Systems biology

Current and future directions in network biology

Marinka Zitnik [1,†], Michelle M. Li [1,†], Aydin Wells [2,3,4,†], Kimberly Glass [5,#],
Deisy Morselli Gysi [5,6,7,#], Arjun Krishnan [8,#], T.M. Murali [9,#], Predrag Radivojac [10,#],
Sushmita Roy [11,12,#], Anaïs Baudot [13], Serdar Bozdag [14,15], Danny Z. Chen [2], Lenore Cowen [16],
Kapil Devkota [16], Anthony Gitter [11,17], Sara J.C. Gosline [18], Pengfei Gu [2], Pietro H. Guzzi [19],
Heng Huang [20], Meng Jiang [2], Ziynet Nesibe Kesimoglu [14,21], Mehmet Koyuturk [22], Jian Ma [23],
Alexander R. Pico [24], Nataša Pržulj [25,26,27], Teresa M. Przytycka [21], Benjamin J. Raphael [28],
Anna Ritz [29], Roded Sharan [30], Yang Shen [31], Mona Singh [28,32], Donna K. Slonim [16],
Hanghang Tong [33], Xinan Holly Yang [34], Byung-Jun Yoon [31,35], Haiyuan Yu [36], Tijana Milenković [2,3,4,*]

[1]Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115, United States
[2]Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN 46556, United States
[3]Lucy Family Institute for Data and Society, University of Notre Dame, Notre Dame, IN 46556, United States
[4]Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN 46556, United States
[5]Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, United States
[6]Department of Statistics, Federal University of Paraná, Curitiba, Paraná 81530-015, Brazil
[7]Department of Physics, Northeastern University, Boston, MA 02115, United States
[8]Department of Biomedical Informatics, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, United States
[9]Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, United States
[10]Khoury College of Computer Sciences, Northeastern University, Boston, MA 02115, United States
[11]Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI 53715, United States
[12]Wisconsin Institute for Discovery, Madison, WI 53715, United States
[13]Aix Marseille Université, INSERM, MMG, Marseille, France
[14]Department of Computer Science and Engineering, University of North Texas, Denton, TX 76203, United States
[15]Department of Mathematics, University of North Texas, Denton, TX 76203, United States
[16]Department of Computer Science, Tufts University, Medford, MA 02155, United States
[17]Morgridge Institute for Research, Madison, WI 53715, United States
[18]Biological Sciences Division, Pacific Northwest National Laboratory, Seattle, WA 98109, United States
[19]Department of Medical and Surgical Sciences, University Magna Graecia of Catanzaro, Catanzaro, 88100, Italy
[20]Department of Computer Science, University of Maryland College Park, College Park, MD 20742, United States
[21]National Center of Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20814, United States
[22]Department of Computer and Data Sciences, Case Western Reserve University, Cleveland, OH 44106, United States
[23]Ray and Stephanie Lane Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, United States
[24]Institute of Data Science and Biotechnology, Gladstone Institutes, San Francisco, CA 94158, United States
[25]Department of Computer Science, University College London, London, WC1E 6BT, England
[26]ICREA, Catalan Institution for Research and Advanced Studies, Barcelona, 08010, Spain
[27]Barcelona Supercomputing Center (BSC), Barcelona, 08034, Spain
[28]Department of Computer Science, Princeton University, Princeton, NJ 08544, United States
[29]Department of Biology, Reed College, Portland, OR 97202, United States
[30]School of Computer Science, Tel Aviv University, Tel Aviv, 69978, Israel
[31]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, United States
[32]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, United States
[33]Department of Computer Science, University of Illinois Urbana-Champaign, Urbana, IL 61801, United States
[34]Department of Pediatrics, University of Chicago, Chicago, IL 60637, United States
[35]Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973, United States
[36]Department of Computational Biology, Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY 14853, United States

*Corresponding author. Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN 46556, United States.
E-mail: tmilenko@nd.edu
†Co-first authors.
#Co-second authors: coordinators for Sections 2–7, listed in their alphabetical order by last name.

All other authors are listed in their alphabetical order by last name.

Abstract

Summary: Network biology is an interdisciplinary field bridging computational and biological sciences that has proved pivotal in advancing the understanding of cellular functions and diseases across biological systems and scales. Although the field has been around for two decades, it remains nascent. It has witnessed rapid evolution, accompanied by emerging challenges. These stem from various factors, notably the growing complexity and volume of data together with the increased diversity of data types describing different tiers of biological organization. We discuss prevailing research directions in network biology, focusing on molecular/cellular networks but also on other biological network types such as biomedical knowledge graphs, patient similarity networks, brain networks, and social/contact networks relevant to disease spread. In more detail, we highlight areas of inference and comparison of biological networks, multimodal data integration and heterogeneous networks, higher-order network analysis, machine learning on networks, and network-based personalized medicine. Following the overview of recent break-throughs across these five areas, we offer a perspective on future directions of network biology. Additionally, we discuss scientific communities, educational initiatives, and the importance of fostering diversity within the field. This article establishes a roadmap for an immediate and long-term vision for network biology.

Availability and implementation: Not applicable.

1. Introduction

A network (or graph) comprises a set of nodes (or vertices) that are connected by a set of edges (or links); see InfoBox 1. Networks allow us to study the properties of a complex system that emerge from interactions between its individual components. Networks have been a powerful way to represent a variety of real-world phenomena, including technological, information, transportation, social, financial, software, ecological, chemical, and biological systems (Barabási 2016, Newman 2018). Our focus is on biological networks, which offer the understanding of complex functions at the levels of genes, proteins, cells, tissues, organs, etc., by representing a given biological system as an interconnected entity rather than a collection of individual components. In a biological network, nodes typically represent biomolecules (e.g. amino acid residues within a protein, proteins within a cell, or cells within a tissue), and edges typically indicate interactions between the biomolecules (e.g. physical, functional, or chemical). While the main focus of our article is on such biological networks that model relationships between biomolecules, i.e. on molecular/cellular networks, our article touches on other types of biological networks, such as biomedical knowledge graphs (BKGs), ontologies, patient similarity networks modeling, e.g. electronic health record data, brain networks constructed from medical imaging data, and even social and contact networks relevant for the spread of disease. We acknowledge that other types of biological networks exist that are not the focus of our article and that we thus do not cover, such as ecological ones.

Network biology (Fig. 1) is an interdisciplinary field spanning computational (e.g. algorithms, graph theory, network science, data mining, and machine learning) and biological sciences. While the field has existed for nearly two decades, it has undergone numerous rapid changes and new computational challenges have arisen. This is caused by many factors, including increasing data complexity, such as multiple types of data becoming available at different levels (or scales) of biological organization, as well as growing data size. Ironically, despite the massive increase in available data, the data remain incomplete and noisy. This means that the research directions in the field also need to evolve.

This article discusses the current state as well as the future of the field. Its goal is to identify pressing challenges with well-established as well as emerging topics in network biology, which are shown in Fig. 1: inference and comparison of biological networks (Section 2), multimodal data integration

and heterogeneous networks (Section 3), higher-order network analysis (Section 4), machine learning on networks (Section 5), and network-based personalized medicine (Section 6). We comment on why these topics have been strategically chosen for discussion in this article.

Noting again that a key focus of our article is on molecular/cellular (i.e. -omics) data, certain types of -omics data are explicitly captured as networks. That is, interactions between biomolecules are measured explicitly by biotechnological data collection platforms. A prominent example is protein–protein interaction (PPI) networks. In these networks, nodes are proteins and edges correspond to physical bindings between the proteins. In human and some model organisms, extensive high-throughput yeast two-hybrid and other experimental efforts have resulted in large sets of "reference" PPIs (such as HURI for humans), along with substantial knowledge about protein binding specificities (Stark *et al.* 2006, Luck *et al.* 2020).

Other types of -omics data are not captured as networks explicitly, but interactions between biomolecules can be inferred computationally, resulting in, e.g. association, correlation, regulatory, or knowledge graphs (InfoBox 1). Section 2 addresses several aspects of the task of inferring a homogeneous network, including a condition-specific network, typically from up to a couple of -omics data types/modes, along with a related topic of differential network analysis, which is one type of network comparison. Section 3 addresses the task of inferring a heterogeneous network, typically from diverse -omics or other multimodal data types (InfoBox 1), along with several other tasks related to multi-omics data integration, including network alignment, which is another type of network comparison. By a homogeneous network, we mean a network with a single node type and a single edge type, while by a heterogeneous network, we mean any nonhomogeneous network (i.e. multiple node types or multiple edge types or both); see InfoBox 1 and Section 3 for details.

Given (explicitly captured or inferred) network data, the next step is to analyze the data. While Sections 2 and 3 already address network analysis from the perspective of network comparison and several other tasks, Sections 4 and 5 further discuss prominent tasks related to network analysis. Namely, Section 4 discusses topics of capturing higher-order network structures called graphlets (subgraphs) in traditionally used pairwise graphs, which capture interactions between pairs of nodes, as well as shifting from pairwise graphs to hypergraphs, which are capable of capturing interactions

---

InfoBox 1. Basic terminology used in the article. Note that distinct scientific communities in network biology, including graph theory, network science, data mining, machine learning, and artificial intelligence, may use varied terminology for the same concepts or identical terms for different concepts.

- A (pairwise, homogeneous) **graph** (or network) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is defined by a set of nodes (or vertices) $\mathcal{V}$ and a set of edges (or links) $\mathcal{E}$. All nodes $v \in \mathcal{V}$ are of the same type. An edge $e_{u,v} \in \mathcal{E}$ indicates a relationship between exactly two nodes $u, v \in \mathcal{V}$.
- In a **PPI network**, nodes are proteins and edges correspond to physical bindings between proteins. Such a network of physical PPIs is also referred to as **interactome**.
- A (physical) PPI network is a special type of an **association network** between proteins. In addition to physical PPIs, an association network may contain links between proteins derived from sequence or 3D structural similarities, genetic interactions, literature-mined edges, or other protein association types.
- **Correlation networks** are calculated from -omics data collected across multiple samples. A prominent type isgene co-expression networks, where nodes (genes) are linked by undirected edges if the genes' expression levels are correlated strongly enough across the samples.
- **Regulatory networks** capture directed relationships between regulators and their targets and describe causal (rather than correlative) relationships between biomolecules. A prominent type is gene regulatory networks where the regulators are transcription factor (TF) proteins (or other molecules that impact gene expression, such as microRNAs) and the targets are genes.
- **BKGs** describe semantic relationships between diverse biomedical entities (e.g. genes, diseases, and patients, as well as associated measurements). They represent facts using "subject–predicate–object" triples as the fundamental unit; the subject and object are nodes in the graph and the predicate (or relation) corresponds to a directed edge between the nodes.
- A **condition-unspecific (or context-unaware) network** spans multiple conditions/contexts such as diseases, ages, cell types, tissues, etc., and ultimately, individuals.
- A **condition-specific network** is inferred by integrating a context-unaware network with condition-specific node measurement (e.g. gene expression or mutation) data. The outcome of the data integration is the identification of network regions that are "active" in the given condition, which can be seen as **condition-specific or disease-dysregulated pathways** (sparse, tree-like subnetworks) or **functional modules** (dense, clique-like subnetworks).
- A **heterogeneous graph** contains multiple types of nodes and/or edges.
- A **multiplex graph** is a heterogeneous graph with multiple edge types between the same nodes, possibly nodes of a single type, in which case the heterogeneity comes from the different edge types.
- A **network-of-networks** is a heterogeneous graph in which different node types exist at different scales (or levels) and nodes at a higher level are graphs themselves at the lower level.
- **Multimodal data** that are represented as a heterogeneous graph in network biology include **multi-omic** data such as epigenomic, transcriptomic, proteomic, and metabolomic molecular measurements as well as nonmolecular data such as text and images from, e.g. patients' electronic health records.
- A **hypergraph** is a generalization of a (pairwise) graph in which an edge (also called a hyperedge) can connect any number (including more than two) of the nodes.
- A **subgraph** (or subnetwork) $\mathcal{G}_S = (\mathcal{V}_S, \mathcal{E}_S)$ of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a set of nodes $\mathcal{V}_S \subseteq \mathcal{V}$ and a set of edges $\mathcal{E}_S \subseteq \mathcal{E}$ such that for each edge $e \in \mathcal{E}_S$, both of its end nodes must be in $\mathcal{V}_S$.
- A subgraph is **induced** if and only if all edges between the nodes in $\mathcal{V}_S$ that exist in $\mathcal{E}$ are in $\mathcal{E}_S$.
- **Graphlets** are connected, nonisomorphic, induced subgraphs of a (pairwise) graph.
- **Hypergraphlets** are graphlet extensions from (pairwise) graphs to hypergraphs.
- A **cluster or community** in a graph is a set of topologically related nodes, typically densely connected to each other and loosely connected to nodes in other clusters.

---

between more than two nodes (InfoBox 1). Section 5 discusses machine learning advances in network biology, which have grown exponentially in the last decade. Key topics discussed include graph representation learning, incorporating knowledge into machine learning models, generative graph modeling, and transfer learning.

Section 6 complements the other, computationally focused sections by discussing an applied aspect of network biology: network-based personalized (or precision) medicine. Precision medicine aims to provide tailored treatment strategies for individuals (Aronson and Rehm 2015, Kaiser 2015). This personalized characterization may include molecular, environmental, lifestyle, and other factors. Integrating such different data types via network approaches can expand the potential for precision therapeutics while providing robustness to various types of data noise (Wang *et al.* 2014).

The five topics are not mutually exclusive. For example, multimodal (including multi-omics) data integration is a topic relevant to almost all of Sections 2–6. After the current research network biology advances are presented in these five sections, Section 7 discusses future research directions in the field, and Section 8 provides additional discussion on scientific communities, education/training, and diversity in computational (including network) biology.

2. Inference and comparison of biological networks

Inference of a network from nonnetwork data. Biological networks that are computationally inferred from nonnetwork -omics data can be categorized into three broad types:

Figure 1. Overview of the network biology field and five research topics discussed in this article. The word cloud in the center, generated using WordClouds.com, contains the top 30 most representative words from this article. Note that each word's rank is based on the sum of the weights of the core word (e.g. learn) and its derived words (e.g. learns, learning, learned).

association networks, correlation networks, and regulatory networks. The three network types are defined briefly in InfoBox 1, discussed in detail in the following text, and illustrated in Fig. 2A.

Association networks typically capture undirected and unsigned relationships between biological molecules; while they might contain experimentally derived interactions, they may also contain interactions derived computationally from a variety of possible data sources. One of the most common types of association networks is (physical) PPI networks, which are explicitly derived via high-throughput experiments (Section 1) (Luck *et al.* 2020). These experiments, primarily co-immunoprecipitation and yeast two-hybrid, differ in their estimated error rates and can produce both false positives and false negatives (Von Mering *et al.* 2002, Sprinzak *et al.*

2003, Bader *et al.* 2004). In addition, for all but the yeast interactome, where a substantial fraction of pairs of proteins have been assayed, even in most model organisms, the majority of pairs of proteins have not been tested for interaction (Sledzieski *et al.* 2021). Thus, even across all the myriad sources of PPI networks, there is much missing data (Sledzieski *et al.* 2021). In addition to physical PPIs, many public resources curate associations between biomolecules from many data sources (Bajpai *et al.* 2020, Wright *et al.* 2024). For example, the widely used STRING association network (Szklarczyk *et al.* 2023) contains interactions between proteins derived from sequence or 3D structural similarities, genetic interactions, literature-mined edges, or other types of pairwise protein associations that are distinct from physical binding between proteins.

In an association network composed of genetic interactions (also known as a genetic interaction network), an edge between nodes (genes/proteins) indicates that mutations or other perturbations to the two nodes produce an unexpected cellular phenotype (Baryshnikova et al. 2013). An example of a genetic interaction is when mutations in both genes/proteins result in cell death, i.e. they are lethal, while the cell remains viable when there is a mutation in just one of them. A weighted version of a genetic interaction network also exists, in which edge weights indicate how strong or weak the observed double mutant phenotype, such as cell growth rate, is compared to the expected phenotype (Costanzo et al. 2016).

Challenges with association networks are that they are generally not condition-specific and contain interactions derived from multiple types of evidence, with different evidence sources having different quality levels and representing different types of biological relationships. Additional investigation of how different evidence sources influence network analysis results is often required (Kim et al. 2021). Although the biological relationships represented in PPI networks and genetic interaction networks are easier to interpret, these networks tend to be incomplete and noisy and only exist for a limited number of species and biological conditions, limiting their use (Rolland et al. 2014, Zitnik et al. 2019b).

Correlation networks are typically calculated from -omics data collected across multiple samples (time points, tissues, patients, ages, drugs, or other conditions); relationships in correlation networks are typically undirected and signed, depending on how the network is inferred. Among the most prominent types of correlation networks are gene co-expression networks. Namely, given transcriptomics data containing the expression (i.e. mRNA abundance) levels of genes across multiple samples, a gene co-expression network can be constructed by linking nodes (genes) via edges if the genes' expression levels are correlated strongly enough across the samples. In addition to being used to capture gene co-expression, correlation networks have been applied in biomedicine to study relationships between many other types of elements, such as metabolites (Perez De Souza et al. 2020), disease biomarkers (Hwa Chu et al. 2014, Nishihara et al. 2017, Huang et al. 2019), and even foods (Kim et al. 2015, Samieri et al. 2020). Correlation networks are widely used in biomedical applications due to their simplicity and the ease with which they can be generated and interpreted (Pierson et al. 2015, Huang et al. 2019, Samieri et al. 2020, Lee et al. 2021). Pearson correlation is the most common measure for calculating correlation networks, i.e. determining which gene pairs should be linked by edges, although other measures, such as Spearman correlation or mutual information, are also used, depending on the nature of the data and nonlinearity of the relationships being captured (Reshef et al. 2011). Multiple algorithms and tools have been developed for inferring correlation networks, including ARACNe (Margolin et al. 2006), which calculates the mutual information between pairs of nodes and then removes indirect relationships; CLR (Faith et al. 2007), which calculates the mutual information between pairs of nodes and then z-score normalizes; WGCNA (Zhang and Horvath 2005), which scales the Pearson correlation to generate a scale-free network topology (or network structure); and wTO (Gysi et al. 2018), which normalizes the chosen correlation by all other correlations and calculates a probability for each edge.

