In-silico prediction of the complete ataxin-3 protein network relevant for Spinocerebellar Ataxia type 3 (SCA3)

Paulo Batista[1,2,*], Jorge Vieira[1,3], Cristina P. Vieira[1,3], Brígida Mónica Faria [1,3]

1. I3S Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto Portugal; ESS Escola Superior de Saúde do Politécnico do Porto, Porto, Portugal; **2.** IBMC Instituto de Biologia Molecular e CelularI3S, Porto, Portugal; **3.** LIACC – Laboratório de Inteligência Artificial e Ciência de Computadores

* Corresponding author email: 10220893@ess.ipp.pt

Introduction: Spinocerebellar ataxia type 3, also known as Machado Joseph disease (SCA3/MJD), is the most common inherited ataxia worldwide and is caused by a pathogenic expansion of the polyglutamine (polyQ) tract, located at the C-terminal region of the ataxin-3 protein (1). The polyQ region is involved in the stabilization of protein- protein interactions (PPIs). Abnormal polyQ expansion results in structural changes of the ataxin-3 (2,3), implying different accessibility at specific interacting residues, needed for the normal protein activity. PolyQ proteins have large protein networks. Mapping of PPIs has been performed using high-throughput methods, that are known to produce false interactions (4). Therefore, the use of multiple interactomes comparisons (conserved interactions between pairs of proteins which have interacting homologs in another organism, as well as proteomic data from cell lines, patients, mutants expressing a human protein, and cross-species genetic screens (modifier screens), available at EvoPPI3 (5)), together with in-silico analyses, can be used to support PPIs, as well as identify novel interactors. **Objectives:** In this work we will: 1- characterize ataxin-3 network (validating the proteins identified in main databases, as well as identify new putative interactors); 2- identifying the interactors that behave differently in the presence of an expanded polyQ using different 3D structure prediction methods and protein docking methods. **Methods:** Using EvoPPI3 and protein expression in tissues that matter to SCA3 for PPI retrieval and validation, as well as identification of new interactors. *In-silico* approaches for predicting protein binding differences between wildtype and expanded ataxin-3 forms will be performed, using different *a)* 3D protein structure predictions (namely ITASSER (6), AlphaFold (7), and D-ITASSER (8)) and *b)* protein docking methodologies (such as HADDOCK (9) and ClustPro (10). **Results:** Using EvoPPI3, there are 422 ataxin-3 interactors in human main databases. From this, 250 proteins have been previously studied. Of the remaining 172 proteins, 158 have been reported from proteomic analyses of human cell lines and ataxin-3 patients (*H. sapiens* polyQ_22 database), and these could be true interactors. 28 proteins are in common when considering the polyQ, *Mus musculus* interlogs and *Danio rerio* interlogs, and these could be novel interactors to study. From the 158, 73 proteins bind more to the expanded form of ataxin-3 using AlphaFold, to confirm these results we used ITASSER, where we obtained 46 of the 73 that bind more to the expanded form. **Conclusion:** This study contributes significantly to understanding SCA3

pathology by delineating a network of ataxin-3 interactors and analysing their behaviour in the presence of an expanded polyQ stretch.

Keywords: ataxin-3, SCA3, polyQ protein-protein interactions, *in-silico* methodology

References:

1. McLoughlin HS, Moore LR, Paulson HL. Pathogenesis of SCA3 and implications for other polyglutamine diseases. Neurobiol Dis [Internet]. 2020 Feb 1 [cited 2024 Jan 17];134. Available from: https://pubmed.ncbi.nlm.nih.gov/31669734/
2. Lim J, Hao T, Shaw C, Patel AJ, Szabó G, Rual JF, et al. A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. Cell [Internet]. 2006 May 19 [cited 2024 Jan 17];125(4):801–14. Available from: https://pubmed.ncbi.nlm.nih.gov/16713569/
3. Rocha S, Vieira J, Vázquez N, López-Fernández H, Fdez-Riverola F, Reboiro-Jato M, et al. ATXN1 N-terminal region explains the binding differences of wild-type and expanded forms. BMC Med Genomics [Internet]. 2019 Oct 26 [cited 2024 Jan 17];12(1):1–14. Available from: https://bmcmedgenomics.biomedcentral.com/articles/10.1186/s12920-019-0594-4
4. Sousa e Silva R, Sousa AD, Vieira J, Vieira CP. The Josephin domain (JD) containing proteins are predicted to bind to the same interactors: Implications for spinocerebellar ataxia type 3 (SCA3) studies using Drosophila melanogaster mutants. Front Mol Neurosci. 2023 Mar 15;16:1140719.
5. Sousa A, Rocha S, Vieira J, Reboiro-Jato M, López-Fernández H, Vieira CP. On the identification of potential novel therapeutic targets for spinocerebellar ataxia type 1 (SCA1) neurodegenerative disease using EvoPPI3. J Integr Bioinform [Internet]. 2023 Jun 1 [cited 2024 Jan 17];20(2). Available from: https://pubmed.ncbi.nlm.nih.gov/36848492/
6. Zhou X, Zheng W, Li Y, Pearce R, Zhang C, Bell EW, et al. I-TASSER-MTD: a deep-learning- based platform for multi-domain protein structure and function prediction. [cited 2024 Apr 20]; Available from: https://doi.org/10.1038/s41596-022-00728-0
7. David A, Islam S, Tankhilevich E, Sternberg MJE. The AlphaFold Database of Protein Structures: A Biologist's Guide. J Mol Biol [Internet]. 2022 Jan 30 [cited 2024 Jan 17];434(2). Available from: https://pubmed.ncbi.nlm.nih.gov/34757056/
8. D-I-TASSER: deep learning-based protein structure prediction [Internet]. [cited 2024 Jan 19]. Available from: https://zhanggroup.org/D-I-TASSER/
9. Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc [Internet]. 2003 Feb 15 [cited 2024 Jan 17];125(7):1731–7. Available from: https://pubmed.ncbi.nlm.nih.gov/12580598/
10. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein-protein docking. Nat Protoc [Internet]. 2017 Feb 1 [cited 2024 Jan 21];12(2):255–78. Available from: https://pubmed.ncbi.nlm.nih.gov/28079879/