Resource Report

Cancer Research

CrosstalkNet: A Visualization Tool for Differential Co-expression Networks and Communities

Venkata Manem[1,2], George Alexandru Adam[1,2], Tina Gruosso[3], Mathieu Gigoux[3], Nicholas Bertos[3], Morag Park[3], and Benjamin Haibe-Kains[1,2,4,5]

Abstract

Variations in physiological conditions can rewire molecular interactions between biological compartments, which can yield novel insights into gain or loss of interactions specific to perturbations of interest. Networks are a promising tool to elucidate intercellular interactions, yet exploration of these large-scale networks remains a challenge due to their high dimensionality. To retrieve and mine interactions, we developed CrosstalkNet, a user friendly, web-based network visualization tool that provides a statistical framework to infer condition-specific interactions coupled with a community detection algorithm for bipartite graphs to identify significantly dense subnetworks. As a case study, we used CrosstalkNet to mine a set of 54 and 22 gene-expression profiles from breast tumor and normal samples, respectively, with epithelial and stromal compartments extracted via laser microdissection. We show how CrosstalkNet can be used to explore large-scale co-expression networks and to obtain insights into the biological processes that govern cross-talk between different tumor compartments.

Significance: This web application enables researchers to mine complex networks and to decipher novel biological processes in tumor epithelial-stroma cross-talk as well as in other studies of intercompartmental interactions. *Cancer Res; 78(8); 2140–3. ©2018 AACR.*

Introduction

Co-expression networks are a representation of putative gene–gene relationships, built by investigating the similarity of transcriptomic profiles across samples. During the last decade, co-expression networks have grown in popularity as they enable the integration of large-scale transcriptomic datasets, allowing identification and exploration of multiple genes with similar expression patterns across conditions (1–6). Given the relevance of these co-expression networks in cancer research, we developed *CrosstalkNet*, a web-based tool that enable users to explore and mine co-expression networks. We present here an application of *CrosstalkNet* in the context of tumor epithelial-stromal cross-talk, although the tool is widely applicable to the analysis of bipartite networks in general.

One of the most important signaling networks in tumor biology is the communication between tumor epithelial (epi) cells and neighboring stromal cells (7), which is mediated through regulatory loops between compartments. A promising approach to shed light on such a cellular cross-talk lies in using the transcriptional profiles of epithelial and stromal sample pairs obtained from laser capture micro-dissection (Supplementary Fig. S1) to construct co-expression networks (3). Given the computational and biological complexity of the co-expression network, it is challenging to explore and comprehend these large-scale networks. Epi–stromal co-expression networks can be represented as a bipartite network where each node represents a gene and each edge denotes a co-expression relationship between the epithelium and stroma (Fig. 1A). Differential networks have recently been used to characterize the rewiring of condition-specific interactions (8). We present here the first application of this methodology in the context of epi–stroma cross-talk to uncover the differential interactions gained and lost by tumor (Supplementary Fig. S2).

Most common visualization tool is Cytoscape, but, BiLayout is the only layout in Cytoscape that allows for visualizing bipartite networks. However, it requires manually selecting the nodes of interest as well as the neighbors to be plotted. In addition, it is difficult to visualize second neighbors. *CrosstalkNet* overcome the limitations the existing web-based visualization tools (Supplementary Information) to better explore large-scale bipartite networks (see Supplementary Video S1). We believe that this web-based application will assist researchers across disciplines in mining bipartite biological co-expression networks, as exemplified by our study of the epithelial–stromal cross-talk that occurs in breast tumors.