One advantage of correlation networks compared to association networks, especially PPI networks resulting from high-throughput experiments, is that correlation networks are explicitly derived from condition-specific -omics data, while association networks generally do not capture condition-specific information (Sonawane et al. 2019). However, despite their popularity, correlation networks have multiple known limitations. One limitation is difficulty translating to biological mechanisms (Larsen et al. 2019). Another limitation is that different network inference methods yield significant dissimilarities in the topology as well as functional content between the resulting correlation networks (Rider et al. 2014). For example, when multiple methods are applied to infer gene co-expression networks based on the same underlying data, the resulting networks tend to capture different sets of edges between the same nodes; furthermore, when those networks are used to predict genes' functional annotations such as Gene Ontology (GO) terms, the results often differ (Li et al. 2023b). Sometimes it might be helpful to combine networks inferred using different methods into a consensus network (Gysi et al. 2018, Li et al. 2023b), where edges are reweighted so that the more networks support an edge and the more strongly they support it, the higher its consensus weight or probability. A further limitation of gene co-expression networks is that co-expression between two genes occurs when one gene regulates another or when two genes are targeted by the same regulator (Ku et al. 2012, Yin et al. 2021). However, these two distinct biological scenarios are represented in the same way in a co-expression network, by linking the two genes with an undirected edge. Instead, regulatory networks can distinguish between the different scenarios, as discussed next.

Regulatory networks capture directed relationships between regulators and their targets and describe causal (rather than just correlative) relationships between biomolecules; although these networks in theory should be signed, in practice deriving the sign of regulatory relationships from high-throughput biological data is challenging. There are many types of regulatory networks in biology. However, for most inferred regulatory networks, the regulators are TF proteins (or other molecules that impact gene expression such as microRNAs) and the targets are genes; these are commonly referred to as gene regulatory networks. There are many approaches to infer gene regulatory networks. For example, TF–gene relationships can be measured experimentally through ChIP-sequencing. In this case, the presence of a TF binding in the regulatory region(s) of a gene can be used to infer an edge from that TF to the gene. However, the cost and experimental limitations make it impossible to infer a complete gene regulatory network in this way. Therefore, many computational approaches have been developed to infer gene regulatory networks. For example, the DNA sequence of gene regulatory regions can be scanned to identify matching patterns (known as sequence motifs) that indicate a potential TF binding site; however, linking TFs to genes based on DNA sequence alone does not give a condition-specific network. Thus, methods to infer gene regulatory networks typically use gene expression data, either alone or in combination with computational evidence for TF binding in gene promoters, to infer TF–gene relationships (Marbach et al. 2012a). Popular algorithms of this type include Inferelator (Bonneau et al. 2006), which uses linear regression, L1 shrinkage, and LASSO to identify a set of parsimonious models to predict
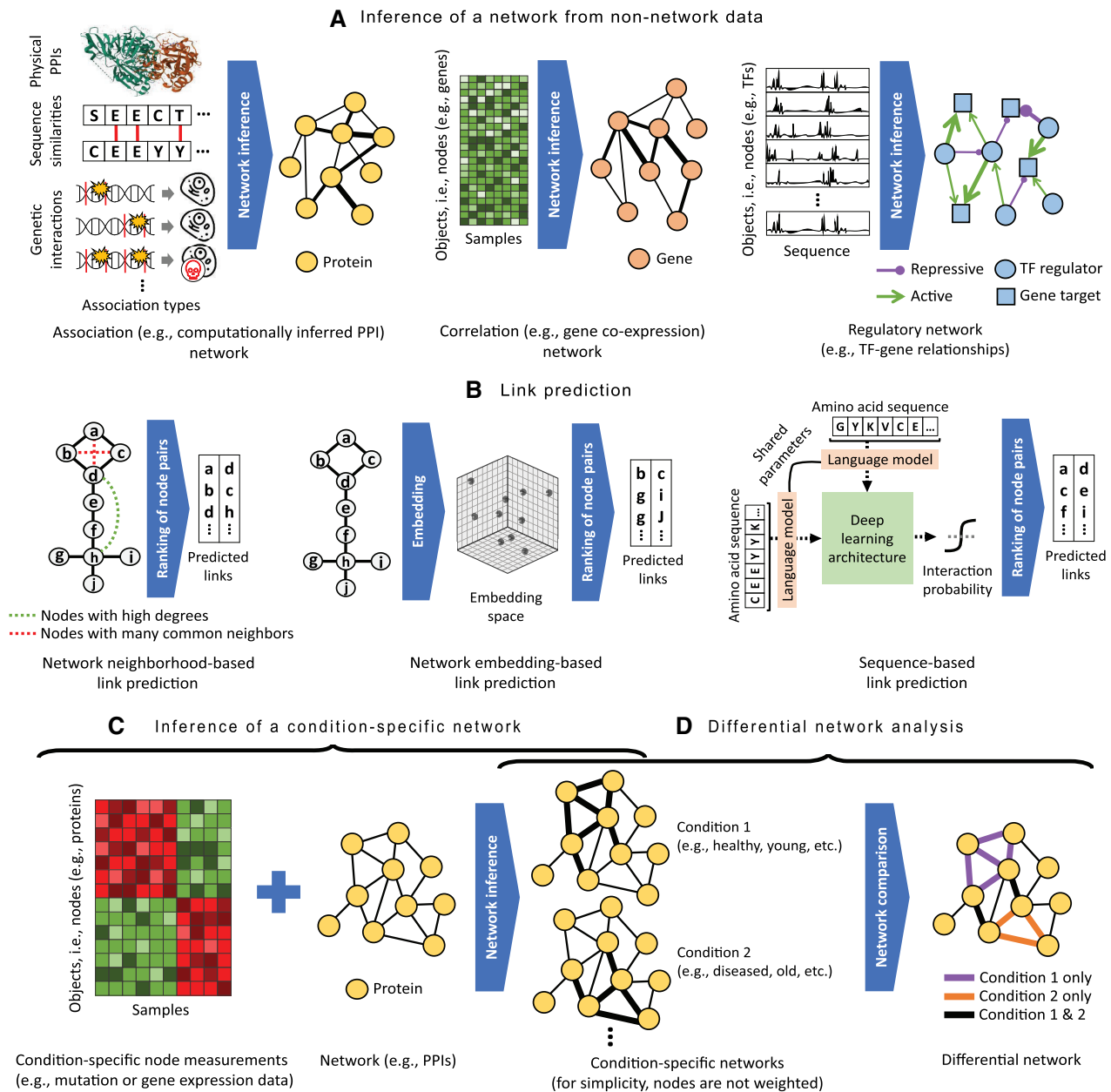
Figure 2. Prominent topics related to network inference and comparison. (A) Inference of an association (*left*), correlation (*middle*), or regulatory (*right*) network from nonnetwork data. (B) Link prediction: inference of new interactions from existing network data via neighborhood- (*left*) or embedding-based (*middle*) approaches, or from sequence data (*right*). For the former, shown are nodes that may be linked by new edges because two given nodes have high degrees (preferential attachment) or share many common neighbors; other neighborhood-based approaches exist, as discussed in the text. (C) Inference of a condition-specific network. The second approach category is illustrated. The thicker an edge in the network for a given condition, the more relevant the edge is for that condition. (D) Differential network analysis. Illustrated is a potential differential network between conditions 1 and 2, containing edges that are highly relevant for condition 1 but not condition 2, edges that are highly relevant for condition 2 but not condition 1, and edges that have consistent relevance patterns in both conditions.

target gene expression levels from TF expression levels (and other factors); GENIE3 (Huynh-Thu *et al.* 2010), which uses tree-based ensemble methods to develop a set of regression problems that predict the expression pattern of each target gene from the expression of a set of input TF genes; and PANDA (Glass *et al.* 2013), which uses message passing to amplify consistent structures across three input data types: TF–TF PPIs, computationally inferred TF–gene relationships, and gene–gene co-expressions. As opposed to Inferelator and GENIE3, PANDA does not consider the expression levels of TFs but instead uses evidence of co-expression in genes as evidence of targeting by the same TF. In contrast, a recent

method NETREX-CF incorporates, among other techniques, a machine learning approach known as collaborative filtering to deal with missing data (Wang *et al.* 2022e).

Other methods to infer regulatory networks incorporate epigenetic data. In particular, chromatin state can indicate whether the DNA is "open" and available to be bound by a TF; thus, computational evidence for TF binding in gene regulatory regions that also overlap with open chromatin can be used to estimate cell type-specific networks (Neph *et al.* 2012). Specific algorithms to infer gene regulatory networks using epigenetic data include TEPIC (Schmidt *et al.* 2017, 2019), which combines TF binding affinities, chromatin state

data, and gene annotation data to predict TF–gene relationships, and SPIDER (Sonawane *et al.* 2021), which uses message passing to infer and amplify consistent structure in an epigenetically pruned gene regulatory network constructed by combining computational evidence for TF binding with open chromatin data. Both TEPIC and SPIDER can also (optionally) incorporate gene expression data. Despite multiple methods in this area (including many beyond those described here), it remains challenging to integrate multiple types of -omics data to effectively infer accurate condition-specific regulatory networks; we elaborate on this challenge in subsection "Inference of a heterogeneous network from multimodal data" of Section 3.

Link prediction: inference of new interactions from existing network data. Link prediction is applicable to any network type, but in network biology, it has prominently been used in association networks containing interactions between proteins. Regardless of the type of data used to construct an association network, the resulting network is often incomplete. For example, many pairs of proteins in an organism may yet to be assayed for physical interaction. However, the "guilt by association" principles that underlie the topological organization of most of these networks (Cowen *et al.* 2017) mean that the patterns of connection of existing links can reliably predict some of the missing edges. We refer to this as network-based link prediction (Fig. 2B). Network-based prediction of new interactions between proteins often uses either a relatively simple rule (e.g. it may be desirable to link nodes that have high degrees, that have many common interacting partners—or neighbors—either direct or extended ones, that share many paths, or that are topologically similar; Hulovatyy *et al.* 2014) or more sophisticated diffusion-based network embeddings (Cowen *et al.* 2017, Hamilton *et al.* 2018, Kovács *et al.* 2019, Devkota *et al.* 2020, Huang *et al.* 2020, Yuen and Jansson 2020, Coşkun and Koyutürk 2021). A mixture of these strategies, where simple rules are employed in the core of the network, and diffusion-based network embeddings are employed outside the core, perform particularly well. However, the set of rules and the embedding used matters (Devkota *et al.* 2020), especially because interaction patterns may be quite different in networks containing physical PPIs versus those containing inferred, non-physical associations between proteins.

Link prediction: other techniques to infer missing interactions. Beyond methods that leverage the topology of the known interactions, the other methods to infer missing interactions will vary based on the underlying type of protein association data used to construct the network. For example, for physical PPI prediction, classical techniques such as docking can also be used when protein 3D structural models are available. With the rise of deep learning methods such as AlphaFold (Jumper *et al.* 2021), ESMFold (Lin *et al.* 2023), and OmegaFold (Wu *et al.* 2022), now a 3D structural model is usually available for most proteins. AlphaFold-Multimer (Evans *et al.* 2021) is a recent deep learning-based extension of AlphaFold that allows for predicting protein complexes, i.e. the quaternary structure of multiple proteins; then, it might be possible to use the confidence score of the predicted structure to predict whether the proteins interact or not. The predicted quaternary structure also provides the interaction interfaces between the proteins.

When the goal is ultrafast prediction (e.g. in order to perform genome-wide scans), there are alternative deep learning methods (Hashemifar *et al.* 2018, Zhang *et al.* 2018, Chen *et al.* 2019, Sledzieski *et al.* 2021) that have had success in sequence-based prediction of PPIs (Fig. 2B). These methods focus on computational speed. That is, like the network-based methods, they seek to predict only *whether* (rather than also *how*, which is more challenging) two protein sequences interact, so that it is tractable to make predictions for all the protein pairs in the network. However, we note that some of these sequence-based methods manage to implicitly incorporate information about protein 3D structures. For example, D-SCRIPT (Sledzieski *et al.* 2021) uses a pretrained protein language model (Bepler and Berger 2021) and implicitly learns a fuzzy contact map representation.

How to simultaneously leverage network- and sequence-based link prediction for physical PPI data remains an open problem, with valuable initial work (Bepler and Berger 2021). Also, evaluating link prediction methods and especially hybrid methods is tricky. This is because existing ground-truth networks (other than HURI; Luck *et al.* 2020) are biased by the portions of the networks containing well-studied proteins and pathways (Schaefer *et al.* 2015). So, it is difficult to come up with fair performance measures that are not biased by node degrees, and that do not advantage network-based methods while disadvantaging sequence-based methods (Singh *et al.* 2022, Wang *et al.* 2023d). On the other hand, sequence-based approaches do better on close homologs of known interacting protein pairs (Sledzieski *et al.* 2023).

Other researchers have noted that databases that amalgamate physical PPI data have not always kept up with the literature, and have proposed text-mining approaches to predict these "missing" links (Kim *et al.* 2008, van Haagen *et al.* 2009, Papanikolaou *et al.* 2015).

Inference of a condition-specific network. While existing biological network data resulting from extensive experimental efforts are an incredible resource, they typically do not capture how interactions in biological networks differ across conditions, i.e. they are context-unaware. By conditions, we mean diseases, ages, cell types, tissues, etc., and ultimately, individuals. Indeed, while human genomes in both healthy and disease populations are rapidly being sequenced, the corresponding condition-specific networks remain largely unknown. Moreover, the substantial amount of genetic variation across populations makes it infeasible in the near term to experimentally determine the full impact of this variation on interactions. So, computational methods have played and will continue to play a major role in inferring condition-specific networks.

We divide computational approaches for inferring condition-specific networks into several broad categories. (i) The first category is approaches that assess whether mutations observed in disease alter protein interactions. (ii) The second category is approaches that combine mutation data (e.g. on how many patients with a disease have genes containing significantly associated single nucleotide polymorphisms, indels, etc.) or condition-specific gene expression data (e.g. information on which genes are significantly expressed—or active—in a given condition; here, typically multiple data samples are needed per condition) with a PPI network. The goal is to identify PPIs that are dysregulated in a given disease or active in a given condition, i.e. to infer a condition-specific PPI network (Fig. 2C). (iii) The third category is approaches that use gene expression data to infer a correlation network

specific to the condition or sample of interest. (iv) The fourth category is analogs of the previous approaches but applied to regulatory networks rather than PPI or correlation networks.

Regarding the first approach category, significant computational efforts have focused on characterizing whether mutations observed in disease and variants across populations alter protein interactions. Early work mapping mutations observed in Mendelian diseases onto protein structures demonstrated that there is a statistically significant enrichment of Mendelian disease mutations in protein interaction interfaces, as compared to neutral polymorphisms observed across populations (Gao *et al.* 2015b). Homology modeling and domain-based approaches to identify sites that participate in interactions with DNA, RNA, peptides, ions, and small molecules have revealed that missense mutations observed in Mendelian diseases and somatic missense mutations in cancer are both enriched in these sites, with the strongest enrichments for DNA-binding sites, while common variants are depleted from these sites (Ghersi and Singh 2014, Kobren and Singh 2019). Further, these enrichments can be leveraged to identify cancer-relevant genes by developing statistical approaches to uncover proteins with more somatic missense mutations in their binding sites than expected (Ghersi and Singh 2014, Kobren *et al.* 2020). Protein interaction interfaces, as identified by homology modeling (Mosca *et al.* 2013) and machine learning (Meyer *et al.* 2018), have also been shown to be enriched in somatic missense mutations as compared to noninterface residues, and specific protein interactions relevant for cancer have been identified (Cheng *et al.* 2021). High-throughput experimental screens have led to estimates that two-thirds of disease-causing polymorphisms perturb protein interactions, with about half of these interrupting specific protein interactions while leaving other interactions unaffected (Sahni *et al.* 2015).

Regarding the second approach category, numerous computational efforts have focused on integrating condition-specific molecular measurements, mainly gene mutation or expression data (also referred to as gene activity data), with PPI network data (which is generally not condition-specific, i. e. context-unaware). They do so by mapping the gene activities onto the corresponding proteins in the PPI network, in order to assign condition-specific weights to the proteins or PPIs (or both) in the network (Fig. 2C). Then, highly weighted PPI network regions are hypothesized to be pathways dysregulated in disease (if using mutation data) or condition-specific subnetworks (if using expression data) (Leiserson *et al.* 2015, Newaz and Milenković 2022). The set of all such PPIs/pathways/subnetworks is a condition-specific PPI network. The data integration step is often performed via network propagation (Cowen *et al.* 2017), which diffuses the gene activities through the PPI network via random walks. Nonetheless, other approach types exist such as kernel, Bayesian, or nonnegative matrix factorization methods (Newaz and Milenković 2022).

