Bioinformatics, 37(21), 2021, 3976–3978
doi: 10.1093/bioinformatics/btab570
Advance Access Publication Date: 5 August 2021
Applications Note

OXFORD

Systems biology

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

Victor-Bogdan Popescu [1], José Ángel Sánchez-Martín[2], Daniela Schacherer [3], Sadra Safadoust[4], Negin Majidi[5], Andrei Andronescu[6], Alexandru Nedea[6], Diana Ion[6], Eduard Mititelu[6], Eugen Czeizler[1,7], and Ion Petre[7,8,*]

[1]Computer Science Unit, Åbo Akademi University, Turku 20520, Finland, [2]Department of Information Systems, Polytechnic University of Madrid, 28040 Madrid, Spain, [3]Faculty of Mathematics and Computer Science, Heidelberg University, 69120 Heidelberg, Germany, [4]Department of Computer Science and Engineering, Koç University, Rumelifeneri Yolu 34450 Sarıyer, İstanbul, Turkey, [5]Department of Computer Science and Engineering, University of California, Santa Cruz, CA 95064, USA, [6]Faculty of Electronics, Telecommunications and Information Technology, POLITEHNICA University of Bucharest 061071, Romania, [7]Department of Bioinformatics and Biostatistics, National Institute for Research and Development for Biological Sciences, Bucharest 060031, Romania and [8]Department of Mathematics and Statistics, University of Turku, Turku 20014, Finland

*To whom correspondence should be addressed.
Associate Editor: Pier Luigi Martelli

Received on May 6, 2021; revised on July 12, 2021; editorial decision on July 30, 2021; accepted on August 3, 2021

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.

Availabilityand implementation: 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.*, 2012) 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 (Biane *et al.*, 2019; Murrugarra *et al.*, 2016; Zañudo *et al.*, 2015) and on target structural controllability (Kanhaiya *et al.*, 2017; Wei-Feng *et al.*, 2017).

We introduce NetControl4BioMed, a free open-source web-based software, aimed at applications in biomedicine and allowing for: (i) constructing directed protein–protein interaction networks, (ii) structural target network controllability analysis focused on identifying effective drug-combinations and (iii) sharing networks

© The Author(s) 2021. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from academic.oup.com/bioinformatics/article/37/21/3976/6342407 by Koç University Suna Kraç Library user on 29 October 2026

and analyses between users. It is a re-engineering of the first version of the software (Kanhaiya *et al.*, 2019), from the algorithms and the implementation to the interface and the functionality. NetControl4BioMed allows multi-user collaborative access to building directed protein–protein interaction networks around a set of seed proteins of interest, to upload such networks from external platforms, and to perform structural target controllability analyses with a focus on drug combination identification. Several approaches exist for Boolean network controllability (Biane *et al.*, 2019; Lin *et al.*, 2012; Murrugarra *et al.*, 2016; Su *et al.*, 2021; Zañudo *et al.*, 2015). For structural target controllability the only other tool that we are aware of is the Cytoscape app CytoCtrlAnalyser (Wu *et al.*, 2018). In comparison to CytoCtrlAnalyser, NetControl4BioMed offers the ability to generate directed protein–protein interaction networks, a much more customizable search, integration with multiple external databases, including drug data and several cell line gene essentiality data, a cloud-based approach to network analysis, independent of the performance of the user's own system, and the possibility for multi-user collaboration.

We discuss in the next sections the data that NetControl 4BioMed integrates and its usability for network generation and network analysis.

2 Data

We use pre-compiled protein data from the public online HGNC (Braschi *et al.*, 2019), Ensembl (Yates *et al.*, 2019), UniProt (Consortium, 2021), NCBI (Brown *et al.*, 2015) and InnateDB (Breuer *et al.*, 2013) databases, with all the corresponding unique identifiers being integrated by the application. The interaction data uses experimentally validated information from the Omnipath (Türei *et al.*, 2016), InnateDB (Breuer *et al.*, 2013) and SIGNOR (Licata *et al.*, 2020) databases. The data contains 42 152 proteins and 46 942 interactions. The application also provides a set of 1578 pre-compiled protein collections, consisting of 52 sets of disease-specific survivability-essential genes for several cancer cell-lines from COLT (Koh *et al.*, 2012), the 1526 sets of mutated genes for several cancer cell-lines from DepMap (Boehm *et al.*, 2021), and the 9 sets of drug–target genes from DrugBank (Wishart *et al.*, 2018).

