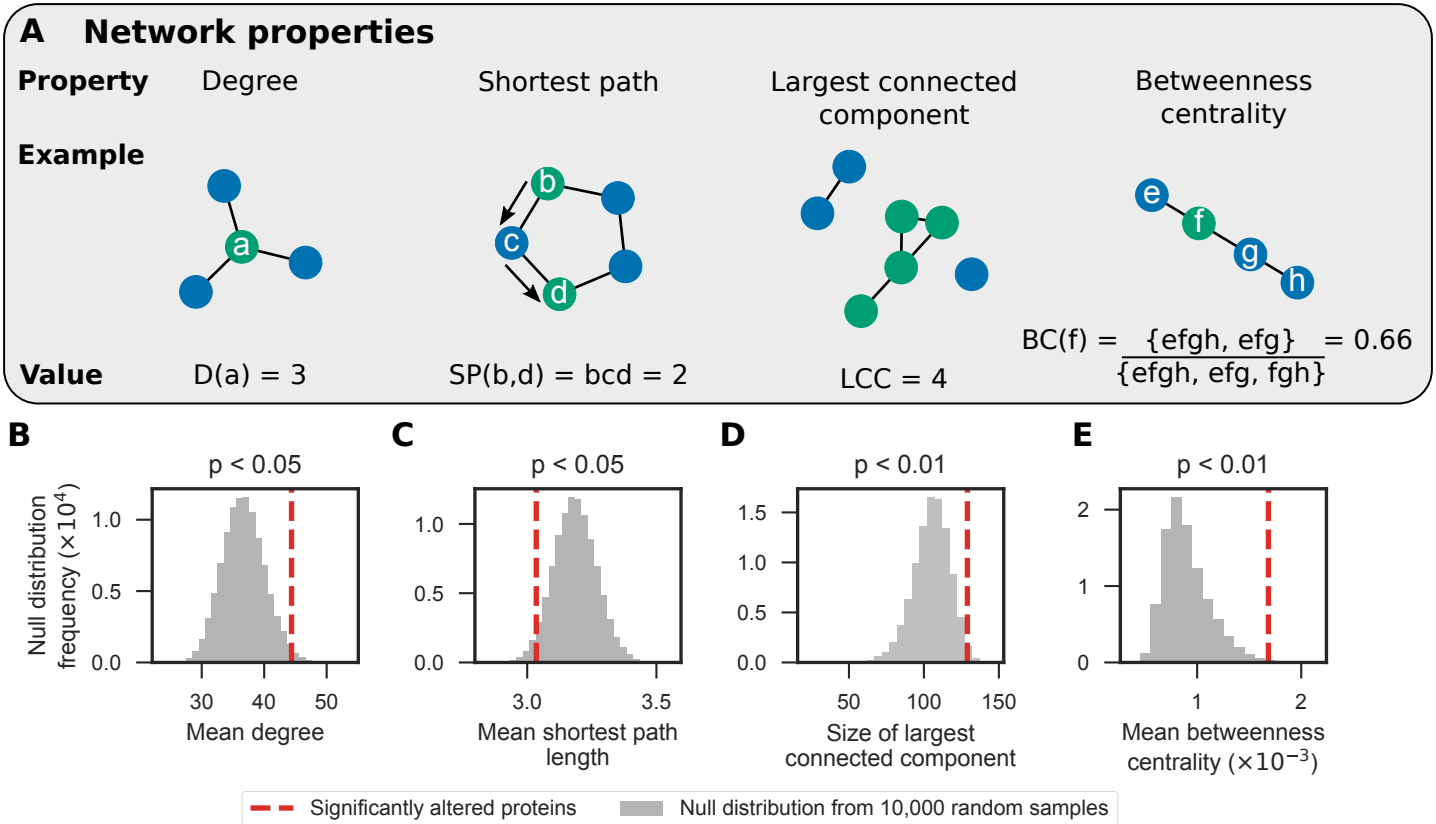
89. Ewald CY, Li C. Understanding the molecular basis of Alzheimer's disease using a Caenorhabditis elegans model system. Brain Struct Funct. 2010;214: 263–283. doi:10.1007/s00429-009-0235-3

