Opinion

CellPress

Signaling hypergraphs

Anna Ritz[1*], Allison N. Tegge[1*], Hyunju Kim[1,†], Christopher L. Poirel[1,‡], and T.M. Murali[1,2]

[1] Department of Computer Science, Virginia Tech, Blacksburg, VA, USA
[2] ICTAS Center for Systems Biology of Engineered Tissues, Virginia Tech, Blacksburg, VA, USA

Signaling pathways function as the information-passing mechanisms of cells. A number of databases with extensive manual curation represent the current knowledge base for signaling pathways. These databases motivate the development of computational approaches for prediction and analysis. Such methods require an accurate and computable representation of signaling pathways. Pathways are often described as sets of proteins or as pairwise interactions between proteins. However, many signaling mechanisms cannot be described using these representations. In this opinion, we highlight a representation of signaling pathways that is underutilized: the hypergraph. We demonstrate the usefulness of hypergraphs in this context and discuss challenges and opportunities for the scientific community.

Signaling pathways and their representations

Signaling pathways mediate the responses of a cell to its environment, starting with recognition of an external stimulus at receptors, proceeding through intracellular protein interactions and activation of transcription factors, and culminating in perturbation of the expression of target genes. Owing to their importance in cellular communication, signaling pathways are often perturbed in diseases. Numerous publicly available and often manually curated databases store information about signaling pathways [1–6]. Despite growing knowledge of signaling pathways gained from experimental data, these databases face a number of obstacles for storing and conveying this information. Databases representing signaling pathways from manual curation of the literature produce high-quality interactions, but are time-consuming to construct, are often incomplete or outdated, and might be biased according to the curators' expertise [7–10]. Databases that use automated methods for literature searches, such as predictive text mining, are relatively easy to maintain but tend to have many erroneous entries [11]. Different databases may represent the same biological event in different ways, making them difficult to standardize for computational use.

In this opinion, we describe the common representations that have been used in computational analyses of signaling pathways. After examining the limitations of these representations, we encourage the use of hypergraphs as models that better capture the complex relationships in underlying biological mechanisms. We describe three applications to motivate more powerful representations of signaling pathways. Pathway enrichment assesses whether discovered proteins are significantly enriched for proteins/interactions in a pathway of interest. Pathway reconstruction explicitly reconstructs and discovers missing proteins and interactions in a pathway of interest. Finally, pathway crosstalk captures how stimulation of one pathway may result in alternative downstream responses.

Current representations of signaling pathways

Signaling pathways as sets of proteins

The simplest representation of a pathway is a list of its members, that is, the set of proteins involved in the pathway (Figure 1). Catalogs such as the Gene Ontology [1] and the Molecular Signatures Database [12] provide signaling pathways in this format. For this representation, pathway enrichment identifies pathways whose members occur surprisingly often in a set of experimentally identified proteins (e.g., from analysis of differential gene expression) [13]. However, such set-based approaches ignore the relationships between proteins within a pathway, and thus provide no clues as to how interactions may alter gene expression [8]. These methods can correct and adjust for proteins shared among multiple pathways [14,15], and thus account for crosstalk to some extent. By definition, purely set-based methods can reconstruct only the proteins in a pathway and not the interactions among them [16].

Signaling pathways as directed graphs

Signaling pathways are also conceptualized as graphs in which nodes represent proteins and edges represent

Glossary

Node: an element (protein or compound).
Undirected edge: an unordered pair of nodes (physical interaction between two proteins).
Directed edge: an ordered pair of nodes (kinase phosphorylates a substrate). In a directed graph, an undirected edge between nodes u and v is replaced by two directed edges (u,v) and (v,u).
Hypernode: a set of node(s) (protein; protein complex).
Directed hyperedge: an ordered pair of sets of hypernodes (complex assembly).
Regulated hyperedge: a directed hyperedge regulated by a hypernode (kinase phosphorylates a protein complex, thereby activating it).

Corresponding author: Murali, T.M. (murali@cs.vt.edu).
Keywords: signaling pathways; graphs; hypergraphs.
[*] These authors contributed equally to this work.
[†] Current address: BEYOND: Center for Fundamental Concepts in Science, Arizona State University, Tempe, AZ, USA.
[‡] Current address: RedOwl Analytics, Baltimore, MD, USA.