Prominent applications of approaches from the second category have been studying cancer (Leiserson *et al.* 2015, Silverbush *et al.* 2019), tissue-specificity (Basha *et al.* 2020), aging (Li *et al.* 2022c), and genome-wide associations (Vanunu *et al.* 2010, Carlin *et al.* 2019). As an example, cancer-related gene mutation data was integrated with PPI data using the HotNet2 algorithm to identify the parts of the PPI network that are likely to be active in cancer (Leiserson *et al.* 2015). Such a cancer-specific network is not necessarily

connected, i.e. it might consist of multiple connected components, each of which can be thought of as a cancer-specific pathway or subnetwork. As another example, a general framework was proposed for assessing the ability of condition-specific PPI network inference approaches to illuminate tissue-specific processes and disease genes (Basha *et al.* 2020). This framework integrated RNA-sequencing profiles for 34 human tissues with a PPI network to create 34 tissue-specific PPI networks. Here, all tissue-specific PPI networks contained the same nodes and interactions, and they differed "only" in the weights associated with them. Then, given data associating GO biological processes to their relevant human tissues, this framework allows different condition-specific PPI network inference approaches to be benchmarked via enrichment tests in terms of their ability to recover tissue-specific processes. As a final example, unlike in the above applications where the inferred cancer- and tissue-specific networks were static, when studying human aging, which is a dynamic biological process, it is desired to infer a dynamic aging-specific network. Of the pioneering approaches towards this goal (Li *et al.* 2021, 2022c, Li and Milenković 2022, Newaz and Milenković 2022), a recent finding is that inferring an aging-specific PPI network that is both weighted and dynamic (as opposed to unweighted or static) results in the most accurate prediction of aging-related genes (Li *et al.* 2021). To infer this network, network propagation was used to map gene expression-based weights at different ages onto nodes in a PPI network. This resulted in a weighted network snapshot for each age, where the different snapshots had the same nodes and PPIs and "only" differed in their age-specific weights. The collection of all age-specific snapshots formed a weighted dynamic aging-specific PPI network. Then, aging-related genes can be predicted from this network, as discussed below (Li *et al.* 2021, 2022c).

An important issue in identifying condition-specific networks and especially disease-altered subnetworks via the above approaches is to determine whether the resulting (sub) networks are due to the molecular measurements (i.e. mutation or expression data) alone, the PPI network topology alone (e.g. due to ascertainment bias in PPI network data), or a combination of molecular measurement and network data. Recent work has shown that in some applications there may be a narrow regime where both molecular data and network information contribute to the identification of disease-dysregulated subnetworks (Reyna *et al.* 2021, Chitra *et al.* 2022).

Regarding the third approach category, condition-specific correlation networks are most often derived by applying a correlation measure to subsets of related samples (Pierson *et al.* 2015). However, since correlation measures rely on defining a distribution, this approach is inappropriate when a specific condition is represented by only a few (or even a single) sample(s). However, recently methods have been developed to infer "sample-specific correlations." That is, given a set of gene expression samples (across which correlation can be measured), these approaches can estimate one network for each individual sample in the input dataset. In particular, both SSN (Liu *et al.* 2016) and LIONESS (Kuijjer *et al.* 2019a, 2019b) work by computing two correlation networks, one with all samples and one with all samples except an individual sample of interest. Then, they use the difference between the two networks to estimate a correlation network specific to the sample of interest.

Finally, regarding the fourth approach category, genetic variants can impact gene regulatory networks by, e.g. altering TF binding or allele-specific expression (Przytycki and Singh 2020). Recall that missense mutations are enriched in sites that participate in interactions with DNA, RNA, peptides, ions, and small molecules, with the strongest enrichments for DNA-binding sites (Ghersi and Singh 2014, Kobren and Singh 2019). Also, recall that statistical approaches to identify proteins with more somatic missense mutations in their binding sites than expected by chance have identified cancer-relevant genes (Kobren and Singh 2019, Kobren et al. 2020). Deep learning approaches trained on DNA binding data from ENCODE (Moore et al. 2020) have also been used to assess whether DNA mutations impact TF binding in a tissue-specific manner (Zhou and Troyanskaya 2015). For some TFs, altered DNA-binding specificities can be predicted de novo using machine learning (Christensen et al. 2012, Persikov and Singh 2014, Sahni et al. 2015, Wetzel et al. 2022). However, if a DNA-binding protein's specificity is known a priori, then it is more accurate to instead predict how mutations alter that specificity rather than predict specificities de novo. For example, accurate predictions about how mutations alter DNA-binding specificities for homeodomain proteins were made by simultaneously learning interaction interfaces between DNA-binding proteins and their binding sites together with a predictive approach for DNA-binding specificity (Wetzel et al. 2022). Extending this approach to all DNA-binding proteins represents an important avenue for future work.

There has been also significant work done to infer condition-specific regulatory networks from various types of -omics data, as has been extensively reviewed in Baur et al. (2020). As one example, PANDA was applied to subsets of GTEx gene expression data to infer 38 tissue-specific gene regulatory networks (Sonawane et al. 2017); then, it was found that changes in TF targeting patterns led to the creation of new regulatory paths, giving them transcriptional control of tissue-specific processes. There also exist approaches that can be used to infer sample-specific networks for different -omics data types. For example, EGRET integrates predicted TF binding sites with genotype and expression quantitative trait loci data to create individual genotype-specific regulatory networks (Weighill et al. 2022). The SPIDER (Sonawane et al. 2021) and TEPIC (Schmidt et al. 2017, 2019) methods (described above) can be applied to individual epigenetic profiles to generate sample-specific regulatory networks. PSIONIC learns patient-specific TF regression weights by using chromatin-filtered TF–gene relationships to predict gene expression. Finally, the LIONESS method (Kuijjer et al. 2019b) can be used together with existing gene regulatory network reconstruction approaches that leverage gene expression data. When applying it in the same way as already described for correlation networks (the third approach category above), the LIONESS framework uses two estimated gene regulatory networks, one inferred with all gene expression samples and one inferred with all samples except one, to estimate a gene regulatory network specific to that sample (Kuijjer et al. 2019b).

Differential network analysis: comparison of condition-specific networks. Condition-specific networks often have the same set of nodes and differ only in terms of their edges. Many approaches have been developed to identify network regions that differ the most between condition-specific

networks; such regions have been shown to be responsible for the underlying biological differences between, e.g. healthy and disease conditions, between different tissues, or between young and old ages (Lichtblau et al. 2017, Basha et al. 2018), as discussed in more detail below. In general, approaches for this task can be characterized in several ways.

One category is based on the stage of network analysis, i.e. when differences between condition-specific networks are measured. Given condition-specific networks, one option is to first compute some topological property of a network region (at the level of a node, edge, network cluster—group of highly interconnected nodes—or entire network; see below) in each condition-specific network and then measure the extent of change in that property across the networks/conditions; the goal is to identify network regions that change the most (Zhu et al. 2016, Lichtblau et al. 2017). By a topological property, we mean a quantifiable measure of network structure such as the degree distribution of a network (the percentage of nodes in the network that have a given number of neighbors, i.e. degree), or centrality measures that rank nodes in a network from most to least central/important (examples are degree centrality according to which nodes with high degrees are central, and betweenness centrality according to which nodes that are on many shortest paths are central) (Barabási 2016, Newman 2018, Newaz and Milenković 2019).

A potential issue is that some topological properties, and especially centrality measures, are meaningful when used within a network but not necessarily when compared across networks (Newman 2018). As an alternative, approaches exist that first use the condition-specific networks to infer a single differential network that intuitively captures edges that differ between the conditions (Fig. 2D); only then, a desired topological property (e.g. centrality of each node) in the differential network is computed to identify network regions that are the most relevant (e.g. central/important) for the underlying condition-specific differences (Ruan et al. 2015).

The other category is based on the level of topology, i.e. where differences between condition-specific networks are measured: at the node (Weighill et al. 2021), edge (Glass et al. 2015), cluster (Padi and Quackenbush 2018), or entire network level (Newaz and Milenković 2022). At the node level, differences in centrality (e.g. degree or betweenness) are often used to identify the biomolecules around which network connectivity varies the most between the compared conditions. For example, "differential targeting," i.e. the difference in gene targeting—or the sum of the weights for all incoming edges to a gene—between two gene regulatory networks was used in combination with standard gene set enrichment tools to identify overrepresented biological processes in pancreatic ductal adenocarcinoma subtypes (Weighill et al. 2021). At the edge level, the goal is typically to determine edges specific to a given condition. This can be done in multiple ways, by taking, e.g. a certain percentage of the highest-weight edges, all edges above a given threshold, edges that have higher weights in one condition compared to others (Sonawane et al. 2017), or a combination of these (Glass et al. 2015). For example, the tissue-specific PPI networks discussed above, which were defined by differential edge scores, were correctly enriched in their respective tissue-associated biological processes; also, when the top 1% of the differential edges were considered, the resulting differential network regions were correctly enriched in genes related to

diseases associated with their respective tissues (Basha *et al.* 2020). Linking this discussion to the first approach category described above, it is important to note that although node centralities are often determined for each condition-specific network and then compared across the networks, they can also be calculated for a network defined by condition-specific edges. For example, degree and betweenness centralities of all genes in 38 tissue-specific gene regulatory networks were used to show that tissue-specific genes tended to assume bottleneck positions in their corresponding networks; in parallel, tissue-specific edges were identified by comparing the weight of each edge in a given tissue to the distribution of that edge's weight across all tissues, and it was found that the tissue-specific edges were enriched for connections between tissue-specific genes and depleted for canonical interactions (Sonawane *et al.* 2017). At the cluster level, e.g. given two condition-specific networks, ALPACA (Padi and Quackenbush 2018) identifies clusters that are shared between networks and distinct to each network. Heterogeneous (specifically, multiplex; Section 3) clustering algorithms (Mucha *et al.* 2010) could be useful for identifying such clusters. At the level of entire networks, typically their pairwise edge overlaps, as measured by, e.g. the Jaccard index, are used to quantify their pairwise (dis)similarities (Newaz and Milenković 2022).

We comment on two additional aspects of differential network analysis. First, while some condition-specific networks are derived from multiple data samples, sample-specific networks have the additional benefit of being able to be compared while accounting for other potentially relevant biomedical information (Kuijjer *et al.* 2019b). For example, the same statistical tools employed for differential gene expression analysis can be used to determine significant changes in the node-, edge-, cluster-, and network-level topological properties between sets of sample-specific networks. Importantly, this allows topological properties to be evaluated in the context of relevant biological and phenotypic variables, as well as potential confounders. For example, limma (Ritchie *et al.* 2015) was applied to compare features between male and female sample-specific gene regulatory networks while controlling for relevant confounders such as body mass index and age; node, edge, and TF-targeting was identified specific to males and females across 29 different tissues (Lopes-Ramos *et al.* 2020), as well as sex-specific targeting of the drug metabolism pathway in colon cancer (Lopes-Ramos *et al.* 2018).

Second, while the above discussion applies to all condition types, including temporal ones, we explicitly wish to comment more on approaches for characterizing how networks change over time (Teschendorff and Feinberg 2021). A prominent application in this context has been studying the change of PPI network topology with age. The process of inferring an aging-specific PPI network has already been discussed above. Here, we comment on how such a network, consisting of network snapshots corresponding to different ages, is analyzed. Original studies asked whether the overall, or global, topology changed with age, by: measuring pairwise edge overlaps between the snapshots; evaluating whether the snapshots' properties such as the average clustering coefficient, diameter, and graphlet degree distributions changed with age; and evaluating the fit of each snapshot to random (e.g. scale-free or geometric) graphs (Faisal and Milenković 2014, Newaz and Milenković 2022). Global topologies of the age-specific snapshots did not significantly change with age. It was then analyzed whether local topological positions of nodes as measured by (normalized) centralities changed with age. Hundreds of such genes were identified and predicted as aging-related; the predictions were validated via functional enrichment analyses (Faisal and Milenković 2014, Newaz and Milenković 2022).

Unlike such unsupervised prediction of aging-related genes, in recent work (Li *et al.* 2021, 2022c), supervised prediction was performed: by relying on knowledge about which genes are aging- versus nonaging-related (de Magalhães *et al.* 2009), new aging-related genes were predicted if their evolving topologies in a dynamic aging-specific PPI network matched topologies of the known aging-related genes. Recall that the state-of-the-art aging-specific dynamic PPI network is weighted. So, weighted node topological measures were used as features for supervised prediction that were simple extensions of unweighted centralities. Also, more advanced measures were proposed, which account for how the distribution of edge weights in the given node's (extended) network neighborhood changes with age, i.e. across the network snapshots (Li *et al.* 2021). A parallel line of work focused on studying how clusters, i.e. community structure, in a dynamic aging-specific human PPI network changed with age, and it was shown that the most prominent changes in the community structure correspond to ages that reflect known shifts from one stage of human lifespan to another (Hulovatyy and Milenković 2016, Crawford and Milenković 2018).

Another prominent point of discussion in the temporal/dynamic context is theoretical studies of molecular networks and observations of cell differentiation (i.e. the transition of a cell from one type to another), which indicate that cellular transitions can be smooth or nonlinear, gradual, or abrupt (Nykter *et al.* 2008, Moris *et al.* 2016). Computational methods to characterize these transitions using single-cell gene expression data include MuTrans (Zhou *et al.* 2021b), QuanTC (Sha *et al.* 2020), and BioTIP (Yang *et al.* 2022). These methods use different statistical approaches (stochastic differential equations, unsupervised learning of cell plasticity, or co-expression) and underlying theories (entropy and energy or tipping-point theory), but converge at the same best-studied bifurcations in six datasets (Yang *et al.* 2022).

Other types of network comparison. Differential network analysis is one type of network comparison, in which networks being compared have the exact same nodes and differ "only" in their edges (or edge weights). In other words, the mapping between the nodes of the compared networks is known. A complementary category of network comparison includes approaches that compare networks when their node mapping is unknown. Here, there are two distinct types: (i) network alignment or alignment-based network comparison and (ii) alignment-free network comparison (Yaveroğlu *et al.* 2015).

Alignment-based network comparison aims to find a mapping between the nodes of the compared networks that optimizes some objective function; this typically means conserving many edges and a large subgraph between the networks (Faisal *et al.* 2015a, Yaveroğlu *et al.* 2015, Guzzi and Milenković 2017). This approach category is useful for comparing biological networks of different species to identify evolutionary conserved parts of the networks. Consequently, network alignment allows for transferring biological knowledge (e.g. proteins' functional annotations or PPIs) between

aligned network regions across the compared species; also, it can complement sequence alignment by allowing for identification of protein orthology relationships based on the proteins' PPI network rather than (just) sequence similarities. Note that even when aligning homogeneous networks, the problem of network alignment can be viewed as integrating these networks into a heterogeneous (specifically, multiplex; Section 3) network representation. For this reason, and because methods have recently been proposed that align heterogeneous networks, we discuss algorithmic aspects of network alignment in the more appropriate Section 3. Here, we mainly aim to contrast general working principles of the different types of network comparison.

In contrast to alignment-based comparison, alignment-free network comparison simply aims to quantify the overall topological similarity between networks, regardless of a node mapping between the networks, and without intending to identify any conserved network regions; this typically means comparing some topological properties between networks, such as their (graphlet) degree distributions (Yaveroğlu et al. 2015, Newaz and Milenković 2019). Alignment-free network comparison is most often used to evaluate the fit of a random graph (e.g. scale-free or geometric) to a real-world network; also, it can identify groups/families of networks that are topologically similar to each other (Yaveroğlu et al. 2015). Given that alignment-free network comparison approaches do not aim to produce a node mapping between the compared networks, while alignment-based approaches do, the former are typically computationally more efficient than the latter (Yaveroğlu et al. 2015).

3. Multimodal data integration and heterogeneous networks

Overview. Network representations of biological systems, from cells to ecosystems, are naturally heterogeneous, consisting of multiple types of nodes and interactions (De Domenico 2023). This section focuses on prominent computational challenges related to inference and analysis of heterogeneous networks. Broadly, a heterogeneous network is defined as a representation of multimodal data where each data mode corresponds to a different node or edge type. In the literature, the term "heterogeneous network" has often been used as a synonym to, e.g. a multiplex, interdependent, multiscale, or multilayer network. The challenge is that sometimes different terminologies are used for the same concept, or the same terminology is used for different concepts; the disparate terminology associated with heterogeneous networks can reflect nuances in their frameworks (Kivelä et al. 2014). Here is the terminology from the existing literature (e.g. Pio-Lopez et al. 2021, Gu et al. 2022) that we use in this article (Fig. 3A).

A heterogeneous network is a network with multiple node types and/or multiple edge types. A multiplex network is a special type of heterogeneous network with multiple edge types between the same nodes, possibly nodes of a single type, in which case the heterogeneity comes from the different edge types. A multiplex network can be viewed as being composed of different network layers sharing the same set (replica) of nodes but each layer having distinct edge types (Kinsley et al. 2020). An example of this type in biology is a molecular network capturing different types of relationships, such as physical interactions, functional relationships, and

sequence similarities between proteins. A typical heterogeneous network, including those discussed in this section, contains both distinct node types and (by definition) distinct edge types. An example of this type is a molecular network representing relationships among heterogeneous node types such as genes, transcripts, proteins, and metabolites. Another example is a knowledge graph representing semantic relationships between node types such as genes, patients, drugs, and diseases. Another level of complexity is handling distinct node types at different scales (or levels) of biological organization, e.g. node types resulting from data modalities that capture molecular measurements in epigenomic, transcriptomic, proteomic, and metabolomic assays and from nonmolecular text and imaging data. Here, a network-of-networks is a special case in which a node at a given scale is a network at the lower scale. For example, a node (protein) in a PPI network can be represented as a protein structure network in which nodes are the protein's amino acids and edges link amino acids that are close enough in the protein's 3D-fold (Gu et al. 2022).

The broad definition of a heterogeneous network that we use subsumes any network type that is not a homogeneous (single node type and single edge type) network. Note that in some scientific fields, such as physics, while a multiplex network typically has the same meaning as above, heterogeneous network is a rarely used term. Instead, a heterogeneous network is often referred to as a multilayer network, and a network-of-networks is sometimes used as a synonym for a multilayer network (De Domenico et al. 2013, Kivelä et al. 2014, De Domenico 2023).

Heterogeneous networks are a powerful framework for the representation, integration, and analysis of diverse data modalities of a complex system with multiple types of nodes or edges (or both), allowing for reconciling complementary measurements and providing a holistic view of the system. Here, we discuss the following major research directions encompassing heterogeneous networks: inference of a heterogeneous network from multimodal data, pathway reconstruction for interpretation of multi-omic data, network alignment, inference and reasoning with BKGs, and network-of-networks analysis. This is not an exhaustive list of topics on heterogeneous networks, and other sections touch on additional topics. For example, Section 5 touches on graph representation learning including but not limited to learning in heterogeneous networks, and Section 6 talks about integration of multimodal data for the purpose of patient stratification, identification of disease-dysregulated molecular pathways and functional modules, and other precision medicine applications.

