Bioinformatics, 31(20), 2015, 3390–3391
doi: 10.1093/bioinformatics/btv385
Advance Access Publication Date: 24 June 2015
Applications Note

Systems biology

Hyperscape: visualization for complex biological networks

Graham L. Cromar[1,*], Anthony Zhao[1], Alex Yang[1,2] and John Parkinson[1,2,3,*]

[1]Program in Molecular Structure and Function, Hospital for Sick Children, 21-9830 PGCRL, 686 Bay Street, Toronto, ON, M5G 0A4, Canada, [2]Department of Biochemistry and [3]Department of Molecular Genetics, University of Toronto, Toronto, ON, M5S 1A8, Canada

*To whom correspondence should be addressed.

Associate Editor: Jonathan Wren

Received on April 2, 2015; revised on May 29, 2015; accepted on June 18, 2015

Abstract

Motivation: Network biology has emerged as a powerful tool to uncover the organizational properties of living systems through the application of graph theoretic approaches. However, due to limitations in underlying data models and visualization software, knowledge relating to large molecular assemblies and biologically active fragments is poorly represented.

Results: Here, we demonstrate a novel hypergraph implementation that better captures hierarchical structures, using components of elastic fibers and chromatin modification as models. These reveal unprecedented views of the biology of these systems, demonstrating the unique capacity of hypergraphs to resolve overlaps and uncover new insights into the subfunctionalization of variant complexes.

Availability and implementation: Hyperscape is available as a web application at http://www.compsysbio.org/hyperscape. Source code, examples and a tutorial are freely available under a GNU license.

Contacts: john.parkinson@utoronto.ca or graham.cromar@gmail.com

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

1 Introduction

Biological systems captured by protein–protein interaction (PPI) networks are typically represented as graphs in which proteins are projected as nodes connected by edges. Other graphs such as bi-partite, compound and metagraphs have been used to visualize more complex relationships (Hu *et al.*, 2007). However, overlapping complexes have not been well supported because (i) the tracking of interaction data at the level of biological assemblages is relatively new, for example although the BioPAX pathway language supports the definition of such assemblages, only MatrixDB (Chautard *et al.*, 2009, 2010) appears to track interactions across hierarchies; and (ii) existing visualization tools, such as Cytoscape (Shannon *et al.*, 2003) and VisANT (Hu *et al.*, 2007), created to support simple graphs, have design considerations preventing the display of overlapping clusters.

To overcome these limitations, hypergraphs have been proposed as viable alternative to more meaningfully represent biological relationships (Klamt *et al.*, 2009; Ritz *et al.*, 2014). A hypergraph is a set V of vertices and a set of non-empty subsets of V, called hyperedges. Despite their ability to better represent the complexity inherent in many biological datasets, tools enabling the practical implementation of hypergraphs on a wide scale have been lacking.

Here, we provide software to translate and render biomolecule descriptions as hypergraphs and we demonstrate the usefulness of hypergraph visualization in interpreting the organization of elastic fibers, major constituents of large blood vessels and resolving overlaps in a network of chromatin modification.

2 Implementation

Using an existing file format previously defined for the Cytoscape plugin, CyoOG (Royer *et al.*, 2008) we wrote a custom Perl script to map biomolecule definitions and interactions into a succinct

description of the graphic elements to be rendered (Supplementary Methods). The BBL file format elegantly describes the necessary graphical elements and relationships for hypergraphs in the form of NODES, SETS (groups of nodes), IN relationships (describing hierarchical relationships) and their EDGEs. We used the D3 javascript library (http://www.d3js.org) to develop 'Hyperscape' as a proof-of-concept, browser-based renderer. Because of our layering approach, nodes which are shared between multiple complexes do not need to be duplicated as they do in compound graph implementations. In addition to circumventing such legacy constraints, Hyperscape is able to recognize stoichiometric information in the form of multiple IN relationships and correctly represent these by adding additional instances of nodes. The features of this web-based renderer can be explored using the data and tutorial provided (in Supplementary Data File S1).

3 Results and discussion

To illustrate the differences in information that can be extracted from hypergraph, compound graph and bi-partite graph representations of the same underlying data; we used a subnetwork consisting of elastic fiber components and their neighbors (Supplementary Methods). A bi-partite graph representation of this network (Supplementary Fig. S1) contains numerous redundant edges as well as a lack of information relating components, their members and their stoichiometry (see also: Supplementary Results/Discussion). The corresponding compound graph (Supplementary Fig. S2) addresses some of these issues by grouping components into metanodes that encircle components. Nesting these created further difficulties, however, as edges between hierarchical levels were not supported.