0167-7799/
 http://dx.doi.org/10.1016/j.tibtech.2014.04.007

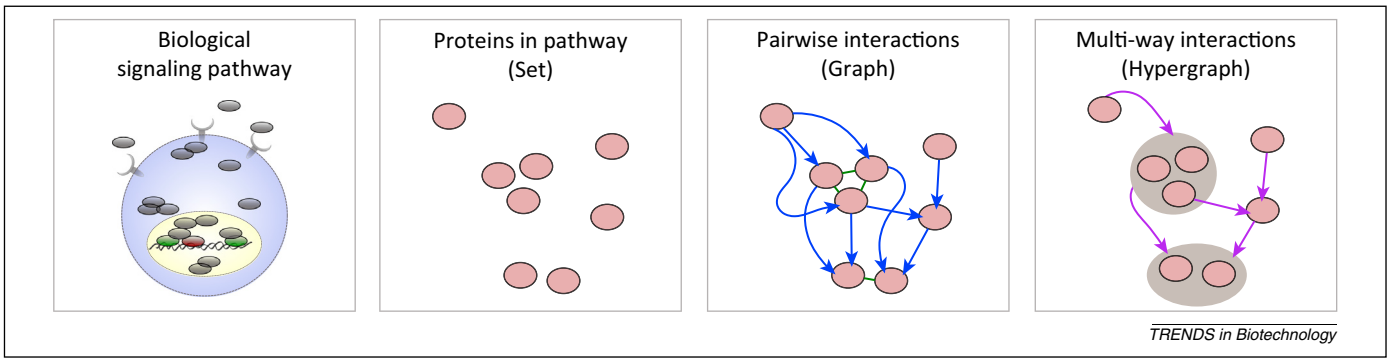

Figure 1. Signaling pathway representations. There are three main ways of representing signaling pathways. A signaling pathway may be simply represented as a set of proteins, with no additional information. Graphs encode pairwise interactions between proteins; these interactions may be undirected (green) or directed (blue). Hypergraphs, the focus of this article, encode multi-way interactions and reactions. Box 1 provides examples of graph and hypergraph representations of reactions in signaling pathways.

pairwise interactions between proteins (Figure 1 and see Glossary). The edges are often directed in signaling pathways, such as when a kinase phosphorylates a substrate. Recent enrichment methods make use of pathway topology in their scoring metrics by taking the interactions among member proteins into account [8]. There is ongoing development of this class of approaches. Pathway reconstruction algorithms for graphs typically use a large background interactome (such as a protein–protein interaction network) and identify pathways as subgraphs of proteins and interactions within the interactome. These approaches often try to find connections between signaling initiators (membrane receptors) and downstream regulators (transcription factors). Pathway reconstruction algorithms use many well-known concepts from graph theory [17–21]. Graph-based approaches to assess pathway crosstalk rely on the notion that two crosstalking pathways (each represented as a set of genes) will have statistically more interactions connecting their members than expected in a random network [22,23]. However, these approaches fail to compute the specific paths of signaling interactions that contribute to such crosstalk.

Graph representations of signaling pathways are an improvement from the 'set of proteins' representation because they capture pairwise relationships between proteins. However, signaling pathways contain more complicated relationships that are problematic for graph representations. For example, graphs often represent a complex by connecting all its members, which can artificially increase the number of edges (Figure IA in Box 1) [24]. More importantly, graphs do not accurately represent several types of molecular reactions, including regulation (e.g., activation and inhibition) or protein complex assembly and disassembly (Figure IB in Box 1). Finally, graphs do not typically distinguish between inactive and active forms of a protein or complex (Figure IC in Box 1).

Other representations of signaling pathways

Although directed graphs have been useful for representing signaling pathways, their limitations are widely recognized. A number of approaches have modified and extended graph representations. Compound graphs [25] and metagraphs [26] represent a complex as a single entity and allow a nested structure among complexes. Factor graphs [27] and Petri nets [28] introduce different types of nodes into a directed graph to represent events involving sets of proteins. Multimodal networks associate four entities with each edge: a head, a tail, a regulator, and a mode [29]. The head, tail, and regulator can each be a set of proteins, and the mode specifies how the regulator controls the transition from head to tail, for example by activation or repression.

These models of signaling pathways seek to address the shortcomings of directed graphs. However, each approach has drawbacks, including an inability to comprehensively model the complexity of signaling pathways, applicability to a limited range of computational problems, and underutilization in systems biology. Nodes in compound graphs and metagraphs focus on protein complexes. Multimodal networks do not support the hierarchical structure of signaling networks. Factor graphs and Petri nets are not ideal for generalizations of common graph-theoretic operations such as paths, connectivity, and random walks. In the next section, we seek to unify these models under the umbrella of signaling hypergraphs.

Signaling pathways as hypergraphs

Hypergraphs are a generalization of graphs that are capable of representing relationships among two or more proteins (Figure 1). Typically, directed hypergraphs consist of a set of nodes and a set of directed hyperedges in which each hyperedge connects two sets of nodes. Directed hypergraphs are an attractive alternative to directed graphs for representing complex facets of cellular processes, especially for metabolic networks [29–33]. They are also advantageous for signaling networks [30,34,35]; however, they remain an underutilized tool.