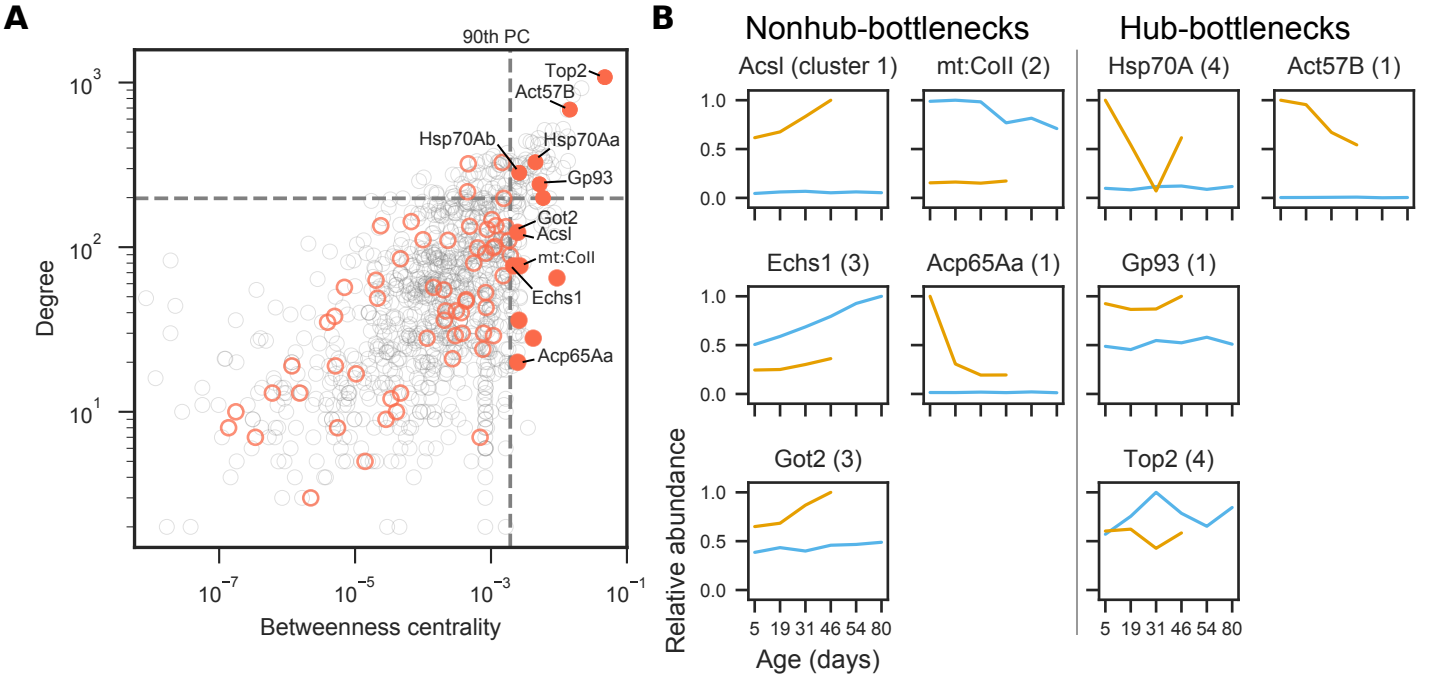
90. Baranello RJ, Bharani KL, Padmaraju V, Chopra N, Lahiri DK, Greig NH, et al. Amyloid-beta protein clearance and degradation (ABCD) pathways and their role in Alzheimer's disease. Curr Alzheimer Res. 2015;12: 32–46. doi:10.2174/1567205012666141218140953