Inference of a heterogeneous network from multimodal data. Heterogeneous network inference is the computational task of inferring the graph connectivity structure from multimodal—to date, typically multi-omic—measurements (Hawe et al. 2019). The vast majority of methods for this task infer connections between nodes corresponding to biomolecules such as genes, proteins, and metabolites (Fig. 3B) using bulk-omic datasets. Single-cell -omic datasets have posed new opportunities for network inference where nodes can represent individual cells. Heterogeneous network inference methods can be grouped into categories based on how much they rely on labeled positive examples of edges.

Probably the simplest category of approaches takes as input labeled examples of edges and nonedges along with

pairwise node feature vectors derived from multimodal data and train binary classifiers to discriminate node pairs with edges from node pairs without edges (Marbach *et al.* 2012b, Greene *et al.* 2015). These binary classification approaches assume that all node pairs are independent of each other and are therefore limited in their ability to exploit the known connectivity structure of the graph. An alternative is embedding methods (discussed in more detail in Section 5) that take as input an incomplete graph and multimodal measurement data as node features and learn an embedding of the nodes based on the (partial) graph structure and measured values, which are then used to infer edges based on link prediction (Lee *et al.* 2019, Yue *et al.* 2020) or matrix completion (Natarajan and Dhillon 2014). Graph embedding methods relax the independence assumption of binary classification methods. As graph embedding methods capture more of the network connectivity, it is conceivable that they need less training data to do as good prediction as simple binary classification. Graph neural networks (GNNs, discussed in more detail in Section 5) offer new ways to incorporate more global information about the network to inform the inference task (Yue *et al.* 2020). The biggest limitation of the above approaches is the need for positive training data (edges) and that negative examples (nonedges) are not truly observed but are assumed to be part of the complement of the positive set.

On the other hand, unsupervised graph structure learning methods take as input node-level measurements and infer the graph structure from these measurements alone, without requiring any labeled examples of edges/nonedges. These approaches can range from correlation-based networks inferring pairwise dependencies between nodes representing different multimodal data (Vasaikar *et al.* 2018, Zhou *et al.* 2021a) to more general approaches based on probabilistic graphical models (Koller and Friedman 2009, Hawe *et al.* 2019). We note that several of these methods were originally developed for transcriptomic datasets and are thus discussed in Section 2. In probabilistic graphical models, nodes are modeled as random variables and edges correspond to statistical dependencies (Koller and Friedman 2009), where each data modality is represented as a different node type (Chen *et al.* 2014, Sedgewick *et al.* 2018). A key modeling challenge when handling multiple types of measurements is to specify the appropriate probability distributions for each data modality (Chen *et al.* 2014, Sedgewick *et al.* 2018). Furthermore, the larger number of variables of multimodal data introduces additional scalability issues for learning the structure of probabilistic graphical models such as general Bayesian networks. Several heuristics such as focusing on promising parents (Friedman *et al.* 1999, Schmidt *et al.* 2007), exploiting modularity of molecular networks (Segal *et al.* 2005), or approximating joint probability distributions as done in dependency networks (Heckerman *et al.* 2000, Greenfield *et al.* 2013, Roy *et al.* 2013) have enabled these models to scale to thousands of variables.

Once the networks have been defined, they can be further clustered into modules to identify potential functional groupings among the nodes (Newman 2006, Mitra *et al.* 2013, Choobdar *et al.* 2019). Unsupervised learning of graph structure from multi-omic data lends itself naturally to the inference of gene regulatory networks (Baur *et al.* 2020), where node types represent target genes and protein regulators. Protein regulators can be further modeled based on their observed mRNA levels or their hidden activity levels (Miraldi *et al.* 2019). While such approaches do not need any edge-level information, if any, potentially noisy, information is available, this can be incorporated as a graph prior to guide the structure learning (Greenfield *et al.* 2013, Siahpirani and Roy 2017, Miraldi *et al.* 2019).

The availability of single-cell multi-omic datasets has also opened up challenges that can be tackled with heterogeneous network inference (Demetci *et al.* 2022, Heumos *et al.* 2023). One such problem is to infer cell–cell networks with nodes corresponding to cells, node types corresponding to different modalities (e.g. scRNA-seq, scATAC-seq) or time points (or both), and edges representing different semantics such as similarity or lineage relationships. Due to the size and sparsity in these data, dimensionality reduction is typically performed prior to inference of network structure. Nonnegative matrix factorization, independent components analysis, and variational autoencoders are common dimensionality reduction approaches for single-cell multi-omic datasets. After dimensionality reduction, graph learning can be done using the k-nearest neighbor approach (Butler *et al.* 2018) or with optimal transport (Schiebinger *et al.* 2019, Demetci *et al.* 2022). Graphs based on k-nearest neighbors, with different distance measures, are straightforward to implement and frequently used in practice, while optimal transport's framework to match probability distributions of cells can be used to capture fine-grained cell dynamics.

Pathway reconstruction for interpretation of multi-omic data. Heterogeneous networks offer a powerful framework to integrate, interpret, and reconcile missing and noisy measurements commonly seen in multi-omic experiments (Haque *et al.* 2017, Peck Justice *et al.* 2021). The task of pathway reconstruction takes as input multi-omic measurements of different biomolecules represented as node types and a large background molecular network. It outputs a sparse subnetwork with high-quality connections among the relevant biomolecules (Garrido-Rodriguez *et al.* 2022) (Fig. 3C). The background networks typically contain PPIs and may also include protein–DNA, protein–RNA, or protein–metabolite interactions to match the available -omic data. Paths from one relevant biomolecule to another in the background network can help prune irrelevant biomolecules and identify those that may play critical roles in the overall biological process but were missed in the -omic measurements (Paull *et al.* 2013, Pirhaji *et al.* 2016, Tuncbag *et al.* 2016, Winkler *et al.* 2022). Note that this task also relates to condition-specific network inference discussed in Section 2 and multi-omic module discovery discussed in Section 6 for discovery of dysregulated pathways in diseases such as cancer.

The sparse subnetwork obtained depends on the choice of optimization algorithm and its parameters. Some pathway reconstruction algorithms are computationally efficient, based on shortest paths (Ritz *et al.* 2016) or network flow (Yeger-Lotem *et al.* 2009). Despite their algorithmic simplicity, these methods can still effectively prioritize biologically relevant nodes and interactions. Network flow-based methods can scale across multiple experiments by relying on the multicommodity flow approach, which identifies nodes and edges that are unique and shared across conditions (Gosline *et al.* 2012). General integer linear programming approaches (Ourfali *et al.* 2007, Chasman *et al.* 2014) support arbitrary node, edge, and path constraints. These provide the greatest customization for a particular multi-omic dataset but less scalability and reusability across applications. Intermediate
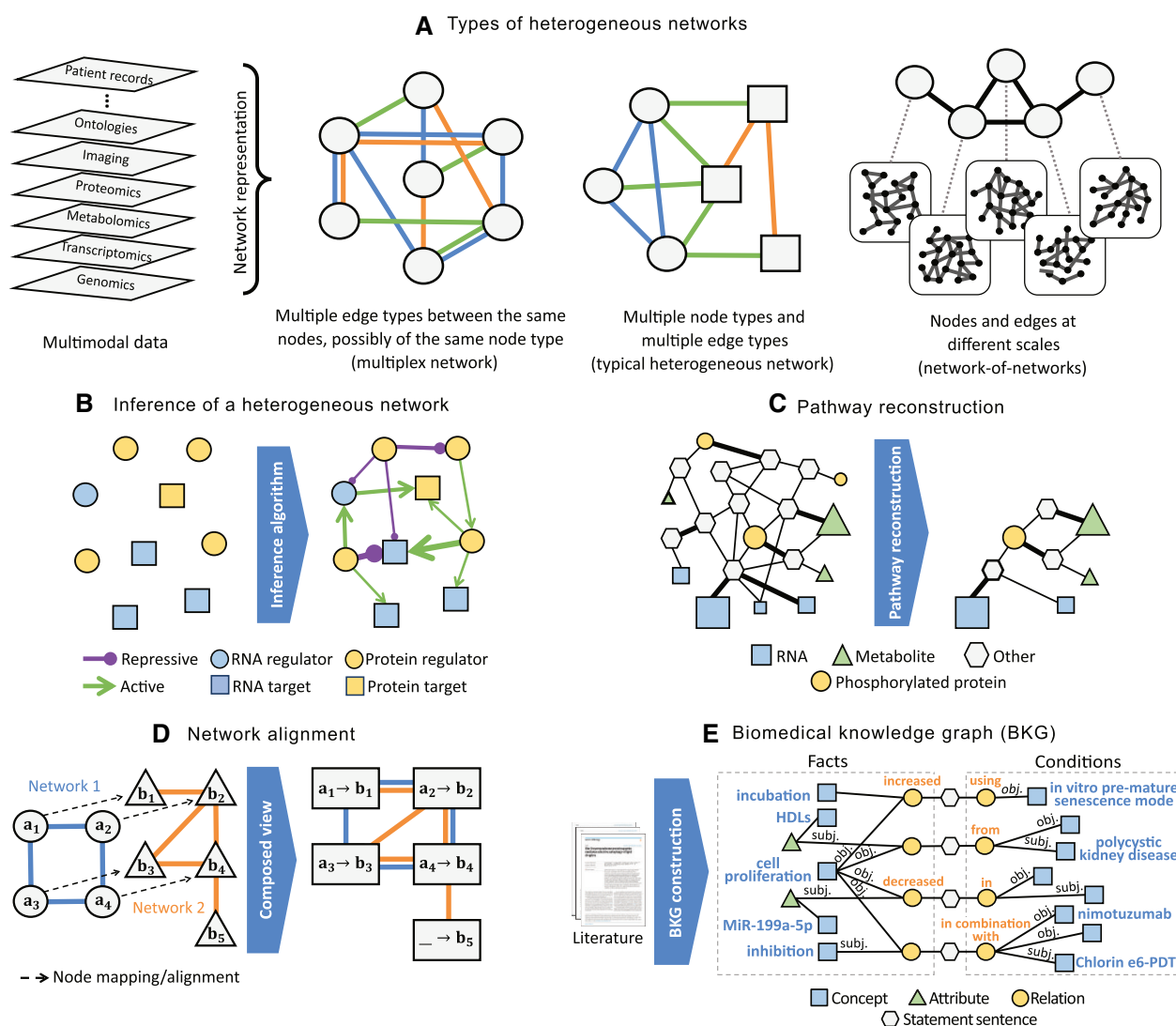
Figure 3. Prominent topics related to multimodal data integration and heterogeneous networks. (A) Heterogeneous networks can naturally represent multimodal data. A heterogeneous network can have only a single node type, with different data modalities representing multiple edge types. Or, there can exist both multiple node and edge types. Different node types can exist at different biological scales; e.g. in a network-of-networks, nodes at a given scale are networks at the lower scale. (B–E) Prominent topics related to heterogeneous networks. (B) Inference of a heterogeneous network aims to learn the graph topology from multimodal—to date, typically multi-omic—measurements. (C) Pathway reconstruction for interpretation of multi-omic data: the input is multi-omic data and a background molecular network, and the output is a sparse subnetwork. Typically input biomolecules with higher scores (indicated by node sizes) and higher-quality connections (indicated by edge thickness) are prioritized in the output. (D) Network alignment: input can be individual homogeneous networks (*left*) or heterogeneous networks. Even alignment of homogeneous networks leads to a heterogeneous network (*right*) whose "supernodes" contain mapped nodes and whose edge types indicate which edges of the original networks are conserved (e.g., between supernodes "$a_1 \rightarrow b_1$" and "$a_2 \rightarrow b_2$" where the edge exists in both network 1 and network 2) versus nonconserved (e.g. between supernodes "$a_1 \rightarrow b_1$" and "$a_3 \rightarrow b_3$" where the edge exists in network 1 but not in network 2) under the given node mapping. (E) Inference of and reasoning on BKGs. Shown is a condition-aware BKG. The middle nodes (hexagons) are statement sentences. The layers on their left represent fact tuples and those on their right represent the conditions associated with the facts. The tuples have relation nodes (circles), concept nodes (squares), and optional attribute nodes (triangles).

approaches such as the Prize-Collecting Steiner Forest (Tuncbag *et al.* 2013) are computationally difficult to solve exactly but can be approximated efficiently. For instance, the Omics Integrator software (Tuncbag *et al.* 2016) based on the Prize-Collecting Steiner Forest algorithm adds prizes to nodes that should be included in the sparse subnetwork and costs to edges based on their reliability. Omics Integrator also includes a module to estimate prizes for active TFs based on chromatin accessibility, gene expression, and DNA-binding motifs. Its parameters control the tradeoff between node prizes and edge costs, a penalty for including nodes with high degrees, and a penalty for the number of connected components in the subnetwork.

Heterogeneous pathway reconstruction is especially powerful because network connections between different types of biomolecules can be combined to reveal more complete and explanatory pathways. For instance, a TF that activates differentially expressed genes detected with RNA-seq may be inferred to be regulated by an upstream phosphorylated kinase detected with mass spectrometry. A study of Kaposi's Sarcoma-associated Herpesvirus infection (Sychev *et al.* 2017) illustrates the data types and algorithms involved, and

biological insights gained in multi-omic pathway reconstruction. The authors profiled the proteomic and phosphoproteomic changes in endothelial cells induced by viral infection using mass spectrometry and gene expression changes with RNA-seq. They used TF binding motifs and a statistical enrichment test with the gene expression data to identify potentially relevant transcriptional regulators. Then, they applied Omics Integrator (Tuncbag *et al.* 2016) to combine the transcriptional regulators, proteomic changes, phosphoproteomic changes, and a PPI background network in order to obtain a holistic view of the endothelial cell response to infection. Ultimately, this analysis revealed peroxisome-related proteins to be an important part of the response. This network-based insight was supported with follow-up wet laboratory experiments (Sychev *et al.* 2017).

Network alignment. In network biology, network alignment has traditionally been used to compare species' PPI networks (Sharan and Ideker 2006, Faisal *et al.* 2015a, Emmert-Streib *et al.* 2016, Guzzi and Milenković 2017, Vijayan *et al.* 2020, Ma *et al.* 2022). In this context, network alignment aims to find a node (protein) mapping between the compared networks that uncover regions of high topological (and often sequence) conservation, with the hypothesis that the resulting aligned nodes and network regions are evolutionary conserved or functionally similar. Finding such a node mapping is closely related to the NP-complete subgraph isomorphism problem, making the network alignment problem NP-hard (Faisal *et al.* 2015a).

Even when comparing PPI networks, which are homogeneous, network alignment can be viewed as a multimodal data integration task. This is because an alignment (i.e. node mapping) in a "composed view" results in a heterogeneous (specifically, multiplex) network whose "supernodes" contain mapped nodes from the individual homogeneous networks and whose edges are of distinct types, indicating which one(s) of the compared networks the given edge is present in under the given node mapping (Fig. 3D). More recently, approaches have been proposed for aligning heterogeneous networks in biology (Gu *et al.* 2018, Milano *et al.* 2020) and other domains (Chen *et al.* 2016, Yan *et al.* 2022). Below, we discuss algorithmic principles of traditional alignment of homogeneous networks and then comment on the alignment of heterogeneous networks.

Analogous to sequence alignment, alignment of homogeneous networks can be local or global (Meng *et al.* 2016). Both have (dis)advantages (Guzzi and Milenković 2017). Also, network alignment can be pairwise (between exactly two networks) or multiple (between more than two networks) (Vijayan and Milenković 2018a). The latter has traditionally been expected to lead to deeper biological insights as it aligns all considered networks simultaneously as opposed to one pair at a time; however, a recent evaluation showed that this is not always the case (Vijayan *et al.* 2020). At the same time, multiple network alignment is computationally more complex (Vijayan and Milenković 2018a).

Network alignment has two main algorithmic components (Faisal *et al.* 2015b). First, topological similarity between nodes across the compared networks is computed via some measure of node conservation; graphlet-based measures (Section 4) are among state-of-the-art (Gu *et al.* 2018, Newaz and Milenković 2019). Second, an alignment strategy quickly identifies alignments that optimize some objective function accounting for total node and ideally also edge conservation

under the given node mapping. That is, a good alignment should both map similar nodes to each other and conserve many edges. Original alignment strategies were of the seed-and-extend type (Singh *et al.* 2008, Kuchaiev *et al.* 2010, Sun *et al.* 2015). The extension around highly similar "seed" nodes, by adding mapped nodes incrementally to build the alignment one step at a time, is intended to explicitly improve node conservation of the resulting alignment, but edge conservation only implicitly. To improve edge conservation explicitly as the alignment is constructed, rather than only evaluating it after the fact, another type of alignment strategy—a search algorithm—was introduced. Here, entire alignments are explored, and the one that scores the highest based on the given (e.g. edge conservation-based) objective function is returned, using, e.g. genetic algorithms (Saraph and Milenković 2014, Vijayan *et al.* 2015, 2017, Vijayan and Milenković 2018a) or simulated annealing (Mamano and Hayes 2017).

