



OPEN

## Exploring the conservation of Alzheimer-related pathways between *H. sapiens* and *C. elegans*: a network alignment approach

Avgi E. Apostolakou<sup>1,3</sup>, Xhuliana K. Sula<sup>1,3</sup>, Katerina C. Nastou<sup>1,2</sup>, Georgia I. Nasi<sup>1</sup> & Vassiliki A. Iconomidou<sup>1</sup>✉

Alzheimer disease (AD) is a neurodegenerative disorder with an –as of yet– unclear etiology and pathogenesis. Research to unveil disease processes underlying AD often relies on the use of neurodegenerative disease model organisms, such as *Caenorhabditis elegans*. This study sought to identify biological pathways implicated in AD that are conserved in *Homo sapiens* and *C. elegans*. Protein–protein interaction networks were assembled for amyloid precursor protein (APP) and Tau in *H. sapiens*—two proteins whose aggregation is a hallmark in AD—and their orthologs APL-1 and PTL-1 for *C. elegans*. Global network alignment was used to compare these networks and determine similar, likely conserved, network regions. This comparison revealed that two prominent pathways, the APP-processing and the Tau-phosphorylation pathways, are highly conserved in both organisms. While the majority of interactions between proteins in those pathways are known to be associated with AD in human, they remain unexamined in *C. elegans*, signifying the need for their further investigation. In this work, we have highlighted conserved interactions related to AD in humans and have identified specific proteins that can act as targets for experimental studies in *C. elegans*, aiming to uncover the underlying mechanisms of AD.

### Abbreviations

AD	Alzheimer disease
APP	Amyloid precursor protein
PPIs	Protein-protein interactions
GNA	Global network alignment

Alzheimer disease (AD) is a chronic, progressive, neurodegenerative disorder, characterized clinically by a gradual impact on mental and cognitive functions, affecting a person's ability to perform common daily activities<sup>1</sup>. It is the most common cause of dementia and is estimated that by 2050 the number of patients with this disease could exceed 100 million worldwide if no cure is discovered<sup>2</sup>. The pathological hallmarks of AD are amyloid plaques and neurofibrillary tangles; amyloid plaques are created by the extracellular deposition of fibrils consisting of abnormally folded A $\beta$  peptide—a cleavage product of amyloid precursor protein (APP)<sup>3</sup>—while neurofibrillary tangles consist mainly of intracellular hyperphosphorylated twisted filaments of the microtubule-associated Tau protein (Tau hereafter)<sup>1,4</sup>. Moreover, both proteins have the ability to form fibrils extracellularly in vivo, and are thus characterized as amyloid fibril proteins by the International Society of Amyloidosis<sup>5</sup>. Therefore, it comes to no surprise that the scientific community has placed much emphasis on understanding the role of these proteins in AD onset.

The study of age-related diseases, such as AD, in humans is practically impossible, due to both the prohibitive nature of an observational study in a human's lifespan and the serious ethical issues raised on performing experimental studies in humans<sup>6</sup>. Consequently, researchers use model organisms in order to gain insight into the molecular mechanisms underlying these diseases. Animals with close evolutionary relationships to humans,

<sup>1</sup>Section of Cell Biology and Biophysics, Department of Biology, National and Kapodistrian University of Athens, 15701 Panepistimiopolis, Athens, Greece. <sup>2</sup>Present address: Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark. <sup>3</sup>These authors contributed equally: Avgi E. Apostolakou and Xhuliana K. Sula. ✉email: veconom@biol.uoa.gr

like mouse, pig and non-human primates, can offer great insight for the study of neurodegenerative diseases, and have extensively used as models to study the disease<sup>7–9</sup>. However, in many instances, in vivo experimentation in mammalian model organisms is complex, time- and resource-consuming, therefore other more distant organisms can be used to alleviate these issues. Prominent among them is the nematode *Caenorhabditis elegans*, which is often used as it provides an attractive model for neurodegenerative diseases, due to its distinctive features (see review by Alexander et al.<sup>10</sup>). Most importantly, many human genes (ca. 60–80%) have an ortholog—genes in different species that evolved from a common ancestor gene and possibly retain the same function<sup>11</sup>—in the genome of *C. elegans*<sup>12</sup>. Even though, *C. elegans* has a gene encoding for a protein similar to human Tau (PTL-1)<sup>13</sup> the APP-related gene (APL-1) does not contain the A $\beta$  peptide sequence. Moreover, *C. elegans* does not possess an ortholog for  $\beta$ -secretase, and thus has no  $\beta$ -secretase activity<sup>10</sup>. For this reason, many transgenic *C. elegans* strains that express human A $\beta$  and Tau sequences in specific cell types have been created<sup>10,13</sup> and are used as models for the study of AD and other neurodegenerative diseases.

Protein–protein interactions (PPIs) govern most biological processes and investigating them in the context of human diseases is essential to fully comprehend their underlying mechanisms. Networks of PPIs are powerful tools used to conceptualize models of molecular interactions in various biological systems<sup>14,15</sup>. One of the main benefits of PPI networks is that they allow the conversion of a wealth of raw data into reasonably structured visual representations. Nowadays, due to high-throughput techniques, the growth of available PPI data is exponential. Utilization of this data has allowed the study of PPI networks for entire organisms (e.g. *C. elegans*<sup>16</sup>) or for specific diseases (e.g. AD<sup>17</sup>). Moreover, the availability of organism-wide PPI networks has made cross-species network comparisons possible. A popular method for comparing networks is network alignment, which, analogously to genome alignment, aims at mapping the nodes of two or more networks, and thereby, determining topologically and functionally similar regions<sup>18,19</sup>. Conserved network regions can be used to transfer biologically relevant information between humans and model organisms<sup>18,20</sup>. Previous studies<sup>21,22</sup> have used PPI networks to explore AD mechanisms using the model organism *C. elegans*. To our awareness this is the first effort to employ network alignment for the transfer of knowledge between *C. elegans* and human for the study of this disease.

The goal of this work was the in silico construction and comparison of Alzheimer-related protein–protein interaction networks in humans and *C. elegans*. Our main aim was the discovery of common biological pathways that are conserved in both organisms and are potentially implicated in AD. Study of such pathways will guide experimental studies on the model organism *C. elegans*, and help in the elucidation of mechanisms involved in the pathogenesis of AD.

## Methods

**Protein and interaction datasets.** Three network datasets were created for this study, each consisting of a *H. sapiens* PPI network and a *C. elegans* PPI network. Every dataset differs in the method of data collection and the purpose it serves. Cross-species network comparison for each pair of networks in the three datasets was achieved by using one or more network alignment algorithms.

**APP and Tau network from the amyloid interactome.** The first human network we used in this study was isolated from the Amyloid Interactome<sup>23</sup>. The Amyloid Interactome is a PPI network of amyloidogenic proteins and their experimentally verified interaction partners. APP, Tau, their interaction partners, as well as any interactions between them, were extracted from this network. The interactions from the Amyloid Interactome are experimentally validated and extracted from the IntAct database<sup>24</sup>. For the *C. elegans* protein dataset, predicted orthologs of the aforementioned human proteins were recovered using OrthoList<sup>25</sup>, a compendium of predicted *C. elegans*–human orthologs. The predicted orthologs were further verified through WormBase<sup>26</sup>, a curated database about the genetics, genomics and biology of *C. elegans*. An attempt was initially made to collect interactions between the predicted *C. elegans* orthologs from IntAct, in a manner similar to the Amyloid Interactome, however, only a very small number of such interactions were available. Instead, the STRING database<sup>27</sup>—a database of known and predicted PPIs—was selected, since it integrates data from various sources and covers a vast number of organisms. STRING primarily uses 4 types of evidence—co-expression, experiments, databases and text mining—that it also propagates according to homology; all evidence contribute to a confidence score provided for each PPI. A cutoff interaction score of 0.7 was selected representing interaction partners of high confidence. This dataset was used to evaluate the network alignment algorithms and determine the best parameters that would be used for the subsequent network alignments.