In our definition, a signaling hypergraph consists of hypernodes, directed hyperedges, and regulated hyperedges. Each hypernode represents an individual protein or a set of proteins, each directed hyperedge connects one set of hypernodes to another, and each regulated hyperedge is a directed hyperedge with one or more hypernodes that act as regulators. Box 1 illustrates this definition of signaling hypergraphs using three biological events in the canonical Wnt signaling pathway. These biological events (protein complexes, assembly of protein complexes, and regulation of proteins and complexes) commonly occur in signaling pathways. Each event may be represented as a graph consisting of multiple edges (Figure I left) or as a

Box 1. Comparison of representations for events that occur in signaling pathways

1. **Protein complexes** (Figure IA). A complex is a set of proteins that bind together to carry out a biological function. On stimulation of FZD5 by the WNT3A ligand, a four-protein complex (FZD5–WNT3A–LRP6–DVL) assembles at the plasma membrane. In a graph, complexes are commonly represented as cliques; therefore, it is unclear whether the proteins form a complex or if they interact in pairs. A signaling hypergraph represents the complex as a single hypernode.
2. **Complex assembly** (Figure IB). Many biochemical reactions involve complex assembly and disassembly. For example, the four-protein complex FZD5–WNT3A–LRP6–DVL sequesters GSK3 and Axin1, thereby disassembling the destruction complex GSK3–Axin1–APC. A graph representation connects every pair of proteins within each complex; in addition to the ambiguity arising from complexes, the exact complex reconstitution is unclear. In a signaling hypergraph, complex rearrangement can be represented as a single directed hyperedge. Unlike the graph representation, the hyperedge representation clarifies the reactants, the products, and the 'direction' of the reaction.
3. **Regulation** (Figure IC). Many cellular reactions are regulated by proteins, small molecules, or complexes. For instance, phosphorylation (and inactivation) of β-catenin is regulated by the destruction complex GSK3–Axin1–APC. A graph can represent this regulation using directed edges from every protein in the regulator (in this case, a complex) to β-catenin, creating the misleading appearance that each protein in the complex can independently regulate β-catenin. Moreover, graphs do not typically distinguish between inactive and active forms of a protein. Instead, hypergraphs can represent such a reaction as a regulated hyperedge, which indicates that the complex is required for inactivation of β-catenin.