A recent algorithmic shift in network alignment has been from unsupervised to supervised, data-driven alignment (Gu and Milenković 2020, 2021). Traditional network alignment uses the notion of topological similarity to quantify how close to isomorphic two nodes' extended network neighborhoods are. A major issue is that regardless of the considered similarity measure, aligned nodes often do not correspond to nodes that should actually be mapped, i.e. that are functionally related (Gu and Milenković 2020). Specifically, when comparing species' PPI networks, aligned nodes do not correspond to proteins that are involved in same biological processes. This is why a move was made from optimizing topological similarity to learning from the data what kind of topological relatedness corresponds to functional relatedness, without assuming that topological relatedness means topological similarity (Gu and Milenković 2020). For example, topological similarity will aim to match a triangle in one network to a triangle in another network, and a square in the former to a square in the latter. Yet, due to biological variation or noise in PPI data, perhaps it is the triangle in the first network that is functionally related and should thus be matched to the square rather than the triangle in the second network, which is what topological relatedness would aim to learn from the data. This resulted in moving from traditional unsupervised alignment (functional labels of nodes, e.g. biological processes of proteins in PPI networks, being used to evaluate an alignment only after it is produced) to supervised, data-driven alignment (functional labels of nodes being used during the process of constructing an alignment, to learn patterns of topological relatedness). A pioneering data-driven network alignment method used traditional machine learning, i.e. user-predefined (graphlet-based) features (Gu and Milenković 2020, 2021) and standard classifiers (e.g. logistic regression), while a follow-up effort used deep learning and specifically GNNs (Ding *et al.* 2023).

Finally, going back to alignment of heterogeneous networks, an earlier attempt in biology was still to align homogeneous networks to each other, where the heterogeneity came from the fact that the individual homogeneous networks being compared were of different types: one was a human PPI network whose nodes were proteins, and the other was a disease–disease association network whose nodes were diseases (Wu *et al.* 2009). Then, the goal of aligning the two networks was to identify causative genes/proteins and their pathways underlying disease families. But, because each of

the compared networks was homogeneous, a homogeneous network alignment approach sufficed for their comparison. A more recent effort towards actually aligning one heterogeneous network to another, each with different node and edge types (or colors), was extending the existing notions of homogeneous graphlet-based node similarity/conservation as well as homogeneous edge conservation (discussed above) into their heterogeneous (or colored) counterparts, and then extending the existing seed-and-extend or search alignment strategies (discussed above) to find high-scoring alignments with respect to the new heterogeneous conservation measures (Gu *et al.* 2018). In evaluations on synthetic and real biological networks, the heterogeneous methods led to higher-quality alignments and better robustness to noise in the data than their homogeneous counterparts (Gu *et al.* 2018). Two types of heterogeneous biological networks were considered: first, PPI networks were aligned to each other, where nodes (proteins) were colored according to whether they were involved in aging, cancer, and/or Alzheimer's disease; second, protein-GO term networks were aligned to each other, where such a network had two types of nodes—proteins and GO terms—and three types of edges—PPIs, protein-GO term annotations, and GO term-GO term semantic similarity associations (Gu *et al.* 2018). This effort (Gu *et al.* 2018) aligned heterogeneous networks globally. In parallel, an approach for their local alignment was proposed (Milano *et al.* 2020).

Ideas from machine learning-based embedding of heterogeneous networks (Section 5) in biology (Pio-Lopez *et al.* 2021) and other domains (Wang *et al.* 2022d, 2023c) could be extended to heterogeneous network alignment. However, to our knowledge, such extension has not yet been carried out in biology but it has been carried out in other domains such as social, information, or technological networks (Zheng *et al.* 2018, Zhang *et al.* 2019b, 2020c, Xiong *et al.* 2021, Wang *et al.* 2022f, Cai *et al.* 2023). Note that in Zhang *et al.* (2019b, 2020c), the heterogeneity of considered networks came from node/edge attributes rather than explicit node/edge types. In these two studies, GNNs were used to first find an embedding of nodes of the compared networks, and then the network alignment problem was viewed as a point registration problem (Zhang *et al.* 2019b) or a neural network transformation problem (Zhang *et al.* 2020c).

Inference of and reasoning on BKGs. BKGs, which describe semantic relationships between biomedical entities, are among the richest examples of heterogeneous networks (Nicholson and Greene 2020). BKGs aim to combine facts about diverse biomedical entities, which can range from genes to individual patients as well as measurements associated with them. BKGs represent biological facts using "subject–predicate–object" triples as the fundamental unit, with the subject and object corresponding to nodes in the graph and the predicate (also called a relation) corresponding to a directed edge, possibly of different types, between the nodes. For example, Chlorin e6-PDT (subject) reduced (predicate) cell proliferation (object); Fig. 3E. Exemplar active BKG projects, each taking a unique approach, include Scalable Precision Medicine Knowledge Engine (SPOKE) (https://spoke.rbvi.ucsf.edu) (Morris *et al.* 2010), BioThings Explorer (https://explorer.biothings.io) (Fecho *et al.* 2022, Lelong *et al.* 2022), biomedical "corner" of Wikidata (https://www.wikidata.org) (Manske *et al.* 2019, Waagmeester *et al.* 2020, Page 2022), and PrimeKG (Chandak *et al.* 2023).

BKGs have emerged as powerful frameworks for diverse biomedical applications (Nicholson and Greene 2020) including drug repurposing (e.g. Hetionet; Himmelstein *et al.* 2017 and SPOKE; Morris *et al.* 2010), rare disease diagnosis (Alsentzer *et al.* 2022), and biomarker discovery (e.g. SPOKE; Himmelstein and Baranzini 2015). BKGs leverage graph databases like Neo4j and Virtuoso, and semantic web standards like the Resource Description Framework for their backend. BKGs leverage over a hundred years of graph theory to enable operations on first neighbors, paths, centralities, and other network components, as well as semantics, inference, and reasoning. There are a number of computational challenges that emerge to maximally extract the information encoded in BKGs for diverse biomedical applications ranging from construction of BKGs to reasoning with BKGs (Peng *et al.* 2023). For example, advanced, multi-hop queries specifying node and edge types are essential to navigating heterogeneous network representations of biomedical knowledge; "multi-hop" refers to having to traverse at least two edges in the graph. Many of these challenges have been approached using similar methods of network inference as previously described (e.g. link prediction) as well as more recently with graph representation learning approaches discussed in Section 5.

Equally important is the question of the representation of biomedical and biological literature to enable advanced queries and reasoning. Traditional BKGs assume that all knowledge can be represented as subject–predicate–object tuples and are constructed using tuple extraction techniques based on machine learning. A simple postprocessing algorithm can extract the tuples from any sentence and represent them as links between nodes on the BKGs. However, traditional BKGs have ignored the conditions (e.g. patient age or environment) of the facts, which capture essential contexts for knowledge exploration and inference. Recently, a new type of BKG, Condition-aware BKG (CondBKG; Jiang *et al.* 2021), has been introduced, which considers both facts and their conditions in the biomedical statements. Unlike traditional BKGs which have only one layer of subject–predicate–object tuples, CondBKG is a three-layered information-lossless representation of BKGs. The first layer has biomedical concept and attribute nodes; the second layer represents both biomedical fact and condition tuples by nodes of the predicate phrases, connecting to the subjects and objects in the first layer; the third layer has nodes that represent statement sentences as their textual attributes and connect to fact and/or condition tuples in the second layer (Fig. 3E). CondBKG is constructed from a machine learning model's output tuples. Given a statement sentence and its context (e.g. nearby sentences) in a scientific article, the model learns from multiple types of input signals of sentence (e.g. word embeddings and part of speech tags) and predicts one or multiple tuples. CondBKG has 18.1 million fact tuples, 7.5 million condition tuples, 10.9 million concept nodes, and 703 000 attribute nodes. CondBKG preserves more knowledge from unstructured text than traditional flat BKGs and can be used to answer tailored queries, such as what factors increase or reduce cell proliferation and their conditions (Fig. 3E). CondBKG can provide a good understanding of biomedical and biological statements and supports diverse applications for biomedical knowledge discovery.

Network-of-networks analysis. Biological systems function at different scales of organization. Thus, network-of-

networks analysis (Fig. 3A) is an exciting, still relatively unexplored area of research. This topic has received an increasing amount of attention only in recent years. This is likely because it has been increasingly recognized that network-of-networks representations of various biological data can be obtained: (i) given that different diseases tend to manifest in different tissues, nodes (diseases) in a disease similarity network can be represented as their associated tissue-specific PPI networks (Ni *et al.* 2016); (ii) nodes in a PPI network can be represented as protein structure networks (Gu *et al.* 2022, Gao *et al.* 2023); (iii) nodes in a network of interacting molecules can be represented as molecular graphs (Wang *et al.* 2021a, 2022a); (iv) nodes in a bipartite graph containing interactions between drugs and their target proteins can be represented as drug molecule graphs and target protein structure networks, respectively (Chu *et al.* 2022). Note that not all existing network-of-networks studies originate in the biology domain. Some have been proposed and evaluated in other domains, such as on text and social network datasets (Li *et al.* 2022a).

The studies that have analyzed biological network-of-networks data typically perform different network analysis and application tasks, as follows. The task of node ranking was applied to candidate disease gene prioritization from the network-of-networks of type (i) above (Ni *et al.* 2016). The task of link prediction was applied to predicting interactions between proteins from the network-of-networks of type (ii) above (Gao *et al.* 2023), between molecules such as drugs from the network-of-networks of type (iii) above (Wang *et al.* 2021a, 2022a), or between drugs and their target proteins from the network-of-networks of type (iv) above (Chu *et al.* 2022). A new task was introduced—that of entity label prediction—which merges the two traditionally isolated tasks of node (protein) classification at the higher scale containing a PPI network and graph (also protein) classification at the lower-scale containing protein structure networks (Gu *et al.* 2022). This task was applied to predicting protein functions from the network-of-networks of type (ii) above (Gu *et al.* 2022). Given that the different approaches were proposed for different tasks/applications, they have typically not been evaluated against each other. It remains unclear whether the different approaches can be effectively used in tasks/applications other than those they were proposed for, as well as what (dis)advantages of each approach are on the methodological level. With the increasing availability of network-of-networks data and the increasing number of approaches for network-of-networks analysis, the need for proper method evaluation will only continue to gain importance. This will require all studies to make their data and code publicly available and easy to use. According to our exploration of the existing network-of-networks studies discussed above, this is not always true.

4. Higher-order network analysis

Need for higher-order graph representations of biological systems. This article, unless explicitly noted otherwise, deals with traditional pairwise graphs (or simply graphs). Such a graph represents the organization of a biological system as a network of pairwise interactions between biomolecules (e.g. a PPI is represented as an edge connecting two proteins, and a transcriptional regulatory interaction is represented as a directed edge from a TF to a gene). However, these interactions often involve additional components and the interactions themselves can be regulated by other components (Battiston *et al.* 2020). In other words, there is often a need to capture interactions between multiple (two or more) nodes rather than between exactly two nodes (as is the case with pairwise graphs). Several higher-order graph ideas have been proposed in the literature to overcome the limitations of pairwise graphs. There are two general categories of such ideas.

The first category still works with pairwise graphs but relies on either higher-order dependencies between two nodes (Xu *et al.* 2016) or small subgraphs (Newaz and Milenković 2019), as follows. Regarding higher-order dependencies, it was shown that when representing sequential data such as global shipping traffic as networks, assuming the first-order dependency, i.e. that the next movement of traffic depends only on the current node, and thus discounting the fact that the movement may depend on several previous steps, can yield inaccurate network analysis results (Xu *et al.* 2016). This is because data derived from many complex systems can show up to fifth-order dependencies between two nodes. Consequently, an approach was proposed for capturing variable orders of dependencies between pairs of nodes (Xu *et al.* 2016). Regarding subgraphs, these can be viewed as "higher-order coordinated patterns" between two or more nodes of a pairwise graph (Battiston *et al.* 2020); a subgraph captures first-order dependencies [as discussed above and defined in Xu *et al.* (2016)] between multiple nodes in a pairwise graph. Examples of subgraph types are cycles (e.g. a triangle or a square) or cliques (the densest of all subgraph types, containing all possible edges between their nodes) (Battiston *et al.* 2020). Two general categories of subgraphs exist: graphlets (Pržulj 2007) and network motifs (Milo *et al.* 2004). Two key differences exist between them: graphlets are induced subgraphs while network motifs are not, and network motifs need to be statistically significantly overrepresented in a pairwise graph compared to a null (i.e. random graph) model while graphlets do not rely on a null model.

Both higher-order dependencies and subgraphs in pairwise graphs from the first category fail to directly account for interactions between more than two nodes in a network. An alternative is the second category of higher-order graph ideas—to explicitly consider higher-order graph structures. Here, while simplicial complexes are a theoretic possibility, they have assumptions that are practically too strong in some systems (Battiston *et al.* 2020). The next most general idea of higher-order interactions that is at the same time less constraining and thus more practical are hypergraphs (Battiston *et al.* 2020).

Higher-order dependencies [as discussed above and defined in Xu *et al.* (2016)] have not yet received attention in the biology domain, which is why we do not discuss this idea further. Graphlets in pairwise graphs (or simply graphlets), hypergraphs, and graphlets in hypergraphs (i.e. hypergraphlets) have received significant attention in the biology domain, which is why the following sections discuss these topics in more detail. While network motifs have also received attention, it remains unclear which random graph model fits real-world networks the best and should thus be used for network motif identification (Artzy-Randrup *et al.* 2004, Newaz and Milenković 2019), which is why we do not discuss network motifs further.

Graphlets. Graphlets, small subgraphs, are Lego-like building blocks of a network. More formally, they are connected,

nonisomorphic, induced subgraphs of a graph (Pržulj *et al.* 2004). Because counting of large graphlets in a large network is time-consuming, in practice, graphlets on up to five nodes have typically been studied. Graphlets were originally proposed as subgraphs of undirected, homogeneous, static, unordered, and pairwise graphs (Newaz and Milenković 2019). More recently, they were extended to their directed (Lugo-Martinez and Radivojac 2014, Sarajlić *et al.* 2016), heterogeneous (Gu *et al.* 2018), dynamic (Hulovatyy *et al.* 2015), ordered (Malod-Dognin and Pržulj 2014, Faisal *et al.* 2017), or hypergraph (Gaudelet *et al.* 2017, Lugo-Martinez *et al.* 2021) counterparts, respectively; the latter are called hypergraphlets and are discussed more below after hypergraphs are introduced. The following concepts are discussed for original graphlets, but they generalize to the more data-rich counterparts as well.

In a graphlet, nodes can correspond to different symmetry groups called automorphism orbits (or just orbits for simplicity) (Pržulj 2007). For example, in a graphlet corresponding to the 3-node path (e.g. $a-b-c$), the two outer nodes (a and c in our illustration) are symmetric to each other and thus belong to the same orbit, while the middle node (b) is in its own orbit. As another example, in a clique, all nodes are symmetric to each other and thus belong to the same orbit. There are 15 orbits for 2- to 4-node graphlets and 73 for 2- to 5-node graphlets. This concept of graphlet orbits can be used to quantify a node's extended network neighborhood into a 15D or 73D embedding, often called the node's graphlet degree vector (GDV) (Milenković and Pržulj 2008). This vector counts how many times a node of interest touches (or participates in) each of the considered graphlets at each of their orbits. By computing GDV for each node in a network, one can obtain the network's GDV matrix, whose entry (i,j) contains the information of how many times node i touches orbit j (Milenković and Pržulj 2008, Newaz and Milenković 2019). Note that there exists an analogous concept of edge (rather than node) as well as node pair orbits, GDVs, and GDV matrices (Solava *et al.* 2012, Hulovatyy *et al.* 2014).

GDV matrices of networks have been used as features to compare extended neighborhoods of nodes (edges, node pairs) in the same network, extended neighborhoods of nodes (edges, node pairs) across different networks, or structures of entire networks (Newaz and Milenković 2019). These, in turn, have been used in numerous computational tasks, such as network alignment, alignment-free network comparison, graph classification, node classification, network denoising via link prediction, inference of a condition-specific network or pathway reconstruction, network clustering, and node centrality computation, as well as for various application problems, such as studying human aging, protein folding and function, cancer and other diseases, pathogenicity, or mental health (e.g. depression and anxiety), as briefly discussed in other sections (Solava *et al.* 2012, Newaz and Milenković 2019, Liu *et al.* 2020, 2021a, Magnano and Gitter 2021, Newaz *et al.* 2022, Arici and Tuncbag 2023).

Hypergraphs. Hypergraphs provide powerful representations by generalizing edges between exactly two nodes to hyperedges that involve multiple nodes (Berge 1985). For example, protein complexes, which involve simultaneous interactions among multiple proteins that carry out function only as a group, are effectively represented using undirected hypergraphs, where each node is a protein and each undirected hyperedge (a set of nodes) is a complex (Klamt *et al.* 2009).

Under this representation, complexes that share interactors can be disambiguated, thus allowing more flexibility to capture multiple functionalities on the same set of nodes. Signaling pathways, on the other hand, are represented using directed hypergraphs in which proteins are represented by nodes and reactions are represented by directed hyperedges (Ritz *et al.* 2014).

Fig. 4 shows an example of nine reactions from the transforming growth factor-beta (TGFβ) signaling pathway (Gillespie *et al.* 2022) and their representation using higher-order graph frameworks. In this example, TGFβ1 binds to the TFGβ receptor and phosphorylates SMAD2/3, which in turn binds to SMAD4; SMAD2/3 is subsequently dephosphorylated by MTMR4. The signaling reactions are captured by a directed hypergraph with nine hyperedges connecting proteins (which may be phosphorylated) and protein complexes (Fig. 4A). Without the directed hyperedges, we have a series of overlapping protein complexes, the structure of which provides some insights into how the protein complexes form (Fig. 4B). Directed and undirected hypergraphs offer more information than a graph that only captures pairwise physical interactions in this cascade (Fig. 4C). If dealing with the pairwise graph representation in Fig. 4C, graphlets can help characterize the local topology of a specific node (Fig. 4D) or an entire network, as discussed above. If dealing with the hypergraph representation from Fig. 4A and B, hypergraphlets, discussed below, can be used to quantify topology (Fig. 4E).