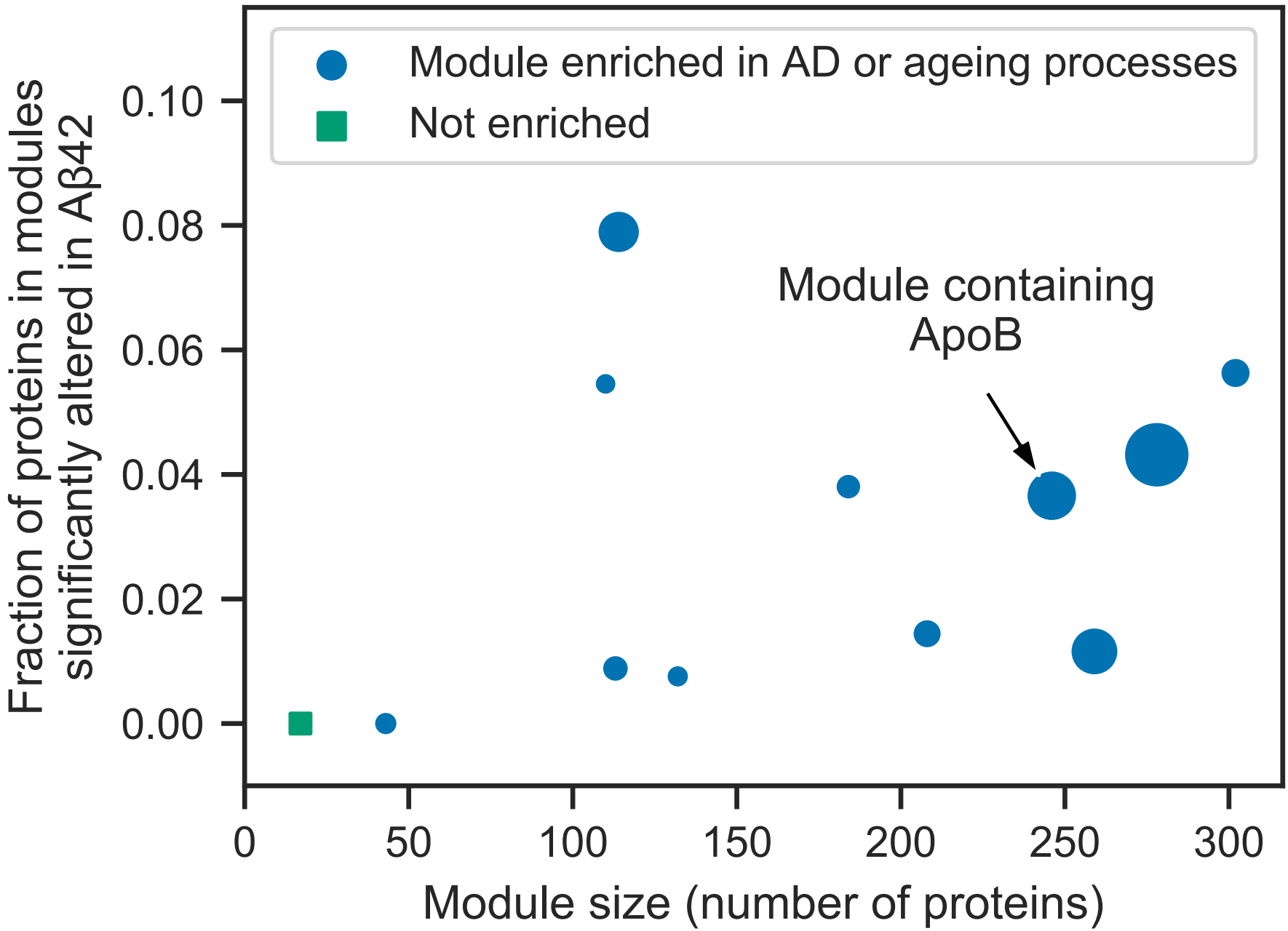
91. Zhang X, Le W. Pathological role of hypoxia in Alzheimer's disease. Exp Neurol. 2010;223: 299–303. doi:10.1016/j.expneurol.2009.07.033