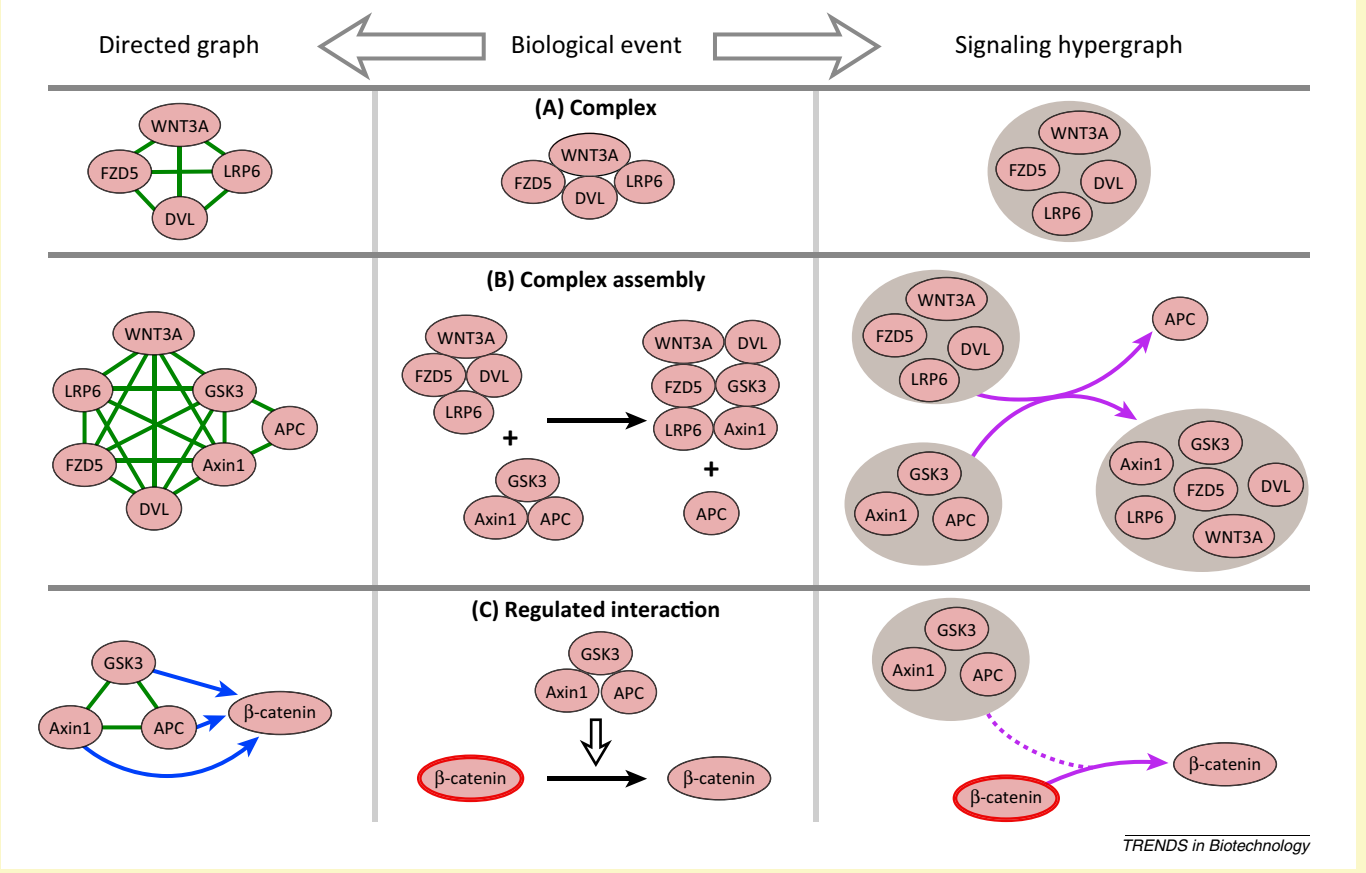

Figure I. Representations of three types of events in signaling pathways. In each panel, the biological representation (middle) can be converted into a graph (left) or a hypergraph (right). Green, undirected edge; blue, directed edge; gray circle, hypernode; purple, hyperedge; red outline, active form; broken purple, regulatory component of a hyperedge.

single hyperedge in a signaling hypergraph (Figure I right). Thus, each of these biological events has a more direct interpretation when represented as a signaling hypergraph rather than as a directed graph.

When these biological events are considered as components of the larger Wnt signaling pathway, it becomes clear that a typical graph-based representation does not have sufficient power to accurately model sequences of signaling events (Figure 2). The first five hyperedges in Figure 2 represent the canonical Wnt signaling pathway (we discuss hyperedges 6 and 7 in the context of pathway crosstalk in the next section). The primary function of Wnt signaling is to control the activity status of β-catenin (hyperedge 1); active β-catenin regulates the transcription of target genes, and inactive β-catenin is degraded by the proteasome. Maintenance of active β-catenin starts with binding of WNT3A to the FZD5–LRP6 complex (hyperedge 2). Next, DVL binds to the FZD5–LRP6–WNT3A complex to form FZD5–LRP6–WNT3A–DVL (hyperedge 3), which the small molecule PIP subsequently activates (hyperedge 4). Activated FZD5–LRP6–WNT3A–DVL interferes with the destruction complex Axin1–APC–GSK3 by sequestering Axin1 and GSK3, and releasing APC (hyperedge 5). However, in the absence of Wnt signaling, the activated destruction complex persists and marks β-catenin for degradation via phosphorylation (hyperedge 1).