A shortcoming of pairwise graphs in representing multi-component interactions is that some paths may be lost (Murgas *et al.* 2022) or ghost paths can be created (Pandey *et al.* 2007) while contracting a multi-way interaction into a set of pairwise interactions. For example, as seen in Fig. 4A, the interaction between TGFβ1 and SMAD2/3 occurs when TGFβ1 is part of the TGFβ complex that is phosphorylated, but this information is lost in the pairwise graph representation shown in Fig. 4C. In addition, contracting multi-way interactions into pairwise interactions results in the replication of interactions between multiple components, inflating subgraph density, multiplicity of paths, and node degrees; while also shortening paths. Generalization of notions such as density or centrality to hypergraphs can therefore provide more reliable insights into the topology and dynamics of biological networks (Feng *et al.* 2021).

In addition to reducing representation loss, hypergraphs also offer meaningful algorithmic advantages. Owing to the graph duality property where each graph can be represented as a hypergraph by inverting nodes and edges of the original graph into hyperedges and nodes, respectively, of a dual graph, hypergraph representations offer a possibility to unify methodology. For example, node classification, edge classification, and link prediction on pairwise graphs can all be seen as node classification on (extended) dual hypergraphs (Lugo-Martinez *et al.* 2021). This allows for the development of general methodologies and software that could support statistical inference tasks on biological networks.

To date, the application of hypergraphs in biological network analysis is limited because of constraints posed by the availability of data and annotations (or lack thereof). In cellular signaling, posttranslational modifications play a central role in multi-way interactions among cellular components, yet only a small fraction of posttranslational modifications are well-characterized (Needham *et al.* 2019). As

biotechnology advances and more data are generated, the availability of algorithms that solve fundamental problems on hypergraph representations, therefore, has the potential to guide data generation and curation of annotations.

Hypergraph algorithms. In the broader computer science community, hypergraph algorithms exist for several problems including shortest paths, random walks, and clustering (Cambini *et al.* 1997, Zhou *et al.* 2006, Ducournau and Bretto 2014, Gao *et al.* 2015a, Ausiello and Laura 2017). Within the context of network biology, hypergraphs have been used to study metabolic networks (Klamt *et al.* 2009), clusters in PPI networks (Ramadan *et al.* 2004), and shortest paths in signaling pathways. This final application is the best-developed use of directed hypergraphs in network biology. Hence, we focus our discussion on it.

Defining reachability in directed hypergraphs is significantly more complex than in pairwise graphs. A key principle is that the nodes in the head of a hyperedge are reachable from some source only if *all* the nodes in the tail are themselves reachable from that source. This principle expresses the natural concept that for any product of a reaction to form, all the reactants must be present. The notion of B-reachability formalizes this idea (Ritz *et al.* 2014, Ausiello and Laura 2017). The challenge now is that computing B-hyperpath with the smallest number of edges is an NP-complete problem, even when the tail and head of each hyperedge contain at most two nodes and we are interested only in acyclic hyperpaths (Ritz *et al.* 2014). An initial approach proposed a mixed-integer linear program to compute optimal hyperpaths (Ritz *et al.* 2014), applying it with success to the Wnt signaling pathway in the NCI Pathway Interaction Database. In practice, a drawback of this method was that a very large number of nodes without any incoming hyperedge had to be included among the sources for any meaningful hyperpath to exist. A later technique relaxed the definition of B-hyperpath (Franzese *et al.* 2019) to address this problem. As another alternative, an efficient heuristic approach can handle cyclic hyperpaths and compute optimal ones in practice (Krieger and Kececioglu 2022b). An exact cutting-plane algorithm can also compute the shortest hyperpaths with cycles while being efficient in practice on both the NCI Pathway Interaction Database and Reactome (Krieger and Kececioglu 2023). Finally, similar problems have been studied in the context of metabolic networks. Here, the notion of the shortest path is generalized to a factory, which also takes reaction stoichiometry into account. A mixed-integer linear program can find factories with the fewest reactions and accommodate negative regulation (Krieger and Kececioglu 2022a).

Statistical learning on hypergraphs. Hypergraphs can be approximated by pairwise graphs (e.g. star expansion, clique expansion; Agarwal *et al.* 2006), but such approximations do not retain all properties of the original hypergraphs (e.g. the cut properties; Ihler *et al.* 1993). Therefore, methods directly developed for learning on hypergraph data can offer practical advantages. A number of such approaches have emerged (Cong *et al.* 1991, Wachman and Khardon 2007, Leordeanu and Sminchisescu 2012, Chitra and Raphael 2019, Lugo-Martinez *et al.* 2021, Maleki *et al.* 2022, Antelmi *et al.* 2023); however, accurate learning on hypergraphs is often hindered by NP-hardness issues (Gärtner *et al.* 2003, Hein *et al.* 2013, Purkait *et al.* 2017) and, thus, methods developed

to directly deal with hypergraph data often trade accuracy for scalability.

A common theme in statistical learning on hypergraphs is finding a typically high-dimensional representation, or an embedding, of the data, and subsequently applying traditional machine learning to learn some concept; see Section 5 for more details. These methods can work at the level of entire graphs for graph classification, or at the level of nodes (edges), for node (edge) classification and link prediction. A well-known graph classification problem is the prediction of toxicity of chemical molecules (Vishwanathan *et al.* 2010), where the nodes are atoms, and the edges are bonds, both of different types, or prediction of protein function (Borgwardt *et al.* 2005). Examples of popular node/edge classification problems are function prediction for proteins/protein complexes in PPI networks or for amino acid residues in protein structure networks (Vacic *et al.* 2010, Lugo-Martinez *et al.* 2016). An example of a link prediction problem is the task of denoising and completion of the PPI network itself, as also discussed in Section 2.

Embeddings are often formalized via kernel-based approaches or representation learning (Section 5), thus allowing the practitioners to use both finite- and infinite-dimensional representations. Well-performing kernel approaches (kernels are symmetric, positive semidefinite similarity functions defined on pairs of objects, that allow efficient learning; Shawe-Taylor and Cristianini 2004) include random walks (Wachman and Khardon 2007) and hypergraphlet counting (Lugo-Martinez *et al.* 2021). Hypergraphlets are typically defined as small, connected, (rooted) hypergraphs, often with a finite number of node and edge types (Lugo-Martinez *et al.* 2021). They are a nontrivial extension of (pairwise) graphlets discussed above (Pržulj *et al.* 2004, Pržulj 2007, Milenković and Pržulj 2008, Shervashidze *et al.* 2009, Vacic *et al.* 2010, Lugo-Martinez and Radivojac 2014), with both illustrated in Fig. 4D and 4E. As with graphlets, the appeal for counting hypergraphlets derives from the graph reconstruction conjecture (Bondy and Hemminger 1977). Though proved only for certain types of graphs (e.g. trees), the graph reconstruction conjecture postulates that a large graph of size n can be reconstructed up to isomorphism from the counts of all subgraphs up to the size of $n - 1$. A stronger version of the conjecture allows for such reconstruction for subgraphs up to the size of some $k < n - 1$. Under these conditions, hypergraphlet counting approaches can lead to embeddings that allow universal approximation on hypergraph data. Another approach, relying on neural-network graph embeddings, allows for scaling hypergraph-based approaches to very large graphs (Maleki *et al.* 2022).

Additional approaches for hypergraphs exist, which are based on deep learning (Gui *et al.* 2016, Tu *et al.* 2018). Among these, a prominent example utilizes a GNN based on self-attention to effectively learn embeddings of the nodes and predict hyperedges for non-k-uniform heterogeneous hypergraphs, enhancing the generalizability (Zhang *et al.* 2020b). This approach and its extensions have been applied to studying chromatin biology (Zhang and Ma 2020, Zhang *et al.* 2022a) and predicting genetic interactions for a group of genes, specifically trigenic interactions, thereby significantly expanding the quantitative characterization of higher-order interactions (Zhang *et al.* 2020a).

Limitations. Three major issues confront the wide adoption of hypergraph-based representations in network biology.
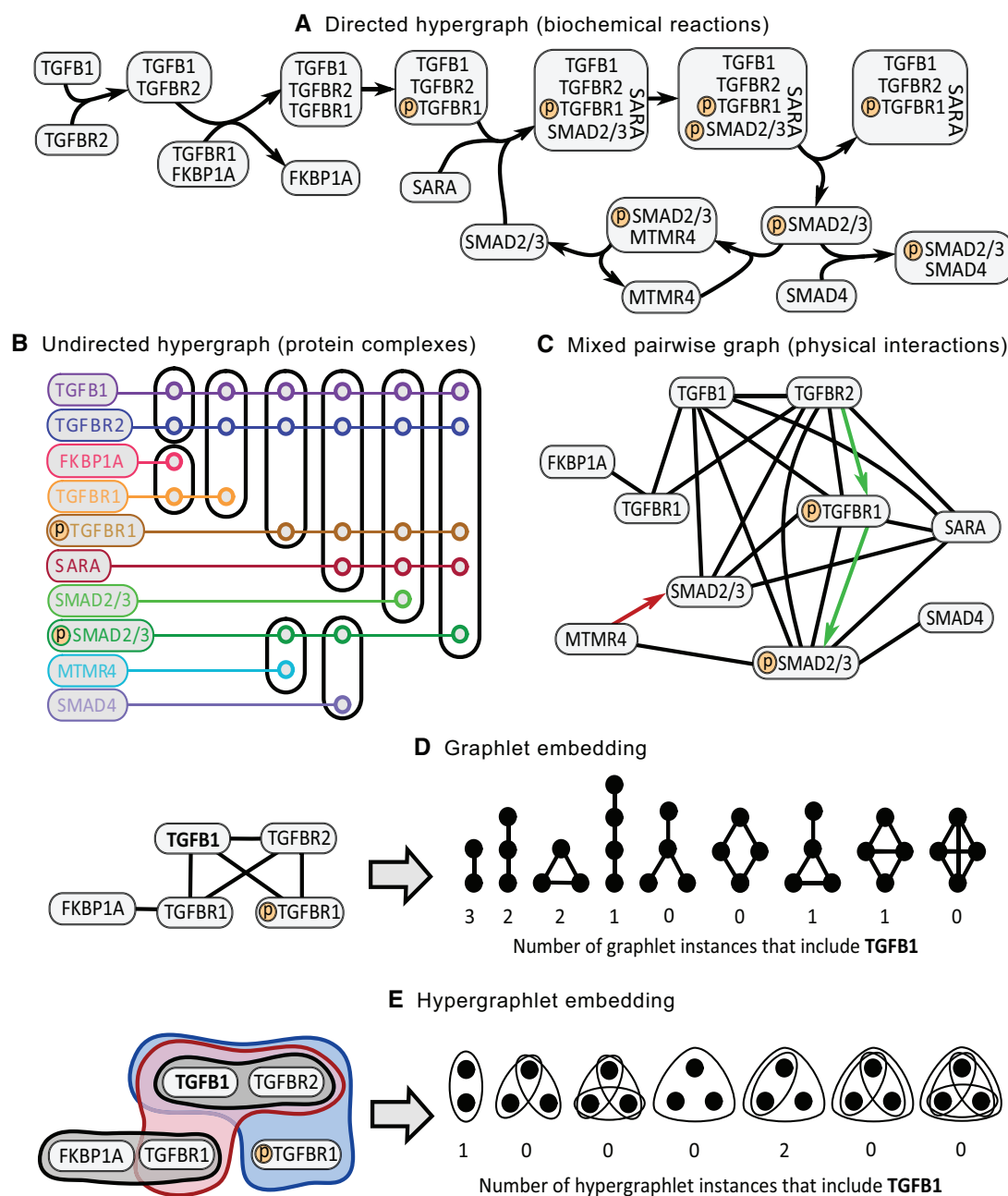
Figure 4. Graph representations of nine reactions from Reactome's TGFβ signaling pathway. (A) In a directed hypergraph, each hyperedge captures a reaction ("p" denotes phosphorylation). (B) In an undirected hypergraph, each hyperedge captures a protein complex. (C) In a (mixed) pairwise graph, each edge captures a pairwise interaction. "Mixed" refers to having both directed and undirected edges in the graph. Undirected edges denote physical interactions; directed edges denote either phosphorylation (the two right-most directed edges) or dephosphorylation (the left-most directed edge). (D) A node in a pairwise graph can be represented as a vector of graphlet counts. The number of 2-, 3-, and 4-node graphlet instances that include TGFB1 in the graph on the left are shown. (E) A node in an undirected hypergraph can be represented as a vector of hypergraphlet counts. The number of 2- and 3-node hypergraphlet instances that include TGFB1 in the hypergraph on the left are shown. In panels (D and E), only the (hyper)graphlet-level counts are shown for simplicity, i.e. (hyper)graphlet orbits are not shown nor considered when doing the counting. However, in practice, the more detailed orbit-level counts are computed rather than the (hyper)graphlet-level counts.

Databases such as Reactome (Gillespie *et al.* 2022) contain well-curated reaction networks that are amenable to representations as generalizations of directed hypergraphs. The first issue is that these resources remain incomplete and rely on manual curation. One promising direction of research is to analyze pairwise graphs to automatically infer reactions. An elegant example is an approach that uses properties of chordal graphs to convert a graph representation of a signaling pathway as a nested tree of protein complexes (Zotenko *et al.* 2006). A graph is chordal if every pair of nodes in every

cycle of length four or more is connected by an edge. Since PPI networks are not necessarily chordal, the authors augment them with additional edges, e.g. those that connect weak siblings, i.e. pairs of nodes that have identical neighbor sets but are themselves not connected by an edge. If the resulting graph is chordal, it admits a representation as a tree of cliques, which can be converted into a tree of complexes in the original graph by deleting the artificially added edges. This method was applied to the TNF-α/NF-κB and phoromone signaling pathways (Zotenko *et al.* 2006). To further

the use of hypergraphs in network biology, it will be important to generalize this method to apply to larger classes of graphs and to unify these methods of automated reconstruction with the results of manual curation. It may also be valuable to formulate hybrid network representations that combine the features of pairwise graphs and hypergraphs. A caveat here is that the need to develop a novel set of algorithms for every new representation might prevent its wide adoption in the community.

The second issue is that the theory for (directed) hypergraphs is much less well-developed than for pairwise graphs. Problems that have well-established and simple polynomial-time solutions on pairwise graphs, e.g. shortest paths, turn out to be computationally intractable on directed hypergraphs (Ritz *et al.* 2014), as discussed above. Incorporating regulation into the definitions of shortest paths continues to be challenging (Krieger and Kececioglu 2022a). Moreover, graph-theoretic concepts such as clusters, flows, random walks, or convolutions that have been employed fruitfully in network biology are either challenging to generalize to hypergraphs or have found limited applications in biology.

The third issue is that it is not clear under what conditions or for which applications a higher-order representation is better than a pairwise graph representation. Arguments often appeal to visual and qualitative reasoning (Fig. 4). We encourage the community to come forward with well-established datasets, evaluation measures, and benchmark frameworks that can pose these questions formally and develop generalizable standards.

5. Machine learning on networks

Overview. Machine learning has emerged as a powerful paradigm for creating predictive models specified as parameterized functions with tunable parameters that operate on structured data, such as graphs, spatial geometries, relational structures, and manifolds. Applying machine learning methods to network data has demonstrated potential in a myriad of biological network analysis tasks (Yue *et al.* 2020, Hetzel *et al.* 2021, Li *et al.* 2022b, Theodoris *et al.* 2023). Recent methods are designed to produce graph representations as compact numerical vectors (or embeddings) corresponding to various graph elements, such as nodes, edges, subgraphs, and entire graphs, and capture essential information about the topology of these elements. These learned representations can be fed into models trained toward a vast array of downstream analytic tasks.

Predictive models on graphs include models for predicting node labels (node classification), edge-level relationships (link prediction), subgraph-level labels (subgraph classification), and graph-level labels (graph classification) (Fig. 5). These models can be created through unsupervised, self-supervised, and supervised learning on all types of networks, including homogeneous, heterogeneous, temporal, and spatial networks, and with additional constraints and domain knowledge imposed on the models. By leveraging deep graph learning models pretrained on large-scale general graph datasets, it is possible to adapt (or fine-tune) pretrained representations for diverse use cases in predictive and generative modeling (Gainza *et al.* 2020, 2023). As machine learning on graphs continues to be developed, appropriate model benchmarking is necessary to ensure that task-specific evaluation measures are well-defined and predictions are fair and robust.

The rest of this section discusses these topics, which are also summarized in Fig. 5.

Unsupervised, self-supervised, and supervised graph learning. Unsupervised learning of graph representations involves optimizing parameterized strategies, such as GNNs, graph transformers, or multi-layer neural message-passing models, to aggregate information from a node's (e.g. a gene in a gene co-expression network or a patient in a patient similarity network) neighbors in the network. The goal is to optimize the representations so that the proximity between entities in the embedding space mirrors their proximity in the network (Cao *et al.* 2020, Atz *et al.* 2021). Prevalent strategies for sampling neighbors in the network vicinity of nodes that get embedded in the latent space include biased and unbiased random walks as well as adaptive neighbor sampling (Hamilton *et al.* 2017a, Veličković *et al.* 2019). Objective functions of these methods aim to maximize embedding similarity in the latent space for neighboring nodes in the network (Perozzi *et al.* 2014, Tang *et al.* 2015, Hamilton *et al.* 2017b, Hamilton 2020). For instance, nodes connected by edges should be embedded closer together in the latent space (i.e. have more similar embeddings) than nodes that are not connected (Grover and Leskovec 2016, Liu *et al.* 2022, Xie *et al.* 2022b, Wu *et al.* 2023).