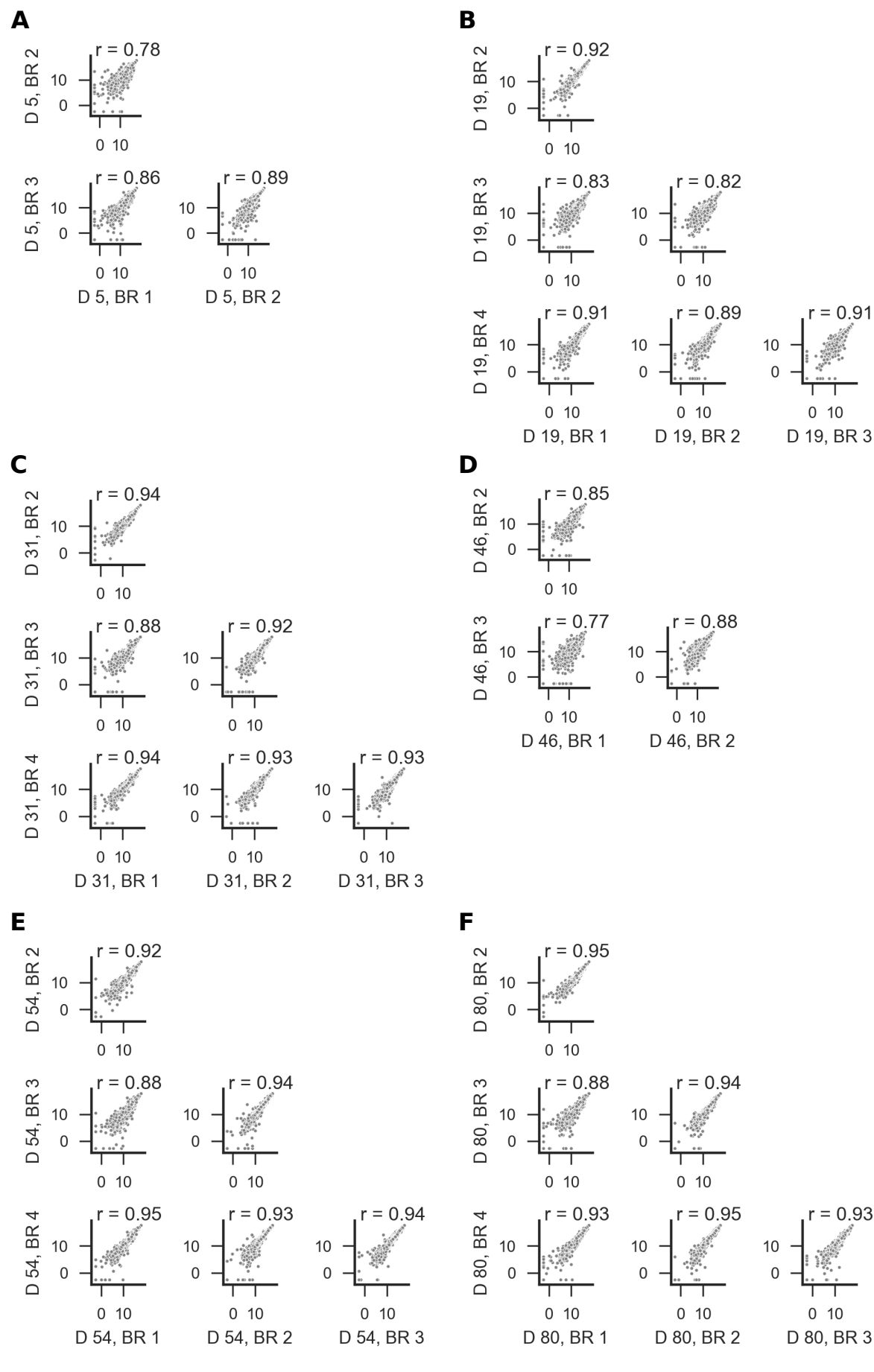

A
TgAD fly line
Mifepristone inducer
elav
pan-neuronal
promoter
GAL4
UAS
Aβ42
B
Survival rate (proportion of total)
Age (days)
Healthy
Aβ42
5
19
31
46
54
80
C
Healthy
Aβ42
Induce Aβ42 flies at 2 days
5···46 54 80
5 19 31 46
Age (days)
Collect samples, dissect brains
Trypsin digest, nanoscale UPLC
Intensity
m/z
Label-free quantitative IM-DIA-MS
Abundance
Data analysis
D
Proteins identified by MS
607
1854
632
Healthy
Aβ42
E
5 (age in days)
PC2 (8.6%)
PC1 (70.6% variance explained)
F
Abundance normalised to healthy flies at 5 days
Higher
Lower
Age (days) Healthy
Age (days) Aβ42
Proteins