**APL-1 and PTL-1 network from STRING.** Afterwards, the interaction partners of APL-1 and PTL-1—the *C. elegans* orthologs of APP and Tau—were extracted from STRING. A cutoff interaction score of 0.7 was selected representing interaction partners and interactions of high confidence. The predicted human orthologs of those interaction partners were found and constitute the human protein dataset for this pair of networks. For the human network an interaction score of 0.9, representative of the highest confidence level in STRING, was utilized<sup>28</sup>. Furthermore, interactions without experimental validation were filtered out at this stage, in order to increase the reliability of collected data.

**Top 100 interaction partners for APP and Tau & APL-1 and PTL-1 from STRING.** The final dataset is the least biased and therefore was the one primarily used to transfer biological information between the two organisms. The 100 interaction partners of APL-1 and PTL-1 with the best confidence score were collected from STRING, to create the *C. elegans* network. A cutoff interaction score of 0.5 was used for interactions between these proteins, to collect at least 100 additional proteins. Using the same approach, the top 100 interaction partners of APP and Tau were collected from STRING to create the human network. In this case a higher interaction cutoff (0.9),

allowed the retrieval of the top 100 interaction partners and the need to reduce it did not emerge. Once again, interactions without experimental validation were filtered out to increase data reliability.

**Network visualization, alignment and comparison.** All collected data was visualized as PPI networks using Cytoscape 3.7.2<sup>29</sup>. Cytoscape is a freely available platform for biological network visualization and analysis that provides a vast array of applications for specialized functions. Two such applications used in the current study are the stringApp<sup>30</sup>, which allowed the direct import of networks from the STRING database into Cytoscape and Omics Visualizer<sup>31</sup>, that was used for visualization purposes in this work. Additionally, Cytoscape.js<sup>32</sup> was used to create interactive networks, available via a web interface at ([http://thalis.biol.uoa.gr/celegans\\_human\\_AD/](http://thalis.biol.uoa.gr/celegans_human_AD/)), where detailed information about the proteins in the aforementioned networks can be retrieved. Detailed description of the functionalities offered by the web application is available at Supplementary File 1.

Comparison between human and *C. elegans* networks was done using Global Network Alignment (GNA), which aims to locate similarities across entire networks. An array of tools is available to perform GNA, the majority of which are based on a cost function for node similarity and attempt to locate the alignment with a maximum node similarity score<sup>33</sup>. Three algorithms for GNA, namely MAGNA++<sup>34</sup>, CytoGEDEVO<sup>35</sup> and NETAL<sup>36</sup>, were selected and tested to determine the best performing. All three algorithms produce one-to-one node alignments, where each node in one network is uniquely mapped to a single node in the other network.

MAGNA++ uses a genetic algorithm to simulate a population of alignments that evolves in time, until the alignment can no longer be improved. A major advantage of MAGNA++ lies in its effort to maximize edge conservation during the alignment process. CytoGEDEVO is a Cytoscape application that employs the GEDEVO algorithm for network alignment. GEDEVO is based on the graph edit distance between two networks, and attempts to transform one network into the other by applying the minimum number of edge additions and removals. NETAL uses a greedy search algorithm to progressively align node pairs with the best score. An advantage of this tool is that topological information is being renewed during the search for the best alignment.

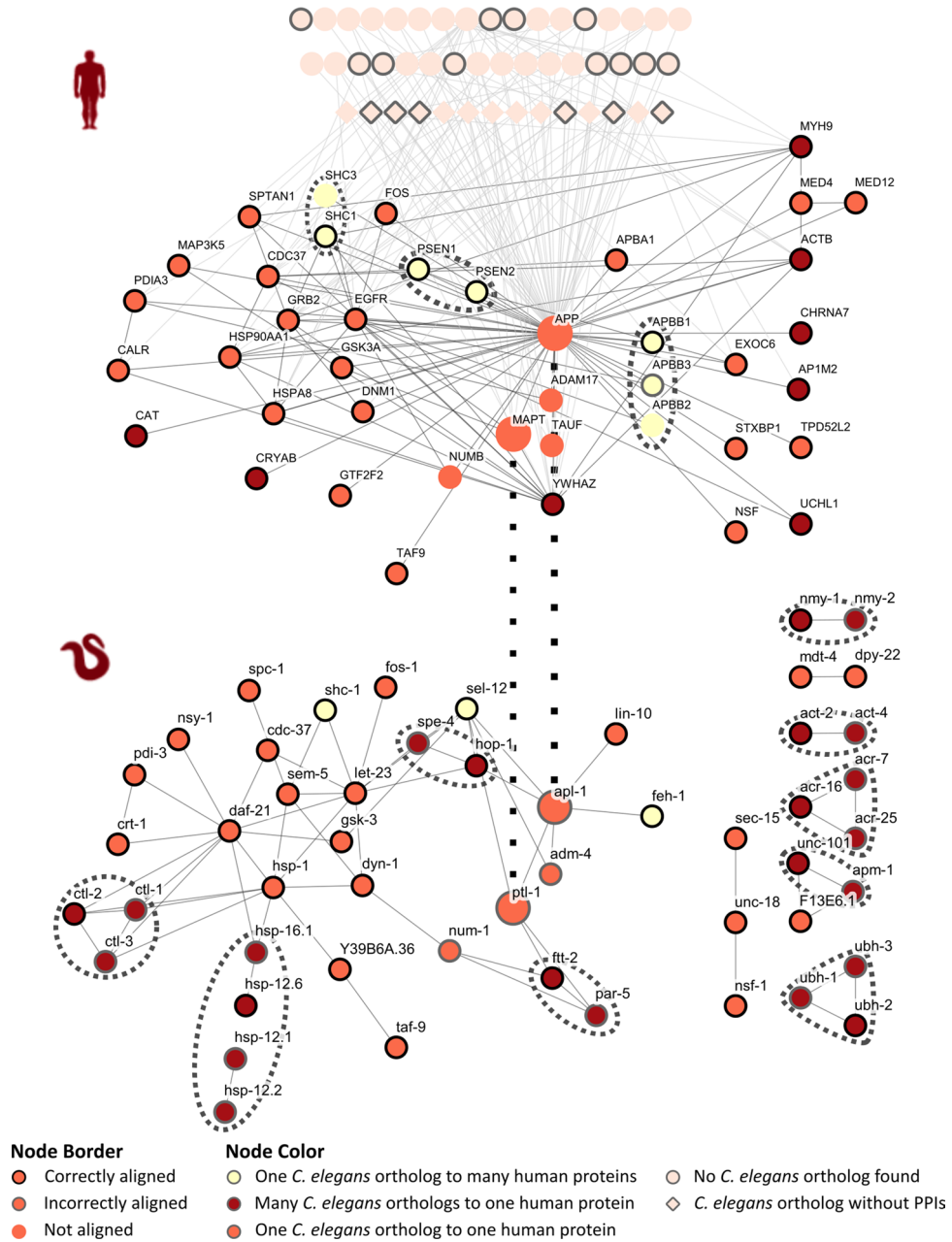
To optimize each comparison, GNA traditionally uses biological information on top of topological parameters<sup>37</sup>. MAGNA++ and CytoGEDEVO, but not NETAL, accept node similarity inputs allowing for the use of biological information to achieve a better network alignment. Protein sequence similarity was calculated and used as a biological information input, where applicable (Supplementary Table 1). To this end, pairwise sequence alignment was done for every human protein against all *C. elegans* proteins in the dataset using the Needleman-Wunsch algorithm<sup>38</sup>, which aims at finding the best alignment across the entire protein length. The protein sequence similarity was determined as the percentage of aligned residues that matched, i.e. have similar physicochemical properties. Evaluation of the algorithms was done based on their ability to correctly align the nodes of the networks according to their predicted orthologs. Details regarding the evaluation of the algorithms and a flowchart with the step-by-step procedure followed for the creation and alignment of the “Top 100 interaction partners” networks are available in Supplementary File 2.