Self-supervised graph representation learning, the predominant approach for machine learning on graphs, leverages not only the network structure but also additional context or auxiliary tasks to generate informative embeddings. Unlike unsupervised methods that solely rely on the network structure for optimization, self-supervised techniques utilize auxiliary (pretext) tasks, such as predicting node attributes or reconstructing graph substructures, to enhance the learning process and create more robust embeddings (Zitnik and Leskovec 2017, Zitnik *et al.* 2018, Hassani and Khasahmadi 2020, Li *et al.* 2022b). An example of a self-supervised node-level auxiliary task is predicting each node's degree. Link prediction is a self-supervised edge-level task that predicts whether an edge exists between a pair of nodes (Kipf and Welling 2016, Li *et al.* 2022b) based on a self-supervised objective (Liu *et al.* 2021b), which can be formulated using contrastive learning (You *et al.* 2020), node or edge masking (Agarwal *et al.* 2023), and generative denoising (Yi *et al.* 2024). Examples of self-supervised subgraph and graph tasks include predicting subgraph and graph properties, such as distributional statistics of shortest path lengths, network diameter, and the presence or absence of specific higher-order structures and graphlets (Alsentzer *et al.* 2020, You *et al.* 2020, Luo *et al.* 2022b).

Graph representation learning, whether unsupervised or self-supervised, can be applied to any type of network, including but not limited to homogeneous, heterogeneous, temporal, spatial, and physical networks. For example, in heterogeneous networks, GNN and graph transformer models leverage node- and edge-based attention weights to aggregate neighborhood information depending on node and edge types (Wang *et al.* 2019, Zhang *et al.* 2019a, Xie *et al.* 2020, Fu *et al.* 2022, Kesimoglu and Bozdag 2023b). Other approaches treat each edge type as a homogeneous graph, apply a graph representation learning model to it, and then integrate edge-type specific node representations into final representations (Wang *et al.* 2021b, Fu *et al.* 2022, Kesimoglu and Bozdag 2023a, 2023b). In a heterogeneous network, subgraphs can be sampled via metapaths (Sun *et al.*

2011), which are defined by sequences of relationships (or edge types) connecting different types of nodes to model semantic nuances underlying the network in a self-supervised manner, such as through contrastive learning (Dong *et al.* 2017, Zhao *et al.* 2021). These advancements in graph representation learning have impacted areas like cancer biology, drug discovery, and disease diagnosis (Esteva *et al.* 2019, Stokes *et al.* 2020, Gysi *et al.* 2021, Huang *et al.* 2022, 2023a).

Supervised graph representation learning uses networks with additional expert-curated or experimentally derived labeled data to directly optimize models for specific prediction tasks (Fig. 5A). In this paradigm, nodes, edges, subgraphs, or entire graphs are associated with labels, and the learning process minimizes the discrepancy between the model's predictions and these labels (Schlichtkrull *et al.* 2018, Veličković *et al.* 2018). Common applications include node classification, where individual nodes are assigned to predefined categories, and graph classification, wherein entire graphs are categorized based on their topological features (Gilmer *et al.* 2017, Eyuboglu *et al.* 2023). Unlike unsupervised and self-supervised models, supervised graph learning directly uses label information, often leading to more task-specific and accurate representations, albeit at the cost of requiring labeled data.

Incorporating knowledge into machine learning models through knowledge graphs, spatial constraints, equivariances, and symmetries. In numerous biological and medical applications, standard graph representation learning often falls short of requirements. In these cases, the model's predictive accuracy can be enhanced by imposing constraints drawn from pre-existing knowledge. Typical strategies encompass incorporating multimodal data into BKGs, augmenting GNNs with bespoke architectures, and applying domain-specific invariances.

BKGs help model heterogeneous relationships between biomedical entities, as already discussed in Section 3. The resulting latent space, which reflects the topology of the underlying knowledge graph, can be operated on to make inferences about existing and novel relationships. Jointly modeling diverse types of relationships in a BKG, such as integrative modeling of transcription regulation and metabolism (Chandrasekaran and Price 2010, Niu *et al.* 2021), can present unique challenges due to the BKG's incompleteness and potential high-order relationships involving heterogeneous entities. Incorporating pathway knowledge, either implicitly as constraints that regularize network embeddings (Niu *et al.* 2021) or directly as a prior placed on the BKG structure and parameters in a Bayesian fashion (Boluki *et al.* 2017), has been shown to improve predictive performance. Supervised machine learning methods often require many samples to identify biologically meaningful patterns, which can limit their applicability in areas such as rare diseases that are inherently limited in clinical cases, leading to few samples to analyze (Banerjee *et al.* 2023). Advances in self-supervised graph learning applied to BKGs have shown promise for rare disease research (Alsentzer *et al.* 2022) and will likely be informative for applications beyond rare diseases for which few samples exist with high-dimensional data.

Temporal and spatial data can be represented as networks, but specialized neural architectures are necessary to learn optimally on temporal/dynamic networks. Temporal graph representation learning methods typically involve two main

components: a GNN architecture to generate embeddings for each time point and a recurrent neural network, such as a long short-term memory network or a transformer network, to perform sequence learning by leveraging temporal relationships between elements in the sequence. Existing approaches use GNNs as feature extractors of nodes and the underlying topology, and recurrent neural networks for temporal learning and to include additional metadata information (Li *et al.* 2018a, Manessi *et al.* 2020, Pareja *et al.* 2020, Peng *et al.* 2020, Zhao *et al.* 2020a). Recently, static GNNs have been extended to handle dynamic graphs by treating time points as hierarchical states (You *et al.* 2022) or applied to irregular time series data by propagating neural messages between time intervals of each sensor as well as between sensors (Zhang *et al.* 2022b). Protein molecular configurations can be depicted as protein structure networks where amino acid nodes are linked by the 3D physical proximity of their residues, and the amino acid spatial coordinate information is encoded as node attributes. Deep learning models, particularly through the use of equivariant GNNs, can both attain high performance and preserve transformations of protein networks under translation, reflection, and rotation of networks in the 3D space (Jumper *et al.* 2021, Batzner *et al.* 2022, Gong *et al.* 2023). For instance, to establish a model that remains invariant to molecular spatial orientation, constraints enforcing rotation invariance ought to be integrated (Jumper *et al.* 2021). Methodologies derived from equivariant neural networks, such as AlphaFold (Jumper *et al.* 2021), can complement sequence-based language models (Lin *et al.* 2023) by harnessing evolutionary data to infer protein structures from primary amino acid sequences, and potentially generate realistic molecular formations.

Generative graph models. Generative graph models are a class of machine learning models specifically designed to generate new graphs, or parts of graphs, that resemble a given set of training graphs in some way. These models learn to capture the underlying patterns and structures in the training graphs and can then be used to produce new graphs with similar properties as the training graphs. For example, in molecular biology, the inherently graph-like nature of molecular structures has made GNNs an ideal tool for generating drug-like molecules, guiding the generation process by learning the underlying patterns and properties from real molecular data (Bilodeau *et al.* 2022). One such method is a variational graph autoencoder that learns embeddings of molecular structures and uses them to generate novel molecular graphs (Kipf and Welling 2016, Jin *et al.* 2018, Li *et al.* 2018b). Other generative models, such as GraphVAE, GraphRNN, and MolGAN, have also been developed to generate realistic graphs (De Cao and Kipf 2018, Simonovsky and Komodakis 2018, You *et al.* 2018). Inspired by generative adversarial networks for image generation, MolGAN pits a generator model (which produces graphs) against a discriminator model (which tries to distinguish between real and generated graphs). Additionally, graph transformer networks have recently been proposed for molecular graph generation, demonstrating the ability to generate molecules with desired properties by training on extensive chemical databases (Bagal *et al.* 2021).

When applied to protein design, GNNs have demonstrated impressive results in designing protein sequences that fold into specific structures (Ingraham *et al.* 2019). Graph-based methods like PotentialNet have shown promise for protein–
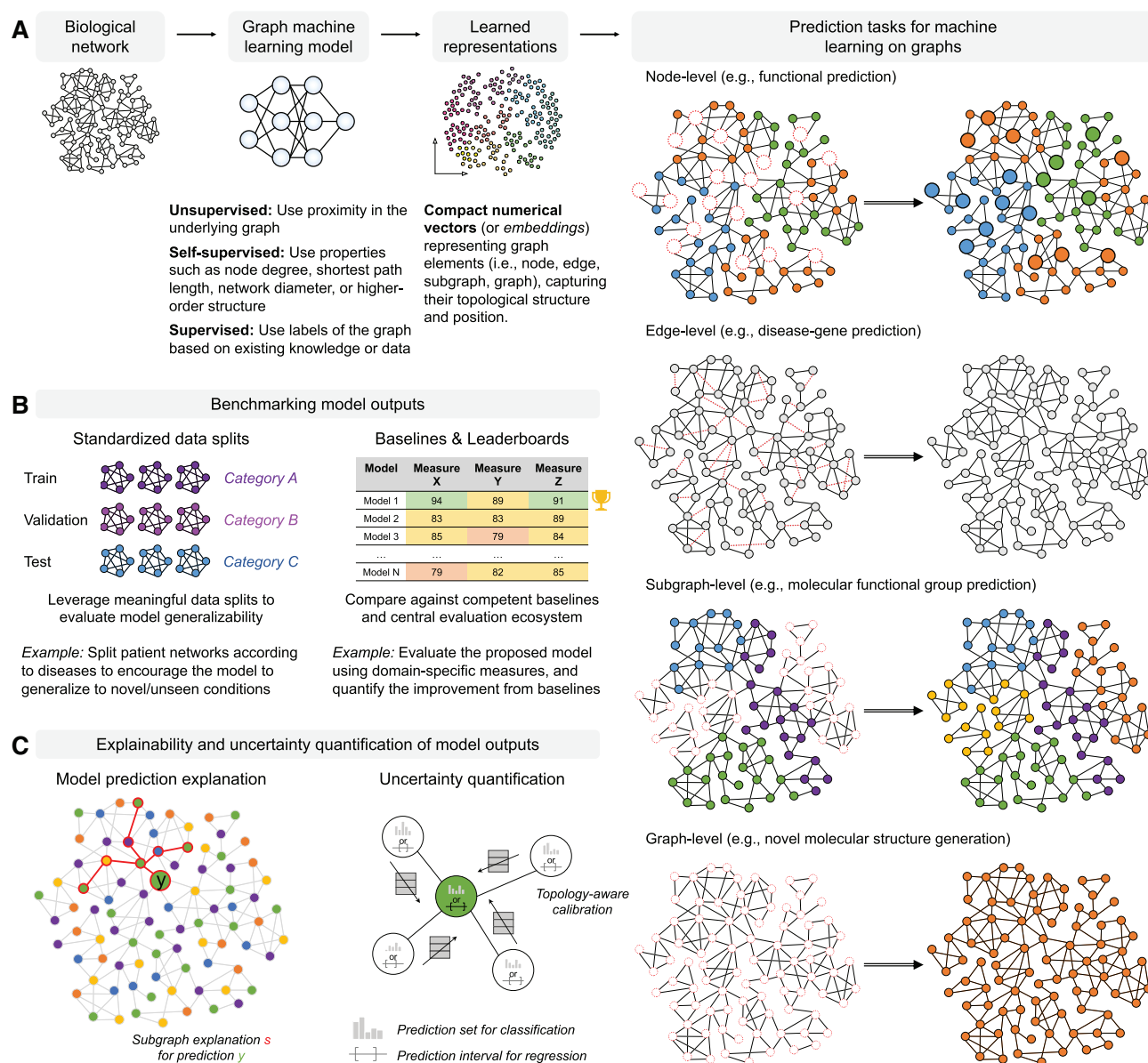
Figure 5. Overview of the components of machine learning on networks. (A) The core of this approach is a machine learning model, typically a neural network, that takes one or more biological networks as input and learns representations (i.e. embeddings) of various graph elements in an unsupervised, self-supervised, or supervised manner. There are four types of prediction tasks (denoted by the red dashed lines): node-, edge-, subgraph-, and graph-level predictions. Colors of nodes for the node-, subgraph-, and graph-level tasks signify the label; white nodes indicate missing labels to be predicted by the model. Examples include functional prediction (node-level), disease–gene prediction or context-specific edge prediction (edge-level), molecular functional group prediction (subgraph-level), and novel molecular structure generation (graph-level). Critical to continued development, wide adoption, and practical utility of network-based machine learning is a parallel improvement in frameworks for (B) rigorous benchmarking via established data splits and baselines, and (C) explainability of model predictions (e.g. identifying a subgraph *s*, denoted by red lines, that best explains the prediction *y* for the query node, denoted in green) and uncertainty quantification (e.g. using the prediction set for a classification task or prediction interval for a regression task; Huang *et al.* 2023b).

ligand binding prediction (Feinberg *et al.* 2018). Similarly, DeepSite uses 3D convolutional neural networks to predict protein–ligand binding sites (Jiménez *et al.* 2017). Moreover, recent generative models, such as ProteinMPNN (Dauparas *et al.* 2022) utilize message-passing neural network architecture to generate protein sequences and structures, further expanding the range of possibilities for protein design.

Diffusion models have recently emerged as powerful tools in protein and drug design (Corso *et al.* 2023, Watson *et al.* 2023, Abramson *et al.* 2024, Yim *et al.* 2024), leveraging their capability to model complex distributions for generating

novel molecular and protein structures. In protein design, diffusion models operate by gradually denoising a random configuration towards a target protein structure, learning the distribution of protein conformations. A notable example is RFDiffusion (Watson *et al.* 2023), a diffusion model that generates protein structures by conditioning on both sequence and structural information, achieving enhanced accuracy in structure prediction. In drug design, these models are adapted to generate molecular graphs by iteratively refining a random molecular graph into a drug-like molecule with desired properties through a learned diffusion process (Pinheiro *et al.* 2024).

Transfer learning. The quality of representations generated by graph representation learning methods is contingent upon the availability of labels. Nevertheless, labels are often in short supply due to the substantial resources required for their curation and validation. A potent solution to addressing this challenge is transfer learning. This approach involves initially training a graph representation learning model on a large reference network via self-supervised pretraining (Hu et al. 2020a, You et al. 2020, Li et al. 2022b, Xie et al. 2022b), followed by adapting the resulting model or its outputs to a different task of interest typically through supervised learning on a small set of labeled examples (fine-tuning). Pretraining a model on a large network followed by fine-tuning of the model using a small labeled dataset allows the model to harness extant information about a network entity (i.e. from the large network utilized for pretraining) in service of diverse tasks with limited task-specific labels.

Transfer learning has shown considerable potential for developing predictive models on condition-specific networks that vary with biological conditions. Networks are typically constructed from context-unaware data (e.g. the human reference PPI network; Luck et al. 2020) or data generated under specific conditions (e.g. a gene co-expression network for a particular disease). Biomedical entities and their interactions can vary across biological conditions, such as tissues, cell types, and disease states. Nevertheless, generalizing knowledge from context-unaware networks to context-specific problems presents considerable challenges. For instance, modeling tissue- or cell type-specific interactions from the human reference PPI network requires the construction of tissue- and cell type-specific networks and the development of multi-scale network models (Greene et al. 2015, Zitnik and Leskovec 2017, Ietswaart et al. 2021, Li and Zitnik 2021). One approach to this challenge involves constructing context-specific networks (as discussed in Sections 2 and 3) and applying independent shallow network embedding layers to learn node representations based on network topology and tissue hierarchical structure (Greene et al. 2015, Zitnik and Leskovec 2017). An alternative strategy is to learn shallow network embeddings on a context-unaware network, such that the embeddings of nodes operating in the same context are more similar to each other than nodes operating in different contexts (Ietswaart et al. 2021). Recent methods incorporate context in a data-driven manner, constructing cell type-specific PPI networks using single-cell transcriptomic data (Li and Zitnik 2021, Li et al. 2024). Unified by a network of cell type and tissue hierarchy, these networks can be harnessed to learn unique protein representations tailored to each cell type context (Li and Zitnik 2021, Li et al. 2024).

Understanding predictive models, benchmarking, and rigorous evaluation across diverse tasks. With the rapid evolution of graph learning methodologies, the need to construct rigorous benchmarks for effectively assessing the performance of these novel techniques is becoming increasingly urgent (Fig. 5B) (Shchur et al. 2018, Hu et al. 2020b). Open-science evaluation platforms such as the Benchmarking GNN (Dwivedi et al. 2022a), Open Graph Benchmark (Hu et al. 2020b, 2021), and others (Table 1) serve as significant assets for general graph benchmarking, while other resources are being curated explicitly for the domain of network biology (Liu and Krishnan 2024).

To provide a comprehensive evaluation, these resources ought to be expanded to include tasks defined at various levels of graphs, including node classification, link prediction, subgraph classification and clustering, and whole-graph classification and regression. In addition to benchmarking models for predictive tasks, evaluation frameworks are needed for generative graph models. They should also encompass diverse types of biological graphs, such as heterogeneous, spatial, and temporal ones. A critical element in this regard is benchmarking the performance of network-based machine learning techniques across multiple dimensions of evaluation beyond accuracy, including robustness, generalizability, and computational efficiency.

Moreover, the explainability of graph-based learning can offer significant insights in the biomedical domain (Fig. 5C) (Ying et al. 2019, Yuan et al. 2021, Xie et al. 2022a, Agarwal et al. 2023). Consequently, it is equally important to examine learned algorithms by examining pretrained graph representations (Forster et al. 2022) and mapping attention mechanisms in attention-based deep learning models (Elmarakeby et al. 2021). As we move towards the broader application of machine learning models in network biology, proper quantification of the uncertainty, error, and utility associated with these models is indispensable. Given the potential for considerable uncertainty in these models, effective techniques for uncertainty quantification are required to fully comprehend the predictive capabilities and limitations of a given model (Abdar et al. 2021).

When the model's objective is specific, such as treatment recommendation, disease diagnosis and prognosis, and steady-state or transient network behavior prediction, an objective-driven approach to uncertainty quantification can be beneficial (Yoon et al. 2013). This approach allows us to quantify uncertainty based on its impact on the expected performance of prediction and intervention tasks. Ultimately, this can pave the way for optimal experimental design techniques (Dehghannasiri et al. 2015a, 2015b) that prioritize experiments to generate the most informative data points selected by active learning strategies, effectively reducing model uncertainty.

