OXFORD

Databases and ontologies

CaNDis: a web server for investigation of causal relationships between diseases, drugs and drug targets

Blaž Škrlj[1,2], Nika Eržen[1], Nada Lavrač[1,2], Tanja Kunej [iD][3] and Janez Konc [iD][4,*]

[1]Department of Knowledge Technologies, Jožef Stefan Institute; [2]Jožef Stefan International Postgraduate School; [3]Department of Animal Science, Biotechnical Faculty, University of Ljubljana and [4]Theory Department, National Institute of Chemistry, SI-1000 Ljubljana, Slovenia

*To whom correspondence should be addressed.

Associate Editor: Lenore Cowen

Abstract

Motivation: Causal biological interaction networks represent cellular regulatory pathways. Their fusion with other biological data enables insights into disease mechanisms and novel opportunities for drug discovery.

Results: We developed Causal Network of Diseases (CaNDis), a web server for the exploration of a human causal interaction network, which we expanded with data on diseases and FDA-approved drugs, on the basis of which we constructed a disease–disease network in which the links represent the similarity between diseases. We show how CaNDis can be used to identify candidate genes with known and novel roles in disease co-occurrence and drug–drug interactions.

Availability and implementation: CaNDis is freely available to academic users at http://candis.ijs.si and http://candis.insilab.org.

Contact: konc@cmm.ki.si

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Protein–protein and protein–drug interaction networks are becoming increasingly important for drug development (Ashburn *et al.*, 2004). Such biological interactions, whether predicted or experimental, are stored in various databases (Konc *et al.*, 2017; Szklarczyk *et al.*, 2019). However, the causal relationships between the interaction partners remain relatively unexplored. A causal relationship occurs when e.g. a transcription factor increases the expression of a particular protein. Knowledge of these relationships could open up new avenues for drug development and improve our understanding of the disease processes.

Tools to explore networks of causal relationships are available (Boué *et al.*, 2015; Licata *et al.*, 2020); however, the fusion of these networks with the knowledge about diseases and drugs into heterogeneous causal networks, should enable even higher understanding of diseases as well as more efficient drug discovery. In particular, such networks could be used for the study of disease co-occurrence, the simultaneous presence of two or more diseases in a patient (Hu *et al.*, 2016).

We developed Causal Network of Diseases (CaNDis), a web server that enables visual exploration of a causal network of protein–protein interactions enriched with annotations on diseases and drugs. We also constructed a new disease–disease network. Here, each node is a subnetwork in the original causal network, in which

all the nodes are linked to the same disease. A score is assigned to a link between a pair of diseases based on the connectivity of the two subnetworks, representing disease similarity. CaNDis is an intuitive web server based on a fast WebGL viewer (see sections 'Software aspects and scaling' and 'Comparison with other tools' in Supplementary Information). It facilitates drug discovery with respect to the identification of disease co-occurrence and drug–drug interactions at the level of drugs, genes, proteins and diseases.

2 The CaNDis web server

The CaNDis heterogeneous causal network is based on the CBN (Boué *et al.*, 2015) and SIGNOR 2.0 (Licata *et al.*, 2020) causal biological networks. It is extended with protein–RNA interactions from the PRD database (Fujimori *et al.*, 2012) and drug–gene interactions from the DGIdb database (Cotto *et al.*, 2018); gene nodes are superimposed with the gene–disease annotations from the DisGeNET database (Piñero *et al.*, 2020). Causal pathways are merged into a single network containing the interacting proteins as nodes and the causal relationships, such as increase or decrease in expression levels, as links (see section 'The CaNDis network' in Supplementary Information).