## Results and discussion

In the course of evolution ortholog proteins occasionally maintain the same function across species, as is often reflected by their sequence similarity<sup>39</sup>. Likewise, PPIs between orthologs can be conserved across organisms and these interactions are referred to as interologs. The aim of this work was to explore both of these concepts in the context of comparing homologous PPI networks in *H. sapiens* and *C. elegans* to uncover conserved pathways between these organisms that could be exploited for the study of AD. Three pairs of networks were assembled and compared via GNA (Supplementary Table 5).

**APP and Tau network from the amyloid interactome.** The Amyloid Interactome<sup>23</sup> reported 86 interaction partners for APP and Tau. This resulted in the human network extracted from it, to consist of 88 proteins and 216 PPIs (Fig. 1). Of the 88 human proteins, 61 were matched to 73 (68 unique) predicted protein orthologs in *C. elegans*; ortholog assignment criteria and mapping can be found in Supplementary File 2 and Supplementary Table 1. A query of IntAct using these 68 proteins returned only 6 interactions amongst the *C. elegans* proteins. Failure to construct a network with data from IntAct, led to the use of STRING instead, and the final *C. elegans* network used in this study consists of 51 proteins and 69 PPIs (Fig. 1).

The aforementioned pair of networks was used as a “gold-standard” for evaluating the network alignment algorithms. The performance of the GNA algorithms was assessed based on their ability to both align nodes according to the mapping of predicted orthologs and also identify interologs. Manual alignment of the networks according to predicted ortholog mapping revealed the existence of 22 interologs. An endeavor was made to align the networks with the algorithms relying exclusively on network topological information. This resulted in complete failure of the algorithms to align the two networks, with both MAGNA++ and NETAL unable to correctly align any nodes and with CytoGEDEVO making only 3 correct node alignments (Table 1). In an effort to optimize the alignments, biological information was introduced to MAGNA++ and CytoGEDEVO; NETAL on the contrary does not allow the inclusion of such information. The combination of topological and biological information was sufficient to allow the two algorithms to successfully align the two networks, as shown in Table 1. MAGNA++ outperformed CytoGEDEVO, both in correctly aligning predicted ortholog pairs and in identifying interologs (Table 1). More detailed information on the evaluation of the GNA algorithms is available in Supplementary File 2. Ultimately, MAGNA++ correctly aligned 33 out of 51 node pairs according to their predicted ortholog mapping as shown in Fig. 1. Due to some proteins mapping to multiple predicted orthologs, and since GNA algorithms work by aligning unique pairs of proteins, it was inherently impossible for MAGNA++ to



**Figure 1.** Alignment of human APP and Tau network from Amyloid Interactome and homologous *C. elegans* network by MAGNA++ with combination of biological and topological information, containing both correct and incorrect alignments. The human network (top) and the *C. elegans* network (bottom) are displayed with the layout of nodes corresponding to the predicted ortholog mapping. Human proteins with no predicted orthologs in the *C. elegans* network are isolated on the upper side of the human network. Two dotted black lines show the positioning of ortholog pairs APP & APL-1 and Tau & PTL-1 that were not correctly aligned; APP and Tau were not mapped to any *C. elegans* protein, while APL-1 and PTL-1 were aligned with RNF32 and PRAM1, respectively. All node alignments are available in Supplementary Table 3. Groups of nodes surrounded by dotted lines represent proteins mapped to the same predicted ortholog. These protein nodes have been colored accordingly, to make it easier to detect. Detailed information about the proteins in these networks can be retrieved via the web application ([http://thalis.biol.uoa.gr/celegans\\_human\\_AD/](http://thalis.biol.uoa.gr/celegans_human_AD/)).

achieve a “perfect” alignment. Based on these results, all subsequent network alignments were performed using MAGNA++ with a combination of topological and biological information.

**APL-1 and PTL-1 network from STRING.** Next, the *C. elegans* network of APL-1, PTL-1 and their interaction partners was extracted from STRING, using the stringApp<sup>30</sup>, and consisted of 61 proteins and 136 PPIs with experimental evidence (Fig. 2a). In a manner similar to above, the *C. elegans* proteins were mapped to

	Alignment based on topological information			Alignment based on topological and biological information	
	MAGNA++	CytoGEDEVO	NETAL	MAGNA++	CytoGEDEVO
Correct node alignments	0	3	0	<b>33</b>	24
Incorrect node alignments	51	48	51	<b>18</b>	27
Aligned interactions	56	46	40	31	<b>39</b>
Interologs identified	0	0	0	<b>15</b>	12

**Table 1.** GNA algorithm evaluation results. Three GNA algorithms were evaluated based on their ability to correctly align predicted ortholog proteins (51) and identify interologs (22). Bold is used to indicate the best value for each metric. All algorithms failed to align the networks using topological information exclusively and combination with biological information proved necessary for the correct alignment of protein orthologs.

their predicted human orthologs. PPIs were collected from STRING for these proteins, resulting in the human network comprising of 127 proteins and 382 PPIs with experimental evidence (Fig. 2a). Alignment of the aforementioned networks, with the previously established procedure, returned 61 node aligned pairs and a unified network of 55 pairs of *H. sapiens* and *C. elegans* proteins and 92 interologs (Fig. 2a). MAGNA++ succeeded in correctly aligning the majority of predicted orthologs (48 out of 55 pairs), thus providing further validation of its ability to accurately align pairs of networks with orthologous proteins.