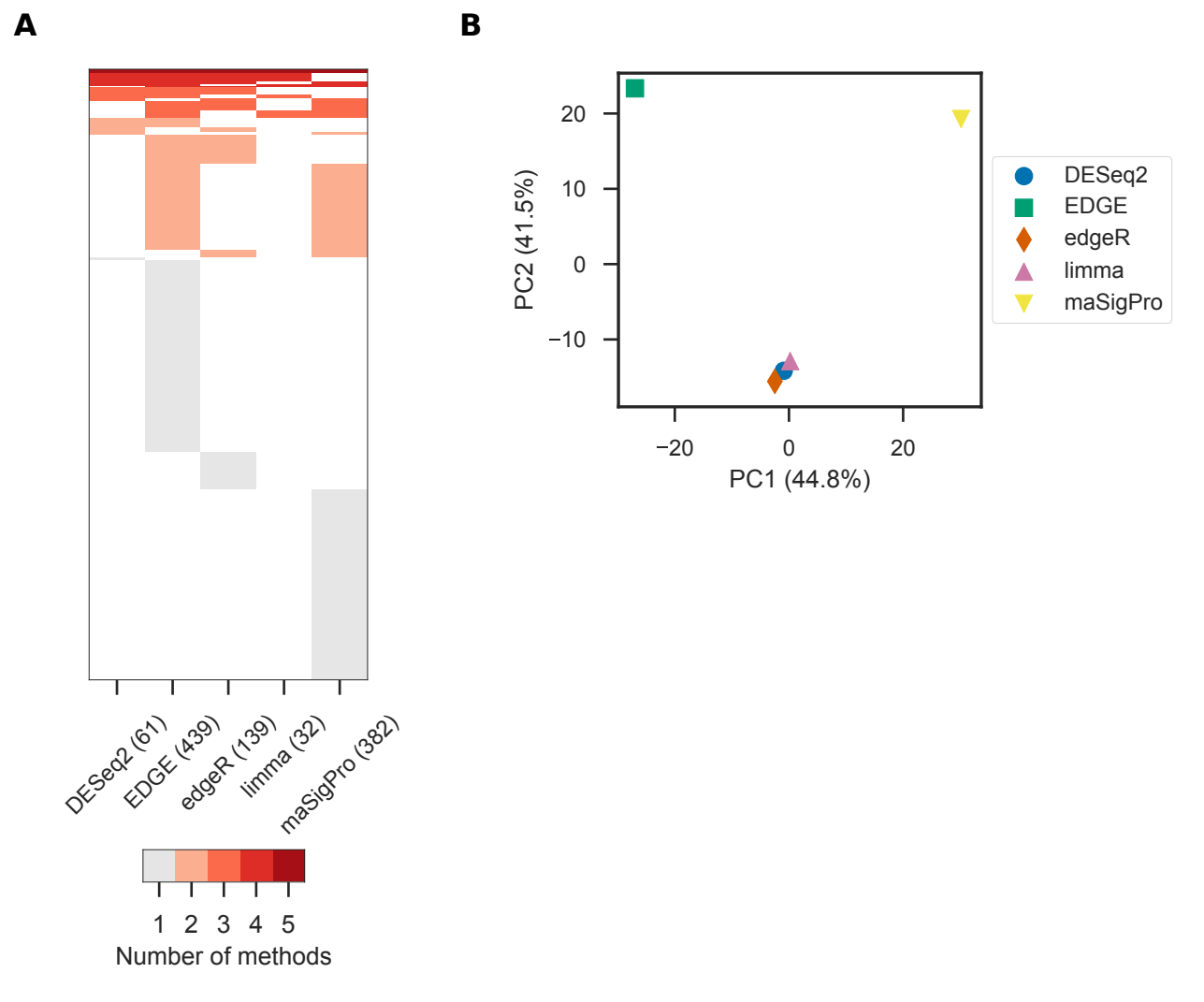

A
600
400
200
0
Proteins
512
228 significantly altered proteins in Aβ42
169
16
38
5
1
2
3
4
5
Number of methods that detect a protein is significantly altered in Aβ42
B
Significantly altered proteins
31
30
198
Ageing
Aβ42
C
Cluster 1 (n = 75)
Cluster 2 (71)
Cluster 3 (39)
Cluster 4 (43)
Relative abundance
2
1
0
−1
−2
5
19
31
46
Age (days)
Healthy
Aβ42

