Proceedings of the ***Conference of the Hellenic Society for Computational Biology and Bioinformatics (HSCBB19)***
FORTH/ICE-HT, Patras, Greece, 8-10 December 2019

Elucidating Alzheimer-related conserved pathways in *Homo sapiens* and *Caenorhabditis elegans*: a network alignment approach

Xhuliana Sula[1,#], Avgi E. Apostolakou[1,#], Katerina C. Nastou[1], Georgia I. Nasi[1], Christos Panagopoulos[2], Ilias Maglogiannis[2,3] and Vassiliki A. Iconomidou[1,*]

[1]Department of Biology, National and Kapodistrian University of Athens, Panepistimiopolis, Athens 15701, Greece,
[2]BioAssist S.A., Patras, Greece, [3]University of Piraeus, Piraeus, Greece

[#]These authors contributed equally to this work, [*]Correspondence to: veconom@biol.uoa.gr

Alzheimer disease (AD) is a neurodegenerative disorder, characterized by the presence of amyloid plaques and neurofibrillary tangles. Experimental study of age-related diseases, like AD, is not possible in humans and therefore the use of model organisms, such as *Caenorhabditis elegans*, is common practice. As a result, the two-way transfer of biological knowledge between model organisms and human systems is of utmost importance. One way to accomplish this goal is protein-protein interaction network alignment. In this work, we implemented such an approach and compared the AD-related networks of *H. sapiens* and *C. elegans*. This allowed the identification of conserved biological pathways that are implicated in AD, which can help guide experimental studies on the nematode model.

I. Introduction

Alzheimer disease (AD) is a chronic, progressive, neurodegenerative disorder, whose pathological hallmarks are amyloid plaques of the abnormally folded Aβ peptide − a cleavage product of amyloid precursor protein (APP) − and neurofibrillary tangles of the microtubule-associated Tau protein (Tau) [1]. Considering the main role of these proteins in AD onset, the scientific community has placed much emphasis on unraveling their role in the disease mechanism.

Model organisms are essential "biological tools" that allow scientists to understand the molecular mechanisms of the underlying diseases. Among model organisms, the nematode *Caenorhabditis elegans* provides an attractive model for neurodegenerative diseases, due to its features [2]. A main feature of *C. elegans* is the presence of APL-1 and PTL-1 proteins, orthologs of the human APP and Tau proteins, respectively.

Protein-protein interactions (PPIs) govern most biological processes and investigating PPIs in the context of human diseases is essential to fully comprehend them. The availability of organism-wide PPI networks has made cross-species network comparison possible in the past few years. A popular method for comparing networks is network alignment, which aims at mapping the nodes of two or more networks, and thereby, determining topologically and functionally similar regions in those [3]. Conserved network regions can be used to transfer biologically relevant information between different organisms, and are commonly used for the transfer of knowledge between human and model organisms' networks.

In this work we performed an *in silico* construction and comparison of Alzheimer-related homologous PPI networks in *H. sapiens* 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. We hope that the 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 this currently incurable disorder.

II. Methods

Three network datasets were created for this study, each consisting of a human PPI network and a *C. elegans* PPI network.

A. APP and Tau network from the Amyloid Interactome

The first human network was extracted from the *Amyloid Interactome* [4] and it included APP, Tau, their interaction partners, as well as any interactions between them. For the *C. elegans* network, orthologs of the aforementioned human proteins were recovered using OrthoList2 [5] and were further verified through WormBase [6]. Their interactions were collected from STRING [7], a database of known and predicted PPIs. This dataset was used to evaluate the network alignment algorithms.

B. APL-1 and PTL-1 network from STRING

Afterwards, the interactors of APL-1 and PTL-1 – the *C. elegans* orthologs of APP and Tau – were extracted from STRING. The human orthologs of those interaction partners were found and were used to create the human protein dataset for this network pair. Furthermore, interactions without experimental validation were filtered out to increase the reliability of collected data.

C. Top 100 interaction partners for APP and Tau & APL-1 and PTL-1 from STRING

Lastly, 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, and the top 100 interaction partners of APP and Tau to create the human network. Once again, interactions without experimental validation were filtered out to increase data reliability. All collected data was visualized as PPI networks using Cytoscape 3.7.2 [8]. Comparison between *H. sapiens* and *C. elegans* networks was done using Global Network Alignment (GNA), a method that aims at locating similarities across entire networks. Three algorithms implemented for GNA, namely MAGNA++ [9], CytoGEDEVO [10] and NETAL [11], were selected and tested to determine the best performing one. To optimize each comparison, biological information in the form of protein sequence similarity was used, in addition to topological

parameters. Evaluation of the algorithms was done based on their ability to correctly align the nodes of the networks according to their orthologs.