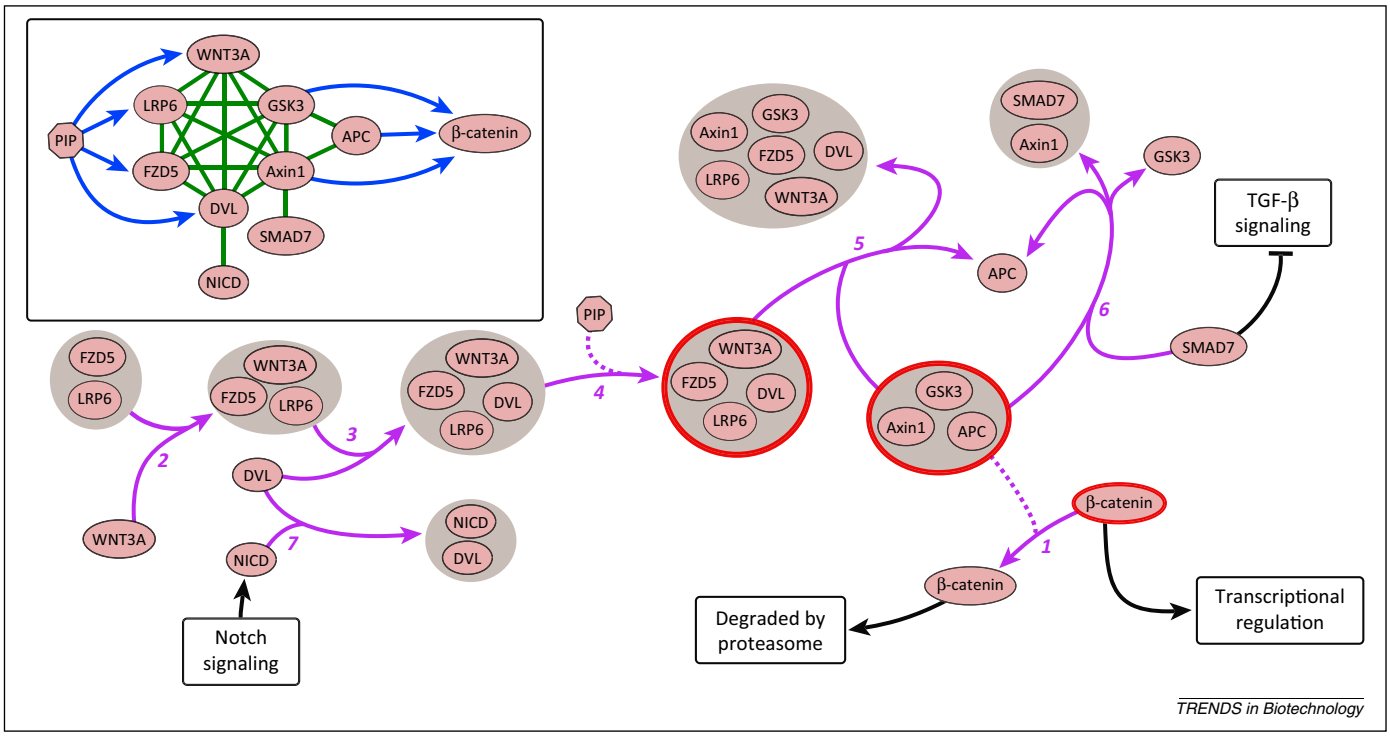

Figure 2. Events following Wnt signaling that lead to release of β-catenin. Black edges denote connections to other signaling pathways and biological processes. The inset shows a possible graph representation of these events. Green, undirected edge; blue, directed edge; purple, hyperedge; broken purple, regulated component of a hyperedge; gray circle, hypernode; red outline, active form.

These examples indicate that signaling hypergraphs offer a more accurate representation than directed graphs of the underlying biological events in signaling pathways. Moving to a hypergraph-based representation brings several computational challenges and opportunities. From what data sets can we build signaling hypergraphs? What algorithmic questions arise and need to be solved for signaling hypergraphs? How exactly can problems in computational systems biology such as pathway enrichment, extension, and crosstalk benefit from this representation? We now turn our attention to these questions.

Building signaling hypergraphs

Although hypergraphs are not widely utilized for analysis of signaling pathways, the increasing popularity of standardized data exchange formats such as BioPAX [36] and SBML [37] could accelerate their adoption. These file formats explicitly support reaction networks [2,36,37], which are in essence hypergraphs. As a result, we can directly build signaling hypergraphs from these representations. In addition, algorithms that operate on signaling hypergraphs can take advantage of the rich information embedded in BioPAX and similar formats.

BioPAX is a format that aims to enable integration and exchange of reactions and biological pathways. It has facilities for representing the diversity of interactions within a signaling pathway, including notions such as complex assembly, biochemical reactions, and control mechanisms. In essence, BioPAX represents each of these notions as a hypergraph, although the notion of a hypergraph is not explicit in BioPAX. Figure 3 shows a scheme for a generic reaction in BioPAX. For example, a complex is simply a set of proteins that we call a hypernode. Complexes in BioPAX can be nested within each other and can be members of reactions. The representation of a reaction such as complex assembly explicitly specifies the reactants and products, each of which can be a set of complexes, proteins, or small molecules, called a directed hyperedge in our nomenclature. BioPAX takes regulation into account as well: a reaction may have a controller, which can be a complex, protein, small molecule, or even another pathway (regulated hyperedge). Thus, it should not be difficult to convert BioPAX-like formats into

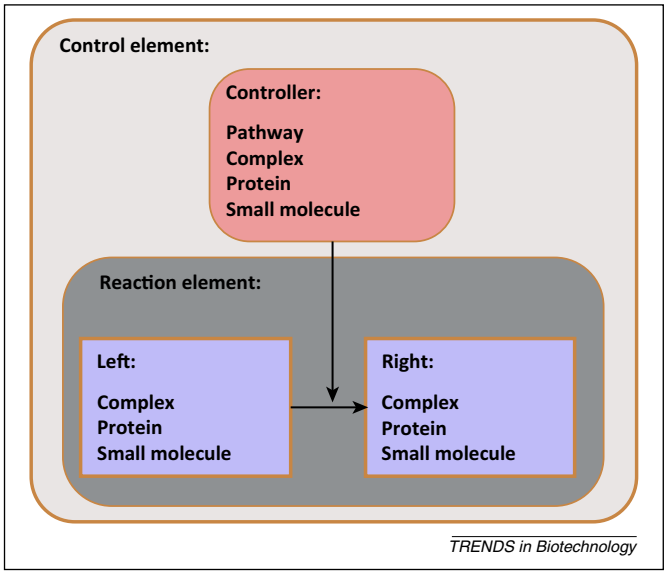

Figure 3. Example of the nested organization of BioPAX pathway elements. This shows the hierarchy of the BioPAX file format. Hypernodes may be complexes, proteins, or small molecules. Reaction elements in BioPAX are equivalent to directed hyperedges, and controlled reactions to regulated hyperedges.

signaling hypergraphs. In addition to representing various elements in a signaling pathway, BioPAX makes explicit the notion that all members of a complex must be present (and bound to each other) before participating in or regulating a reaction. Graphs only model pairwise representations, so they cannot represent such requirements conveniently.

