NetControl4BioMed: A web-based platform for controllability analysis of protein-protein interaction networks

Victor-Bogdan Popescu¹, Jose Angel Sanchez-Martinez², Daniela Schacherer³, Sadra Safadoust⁴, Negin Majidi⁵, Andrei Andronescu⁶, Alexandru Nedea⁶, Diana Ion⁶, Eduard Mititelu⁶, Eugen Czeizler¹,²,⁸ and Ion Petre⁷,⁸,*

¹Åbo Akademi University, Turku, Finland and ²Technical University of Madrid, Spain and ³Heidelberg University, Germany and ⁴Koç University, Istanbul, Turkey and ⁵University of California, Santa Cruz, United States and ⁶University “Politehnica” of Bucharest, Romania and ⁷University of Turku, Finland and ⁸National Institute for Research and Development for Biological Sciences, Bucharest, Romania.

*To whom correspondence should be addressed.

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: There is an increasing amount of data coming from genome-wide studies identifying disease-specific survivability-essential proteins and host factors critical to a cell becoming infected. Targeting such proteins has a strong potential for targeted, precision therapies. Typically however, too few of them are drug targetable. An alternative approach is to influence them through drug targetable proteins upstream of them. Structural target network controllability is a suitable solution to this problem. It aims to discover suitable source nodes (e.g., drug targetable proteins) in a directed interaction network that can control (through a suitable set of input functions) a desired set of targets.

Results: We introduce NetControl4BioMed, a free open-source web-based application that allows users to generate or upload directed protein-protein interaction networks and to perform target structural network controllability analyses on them. The analyses can be customized to focus the search on drug targetable source nodes, thus providing drug therapeutic suggestions. The application integrates protein data from HGNC, Ensemble, UniProt, NCBI, and InnateDB, directed interaction data from InnateDB, Omnipath, and SIGNOR, cell-line data from COLT and DepMap, and drug-target data from DrugBank.

Availability: The application and data are available online at https://netcontrol.combio.org/. The source code is available at https://github.com/Vilksar/NetControl4BioMed under an MIT license.

Contact: ion.petre@utu.fi

1 Introduction

Genome-wide association studies led in the last few years to an increasing availability of data on disease-specific survivability-essential genes (Koh et al., 2011) and on host factors critical to cell infection (Daniloski et al., 2021). Such data can be used in network-based drug repurposing studies (Morselli Gysi et al., 2021). The concept is to trace the cascading signals of drug combinations through directed protein-protein interactions from the drug targets to the essential/critical proteins. One of the promising computational approaches to this problem is target network controllability, that can be used to identify combinations of drug targetable proteins controlling a set of critical targets in a directed network. Several formulations and demonstrations of this approach exist, especially on Boolean network controllability (Murrugarra et al., 2016, Zañudo and Albert, 2015, Biane and Delaplace, 2019) and on target structural controllability (Wei-Feng et al., 2017, Kanhaiya et al., 2017).
4 Network analysis

To run a controllability analysis the user needs to specify the following: (1) the network to be analyzed; (2) (optional) the list of source protein identifiers which would be preferred as control inputs; (3) the list of target protein identifiers which should be controlled; (4) the algorithm for the controllability analysis and its parameters. Two controllability algorithms are available: the greedy algorithm described in Creutz et al., 2018 and the genetic algorithm described in Popescu et al., 2021. Each algorithm requires several specific parameters, and predefined default values for each parameter are available. The output of the analysis consists of one or more sets of control paths, each of them containing the list of control inputs able to control the entire target set (with the drug-targets among them distinctly marked), as well as the list of individual paths between each target and its corresponding control input. These control paths can be individually inspected and downloaded for external use and visualization. The duration of the controllability analysis varies based on the size of the network, the number of target proteins, and the parameters of the algorithm. The analysis runs on the server and the user is notified when the results are available.

5 Conclusions

We present a new web application for network generation and network structural target controllability analysis, with a focus on biomedicine. The software provides a modern and friendly user interface, allowing for sharing and collaboration between users. We provide several already compiled and ready-to-be-used datasets on protein-protein interaction networks, disease-specific survivability-essential and mutated genes, and drug-target genes. We believe that the application will facilitate experimenting and effective application of network analysis techniques in the biomedical domain. It can be potentially useful to researchers for better understanding of interaction networks pathway structure, for identifying novel therapeutic suggestions, and for a patient- and disease-specific personalized approach to treatment.