III. Results & Discussion

The human network extracted from the *Amyloid Interactome* contains 88 proteins and 212 interactions, while the *C. elegans* network from STRING consists of 51 orthologous proteins and 69 interactions between them. These homologous networks were used as a "gold-standard" to evaluate the performance of GNA algorithms that would later allow us to select the best-performing to conduct all subsequent analyses. The use of topological information exclusively, resulted in a complete alignment failure for all pairs of ortholog proteins. On the contrary, the combination of topological with biological information, showed that MAGNA++ and CytoGEDEVO successfully aligned the two networks, with MAGNA++ giving the best results. As a consequence, we resorted to the use of MAGNA++ and biological information for all subsequent alignments.

The *C. elegans* network extracted from STRING, consists of 61 proteins and 136 interactions, while the human network consists of 127 orthologous proteins and 382 interactions. Alignment of the aforementioned networks resulted in a connected unified network composed of 55 pairs of aligned proteins and 92 conserved interactions. MAGNA++ succeeded in correctly aligning the majority of orthologs, further validating its ability to correctly align the pairs of orthologous networks.

The two aforementioned datasets relied on the mapping of ortholog proteins across the two organisms, a procedure that inevitably introduced bias. Conversely, the dataset of the top 100 interaction partners is less biased, and therefore was the one primarily used to transfer biological information between the two organisms. The human network has 91 proteins and 164 interactions, while the *C. elegans* network has 82 proteins and 202 interactions. Alignment of these networks resulted in a unified network comprising of 74 aligned protein pairs and 152 conserved interactions.

Common elements between the unified networks of this and the APL-1 and PTL-1 network from STRING dataset were extracted, and a network consisting of 12 aligned protein pairs, found in both datasets, and 16 conserved interactions emerged. An extensive manual literature search was conducted for the retrieval of information about the function of these proteins in *H. sapiens* and *C. elegans*, focusing specifically on their relation to AD.

Overall, we observed that the APP processing pathway was to a great extend conserved. A notable difference though lays in the absence of beta-secretase; however, this was expected since *C. elegans* does not express this protein. Moreover, many of the common pathways examined involved post-translational modifications, including several conserved interactions implicated in the phosphorylation of Tau, Aβ peptide or both, emphasizing the importance of this pathway in AD. Additionally, proteins and interactions involved in APP neddylation, a process considered to be deregulated in AD, were conserved in *C. elegans*.

IV. Conclusions

Use of model organisms has made the significant advancement in the elucidation of the molecular mechanisms that underlie AD possible. The focus in the study of neurodegenerative disorders has recently shifted to invertebrate animal models, such as *C. elegans*, due to its simple and well-studied nervous system. In this study, conserved protein-protein interactions between *H. sapiens* and *C. elegans* were identified, focusing on proteins associated with APP/APL-1 and Tau/PTL-1. This allowed the investigation of previously unexplored interactions in *C. elegans*, that are involved in human pathways implicated in AD.

V. Acknowledgments

The present work was co-funded 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)

References

[1] C.A. Lane, J. Hardy and J.M. Schott, "Alzheimer's disease," European journal of neurology 2018**,** *25* (1), 59-70.

[2] A.G. Alexander, V. Marfil and C. Li, "Use of Caenorhabditis elegans as a model to study Alzheimer's disease and other neurodegenerative diseases," Frontiers in genetics 2014**,** *5*, 279.

[3] F.E. Faisal, L. Meng, J. Crawford and T. Milenkovic, "The post-genomic era of biological network alignment," EURASIP journal on bioinformatics & systems biology 2015**,** *2015*, 3.

[4] K.V. Biza, K.C. Nastou, P.L. Tsiolaki, C.V. Mastrokalou, S.J. Hamodrakas and V.A. Iconomidou, "The amyloid interactome: Exploring protein aggregation," PloS one 2017**,** *12* (3), e0173163.

[5] W. Kim, R.S. Underwood, I. Greenwald and D.D. Shaye, "OrthoList 2: A New Comparative Genomic Analysis of Human and Caenorhabditis elegans Genes," Genetics 2018**,** *210* (2), 445-461.

[6] R.Y.N. Lee, et al., "WormBase 2017: molting into a new stage," Nucleic acids research 2018**,** *46* (D1), D869-D874.

[7] D. Szklarczyk, et al., "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," Nucleic acids research 2019**,** *47* (D1), D607-D613.

[8] P. Shannon, et al., "Cytoscape: A software environment for integrated models of biomolecular interaction networks," Genome Res 2003**,** *13* (11), 2498-2504.

[9] V. Vijayan, V. Saraph and T. Milenkovic, "MAGNA++: Maximizing Accuracy in Global Network Alignment via both node and edge conservation," Bioinformatics 2015**,** *31* (14), 2409-2411.

[10] M. Malek, R. Ibragimov, M. Albrecht and J. Baumbach, "CytoGEDEVO-global alignment of biological networks with Cytoscape," Bioinformatics 2016**,** *32* (8), 1259-1261.

[11] B. Neyshabur, A. Khadem, S. Hashemifar and S.S. Arab, "NETAL: a new graph-based method for global alignment of protein-protein interaction networks," Bioinformatics 2013**,** *29* (13), 1654-1662.