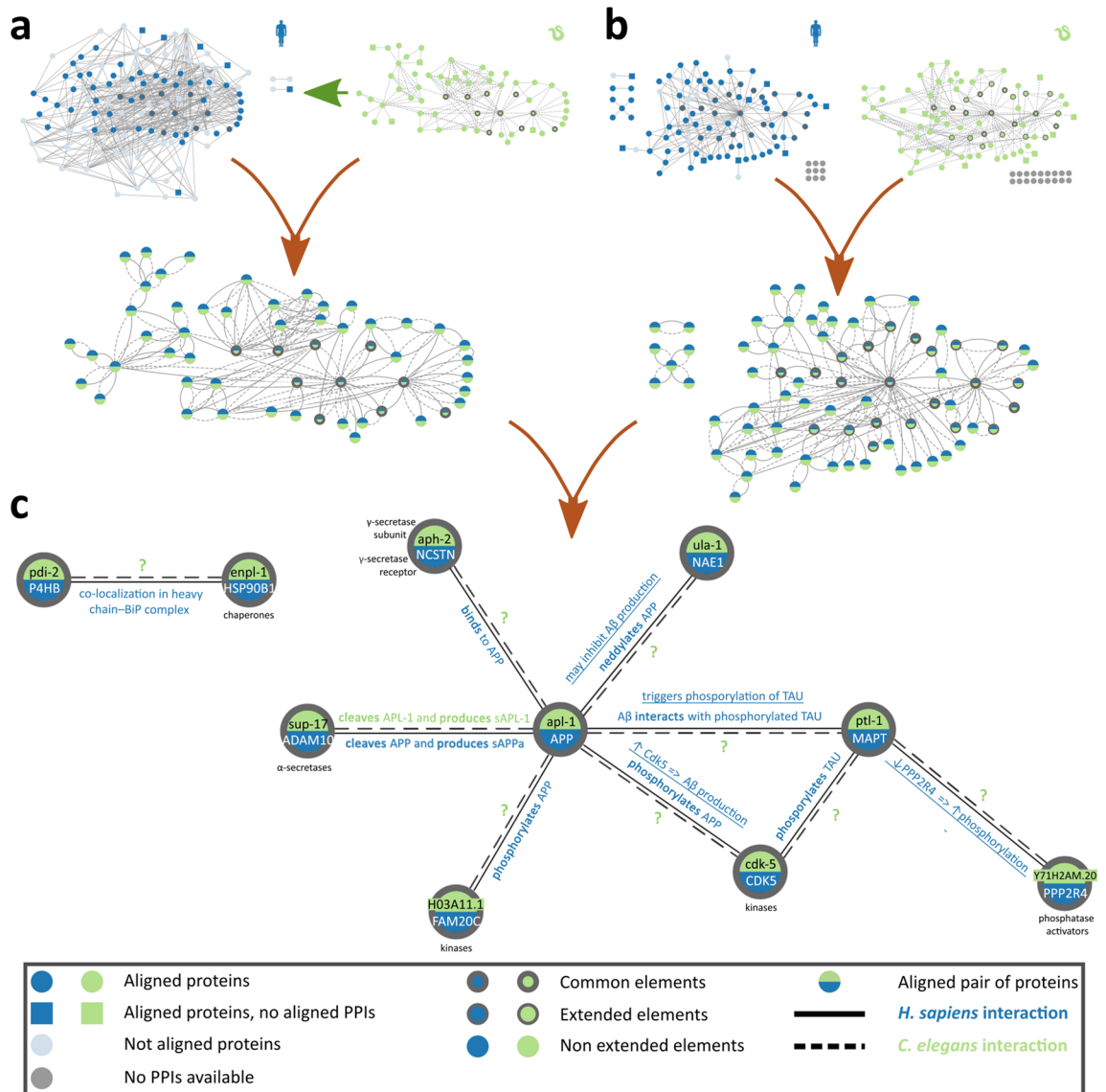
**Top 100 interaction partners for APP and Tau & APL-1 and PTL-1 from STRING.** The two aforementioned datasets relied on mapping predicted ortholog proteins across the two organisms, a procedure that inevitably introduced bias. Conversely, the “Top 100 interaction partners” dataset is the least biased dataset, and was therefore the one primarily used to transfer biological information between the two organisms. The human network had 102 proteins and 164 PPIs, while the *C. elegans* network had 102 proteins and 202 PPIs (Fig. 2b). Because PPIs without experimental validation were excluded, a number of proteins in each network are left with no PPIs. Specifically, interaction data was available for 91 human proteins and 82 *C. elegans* proteins. Alignment of these networks resulted in 82 node aligned pairs and a unified network of 74 aligned protein pairs and 152 conserved interactions (Fig. 2b).

During initial investigation of these networks, the network proteins of both organisms were functionally annotated by searching the literature for information; emphasis was placed on all AD related processes. The vast majority of human proteins in the dataset were found to be AD-associated, thus supporting the relevance of the network in regard to the disease under study. Every association was supported by at least one publication as shown in Supplementary Table 4. Additionally, the evidence provided for the interactions were collected from STRING and the evidence sourced from *C. elegans*, i.e. not transferred from another organism, were extracted (Supplementary Table 2 and Supplementary File 3).

Next, focus was applied to the identification of interactions conserved in both human and *C. elegans*. Common elements between this dataset’s unified network and the unified network of the “APL-1 and PTL-1 network from STRING” dataset were extracted. In total 10 pairs of human and *C. elegans* proteins were commonly aligned in both approaches, forming the *Common network* made up of 10 nodes and 9 edges (Fig. 2c). Aside from being commonly aligned, these protein pairs are also predicted orthologs, thereby confirming the conserved interactions between them as interologs. Most of these conserved interactions are linked to post-translational modifications on APP or Tau in human, while information for their role in *C. elegans* was available for only one interaction (Fig. 2c). Limited *C. elegans* evidence were available for 5 of these interactions, the interactions of APL-1 with proteins ULA-1, APH-2 and CDK-5, and the interactions of PTL-1 with CDK-5 and Y71H2AM.20 (Supplementary Table 2 and Supplementary File 3).

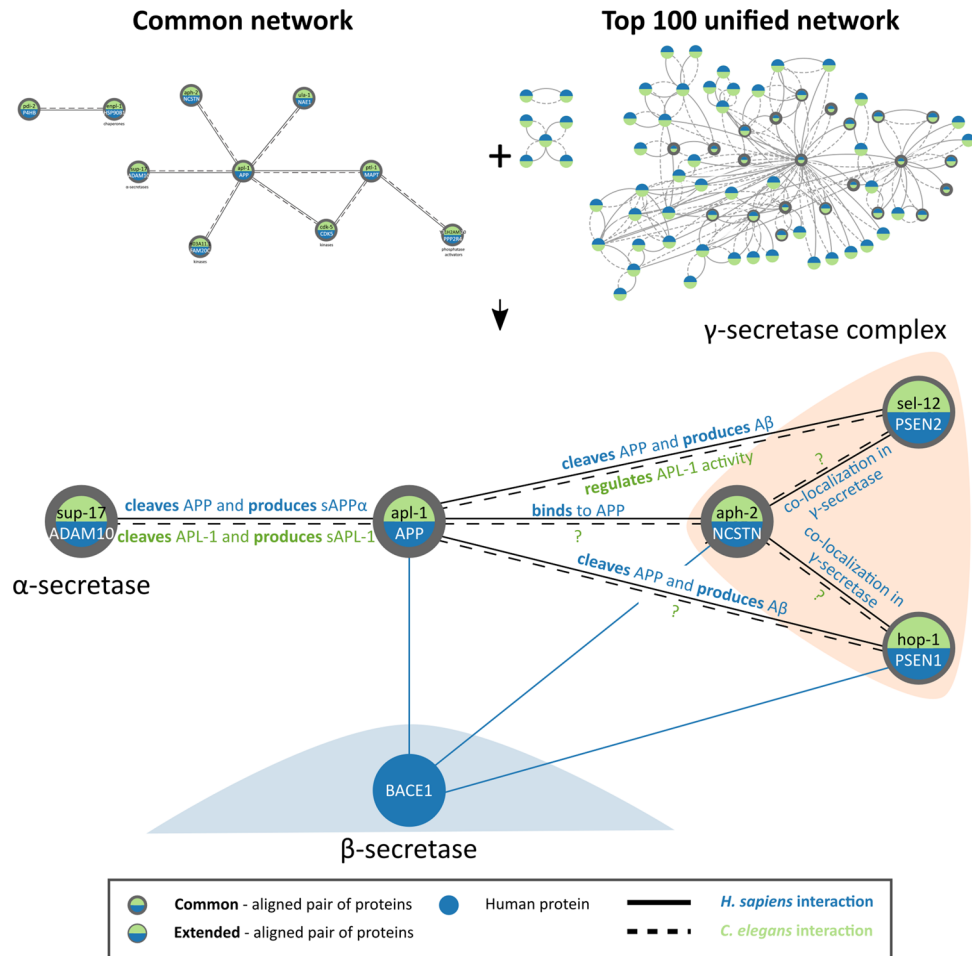
**Functional associations: the APP processing pathway.** Taking a closer look at the *Common network* in Fig. 2c revealed several interesting associations between the network’s proteins. Located in the *Common network* and interacting with APP are two proteins involved in APP processing, namely disintegrin and metalloproteinase domain-containing protein 10 (ADAM10) and nicastrin (NCSTN). ADAM10 is an  $\alpha$ -secretase that proteolytically cleaves APP as part of the non-amyloidogenic APP processing pathway, leading to the release of sAPP $\alpha$ , an important peptide for neuronal function, while also prohibiting the production of A $\beta$ <sup>40</sup>. ADAM10 participates in many important processes, including brain function and development, and is therefore considered a viable drug target; however caution is necessary due to the existence of multiple ADAM10 substrates<sup>41</sup>. Network alignment paired ADAM10 to its ortholog SUP-17, one of two *C. elegans* proteins with  $\alpha$ -secretase activity with the other protein being ADM-4. *C. elegans* APL-1, like APP, is cleaved by an  $\alpha$ -secretase resulting in the release of sAPL-1<sup>42</sup>. Following  $\alpha$ - or  $\beta$ -secretase cleavage, APP is further cleaved by  $\gamma$ -secretase, a complex with four main components: presenilins 1 (PSEN1) and 2 (PSEN2), anterior pharynx-defective-1 (APH1A or APH1B) and nicastrin (NCSTN)<sup>43</sup>. Nicastrin is thought to act as a receptor binding to APP and other  $\gamma$ -secretase substrates<sup>44</sup>. This protein was aligned to its ortholog APH-2, a protein that also participates in the  $\gamma$ -secretase complex of *C. elegans*, functioning similarly to the human nicastrin<sup>45</sup>.