3 Network generation

To generate a network the user needs to specify the following: (i) the list of seed protein identifiers around which the network will be built, (ii) the interaction database(s) to be used by the network and (iii) the algorithm for the network generation. Several algorithms are available: selecting all interactions containing the seed proteins, selecting only direct interactions between the seed proteins, selecting the interactions between seed proteins with at most one to four intermediary proteins. The output consists of a network which can be inspected, downloaded for external use and visualization, or used further in the application for analysis. The size of the generated networks varies based on the number of seed proteins, the number of selected interaction databases and the generation algorithm. Networks with tens of thousands of interactions can easily be handled by the software.

4 Network analysis

To run a controllability analysis the user needs to specify the following: (i) the network to be analyzed, (ii) (optional) the list of source protein identifiers which would be preferred as control inputs, (iii) the list of target protein identifiers which should be controlled and (iv) the algorithm for the controllability analysis and its parameters. Two controllability algorithms are available: the greedy algorithm described in Czeizler *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 Ministry of Education and Research, CCCDI–UEFISCDI (project number PNIII-P2-2.1-PED-2019-2391, within PNCDI III awarded to IP) and by the Academy of Finland (project number 311371 awarded to EC).

Conflict of Interest: none declared.

References

Biane,C. *et al.* (2019) Causal reasoning on Boolean control networks based on abduction: theory and application to cancer drug discovery. *IEEE/ACM Trans. Comput. Biol. Bioinf.*, **16**, 1574–1585.

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

Braschi,B. *et al.* (2019) Genenames.org: the HGNC and VGNC resources in 2019. *Nucleic Acids Res.*, **47**, D786–D792.

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

Brown,G.R. *et al.* (2015) Gene: a gene-centered information resource at NCBI. *Nucleic Acids Res.*, **43**, D36–D42.

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

Czeizler,E. *et al.* (2018) Structural target controllability of linear networks. *IEEE/ACM Trans. Comput. Biol. Bioinf.*, **15**, 1217–1228.

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

Kanhaiya,K. *et al.* (2017) Controlling directed protein interaction networks in cancer. *Sci. Rep.*, **7**, 10327.

Kanhaiya,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.* (2012) Colt-cancer: functional genetic screening resource for essential genes human cancer cell lines. *Nucleic Acids Res.*, **40**, D957–D963.

Licata,L. *et al.* (2020) SIGNOR 2.0, the SIGnaling Network Open Resource 2.0: 2019 update. *Nucleic Acids Res.*, **48**, D504–D510.

Lin,P.-C.K. *et al.* (2012) Application of max-sat-based ATPG to optimal cancer therapy design. *BMC Genomics*, **13**, S5.

Morselli Gysi,D. *et al.* (2021) Network medicine framework for identifying drug-repurposing opportunities for covid-19. *Proc. Natl. Acad. Sci. USA*, **118**, e2025581118.

Downloaded from academic.oup.com/bioinformatics/article/37/21/3976/6342407 by Koç University Suna Kraç Library user on 29 October 2026

Murrugarra,D. *et al.* (2016) Identification of control targets in Boolean molecular network models via computational algebra. *BMC Syst. Biol.*, **10**, 94.
Popescu,V.-B. *et al.* (2021) Identifying efficient controls of complex interaction networks using genetic algorithms. https://arxiv.org/abs/2007.04853 (15 July 2021, date last accessed).
Su,C. *et al.* (2021) Cabean: a software for the control of asynchronous Boolean networks. *Bioinformatics*, **37**, 879–881.
Türei,D. *et al.* (2016) OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nat. Methods*, **13**, 966–967.
Wei-Feng,G. *et al.* (2017) Constrained target controllability of complex networks. *J. Stat. Mech. Theory Exp.*, **2017**, 063402.
Wishart,D.S. *et al.* (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.*, **46**, D1074–D1082.
Wu,L. *et al.* (2018) CytoCtrlAnalyser: a Cytoscape app for biomolecular network controllability analysis. *Bioinformatics*, **34**, 1428–1430.
Yates,A.D. *et al.* (2019) Ensembl 2020. *Nucleic Acids Res.*, **48**, D682–D688.
Zañudo,J.G.T. *et al.* (2015) Cell fate reprogramming by control of intracellular network dynamics. *PLOS Comput. Biol.*, **11**, e1004193.

Downloaded from academic.oup.com/bioinformatics/article/37/21/3976/6342407 by Koç University Suna Kraç Library user on 29 October 2026