Applications of signaling hypergraphs

Although hypergraphs have been an established area of mathematics since the 1960s [38], their application to systems biology may have been hampered by the lack of powerful algorithms. Graphs are a 'special case' of signaling hypergraphs (in which each hypernode contains exactly one node and each hyperedge contains exactly two nodes), so computational problems on hypergraphs are likely to be at least as difficult as the corresponding problems for graphs [30]. In fact, many problems in graphs that can be solved in polynomial time become computationally hard (*NP*-complete) when posed on hypergraphs, such as computing the shortest paths in directed hypergraphs [39]. Nevertheless, as we discuss below, hypergraphs are a promising representation for pathway enrichment, reconstruction, and crosstalk applications. We illustrate these applications with examples from the Wnt signaling pathway.

Pathway enrichment

For the signaling hypergraph in Figure 2, we consider a scenario in which only WNT3A and DVL are differentially expressed. Set-based methods may not identify the Wnt signaling pathway as significantly enriched because WNT3A and DVL are the only two proteins used to compute enrichment. A hypergraphical approach may reveal that WNT3A or DVL participate in seven hypernodes and five hyperedges, and may conclude that Wnt signaling is perturbed. If WNT3A and DVL were downregulated, then the method might conclude that Wnt signaling would not occur, the integrity of the destruction complex would be maintained, and degradation of β-catenin would ensue. Signaling hypergraphs impose an added layer of complexity because they can represent the formation of a complex, both the active and inactive forms of a complex, and the regulation of reactions. We expect that ideas borrowed from factor graphs [40] and pathway enrichment techniques that consider edge structure [8,41] will be useful for developing hypergraph-based approaches for pathway enrichment.

Pathway reconstruction

Existing methods have relied on well-known graph algorithms, but the corresponding theory for hypergraphs is considerably less well developed [42]. However, pathway reconstruction is an excellent application for the development of new mathematical, statistical, and algorithmic research on hypergraphs. Consider the hypergraph representation in Figure 2, in which the FZD5–LRP6 complex represents a signal initiator and the active form of β-catenin represents a transcriptional regulator. Hypergraph-based pathway reconstruction seeks to determine a series of hyperedges that links the FZD5–LRP6 complex with the active form of β-catenin, as shown by hyperedges 1–5 in Figure 2. By using the hyperedges and hypernodes as the units of connection in the pathway reconstruction problem, we can preserve the integrity of complexes and reactions, whereas graph-based methods struggle to discern these entities.

Pathway crosstalk

Crosstalk among pathways is frequently identified as activation of genes downstream of one pathway after a stimulus for another pathway. Existing methods for computational estimation of crosstalk use the intuition that it results from proteins shared by both pathways [15,43]. Hypergraph-based approaches have the potential to unveil connected sequences of interactions that contribute to crosstalk between pathways.

Figure 2 shows crosstalk between the Wnt signaling pathway and the TGF-β (hyperedge 6) and Notch (hyperedge 7) pathways. SMAD7, a negative regulator of the TGF-β signaling pathway, mediates the crosstalk with the Wnt pathway. Specifically, SMAD7 catalyzes dissociation of the destruction complex by binding to Axin1 and releasing APC and GSK3. This event results in increased transcriptional regulation by β-catenin and increased TGF-β signaling [44]. Crosstalk between the Notch and Wnt signaling pathways occurs via an interaction between NICD (NOTCH intracellular domain) and DVL [45]. Once DVL is bound to NICD, Wnt signaling cannot proceed, allowing the destruction complex to mark β-catenin for degradation. Thus, activation of Notch signaling downregulates Wnt signaling. Computational discovery of these sequences of events that lead to crosstalk is an outstanding problem in both graph and hypergraph representations.

Outstanding challenges

Several issues confront both graph- and hypergraph-based approaches for analysis of signaling pathways. Current interaction data sets remain highly incomplete and/or contain many false-positive interactions. The same pathway can have considerably different representations in different databases, sometimes resulting from the focus of the curator. Estimation of the reliability of interactions (from the type of experiment or the curation method) continues to be an important challenge. Although most signaling pathways in databases are organism-specific, interactomes established by combining multiple databases are not tissue-specific an issue that affects the analysis for multi-cellular organisms.

Existing graph-based approaches for pathway reconstruction often analyze the entire interactome (i.e., all pairs of proteins known to interact). The interactome is often constructed from a combination of signaling pathway databases, more general protein–protein interaction databases, and high-throughput experiments. Hence, the quality of solutions from pathway reconstruction algorithms depends on the reliability of this interactome [43]. To use signaling hypergraphs for pathway reconstruction, we must establish a corresponding hypergraph interactome that represents the interactions among multiple proteins. As a start, a hypergraph interactome can be constructed by starting from a graph built from protein–protein interaction data and

adding hyperedges as they are represented in the signaling pathway databases (e.g., using BioPAX). Construction of high-quality and high-coverage hypergraph interactomes is an important research direction because these will ultimately determine the quality of solutions from hypergraph-based algorithms.