Gene–disease annotations are added as attributes to the protein nodes and drugs and RNA are added as new nodes, which are linked
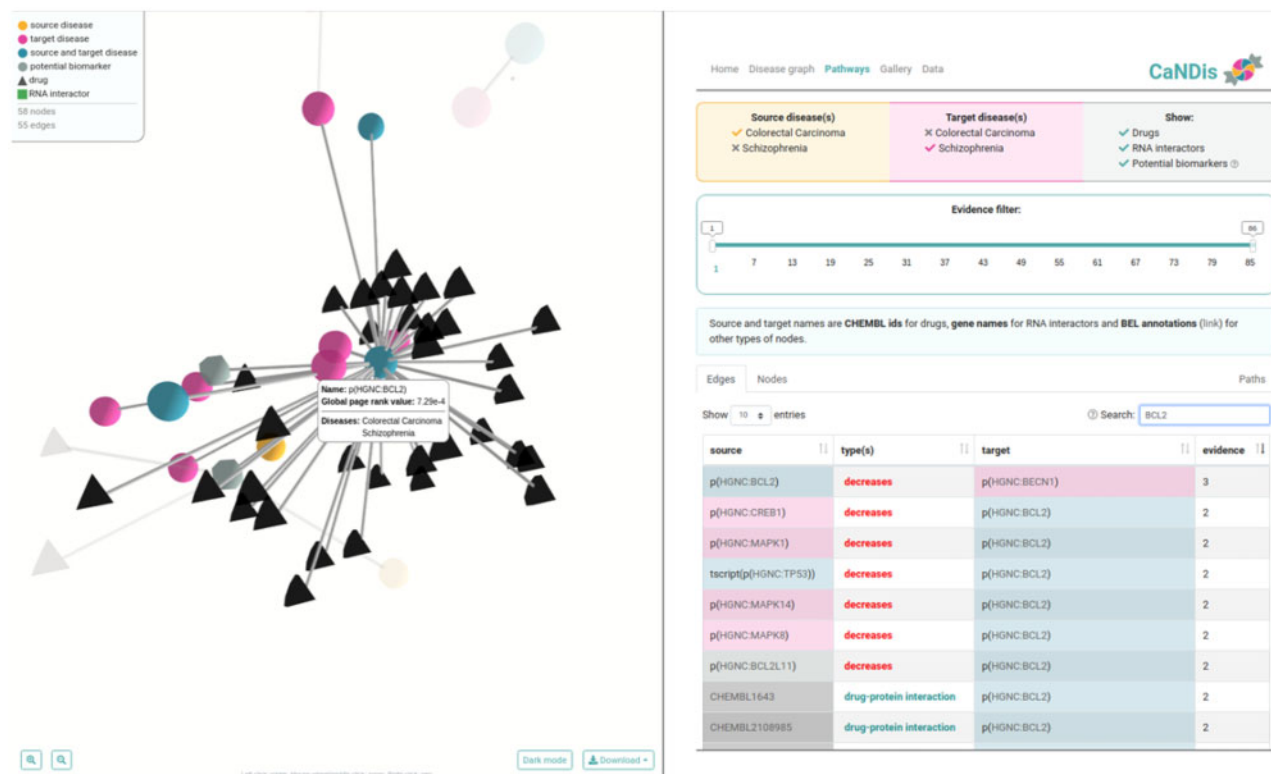
Fig. 1. CaNDis web server causal interaction network for the *BCL2* gene. Left: proteins (identified by their gene names) associated with schizophrenia (yellow circles), cancer (purple circle), both diseases (blue circles) or none of the diseases (grey circles) represented as a 3D interaction network. Drugs are black triangles. Highlighted are the *BCL2* gene in the middle and its neighbors. Right: tabular view of the same interactions labeled with causal relationships (e.g. decreases). Some of the relevant neighbors of *BCL2* gene are *DDIT3*, *MAPK8*, *MAPK14*, *MAPK1*, *CREB1*, *PRKCA*, *CASP3* and *BECN1* (all genes are available on the web server)

to their respective target protein nodes. The obtained heterogeneous causal network contains annotations on more than 3000 complex diseases as well as on the FDA approved drugs to enable the investigation of diseases (i.e. co-occurrence of diseases) and drug–drug interactions (see section 'Use in drug–drug interactions' in Supplementary Information).

A new disease–disease network is also developed from the heterogeneous causal network (sections 'Disease–disease network construction' and 'Normalized disease–disease similarity score' in Supplementary Information), and is available on the web server.