Investigation of the network of “Top 100 interaction partners” revealed additional elements of the APP processing pathway in the unified network (Fig. 3). These are presenilin 1 (PSEN1) and presenilin 2 (PSEN2) that



**Figure 2.** Alignments for the “*APL-1* and *PTL-1* network from STRING” and the “Top 100 interaction partners” datasets and their common elements. **(a)** The *C. elegans* *APL-1* and *PTL-1* network from STRING (green) and the homologous human network (blue), with the unified network resulting from their alignment below. **(b)** The Top 100 interaction partners of *APL-1* and *PTL-1* for the *C. elegans* network (green) and of *APP* and *Tau* for the human network (blue). Below these networks is the unified network resulting from their alignment. **(c)** The common elements of the two unified networks (pairs of proteins and their PPIs aligned similarly in both datasets). Omics Visualizer<sup>31</sup> was used to apply colors on the network’s nodes, with the green half representing the *C. elegans* protein and the blue half representing the human protein. Each half is labeled with their corresponding gene symbol (gene symbol for *Tau* is *MAPT*). All human interactions were annotated with at least one supporting publication (Supplementary Table 4); edges with underlined labels indicate annotations directly associated with AD. On the contrary, only one *C. elegans* interaction could be annotated as AD related based on literature evidence.

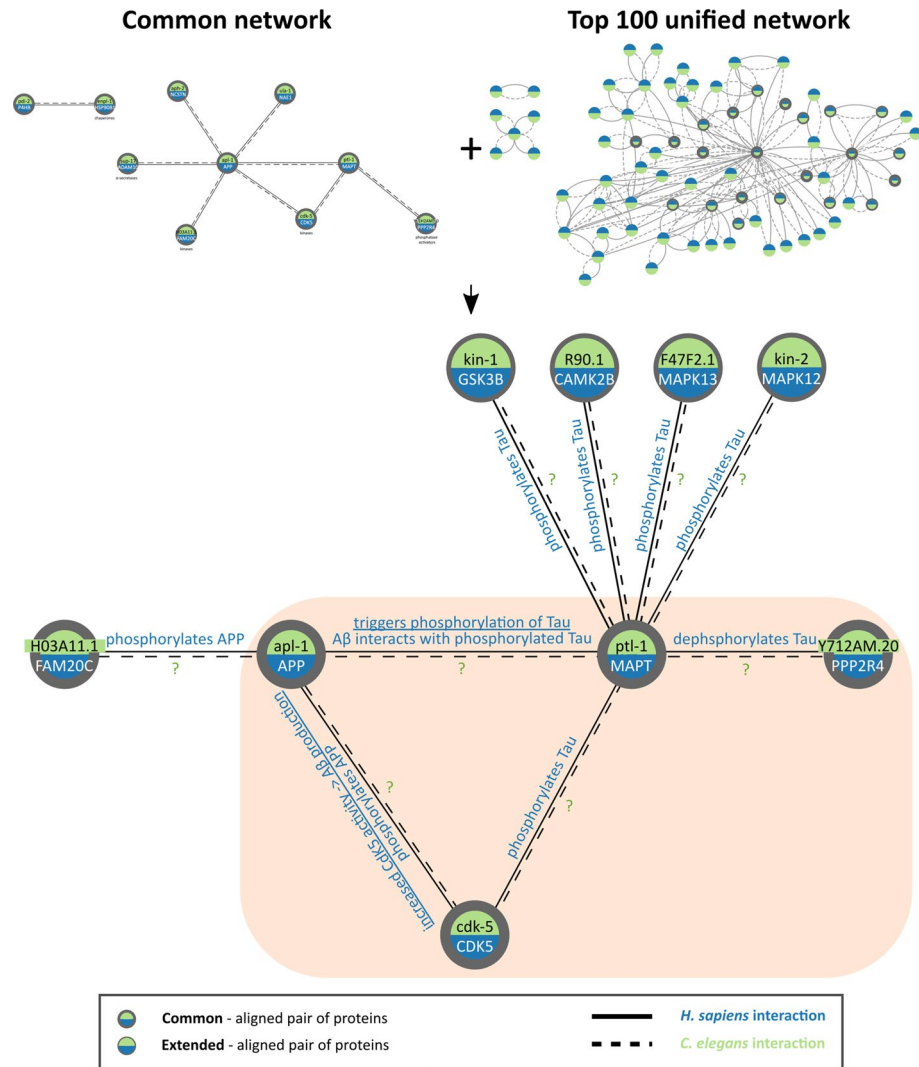
are responsible for the catalytic activity of the  $\gamma$ -secretase complex<sup>46</sup>. Cleavage of *APP* by  $\beta$ -secretase followed by  $\gamma$ -secretase, without any  $\alpha$ -secretase activity, results in the release of  $A\beta$ , according to the amyloidogenic pathway of *APP* processing<sup>47</sup>. Thus, catalysis by the  $\gamma$ -secretase complex is crucial both for the production of  $A\beta$  and the determination of its length and in turn its propensity to aggregation<sup>46</sup>. The catalytic subunits of  $\gamma$ -secretase, *PSEN1* and *PSEN2*, were aligned to *C. elegans* presenilin orthologs *HOP-1* and *SEL-12*, respectively. *HOP-1* and *SEL-12* are parts of the  $\gamma$ -secretase complex, just like their human counterparts, with evidence suggesting that *SEL-12*, directly or indirectly, regulates the activity of *APL-1*<sup>42</sup>. Overall, most of the *APP* processing machinery, and its corresponding orthologs, arose as important common elements during this work and were determined to indeed be conserved in the model organism. A notable exception is  $\beta$ -secretase (*BACE1*) that is necessary for the production of  $A\beta$  peptide and is absent in *C. elegans*, explaining why  $A\beta$ —and not *APP*—transgenes of this organism are used to model for AD<sup>48</sup>.



**Figure 3.** Conservation of the APP processing pathway in *H. sapiens* and *C. elegans*. Shown here are elements of the APP processing pathway found in the *Common Network*, extended with relevant proteins and interactions from the “*Top 100 interaction partners*” dataset. Elements associated with  $\alpha$ - and  $\gamma$ -secretase activity are conserved in both organisms, while  $\beta$ -secretase is missing from *C. elegans*. Even though all interactions that take place in this pathway have been experimentally investigated in human, information related to AD could be retrieved only for two *C. elegans* interactions.

For the interaction between APL-1 and SUP-17, as well as for the interactions between APH-2 and proteins SEL-12 and HOP-1, an abundance of evidence was available that was not transferred from homology (Supplementary File 3, Figure S14). On the contrary, most of the evidence for the interactions between APL-1 and proteins SEL-12, HOP-1 and APH-2, are transferred from homology (Supplementary Table 2 and Supplementary File 3, Figure S14). Given the central role of the  $\gamma$ -secretase complex in APP processing in humans, these three proteins are promising targets to experimentally study their interactions between APL-1 and the *C. elegans*  $\gamma$ -secretase complex or to study their role in transgenic *C. elegans*, models that express the human protein APP.