For any representation, we can ask how the activity (or inactivity) of signaling pathway members affects the rest of the pathway. Boolean networks, for example, allow each node in a graph to take one of two values: 0 (inactive) or 1 (active). Logical models use Boolean gates to represent how a set of proteins may influence another protein [35,46]. Here, each gate is a special case of a directed hyperedge with many nodes in the tail (inputs) and one node in the head (output). Extension of logical models that are appropriate for signaling hypergraphs is an important research direction.

Signaling hypergraphs do not include stoichiometric or kinetic information, and focus more on capturing pathway structure. By contrast, dynamic models of signaling pathways capture both the structure and stoichiometry of signaling pathways [46,47]. However, these models tend to scale poorly because they require extensive prior experimental knowledge to fit their parameters. Extensions of logical models to signaling hypergraphs that incorporate stoichiometric information may provide a scalable alternative to dynamic models.

Concluding remarks

The study of signaling pathways is a cornerstone of modern molecular and cellular biology. Unfortunately, their representation in common signaling pathway databases varies widely in terms of completeness, quality, and standardization. Furthermore, methods that computationally analyze such pathways may often ignore important characteristics of their structure. These gaps necessitate the development of a fresh representational approach. We described the use of signaling hypergraphs, an underutilized representation in systems biology, to overcome these limitations.

Our goal is not to replace graph-based approaches in systems biology. Rather, our intent is to stimulate and encourage the use of hypergraph-based methods for contexts such as signaling pathways for which their unique capabilities could have a considerable impact. We hope that in the future, hypergraph-based analyses can encompass other types of cellular processes in conjunction with methods that infer hypergraphs directly from systems biology data sets [48–51].

We acknowledge that computational analysis of signaling hypergraphs will be challenging because hypergraph theory is much less well developed than the theory for graphs. Nonetheless, the development of pathway enrichment, pathway reconstruction, and pathway crosstalk algorithms for hypergraphs shows promise for better representation of biological signaling pathways. We hope that our advocacy for signaling hypergraphs will stimulate new directions of mathematical, statistical, and algorithmic research. Favorable results could promote increased use of signaling hypergraphs in computational systems biology.

Acknowledgments

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health grant R01-GM095955 (TMM), National Science Foundation (NSF) grant DBI-1062380 (TMM), and an NSF Graduate Research Fellowship (CLP).

References

1 Ashburner, M. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29
2 Kanehisa, M. *et al.* (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40, D109–D114
3 Matthews, L. *et al.* (2009) Reactome knowledgebase of human biological pathways and processes. *Nucleic Acids Res.* 37, D619–D622
4 Kandasamy, K. *et al.* (2010) NetPath: a public resource of curated signal transduction pathways. *Genome Biol.* 11, R3
5 Paz, A. *et al.* (2011) SPIKE: a database of highly curated human signaling pathways. *Nucleic Acids Res.* 39 (Suppl. 1), D793–D799
6 Breuer, K. *et al.* (2013) InnateDB: systems biology of innate immunity and beyond - recent updates and continuing curation. *Nucleic Acids Res.* 41, D1228–D1233
7 Bauer-Mehren, A. *et al.* (2009) Pathway databases and tools for their exploitation: benefits, current limitations and challenges, Mol. *Syst. Biol.* 5, 290
8 Khatri, P. *et al.* (2012) Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput. Biol.* 8, e1002375
9 Korcsmáros, T. *et al.* (2010) Uniformly curated signaling pathways reveal tissue-specific cross-talks and support drug target discovery. *Bioinformatics* 26, 2042–2050
10 Fearnley, L. *et al.* (2013) Extracting reaction networks from databases - opening Pandora's box. *Brief. Bioinform.* http://dx.doi.org/10.1093/bib/bbt058
11 Cusick, M. *et al.* (2009) Literature-curated protein interaction datasets. *Nat. Methods* 6, 39–46
12 Subramanian, A. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550
13 Huang, D. *et al.* (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13
14 Bauer, S. *et al.* (2010) GOing Bayesian: model-based gene set analysis of genome-scale data. *Nucleic Acids Res.* 38, 3523–3532
15 Donato, M. *et al.* (2013) Analysis and correction of crosstalk effects in pathway analysis. *Genome Res.* 23, 1885–1893
16 Sharan, R. *et al.* (2007) Network-based prediction of protein function. *Mol. Syst. Biol.* 3, 88
17 Tuncbag, N. *et al.* (2013) Simultaneous reconstruction of multiple signaling pathways via the prize-collecting Steiner forest problem. *J. Comput. Biol.* 20, 124–136
18 Mostafavi, S. *et al.* (2008) GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol.* 9 (Suppl. 1), S4
19 Yosef, N. *et al.* (2011) ANAT: a tool for constructing and analyzing functional protein networks. *Sci. Signal.* 4, pl1
20 Poirel, C. *et al.* (2013) Top-down network analysis to drive bottom-up modeling of physiological processes. *J. Comput. Biol.* 20, 409–418
21 Navlakha, S. *et al.* (2012) A network-based approach for predicting missing pathway interactions. *PLoS Comput. Biol.* 8, e1002640
22 Dotan-Cohen, D. *et al.* (2009) Biological process linkage networks. *PLoS ONE* 4, e5313
23 McCormack, T. *et al.* (2013) Statistical assessment of crosstalk enrichment between gene groups in biological networks. *PLoS ONE* 8, e54945
24 Schelhorn, S-E. *et al.* (2011) Inferring physical protein contacts from large-scale purification data of protein complexes. *Mol. Cell. Proteomics* 10, M110.004929
25 Fukuda, K. and Takagi, T. (2001) Knowledge representation of signal transduction pathways. *Bioinformatics* 17, 829–837
26 Hu, Z. *et al.* (2007) Towards zoomable multidimensional maps of the cell. *Nat. Biotechnol.* 25, 547–554
27 Gat-Viks, I. and Shamir, R. (2007) Refinement and expansion of signaling pathways: the osmotic response network in yeast. *Genome Res.* 17, 358–367