Network Inference

The datasets used in this study are presented in the Supplementary Table S1. The genome-wide tumor epi–stroma network was constructed by estimating pairwise co-expression using Pearson's correlation coefficient between the tumor epithelium and tumor stroma, whereas the normal network was constructed using the same approach based on normal epithelium and normal stroma. Each node represents a gene and co-expressions between $gene_i$ in the epi cells with $gene_j$, $gene_k$ in the stromal cells are represented by edges, as seen in Fig. 1A and Supplementary

[1]Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada. [2]Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada. [3]Goodman Cancer Research Centre, McGill University, Montreal, Canada. [4]Department of Computer Science, University of Toronto, Toronto, Ontario, Canada. [5]Ontario Institute of Cancer Research, Toronto, Ontario, Canada.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

V. Manem and G.A. Adam are co-first authors for the article.

Corresponding Author: Benjamin Haibe-Kains, Princess Margaret Cancer Centre, University Health Network, 101 College Street, Toronto, Ontario, Canada M5G 1L7. Phone/Fax: 416-581-8626; E-mail: bhaibeka@uhnresearch.ca

doi: 10.1158/0008-5472.CAN-17-1383

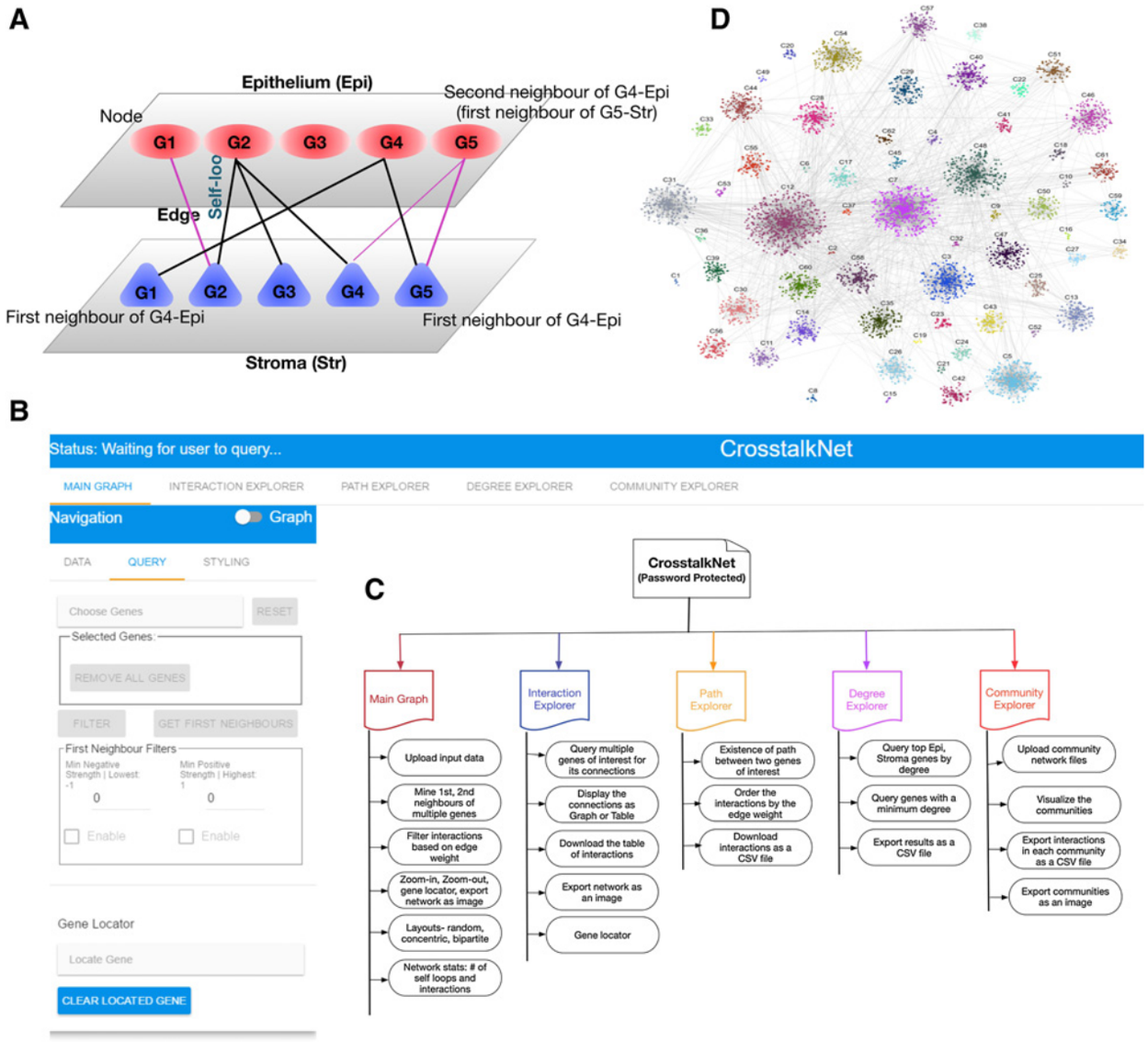

Figure 1.