**Functional associations: phosphorylation pathways in AD.** Aside from APP processing, many proteins recovered from the network were involved in phosphorylation (Fig. 4). One of these is cyclin-dependent kinase 5 (CDK5), a prominent kinase in AD, whose aberrant action is associated with pathological processes, including the formation of amyloid plaques and neurofibrillary tangles in AD<sup>49</sup>. CDK5 phosphorylates APP<sup>50</sup>, promoting the production of A $\beta$  peptide<sup>51</sup>, and mediates Tau hyperphosphorylation, leading to its dissociation from microtubules and its aggregation<sup>52</sup>. Furthermore, A $\beta$  peptide was shown to interact with phosphorylated Tau protein in AD brains, potentially causing neuronal damage<sup>53</sup>, while another study reported an affinity of A $\beta$  peptide for non-phosphorylated Tau, triggering its phosphorylation and promoting its aggregation<sup>54</sup>. Modification of Tau by kinases can however be reversed through the action of phosphatases<sup>55</sup>. Primarily responsible for dephosphorylation of Tau is protein phosphatase 2A (PP2A), that is reportedly downregulated in AD, thereby contributing to the hyperphosphorylation of Tau<sup>56</sup>. Located in the *Common network* is the PP2A activator, PPP2R4, which interacts with Tau. In *C. elegans* the predicted orthologs of the aforementioned human proteins also interact, however no additional information is available concerning these interactions. Additionally, a number of kinases that phosphorylate Tau were found in the human network of the “*Top 100 interaction partners*”



**Figure 4.** AD associated phosphorylation-related proteins conserved in *H. sapiens* and *C. elegans*. Shown here are elements related to phosphorylation of APP and Tau found in the *Common Network*, extended with relevant proteins and interactions from the “*Top 100 interaction partners*” dataset. The interactions appear to be conserved in both organisms, but no information was available for the *C. elegans* interactions. The gene symbol for Tau is MAPT.

dataset and were aligned to *C. elegans* kinases. Glycogen synthase kinase-type 3 $\beta$  (GSK3B) is crucial for neurodevelopment and –like CDK5 –it is a main Tau kinase that contributes to the evident hyperphosphorylation of Tau in AD<sup>57</sup>. Another Tau kinase is Ca<sup>2+</sup>/calmodulin (CaM)-dependent protein kinase II (CaMKII)<sup>58</sup> that is important for synaptic function and is reportedly dysregulated in AD<sup>59</sup>. Lastly, p38 mitogen-activated protein kinases (MAPKs), MAPK12 and MAPK13, were found to interact with Tau; p38 MAPK mediates proinflammatory signaling, is upregulated in AD and is associated with both A $\beta$  and Tau pathology<sup>60</sup>.

Evidence in *C. elegans* was available for the interaction of APL-1 and CDK-5, as well as for the interactions between PTL-1 and proteins CDK-5, F47F2.1, KIN-1, KIN-2 and Y71H2AM.20 was mainly based on homology (Supplementary Table 2L and Supplementary File 3, Figure S15). In the case of APL-1 and CDK-5, considering the central role phosphorylation by CDK5 has in the processing of APP<sup>61</sup>, targeted studies should focus on exploring the phosphorylation potential of the *C. elegans* ortholog or the lack thereof. More specifically, any conserved phosphorylation sites found would help further elucidate the functional role of APP and would be vital to explain the differences in the aggregation mechanisms in the two organisms. Phosphorylation also has a crucial role in the aggregation of Tau and given the relatively recent shift of focus towards Tau-related drug discovery for AD<sup>62</sup>, Tau kinases and phosphatases are subjects of intense study<sup>63</sup>. Despite the importance of Tau, its *C. elegans* ortholog PTL-1 remains understudied with very few available data regarding interactions with other proteins not transferred from homology. This is evident in the results of this study, as only one PTL-1 phosphorylation-related interaction has evidence originated in experiments using *C. elegans*. CDK-5, F47F2.1, KIN-1 and KIN-2 are all kinases aligned to human kinases known to phosphorylate Tau, similarly with Y71H2AM.20 aligned to a human Tau phosphatase. It is therefore possible that these proteins play a role in the phosphorylation



of PTL-1 and further studies focusing on the role of this process in *C. elegans* and how it relates to AD-related processes in humans is necessary.

## Conclusions

AD is a complex disease still under study that involves an array of proteins and biological pathways. Two critical players in AD are A $\beta$  and Tau, the main components of amyloid plaques and neurofibrillary tangles, respectively. Unveiling the pathological events leading to the development of AD is a prerequisite for the discovery of the effective prevention and treatment of AD. To this end research often relies in the experimentation on model organisms. In this study we attempted to bridge the gap between AD-related processes in *H. sapiens* and the model organism *C. elegans*. PPI networks were constructed for the human proteins APP and Tau and for their respective *C. elegans* orthologs APL-1 and PTL-1. To achieve this we used network alignment in order to identify proteins and interactions conserved in both organisms that are potentially implicated in AD. Amongst the most prominent pathways, that emerged through this analysis, were the APP processing pathway and the Tau phosphorylation pathway. We extensively reviewed these, critical for AD, pathways and it was revealed that they are highly conserved between the two organisms. Most importantly, we discovered that while the majority of interactions involved in these pathways have been studied in human and associated with AD, almost no information was available about them in *C. elegans*. Additionally, for some of these interactions the evidence in *C. elegans* was very limited indicating the need for more experiments to expand the known *C. elegans* interactome. We therefore managed to showcase significant targets for the study of AD that have been as of yet unexplored in *C. elegans*. In regard to APP processing, candidates for experimental study are the interactions between the members of the *C. elegans*  $\gamma$ -secretase complex, SEL-12, APH-2 and HOP-1, as well as their respective interactions with APL-1. Also, emerging as a target for exploring the neuroprotective role of the human  $\alpha$ -secretase responsible for the non-amyloidogenic processing of APP, was its ortholog SUP-17 and its interaction with APL-1. Additionally, several *C. elegans* kinases are prime subjects for investigation, namely, KIN-1, KIN-2, R90.1 and F47F2.1 that interact with PTL-1 and CDK-5, a kinase that interacts with both APL-1 and PTL-1. These as well as other pathways can be further investigated to identify more critical proteins for experimental validation in *C. elegans*. To enable this functionality, we have created a web application that allows detailed browsing of the networks presented in this work ([http://thalis.biol.uoa.gr/celegans\\_human\\_AD/](http://thalis.biol.uoa.gr/celegans_human_AD/)). In conclusion, we revealed promising AD-related targets for study in *C. elegans* and provided a framework for the transfer of knowledge between *H. sapiens* and *C. elegans* through the computational study of PPI networks.