28 Ren, D. and Hassane, A. (2010) *Discrete, Continuous, and Hybrid Petri Nets*. (2nd Edition), Springer
29 Heath, L. and Sioson, A. (2009) Semantics of multimodal network models. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 6, 271–280
30 Klamt, S. *et al.* (2009) Hypergraphs and cellular networks. *PLoS Comput. Biol.* 5, e1000385
31 Zhou, W. and Nakhleh, L. (2011) Properties of metabolic graphs: biological organization or representation artifacts? *BMC Bioinformatics* 12, 132
32 Christensen, T. *et al.* (2009) Reconstruction and logical modeling of glucose repression signaling pathways in *Saccharomyces cerevisiae*. *BMC Syst. Biol.* 3, 7
33 Ramadan, E. *et al.* (2010) A hyper-graph approach for analyzing transcriptional networks in breast cancer. In *Proceedings of the 1st ACM International Conference on Bioinformatics and Computational Biology, BCB' 10, ACM*. pp. 556–562
34 Klamt, S. *et al.* (2006) A methodology for the structural and functional analysis of signaling and regulatory networks. *BMC Bioinformatics* 7, 56
35 Samaga, R. and Klamt, S. (2013) Modeling approaches for qualitative and semi-quantitative analysis of cellular signaling networks. *Cell Commun. Signal.* 11, 43
36 Demir, E. *et al.* (2010) *Nat. Biotechnol.* 28, 935–942
37 Hucka, M. *et al.* (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19, 524–531
38 Berge, C. (1989) *Hypergraphs: Combinatorics of Finite Sets (North-Holland Mathematical Library, Vol. 45)*. (1st Edition), North Holland
39 Gallo, G. *et al.* (1993) Directed hypergraphs and applications. *Discr. Appl. Math.* 42, 177–201
40 Vaske, C. *et al.* (2010) Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics* 26, i237–i245
41 Gonçalves, J. *et al.* (2011) TFRank: network-based prioritization of regulatory associations underlying transcriptional responses. *Bioinformatics* 27, 3149–3157
42 Ausiello, G. *et al.* (2001) Directed hypergraphs: problems, algorithmic results, and a novel decremental approach. In *Theoretical Computer Science, Vol. 2202 of Lecture Notes in Computer Science*. pp. 312–328, Springer
43 Kirouac, D. *et al.* (2012) Creating and analyzing pathway and protein interaction compendia for modelling signal transduction networks. *BMC Syst. Biol.* 6, 29
44 Guo, X. and Wang, X-F. (2009) Signaling cross-talk between TGF-β/BMP and other pathways. *Cell Res.* 19, 71–88
45 Ranganathan, P. *et al.* (2011) Notch signalling in solid tumours: a little bit of everything but not all the time. *Nat. Rev. Cancer* 11, 338–351
46 Karlebach, G. and Shamir, R. (2008) Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol.* 9, 770–780
47 Aldridge, B. *et al.* (2006) Physicochemical modelling of cell signalling pathways. *Nat. Cell Biol.* 8, 1195–1203
48 Tian, Z. *et al.* (2009) A hypergraph-based learning algorithm for classifying gene expression and arraycgh data with prior knowledge. *Bioinformatics* 25, 2831–2838
49 Battle, A. *et al.* (2010) Automated identification of pathways from quantitative genetic interaction data. *Mol. Syst. Biol.* 6, 379
50 Sumazin, P. *et al.* (2011) An extensive microrna-mediated network of rna-rna interactions regulates established oncogenic pathways in glioblastoma. *Cell* 147, 370–381
51 Rahman, A. *et al.* (2012) Reverse engineering molecular hypergraphs. In *Proceedings of the ACM Conference on Bioinformatics, Computational Biology, and Biomedicine*. pp. 68–75