Funding

This work was partially supported by the Romanian National Education and Research, CCCDI – UEFISCEDI, project number PNII-P2-1.1-PED-2019-2391, within PNCDI III.

References

Biane, C. et al. (2019). Causal reasoning on boolean control networks based on ablation: Theory and application to cancer drug discovery. IEEE/ACM Trans. Comput. Biol. Bioinformatics, 16(5), 1574–1585.

Boehm, J. S., et al. (2021). Cancer research needs a better map. Nature, 589, 514–516.

Brachi, B., et al. (2019). Genenames.org: the HGNC and VGNC resources in 2019. Nucleic Acids Research, 47(D1), D796–D802.

Breuer, K., et al. (2013). InnateDB: systems biology of innate immunity and beyond--recent updates and continuing curation. Nucleic Acids Research, 41(Database issue), D1228–D1233.

Brown, G. R., et al. (2015). Gene: a gene-centered information resource at NCBI. Nucleic Acids Research, 43(Database issue), D356–D362.

Consortium, T. U. (2021). UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research, 49(D1), D480–D489.

Creutz, E., et al. (2018). Structured target controllability of linear networks. IEEE/ACM Trans. on Computational Biology and Bioinf., 15(4), 1217–1228.

Daniloški, Z., et al. (2021). Identification of required host factors for SARS-CoV-2 infection in human cells. Cell, 184(1), 92–105.e16.

Kanthaiya, K., et al. (2017). Controlling directed protein interaction networks in cancer. Scientific Reports, 7(1), 10327.

Kanthaiya, K., et al. (2019). NetControl4BioMed: a pipeline for biomedical data acquisition and analysis of network controllability. BMC Bioinformatics, 19(185).
Koh, J. L. Y., et al. (2011). Colt-cancer: functional genetic screening resource for essential genes human cancer cell lines. Nucleic Acids Res., 40(D1), D957–D963.
Licata, L., et al. (2020). SIGNOR 2.0, the SIGNaling Network Open Resource 2.0: 2019 update. Nucleic Acids Research, 48(D1), D504–D510.
Liu, P.-C. K. et al. (2012). Application of max-sat-based atpg to optimal cancer therapy design. BMC Genomics, 13(6), S5.
Licata, L., et al. (2020). SIGNOR 2.0, the SIGnaling Network Open Resource 2.0: 2019 update. Nucleic Acids Research, 48(D1), D504–D510.
Lin, P.-C. K. et al. (2012). Application of max-sat-based atpg to optimal cancer therapy design. BMC Genomics, 13(6), S5.
Morselli Gysi, D., et al. (2021). Network medicine framework for identifying drug-repurposing opportunities for covid-19. Proceedings of the National Academy of Sciences, 118(19), e2025811118.
Murrugarra, D., et al. (2016). Identification of control targets in boolean molecular network models via computational algebra. BMC Systems Biology, 10(1), 94.
Popescu, V.-R., et al. (2021). Identifying efficient controls of complex interaction networks using genetic algorithms. Available at https://arxiv.org/abs/2007.04853.
Su, C. et al. (2021). Cbean: a software for the control of asynchronous boolean networks. Bioinformatics, 37(6), 879–881.
Türi, D., et al. (2016). OmniPath: guidelines and gateway for literature-curated signaling pathway resources. Nature Methods, 13, 966–967.
Wei-Feng, G., et al. (2017). Constrained target controllability of complex networks. Journal of Statistical Mechanics: Theory and Experiment, 2017, 063402.
Wishart, D. S., et al. (2017). DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Research, 46(D1), D1074–D1082.
Wu, L., et al. (2017). CytosCtrlAnalyzer: a Cytoscape app for biomolecular network controllability analysis. Bioinformatics, 34(8), 1428–1430.
Yates, A. D., et al. (2019). Ensembl 2020. Nucleic Acids Res., 48(D1), D682–D688.
Zañudo, J. G. T. et al. (2015). Cell fate reprogramming by control of intracellular network dynamics. PLOS Computational Biology, 11(4), e1004193–.