Received: 1 October 2020; Accepted: 11 January 2021

Published online: 25 February 2021

## References

- Lane, C. A., Hardy, J. & Schott, J. M. Alzheimer's disease. *Eur. J. Neurol.* **25**, 59–70. <https://doi.org/10.1111/ene.13439> (2018).
- Brookmeyer, R., Johnson, E., Ziegler-Graham, K. & Arrighi, H. M. Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement.* **3**, 186–191. <https://doi.org/10.1016/j.jalz.2007.04.381> (2007).
- Zhang, Y. W., Thompson, R., Zhang, H. & Xu, H. APP processing in Alzheimer's disease. *Mol. Brain.* **4**, 3. <https://doi.org/10.1186/1756-6606-4-3> (2011).
- Binder, L. I., Guillozet-Bongaarts, A. L., Garcia-Sierra, F. & Berry, R. W. Tau, tangles, and Alzheimer's disease. *Biochim. Biophys. Acta* **1739**, 216–223. <https://doi.org/10.1016/j.bbadis.2004.08.014> (2005).
- Benson, M. D. *et al.* Amyloid nomenclature 2018: recommendations by the International Society of Amyloidosis (ISA) nomenclature committee. *Amyloid* **25**, 215–219. <https://doi.org/10.1080/13506129.2018.1549825> (2018).
- Mitchell, S. J., Scheibye-Knudsen, M., Longo, D. L. & de Cabo, R. Animal models of aging research: implications for human aging and age-related diseases. *Annu. Rev. Animal Biosci.* **3**, 283–303. <https://doi.org/10.1146/annurev-animal-022114-110829> (2015).
- Janus, C. & Welzl, H. Mouse models of neurodegenerative diseases: criteria and general methodology. *Methods Mol. Biol.* **602**, 323–345. [https://doi.org/10.1007/978-1-60761-058-8\\_19](https://doi.org/10.1007/978-1-60761-058-8_19) (2010).
- Holm, I. E., Alstrup, A. K. & Luo, Y. Genetically modified pig models for neurodegenerative disorders. *J. Pathol.* **238**, 267–287. <https://doi.org/10.1002/path.4654> (2016).
- Emborg, M. E. Nonhuman Primate Models of Neurodegenerative Disorders. *ILAR J.* **58**, 190–201. <https://doi.org/10.1093/ilar/ilx021> (2017).
- Alexander, A. G., Marfil, V. & Li, C. Use of *Caenorhabditis elegans* as a model to study Alzheimer's disease and other neurodegenerative diseases. *Front. Genet.* **5**, 279. <https://doi.org/10.3389/fgene.2014.00279> (2014).
- Fitch, W. M. Distinguishing homologous from analogous proteins. *Syst. Zool.* **19**, 99–113 (1970).
- Corsi, A. K., Wightman, B. & Chalfie, M. A transparent window into biology: a primer on *Caenorhabditis elegans*. *Genetics* **200**, 387–407. <https://doi.org/10.1534/genetics.115.176099> (2015).
- Pir, G. J., Choudhary, B. & Mandelkow, E. *Caenorhabditis elegans* models of tauopathy. *FASEB J* **31**, 5137–5148. <https://doi.org/10.1096/fj.201701007> (2017).
- Liu, W., Wu, A., Pellegrini, M. & Wang, X. Integrative analysis of human protein, function and disease networks. *Sci. Rep.* **5**, 14344. <https://doi.org/10.1038/srep14344> (2015).
- Kikuchi, M. *et al.* Identification of unstable network modules reveals disease modules associated with the progression of Alzheimer's disease. *PLoS ONE* **8**, e76162. <https://doi.org/10.1371/journal.pone.0076162> (2013).
- Li, S. *et al.* A map of the interactome network of the metazoan *C. elegans*. *Science* **303**, 540–543. <https://doi.org/10.1126/science.1091403> (2004).
- Chen, J. Y., Shen, C. & Sivachenko, A. Y. Mining Alzheimer disease relevant proteins from integrated protein interactome data. *Pacific Symp. Biocomput.* **50**, 367–378 (2006).
- Faisal, F. E., Meng, L., Crawford, J. & Milenkovic, T. The post-genomic era of biological network alignment. *EURASIP J. Bioinform. Syst. Biol.* **2015**, 3. <https://doi.org/10.1186/s13637-015-0022-9> (2015).
- Kelley, B. P. *et al.* Conserved pathways within bacteria and yeast as revealed by global protein network alignment. *Proc. Natl. Acad. Sci. U S A* **100**, 11394–11399. <https://doi.org/10.1073/pnas.1534710100> (2003).
- Qian, X. & Yoon, B. J. Comparative analysis of protein interaction networks reveals that conserved pathways are susceptible to HIV-1 interception. *BMC Bioinf.* **12**, 19. <https://doi.org/10.1186/1471-2105-12-S1-S19> (2011).