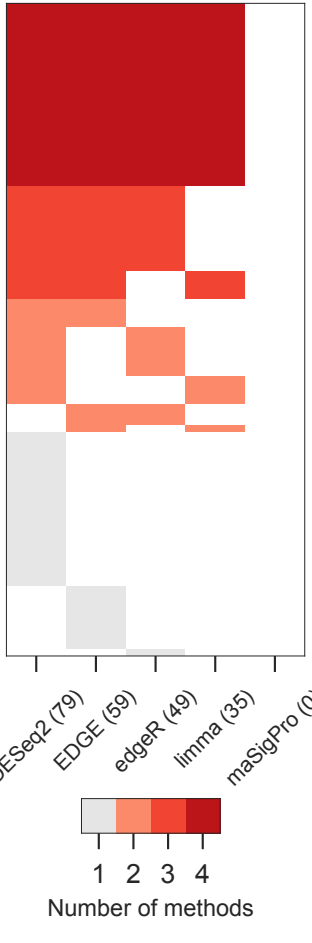

A Network properties
Property
Degree
Shortest path
Largest connected component
Betweenness centrality
Example
a
b
c
d
e
f
g
h
Value
D(a) = 3
SP(b,d) = bcd = 2
LCC = 4
BC(f) = {efgh, efg} / {efgh, efg, fgh} = 0.66
B
p < 0.05
Null distribution frequency (×10^4)
1.0
0.5
0.0
30
40
50
Mean degree
C
p < 0.05
1.0
0.5
0.0
3.0
3.5
Mean shortest path length
D
p < 0.01
1.5
1.0
0.5
0.0
50
100
150
Size of largest connected component
E
p < 0.01
2
1
0
1
2
Mean betweenness centrality (×10^−3)
Significantly altered proteins
Null distribution from 10,000 random samples