CrosstalkNet web-application. Co-expression of genes in tumor epithelial and stromal cells. **A,** Genome wide epi–stroma interaction network. Red and blue nodes represent epi and stroma genes, respectively. The color gradient of the edges denotes the strength of the co-expression relationship. Black and magenta denote positive and negative co-expression relationships between the genes, respectively. **B,** *CrosstalkNet* is divided into five major panels, namely, Main Graph, Interaction Explorer, Path Explorer, Degree Explorer, Community Explorer. **C,** Flow chart representing the functionalities of each panel. **D,** Communities represented in different colors in a large-scale interaction network.

Fig. S3. This approach is extended by statistically comparing the interactions inferred from tumor and matched normal samples (8) and are applied here for the first time to infer epi–stroma networks (Supplementary Information).

Software Architecture

CrosstalkNet uses a node.js server that interacts with Rscript to execute user queries. Rscripts are responsible for interacting with the bipartite network files and determining neighbors and paths in those networks. Communication between node.js and Rscript happens via JavaScript Object Notation (JSON). Once the network files have been accessed, the results are sent back to the user and displayed via cytoscape.js, which is embedded in an AngularJS and Angular Material front-end.

Features

CrosstalkNet is divided into five major panels, namely, Main Graph, Interaction Explorer, Path Explorer, Degree Explorer, Community Explorer (Fig. 1B). Each panel has its own modular workflow, and is shown in Fig. 1C. The functionalities of the panels are described below in detail. Case studies will be further presented to exemplify the use of each of the panels in the Supplementary Material.

Upload. Users have the ability to upload their own adjacency matrix files to the server. Uploaded files must have an Rdata extension and must contain a single dgCMatrix (sparse matrix) from the *Matrix* package saved using the *saveRDS* function. Furthermore, row names and column names have to match, and the file must not contain missing (NA) values. In addition, row and column names and either be gene symbols or entrez gene identifiers. Uploaded files will be private to the user that uploaded them and cannot be accessed by other users, and can be deleted at anytime.

Main graph. The Main Graph is used to view the first and second neighbors of multiple genes of interest at the same time (See Supplementary Video S2). To prevent graphs from getting unmanageably large in size, a filtering option is provided in order to allow filtering of both first and second neighbors individually. This makes it possible to focus on the more relevant interactions. There is a table view showing the interactions between the genes in the corresponding graph, and the table is downloadable as a comma-separated value (CSV) file.

Interaction explorer. The interaction explorer enables this functionality by providing the ability to view arbitrarily many levels of neighbors (See Supplementary Video S3).

Path explorer. The path explorer makes it possible to obtain all paths with a maximum of one hop between two genes of interest (See Supplementary Video S4).

Degree explorer. The degree explorer displays either the top genes based on their degree, or by showing genes having a degree greater or equal to a specified number, or using both the options (See Supplementary Video S5). It is capable of showing an arbitrarily large list of genes, which can be downloaded in a simple CSV format.

Community explorer. Communities of interactors hold the potential to reveal functional processes of a biological system. The community explorer functionality allows the user to upload communities of the co-expression network using their favorite community detection algorithms,visualize them as seen in Fig. 1D, and can be downloaded in a CSV format. See Supplementary Video S6 for a detailed tutorial.