21. Mukherjee, S. *et al.* Systems biology approach to late-onset Alzheimer's disease genome-wide association study identifies novel candidate genes validated using brain expression data and *Caenorhabditis elegans* experiments. *Alzheimers Dement.* **13**, 1133–1142. <https://doi.org/10.1016/j.jalz.2017.01.016> (2017).
22. Godini, R., Pocock, R. & Fallahi, H. *Caenorhabditis elegans* hub genes that respond to amyloid beta are homologs of genes involved in human Alzheimer's disease. *PLoS ONE* **14**, e0219486. <https://doi.org/10.1371/journal.pone.0219486> (2019).
23. Biza, K. V. *et al.* The amyloid interactome: exploring protein aggregation. *PLoS ONE* **12**, e0173163. <https://doi.org/10.1371/journal.pone.0173163> (2017).
24. Orchard, S. *et al.* The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **42**, D358–363. <https://doi.org/10.1093/nar/gkt1115> (2014).
25. Kim, W., Underwood, R. S., Greenwald, I. & Shaye, D. D. OrthoList 2: a new comparative genomic analysis of human and *Caenorhabditis elegans* genes. *Genetics* **210**, 445–461. <https://doi.org/10.1534/genetics.118.301307> (2018).
26. Lee, R. Y. N. *et al.* WormBase 2017: molting into a new stage. *Nucleic Acids Res.* **46**, D869–D874. <https://doi.org/10.1093/nar/gkx998> (2018).
27. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613. <https://doi.org/10.1093/nar/gky1131> (2019).
28. Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368. <https://doi.org/10.1093/nar/gkw937> (2017).
29. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504. <https://doi.org/10.1101/gr.1239303> (2003).
30. Doncheva, N. T., Morris, J. H., Gorodkin, J. & Jensen, L. J. Cytoscape stringapp: network analysis and visualization of proteomics data. *J. Proteome Res.* **18**, 623–632. <https://doi.org/10.1021/acs.jproteome.8b00702> (2019).
31. Legeay, M., Doncheva, N. T., Morris, J. H. & Jensen, L. J. Visualize omics data on networks with omics visualizer a cytoscape. *Appl. Res.* **9**, 157. <https://doi.org/10.12688/fl1000research.22280.1> (2020).
32. Franz, M. *et al.* Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics* **32**, 309–311. <https://doi.org/10.1093/bioinformatics/btv557> (2016).
33. Guzzi, P. H. & Milenkovic, T. Survey of local and global biological network alignment: the need to reconcile the two sides of the same coin. *Brief Bioinform.* **19**, 472–481. <https://doi.org/10.1093/bib/bbw132> (2018).
34. Vijayan, V., Saraph, V. & Milenkovic, T. MAGNA++: maximizing accuracy in global network alignment via both node and edge conservation. *Bioinformatics* **31**, 2409–2411. <https://doi.org/10.1093/bioinformatics/btv161> (2015).
35. Malek, M., Ibragimov, R., Albrecht, M. & Baumbach, J. CytoGEDEVO—global alignment of biological networks with cytoscape. *Bioinformatics* **32**, 1259–1261. <https://doi.org/10.1093/bioinformatics/btv732> (2016).
36. Neyshabur, B., Khadem, A., Hashemifar, S. & Arab, S. S. NETAL: a new graph-based method for global alignment of protein-protein interaction networks. *Bioinformatics* **29**, 1654–1662. <https://doi.org/10.1093/bioinformatics/btt202> (2013).
37. Sharan, R. & Ideker, T. Modeling cellular machinery through biological network comparison. *Nat. Biotechnol.* **24**, 427–433. <https://doi.org/10.1038/nbt1196> (2006).
38. Needleman, S. B. & Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443–453. [https://doi.org/10.1016/0022-2836\(70\)90057-4](https://doi.org/10.1016/0022-2836(70)90057-4) (1970).
39. Dolinski, K. & Botstein, D. Orthology and functional conservation in eukaryotes. *Annu. Rev. Genet.* **41**, 465–507. <https://doi.org/10.1146/annurev.genet.40.110405.090439> (2007).
40. Asai, M. *et al.* Putative function of ADAM9, ADAM10, and ADAM17 as APP alpha-secretase. *Biochem. Biophys. Res. Commun.* **301**, 231–235. [https://doi.org/10.1016/s0006-291x\(02\)02999-6](https://doi.org/10.1016/s0006-291x(02)02999-6) (2003).
41. Endres, K. & Deller, T. Regulation of alpha-secretase ADAM10 in vitro and in vivo: genetic, epigenetic, and protein-based mechanisms. *Front. Mol. Neurosci.* **10**, 56. <https://doi.org/10.3389/fnmol.2017.00056> (2017).
42. Hornsten, A. *et al.* APL-1, a *Caenorhabditis elegans* protein related to the human beta-amyloid precursor protein, is essential for viability. *Proc. Natl. Acad. Sci. U S A* **104**, 1971–1976. <https://doi.org/10.1073/pnas.0603997104> (2007).
43. Takasugi, N. *et al.* The role of presenilin cofactors in the gamma-secretase complex. *Nature* **422**, 438–441. <https://doi.org/10.1038/nature01506> (2003).
44. Shah, S. *et al.* Nicastrin functions as a gamma-secretase-substrate receptor. *Cell* **122**, 435–447. <https://doi.org/10.1016/j.cell.2005.05.022> (2005).
45. Levitan, D., Yu, G., St George, H. P. & Goutte, C. APH-2/nicastrin functions in LIN-12/Notch signaling in the *Caenorhabditis elegans* somatic gonad. *Dev. Biol.* **240**, 654–661. <https://doi.org/10.1006/dbio.2001.0486> (2001).
46. De Strooper, B., Iwatsubo, T. & Wolfe, M. S. Presenilins and gamma-secretase: structure, function, and role in Alzheimer Disease. *Cold Spring Harb. Perspect. Med.* **2**, a006304. <https://doi.org/10.1101/cshperspect.a006304> (2012).
47. O'Brien, R. J. & Wong, P. C. Amyloid precursor protein processing and Alzheimer's disease. *Annu. Rev. Neurosci.* **34**, 185–204. <https://doi.org/10.1146/annurev-neuro-061010-113613> (2011).
48. Link, C. D. C. *C. elegans* models of age-associated neurodegenerative diseases: lessons from transgenic worm models of Alzheimer's disease. *Exp. Gerontol.* **41**, 1007–1013. <https://doi.org/10.1016/j.exger.2006.06.059> (2006).
49. Liu, S. L. *et al.* The role of Cdk5 in Alzheimer's Disease. *Mol. Neurobiol.* **53**, 4328–4342. <https://doi.org/10.1007/s12035-015-9369-x> (2016).
50. Iijima, K. *et al.* Neuron-specific phosphorylation of Alzheimer's beta-amyloid precursor protein by cyclin-dependent kinase 5. *J. Neurochem.* **75**, 1085–1091. <https://doi.org/10.1046/j.1471-4159.2000.0751085.x> (2000).
51. Lee, M. S. *et al.* APP processing is regulated by cytoplasmic phosphorylation. *J. Cell Biol.* **163**, 83–95. <https://doi.org/10.1083/jcb.200301115> (2003).
52. Patrick, G. N. *et al.* Conversion of p35 to p25 deregulates Cdk5 activity and promotes neurodegeneration. *Nature* **402**, 615–622. <https://doi.org/10.1038/45159> (1999).
53. Manczak, M. & Reddy, P. H. Abnormal interaction of oligomeric amyloid-beta with phosphorylated tau: implications to synaptic dysfunction and neuronal damage. *J. Alzheimers Dis.* **36**, 285–295. <https://doi.org/10.3233/JAD-130275> (2013).
54. Guo, J. P., Arai, T., Miklossy, J. & McGeer, P. L. Abeta and tau form soluble complexes that may promote self aggregation of both into the insoluble forms observed in Alzheimer's disease. *Proc. Natl. Acad. Sci. U S A* **103**, 1953–1958. <https://doi.org/10.1073/pnas.0509386103> (2006).
55. Martin, L. *et al.* Tau protein phosphatases in Alzheimer's disease: the leading role of PP2A. *Ageing Res. Rev.* **12**, 39–49. <https://doi.org/10.1016/j.arr.2012.06.008> (2013).
56. Liu, F., Grundke-Iqbal, I., Iqbal, K. & Gong, C. X. Contributions of protein phosphatases PP1, PP2A, PP2B and PP5 to the regulation of tau phosphorylation. *Eur. J. Neurosci.* **22**, 1942–1950. <https://doi.org/10.1111/j.1460-9568.2005.04391.x> (2005).
57. Flaherty, D. B., Soria, J. P., Tomasiewicz, H. G. & Wood, J. G. Phosphorylation of human tau protein by microtubule-associated kinases: GSK3beta and cdk5 are key participants. *J. Neurosci. Res.* **62**, 463–472. [https://doi.org/10.1002/1097-4547\(20001101\)62:3%3c463::AID-JNR16%3e3.0.CO;2-7](https://doi.org/10.1002/1097-4547(20001101)62:3%3c463::AID-JNR16%3e3.0.CO;2-7) (2000).
58. Yamamoto, H. *et al.* Phosphorylation of tau at serine 416 by Ca<sup>2+</sup>/calmodulin-dependent protein kinase II in neuronal soma in brain. *J. Neurochem.* **94**, 1438–1447. <https://doi.org/10.1111/j.1471-4159.2005.03307.x> (2005).
59. Ghosh, A. & Giese, K. P. Calcium/calmodulin-dependent kinase II and Alzheimer's disease. *Mol. Brain.* **8**, 78. <https://doi.org/10.1186/s13041-015-0166-2> (2015).

60. Kheiri, G., Dolatshahi, M., Rahmani, F. & Rezaei, N. Role of p38/MAPKs in Alzheimer's disease: implications for amyloid beta toxicity targeted therapy. *Rev. Neurosci.* **30**, 9–30. <https://doi.org/10.1515/revneuro-2018-0008> (2018).
61. Liu, F. *et al.* Regulation of amyloid precursor protein (APP) phosphorylation and processing by p35/Cdk5 and p25/Cdk5. *FEBS Lett.* **547**, 193–196. [https://doi.org/10.1016/s0014-5793\(03\)00714-2](https://doi.org/10.1016/s0014-5793(03)00714-2) (2003).
62. Brunden, K. R., Trojanowski, J. Q. & Lee, V. M. Advances in tau-focused drug discovery for Alzheimer's disease and related tauopathies. *Nat. Rev. Drug Discov.* **8**, 783–793. <https://doi.org/10.1038/nrd2959> (2009).
63. Martin, L. *et al.* Tau protein kinases: involvement in Alzheimer's disease. *Ageing Res. Rev.* **12**, 289–309. <https://doi.org/10.1016/j.arr.2012.06.003> (2013).

## Acknowledgements

This research has been co-financed by the European Union and Greek national funds through the Operational Program "Competitiveness, Entrepreneurship and Innovation", under the call "RESEARCH-CREATE-INNOVATE" (Project Code:T1EDK-00353).

## Author contributions

Study design: K.C.N., V.A.I.; Conceptualization: A.E.A., K.C.N., V.A.I.; Methodology: A.E.A., K.C.N.; Software: A.E.A., X.K.S., K.C.N.; Visualization: A.E.A., X.K.S., K.C.N.; Data curation: A.E.A., X.K.S.; Formal analysis: A.E.A., X.K.S., K.C.N., G.I.N.; Writing-original draft: A.E.A.; Writing-review and editing: A.E.A., K.C.N., G.I.N., X.K.S., V.A.I.; Funding acquisition: V.A.I.; Supervision: V.A.I..

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-83892-9>.

**Correspondence** and requests for materials should be addressed to V.A.I.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



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

© The Author(s) 2021