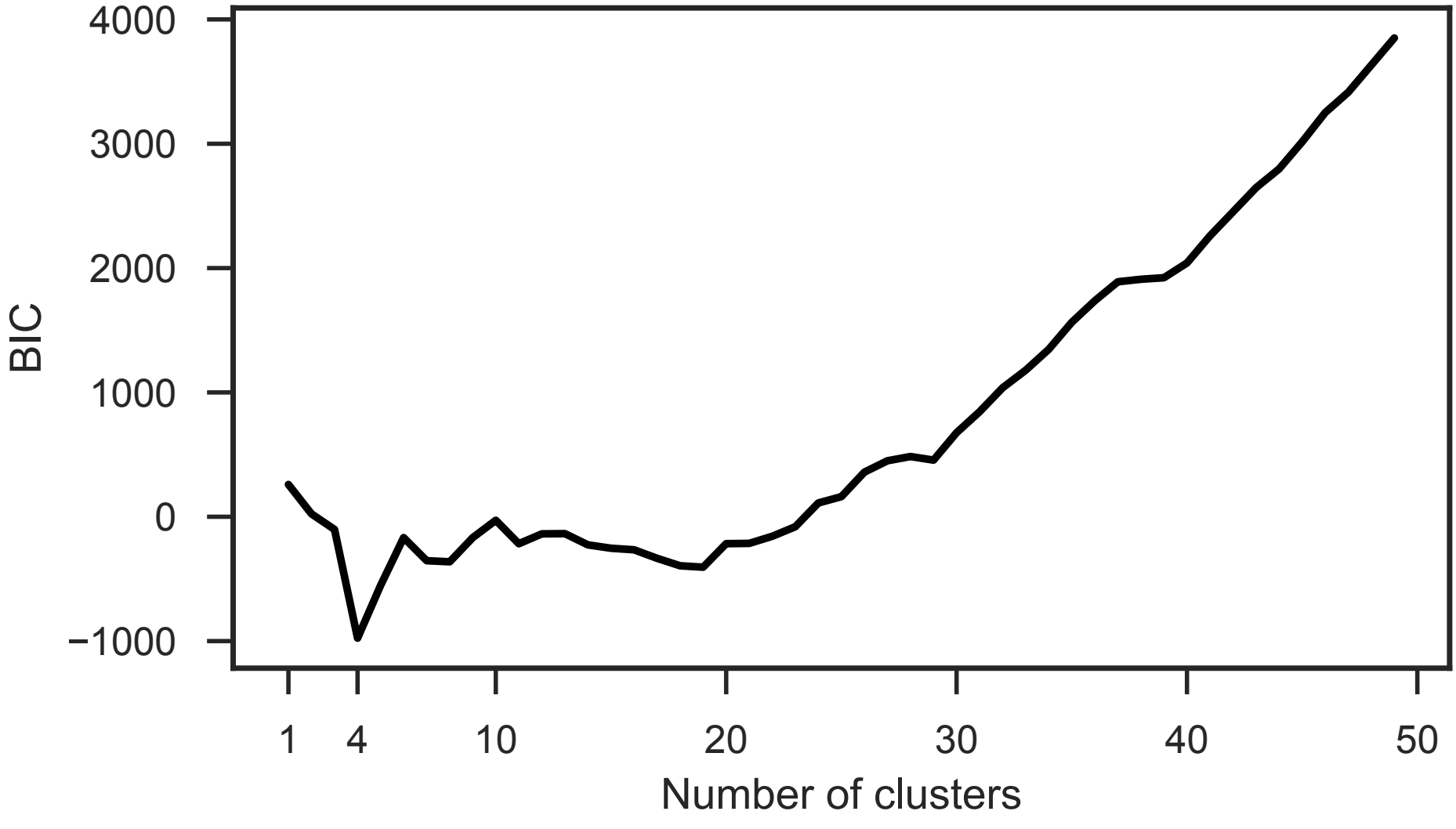

A
90th PC
Top2
Act57B
Hsp70Ab
Hsp70Aa
Gp93
Got2
Acsl
mt:CoII
Echs1
Acp65Aa
Degree
Betweenness centrality
B
Nonhub-bottlenecks
Hub-bottlenecks
Acsl (cluster 1)
mt:CoII (2)
Hsp70A (4)
Act57B (1)
Echs1 (3)
Acp65Aa (1)
Gp93 (1)
Got2 (3)
Top2 (4)
Relative abundance
Age (days)


Module enriched in AD or ageing processes
Not enriched
Module containing ApoB
Fraction of proteins in modules significantly altered in Aβ42
0.10
0.08
0.06
0.04
0.02
0.00
0
50
100
150
200
250
300
Module size (number of proteins)


A
r = 0.78
r = 0.86
r = 0.89
D 5, BR 2
D 5, BR 3
D 5, BR 1
D 5, BR 2
B
r = 0.92
r = 0.83
r = 0.82
r = 0.91
r = 0.89
r = 0.91
D 19, BR 2
D 19, BR 3
D 19, BR 4
D 19, BR 1
D 19, BR 2
D 19, BR 3
C
r = 0.94
r = 0.88
r = 0.92
r = 0.94
r = 0.93
r = 0.93
D 31, BR 2
D 31, BR 3
D 31, BR 4
D 31, BR 1
D 31, BR 2
D 31, BR 3
D
r = 0.85
r = 0.77
r = 0.88
D 46, BR 2
D 46, BR 3
D 46, BR 1
D 46, BR 2
E
r = 0.92
r = 0.88
r = 0.94
r = 0.95
r = 0.93
r = 0.94
D 54, BR 2
D 54, BR 3
D 54, BR 4
D 54, BR 1
D 54, BR 2
D 54, BR 3
F
r = 0.95
r = 0.88
r = 0.94
r = 0.93
r = 0.95
r = 0.93
D 80, BR 2
D 80, BR 3
D 80, BR 4
D 80, BR 1
D 80, BR 2
D 80, BR 3


A
B
DESeq2 (61)
EDGE (439)
edgeR (139)
limma (32)
maSigPro (382)
1 2 3 4 5
Number of methods
PC2 (41.5%)
PC1 (44.8%)
20
10
0
−10
−20
0
20
DESeq2
EDGE
edgeR
limma
maSigPro


DESeq2 (79)
EDGE (59)
edgeR (49)
limma (35)
maSigPro (0)
1
2
3
4
Number of methods


4000
3000
2000
1000
0
−1000
BIC
1
4
10
20
30
40
50
Number of clusters