Availability of data and material

Documentation along with a video tutorial for the web-application are provided within the tool (http://crosstalknet.pmgenomics.ca/app/#/documentation). The examples used in the web application, namely, normal, tumor and differential epi-stroma networks can be downloaded directly from the tool. These examples are provided to guest users as well as registered users. The methodology used to construct differential network as well as the web-based tool can be used to mine networks based on pan cancer transcriptomic profiles. The code is open-source and freely available from GitHub (github.com/bhklab/CrosstalkNet/).

Results

In this study, we show how the *CrosstalkNet* web-application can be used to uncover the epi–stroma co-dependencies in ER^+ breast cancer patients. The details of the network topology are provided in the Supplementary Text. The degree distributions of epi, stroma genes in the tumor network is presented in Supplementary Fig. S4A and S4B, respectively. The number of interactions with respect to their significance (*P* value) is provided in Supplementary Fig. S5. We used a false discovery rate (FDR) of 15% to select the set of significant interactions. Concordant with our previous work (3), we found that the proportion of self-loops to be significantly higher in tumor compared to normal (Supplementary Fig. S6). *CrosstalkNet* enables (i) efficient exploration of a shared neighborhood (first neighbors) of genes of interest (Supplementary Fig. S7A); (ii) exploration of paracrine interactions (Supplementary Fig. S7B); (iii) identification of paths between any functionally similar genes (Supplementary Table S2); and (iv) identification of hubs as potential drug targets (Supplementary Table S3). We show hereafter how *CrosstalkNet* enables efficient exploration of gene interactions through two case studies describing how to mine a differential network, and extract the community structure from the tumor network.

Identification of gain/loss of interactions in the differential network

To identify the epi–stroma interactions that are gained or lost in ER^+ breast tumors, we used the *Delta* option in the main graph panel to explore a differential network. With the tumor-specific epi–stroma (differential) network selected, we explored the interactions involving *S100A7*, a protein in the S100 family known to be a key contributor to the onset of aggressive and invasive tumors with the help of the tumor stroma (9). We identified a large change in the co-expression of *S100A7*-epi and *S100A7*-stroma between normal and tumor networks ($\Delta = 0.99$; Supplementary Fig. S7C), indicating that there is a strong gain of interaction in the tumor. This supports the fact that *S100A7* is one of the key genes that is differentially co-expressed between normal epi-stroma and tumor epi–stroma, and contributes to tumorigenesis (9).

Identification of communities of interactors

Biological networks are known to have specific topology and structure, consisting of nodes that interact with each other and govern specific cellular processes. These groups of nodes are termed a community in the biological network. We used the CONDOR method (10) to obtain communities in the tumor epi–stroma interaction network, resulting in 62 communities, which can be displayed using the Community Explorer (Supplementary Fig. S8A). Each community consists of both epi–genes and stroma–genes (Supplementary Fig. S8B). We then performed a pathway analysis on all the communities that have ≥ 5 using all the genes with Gene Ontology terms (11). A total of 705 pathways are overrepresented with an FDR < 5% (Supplementary Fig. S8C). Broadly, the enriched GO terms belong to cell cycle, DNA synthesis and integrity, metabolism, etc., which are known to regulate tumor progression, resistance in ER^+ breast cancer patients (12–14).

Conclusions

The availability of high-throughput data along with the advancement of web-based technologies has led to the development of network visualization tools to mine large-scale interaction graphs. We developed *CrosstalkNet*, a user-friendly web-application that can be used to explore and mine interactions.

The application gives users the ability to upload their customized files and, to search and highlight a gene of interest in a dense network, along with multiple layout options. Users can also explore their inferred bipartite networks by focusing on a specific subgraph, filtering the network based on interaction strength, investigating the possible paths between two genes, identifying the hubs along with identification of communities in the network.

In the present study, we have described the use of *CrosstalkNet* with biologically relevant case studies reporting relevant epi–stroma interactions in ER^+ breast cancer. We showed how to mine interactions in differential co-expression networks and how to extract the community structure from the epi–stroma interaction networks. Overall, our case studies highlight the importance of differential networks and community structure in a biologically meaningful way. Moreover, these case studies (including the ones in Supplementary Information) reiterate the relevance and versatility of the *CrosstalkNet* web-application in cancer research.