Further, protein fragments were inappropriately represented. COL18A1, for example, has no relevant elastic fiber interactions except through its cleavage product, endostatin, a potent inhibitor of endothelial cell proliferation and angiogenesis (Sertie *et al.*, 2000). Overlap of components such as COL4A5 which appears here as two variants of collagen IV (Supplementary Fig. S3), v2 (composed of COL4A3, COL4A4 and COL4A5) and v3 (composed of COL4A5 and COL4A6), required that duplicate nodes be created to represent each instance of the protein.

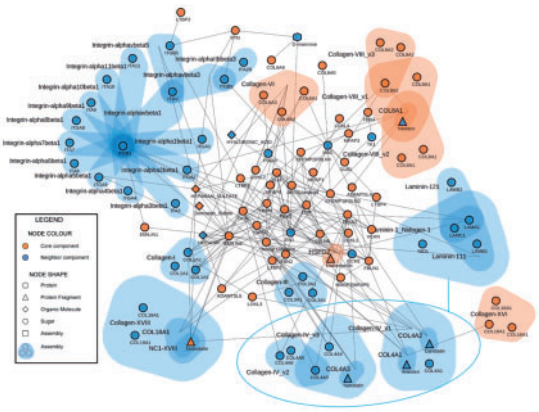

Fig. 1. The elastic fiber interactome and selected neighbors. Note overlapping assemblies as well as an additional hyperedge (blue oval) depicting the interaction between collagen IV (various assemblies) and the laminin-1:nidogen-1 complex. Simple edges represent experimentally determined, physical interactions (Supplementary Methods). Neighbors were selected as needed to connect all elastic fiber components (EFCs) into a single, connected subgraph and expanded to include all additional components of multimers associated with an EFC or nearest neighbor. A complete set of neighbors are included in additional online material (Supplementary Figs S3–S5)

Our hypergraph (Fig. 1 and Supplementary Fig. S4) represents a more intuitive approach that highlights here, several indications of differential wiring. For example, collagen VIII v3 does not contain COL8A1 and hence, unlike collagen VIII v1 and v2, does not interact with ADAMTSL4, whereas collagen IV v1 interacts with LOXL3 and carries a distinct payload of statins—canstatin and arresten as opposed to tumstatin for collagen IV v3. This suggests functionally relevant subfunctionalization of variant complexes via differential inclusion of paralogous components. Together with endorepellin (derived from HSPG2) and vastatin (derived from COL8A1), these fragments suppress pathogenic angiogenesis but act across a variety of target pathways affecting endothelial cell growth and apoptosis (Pupa *et al.*, 2002; Sudhakar *et al.*, 2003).

As a further example, we applied Hyperscape to a network of chromatin modification (CM) as described in Supplementary (Supplementary Fig. S6 and Supplementary Results/Discussion).

4 Conclusion

With increasing interest in applying hypergraphs to study protein complexes, metabolic pathways, regulatory and signaling networks (Hu *et al.*, 2007; Klamt *et al.*, 2009; Ritz *et al.*, 2014) we believe Hyperscape will find wide application in systems biology.

Acknowledgements

The authors gratefully acknowledge the assistance of S. Ricard-Blum (UMR 5086 CNRS—Université Lyon 1, 69367 Lyon Cedex 07, France) for providing updated biomolecule definitions and comments.

Funding

This work was supported by a grant in aid from the Heart and Stroke Foundation of Canada.

Conflict of Interest: none declared.

References

Chautard,E. *et al.* (2009) MatrixDB, a database focused on extracellular protein-protein and protein-carbohydrate interactions. *Bioinformatics*, **25**, 690–691.

Chautard,E. *et al.* (2010) MatrixDB, the extracellular matrix interaction database. *Nucleic Acids Res.*, **39**, D235–D240.

Hu,Z. *et al.* (2007) Towards zoomable multidimensional maps of the cell. *Nat. Biotechnol.*, **25**, 547–554.

Klamt,S. *et al.* (2009) Hypergraphs and cellular networks. *PLoS Comput. Biol.*, **5**, e1000385.

Pupa,S.M. *et al.* (2002) New insights into the role of extracellular matrix during tumor onset and progression. *J. Cell. Physiol.*, **192**, 259–267.

Ritz,A. *et al.* (2014) Signaling hypergraphs. *Trends Biotechnol.*, **32**, 356–362.

Royer,L. *et al.* (2008) Unraveling protein networks with power graph analysis. *PLoS Comput. Biol.*, **4**, e1000108.

Sertie,A.L. *et al.* (2000) Collagen XVIII, containing an endogenous inhibitor of angiogenesis and tumor growth, plays a critical role in the maintenance of retinal structure and in neural tube closure (Knobloch syndrome). *Hum. Mol. Genet.*, **9**, 2051–2058.

Shannon,P. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.

Sudhakar,A. *et al.* (2003) Human tumstatin and human endostatin exhibit distinct antiangiogenic activities mediated by alpha v beta 3 and alpha 5 beta 1 integrins. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 4766–4771.