3 Use for identification of genes with roles in disease co-occurrence

According to epidemiological studies, patients with central nervous system (CNS) disorders, that is, Alzheimer's disease, Parkinson's disease and schizophrenia, are less likely to develop certain cancers (colorectal, lung and prostate), and vice versa. This inverse co-occurrence is attributed to a set of proteins, including the proteins coded by the *TP53* and *BCL2* genes (Behrens *et al.*, 2009; Ibáñez *et al.*, 2014; Tabarés-Seisdedos *et al.*, 2009), both of which are down-regulated in cancers and up-regulated in CNS disorders. Using 'colorectal carcinoma' and 'schizophrenia' as the inputs to the CaNDis web server, we successfully identified the *TP53* tumor suppressor gene and the *BCL2* anti-apoptotic gene as those in common to both the diseases (*BCL2* network is shown in Fig. 1). In addition, we also found proteins that have not yet been associated with either disease (Piñero *et al.*, 2020), but which could play a role in this inverse disease co-occurrence due to their interactions with the protein coded by the *BCL2* gene (the list of proteins is in the caption of Fig. 1). Literature search showed that some of the identified proteins are indeed associated with both diseases. For example, the oncoprotein coded by the *DDIT3* gene (Engström *et al.*, 2006) is associated with lung cancer (Li *et al.*, 2015) and upregulated in schizophrenia

(Umeda-Yano *et al.*, 2013). Similarly, *MAPK8* plays a role in both colorectal cancer and schizophrenia (Chandrasekaran *et al.*, 2012; He *et al.*, 2012; Szatkiewicz *et al.*, 2014). This illustrates the use of the CaNDis web server's innovative network visualization to quickly suggest candidate proteins with potential new roles in co-occurrence of diseases.

Funding

References

Ashburn,T.T. *et al.* (2004) Drug repositioning: identifying and developing new uses for existing drugs. *Nat. Rev. Drug Discov*., **3**, 673–683.

Behrens,M.I. *et al.* (2009) A common biological mechanism in cancer and Alzheimer's disease? *Curr. Alzheimer Res*., **6**, 196–204.

Boué,S. *et al.* (2015) Causal biological network database: a comprehensive platform of causal biological network models focused on the pulmonary and vascular systems. *Database (Oxford)*, **2015**, bav030.

Chandrasekaran,S. *et al.* (2012) A network view on Schizophrenia related genes. *Network Biol*., **2**, 16–25.

Cotto,K.C. *et al.* (2018) DGIdb 3.0: a redesign and expansion of the drug–gene interaction database. *Nucleic Acids Res*., **46**, D1068–D1073.

Engström,K. *et al.* (2006) The myxoid/round cell liposarcoma fusion oncogene FUS-DDIT3 and the normal DDIT3 induce a liposarcoma phenotype in transfected human fibrosarcoma cells. *Am. J. Pathol*., **168**, 1642–1653.

Fujimori,S. *et al.* (2012) PRD: a protein–RNA interaction database. *Bioinformation*, **8**, 729–730.

He,W. *et al.* (2012) Attenuation of TNFSF10/TRAIL-induced apoptosis by an autophagic survival pathway involving TRAF2- and RIPK1/RIP1-mediated MAPK8/JNK activation. *Autophagy*, **8**, 1811–1821.

Hu,J.X. *et al.* (2016) Network biology concepts in complex disease comorbidities. *Nat. Rev. Genet.*, **17**, 615–629.

Ibáñez,K. *et al.* (2014) Molecular evidence for the inverse comorbidity between central nervous system disorders and cancers detected by transcriptomic meta-analyses. *PLoS Genet.*, **10**, e1004173.

Konc,J. *et al.* (2017) GenProBiS: web server for mapping of sequence variants to protein binding sites. *Nucleic Acids Res.*, **45**, W253–W259.

Li,T. *et al.* (2015) DDIT3 and KAT2A proteins regulate TNFRSF10A and TNFRSF10B expression in endoplasmic reticulum stress-mediated apoptosis in human lung cancer cells. *J. Biol. Chem.*, **290**, 11108–11118.

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

Piñero,J. *et al.* (2020) The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.*, **48**, D845–D855.

Szatkiewicz,J.P. *et al.* (2014) Copy number variation in schizophrenia in Sweden. *Mol. Psychiatry*, **19**, 762–773.

Szklarczyk,D. *et al.* (2019) STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.*, **47**, D607–D613.

Tabarés-Seisdedos,R. *et al.* (2009) Chromosome 8p as a potential hub for developmental neuropsychiatric disorders: implications for schizophrenia, autism and cancer. *Mol. Psychiatry*, **14**, 563–589.

Umeda-Yano,S. *et al.* (2013) The regulation of gene expression involved in TGF-β signaling by ZNF804A, a risk gene for schizophrenia. *Schizophr. Res.*, **146**, 273–278.