In conclusion, we developed *CrosstalkNet*, a web-application enabling exploratory network analysis for researchers without requiring computational skills. *CrosstalkNet* will assist researchers across various disciplines along with the clinicians in mining complex networks to decipher novel biological processes in the tumor epithelial–stroma cross-talk, as well as in other studies of intercompartmental interactions.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Gary Bader and Max Franz for their help in using the cytoscape.js framework that has been used to develop the *CrosstalkNet* web-application. V.S.K. Manem was supported by the US National Institutes of Health (subaward #236136) and the Cancer Research Society (grant #21363). G.A. Adam was supported by the Canadian Institutes of Health Research. B. Haibe-Kains was supported by the Gattuso Slaight Personalized Cancer Medicine Fund at Princess Margaret Cancer Center, the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada, the Ministry of Economic Development and Innovation/Ministry of Research & Innovation of Ontario, and a Stand Up To Cancer Canada—Canadian Cancer Society Breast Cancer Dream Team Research Funding, with supplementary support of the Ontario Institute for Cancer Research through funding provided by the Government of Ontario (Funding Award Number: SU2C-AACR-DT-18-15). Stand Up To Cancer Canada is a program of the Entertainment Industry Foundation Canada. Research funding is administered by the American Association for Cancer Research International—Canada, the Scientific Partner of SU2C Canada.

Received June 13, 2017; revised November 23, 2017; accepted February 13, 2018; published first February 19, 2018.

References

1. Yang Y, Han L, Yuan Y, Li J, Hei N, Liang H. Gene co-expression network analysis reveals common system-level properties of prognostic genes across cancer types. Nat Commun 2014;5:3231.
2. Wolfe CJ, Kohane IS, Butte AJ. Systematic survey reveals general applicability of "guilt-by-association" within gene co-expression networks. BMC Bioinformatics 2005;6:227.
3. Oh E-Y, Christensen SM, Ghanta S, Jeong JC, Bucur O, Glass B, et al. Extensive rewiring of epithelial-stromal co-expression networks in breast cancer. Genome Biol 2015;16:128.
4. Ivliev AE, 't Hoen PAC, Sergeeva MG. Co-expression network analysis identifies transcriptional modules related to proastrocytic differentiation and sprouty signaling in glioma. Cancer Res 2010;70:10060–70.
5. Ivliev AE, 't Hoen PAC, van Roon-Mom WMC, Peters DJM, Sergeeva MG. Exploring the transcriptome of ciliated cells using in silico dissection of human tissues. PLoS One 2012;7:e35618.
6. Crow M, Paul A, Ballouz S, Huang ZJ, Gillis J. Exploiting single-cell expression to characterize co-expression replicability. Genome Biol 2016;17:101.
7. Hanahan D, Douglas H, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell 2011;144:646–74.
8. Ideker T, Trey I, Krogan NJ. Differential network biology. Mol Syst Biol 2012;8. Available from: http://dx.doi.org/10.1038/msb.2011.99.
9. Emberley ED, Murphy LC, Watson PH. S100A7 and the progression of breast cancer. Breast Cancer Res 2004;6:153–9.
10. Platig J, Castaldi PJ, DeMeo D, Quackenbush J. Bipartite community structure of eQTLs. PLoS Comput Biol 2016;12:e1005033.
11. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unfication of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25–9.
12. Finn RS, Aleshin A, Slamon DJ. Targeting the cyclin-dependent kinases (CDK) 4/6 in estrogen receptor-positive breast cancers. Breast Cancer Res 2016;18:17.
13. Timms KM, Abkevich V, Hughes E, Neff C, Reid J, Morris B, et al. Association of BRCA1/2 defects with genomic scores predictive of DNA damage repair deficiency among breast cancer subtypes. Breast Cancer Res 2014;16:475.
14. Monteith GR, Prevarskaya N, Roberts-Thomson SJ. The calcium–cancer signalling nexus. Nat Rev Cancer 2017;17:367–80.