6. Network-based personalized medicine

Overview. The overarching goal of precision medicine is to develop diagnostic and treatment strategies tailored to individual patients (Aronson and Rehm 2015, Kaiser 2015, Malod-Dognin et al. 2018), while also taking into account the desired level of precision for each treatment. Personalized characterization of an individual or a group can encompass various data types, including molecular, healthcare, environmental, lifestyle, and behavioral information, commonly modeled and analyzed as networks (Pržulj and Malod-Dognin 2016). By assimilating data from different modalities, precision therapeutics can amplify their potential and bolster resilience against diverse data noise (Gligorijevic et al. 2016a, Huang et al. 2021, 2022). Fusing data from multiple sources has proven effective in advancing precision medicine (Wang et al. 2014, Gligorijevic et al. 2016b, Malod-Dognin et al. 2019, Gaudelet et al. 2021).

Patient stratification. Precision medicine aims to provide individualized diagnostic and treatment strategies. Developing treatments tailored to specific patient groups based on distinct disease subtypes (Fig. 6A) is poised to transform a prevailing one-size-fits-all approach used in healthcare. Network methods can integrate multimodal data to identify patient groups with coherent genetic, genomic,

physiological, and clinical profiles (Gligorijevic *et al.* 2016b, Ektefaie *et al.* 2023, Petti and Farina 2023), even when the underlying data are incomplete and noisy (Pai and Bader 2018). The methods assume that patients with similar clinical signatures and similar -omics profiles have similar clinical outcomes. Similarities between patients can be efficiently represented through patient similarity networks; in these networks, nodes symbolize patients, and weighted edges denote the degree of similarity derived from clinical and biomolecular patient attributes. Each patient data attribute, such as age, sex, mutation status, or gene expression profile, can be used to create a network of pairwise patient similarities. Then, the set of all such networks can be viewed as a multiplex network, with a layer for each of the attributes. Various similarity measures can be employed to assess patient similarity across different datasets corresponding to different attributes. After the multiplex patient similarity network is constructed, patient subtypes can be identified by examining the community (clustering) structure within the network. Communities are characterized as subsets of nodes that are densely connected to each other and loosely connected to nodes in different communities (Fortunato 2010). Communities in a patient similarity network are thus densely/strongly linked patient groups and can shed light on distinct disease subtypes.

Network methods offer distinct advantages over nonnetwork approaches, which often grapple with the complexities of integrated datasets (Gligorijević and Pržulj 2015). Patient stratification has increasingly benefited from network-based methodologies, which can elucidate intricate biological interactions, especially within disease mutation landscapes, such as cancer (Gligorijevic *et al.* 2016b) or rare hereditary diseases (Malod-Dognin *et al.* 2023). By studying different types of gene–gene interactions, encompassing aspects like mutual exclusivity, co-occurrence, and both physical and functional associations, and analyzing personalized gene regulatory networks (Rogers *et al.* 2022), one can better understand interindividual variation in disease driven by differences in interactions caused by each patient's genetic background, environmental exposures, and the proportions of specific cell types involved in disease (Van Der Wijst *et al.* 2018). Such insights can elevate the accuracy of patient stratification, which is typically measured as the ability to classify patients as belonging to known disease subtypes (Pai *et al.* 2019) or the ability to identify disease biomarkers that generalize (maintain performance) when applied to new data that have not yet been seen by the model (Alsentzer *et al.* 2022, Kong *et al.* 2022). These insights can also guide the refinement of therapeutic strategies, ensuring they are optimally tailored to specific patient groups (Gligorijevic *et al.* 2016b, Dao *et al.* 2017, Huang *et al.* 2023a).

Identification of pathways associated with disease subtypes and patient groups. Identifying group-specific mutations provides valuable insights into the underlying biochemical pathways associated with the disease (Fig. 6B). These pathways can be conceptualized as networks, laying the foundation for an in-depth understanding of disease mechanisms. Incorporating individual mutation or expression data into pathway-based (i.e. network-based) methods aids in identifying targetable mutations (Park *et al.* 2019). This approach is especially pertinent in determining functional pathways that play roles in expression responses to disease-propagating mutations, leveraging the concept of pathway centrality (Windels *et al.* 2022a, 2022b).

For instance, by integrating genomic, clinical, and therapeutic data through networks, physicians can categorize patients with treatment-resistant prostate cancer based on specific gene mutations like AR, PTEN, and BRCA2. Recognizing these mutations facilitates the adoption of personalized therapies, targeting the aberrant pathways distinctive to each patient's tumor profile. As a result, this tailored treatment strategy offers the potential for safer and more effective treatments (Mateo *et al.* 2020).

Furthermore, recent research has illuminated the importance of tissue-specific regulatory networks and the pathways they encompass, which frequently manifest genetic mutations in particular patient cohorts. This understanding emerged from the combined analysis of expression and chromatin accessibility data, unveiling a previously unidentified tissue-specific stem-cell-like subtype of treatment-resistant prostate cancer that may be a target for intervention (Tang *et al.* 2022). Similarly, a comparative structural analysis of the chromatin structure network in chronic lymphocytic leukemia and control tissue of origin revealed that genes driving this cancer type are characterized by specific local wiring patterns not only in the chromatin structure network of chronic lymphocytic leukemia cells but also of healthy cells (Malod-Dognin *et al.* 2020). This allows for the successful prediction of new DNA elements related to this cancer type, and importantly, it shows that cancer-related DNA elements can be identified in other cancer types by investigating the chromatin structure network of the healthy cell of origin, a critical new insight paving the road to new therapeutic strategies (Malod-Dognin *et al.* 2020).

Identification of disease-dysregulated functional modules. Studying disease-dysregulated functional modules of genes can advance the understanding of disease beyond isolated mutations or pathway dysregulations. Disease-associated behaviors can materialize in clusters of tightly interacting proteins forming functional modules (Fig. 6B) (Menche *et al.* 2015, Agrawal *et al.* 2018) rather than exclusively via singular gene mutations or perturbed gene expression (Schadt 2009).

The quest to uncover disease-associated functional gene modules from molecular networks is a long-standing challenge with implications for precision medicine (Barabási *et al.* 2011, Mitra *et al.* 2013, Choobdar *et al.* 2019, Gaudelet *et al.* 2020, Eyuboglu *et al.* 2023, Morselli Gysi and Barabási 2023). Prevailing approaches for finding disease modules rely on the assumption that interacting genes tend to associate with similar phenotypes. For instance, gene co-expression network analysis has been employed to pinpoint modules of genes that exhibit analogous co-expression patterns in breast cancer. Notably, these clusters of genes correlate with distinct metastasis progression patterns in patients (Chuang *et al.* 2007). Multi-omic module detection in cancer can consider mutation mutual exclusivity, transcriptional regulation, and gene co-expression alongside PPI connections (Silverbush *et al.* 2019).

Given the complexity of disease circuits in many complex diseases, concentrated efforts have been directed toward identifying disease-associated gene modules that correlate with patient phenotypes (Saelens *et al.* 2018, Choobdar *et al.* 2019). Disease-associated gene modules, identified through computational approaches and various types of gene networks, have been used to refine disease diagnosis (Morselli Gysi *et al.* 2020). They can also forecast the response of individual cell lines to specific anticancer agents and potentially suggest patient-tailored drug combinations (Kim *et al.* 2020,

Table 1. Prominent open-source benchmark datasets for machine learning on biological networks.

Data type	Database	Task type	Prediction tasks
General	Long Range Graph Benchmark (Dwivedi et al. 2022b)	Edge-level	Molecular bond
		Graph-level	Peptide function, peptide structure
General	Open Biomedical Network Benchmark (Liu and Krishnan 2024)	Node-level	Protein function
		Edge-level	Disease–gene association
General	Open Graph Benchmark (Hu et al. 2020b)	Node-level	Protein function
		Edge-level	Protein–protein association, drug–drug interaction, heterogeneous interaction, vessels in mouse brain
		Graph-level	Molecular property, species-specific protein association
General	SubGNN Benchmarks (Alsentzer et al. 2020)	Subgraph-level	Proteins associated with biological process, rare neurological disorders phenotype-based diagnosis, and rare metabolic disorders phenotype-based diagnosis
General	Temporal Graph Benchmark (Huang et al. 2024)	Node-level	Dynamic node affinity prediction
		Edge-level	Dynamic link prediction
Knowledge graph	PrimeKG (Chandak et al. 2023)	Node-level	Identity of protein/gene, disease, drug, biological process, pathway, phenotype, molecular function, cellular component, exposure, and anatomical region
		Edge-level	Protein–protein interaction, disease–drug indication, disease–drug contraindication, disease-drug off-label use, disease–phenotype association, disease–disease association, disease–protein association, disease–exposure association, phenotype–protein association, pathway–gene association, etc.
Knowledge graph	Phenotype Knowledge Translator (Callahan et al. 2024)	Node-level	Identity of tissue, cell, DNA, RNA, gene, miRNA, variant, protein, disease, biological process, pathway, phenotype, molecular function, cellular component, and chemical
		Edge-level	Tissue-/cell-specific gene expression, gene-variant association, variant-disease association, chemical-disease association, chemical-pathway association, etc.
Molecular design	Protein sEquence undERstanding (Xu et al. 2022)	Edge-level	Protein–protein interaction, contact prediction
		Graph-level	Molecular property (e.g. fold classification, secondary structure prediction)
Molecular design	Tasks Assessing Protein Embeddings (Rao et al. 2019)	Edge-level	Protein–protein interaction, contact prediction
		Graph-level	Molecular property (e.g. fold classification, secondary structure prediction)
Molecular design	Graph Explainability Library (Agarwal et al. 2023)	Graph-level	Molecular mutagenic property, molecular functional group (e.g. benzine rings, fluoride carbonyl)
Neurology	NeuroGraph (Said et al. 2023)	Graph-level	Donor demographics (age and gender), task states (emotion processing, gambling, language, motor, relational processing, social cognition, and working memory), cognitive traits (working memory, fluid intelligence)
Therapeutic discovery	AVIDa-hIL6 (Tsuruta et al. 2024)	Edge-level	Antigen–antibody interaction
Therapeutic discovery	Therapeutic Data Commons (Huang et al. 2021)	Edge-level	Drug–target interaction, drug–drug interaction, protein–protein interaction, disease–gene association, drug–response prediction, drug–synergy prediction, peptide-MHC binding, antibody–antigen affinity, miRNA–target prediction, catalyst prediction, TCR–epitope binding, and clinical trial outcomes
		Graph-level	Molecular property (e.g. synthesizability, drug-likeness)

Databases are categorized by data type. The table is organized alphabetically by data type and database names.

Salazar et al. 2021). Supplementing these techniques, differential network analysis (Section 2) can reveal differential connections or rewiring of a molecular network under varying

conditions. This complements traditional differential gene expression analyses, giving a robust framework to investigate diverse conditions and, by extension, different patient groups
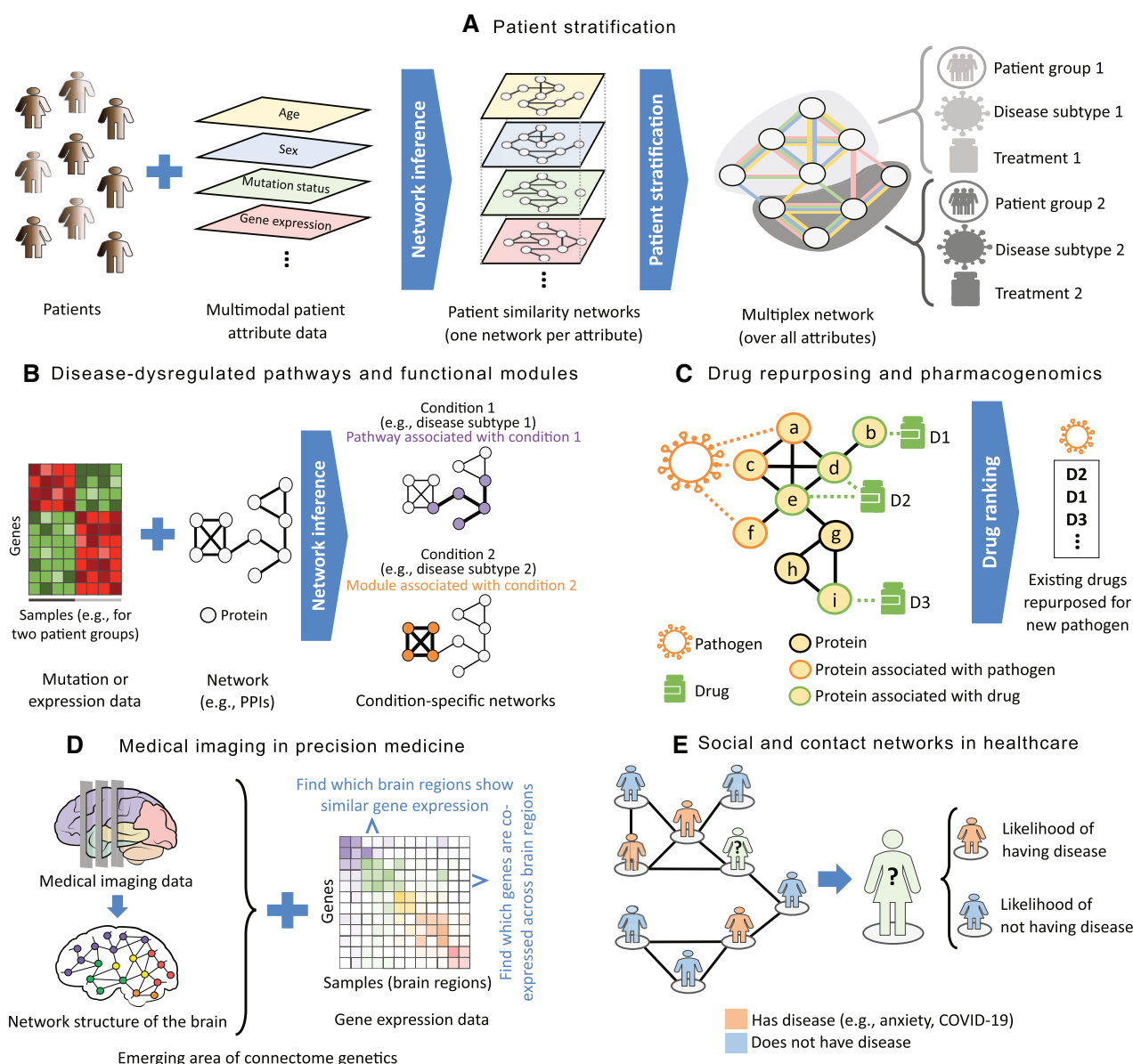
Figure 6. Prominent topics in network-based precision medicine. (A) Groups of patients that correspond to their communities (clusters) in a patient similarity network may shed light on distinct disease subtypes and thus lead to tailored, group-specific therapeutic strategies. (B) Identification of pathways (sparse, tree-like subnetworks) or functional modules (dense, clique-like subnetworks) associated with disease (subtypes) is related to inference of a condition-specific network (Section 2) and pathway reconstruction (Section 3). (C) Drug repurposing evaluates the fit of existing drugs to new diseases based on network "relatedness" between protein targets of the existing drugs and proteins associated with the new diseases, e.g. existing drug D2 may be a good treatment for the new pathogen because D2 targets two proteins (d and e), both of which directly interact with two of the proteins associated with the pathogen (a and c); the four proteins (a, c, d, e) form a clique, which further adds to their "relatedness." (D) An important application of medical imaging lies in brain disorders. In connectome genetics, network structure of the brain meets -omics data. (E) An individual's position in their social/contact network, along with demographic, personality, physical/mental health, etc. information about the other individuals, can give insights into the given individual's health.

(Gysi and Nowick 2020, Morselli Gysi *et al.* 2020, Tu *et al.* 2021).

Precision medicine's applications in identifying candidate anticancer therapeutics have broadened its scope to probe molecular shifts linked with other diseases and aging. Recent research endeavors have used multi-omics strategies to pinpoint innovative therapeutic targets for ulcerative colitis (Voitalov *et al.* 2022) and rheumatoid arthritis (Li *et al.* 2024). As another example, complementing the above discussion of detecting disease-associated modules of genes from a molecular network, modules of diseases have been detected from a heterogeneous disease–disease similarity network

(Halu *et al.* 2019). Other studies have delved into molecular biomarkers, their regulatory pathways, and age-related modifications (Tseng *et al.* 2018). These studies aim to formulate therapies adeptly tailored to diverse age demographics. Complementing the focus on aging, there is a burgeoning interest in discerning patient sex-specific disparities. These lines of inquiry draw motivation from epidemiological data, which delineate differential patterns in the incidence, progression, and prognosis of complex diseases across gender and age brackets (Cannistraci *et al.* 2021).

Drug repurposing and pharmacogenomics. Compared to traditional drug development, drug repurposing (Fig. 6C)

offers significant advantages such as low cost, reduced risk, and faster drug development timelines (Cheng *et al.* 2018, Langhauser *et al.* 2018, Pushpakom *et al.* 2019, Ünsal *et al.* 2023). While early examples of successfully repurposed drugs have been identified through serendipitous discoveries, the availability of massive amounts of -omics and knowledge data and advances in computational techniques have provided opportunities for systematic *in silico* inference of novel indications for existing drugs (Guney *et al.* 2016, Zambrana *et al.* 2021, Huang *et al.* 2023a, Wen *et al.* 2023, Xenos *et al.* 2023). Network science and machine learning models have demonstrated impressive capabilities, but the bar for clinical applications is high. For example, an ensemble network approach has been used to identify drug candidates for repurposing against COVID-19 viral replication (Gysi *et al.* 2021, Patten *et al.* 2022). As another example, a heterogeneous network approach revealed diseases that are most similar to COVID-19, thus reflecting conditions that are risk factors in patients and suggesting the suitability of this approach for use in drug repurposing (Verstraete *et al.* 2020). Validation of the most promising computational predictions in the laboratory yielded an order of magnitude more potent candidates than nonguided experimental screening. In pharmacogenomics, graph convolutional neural networks trained on heterogeneous networks of drug–drug interactions identified adverse events due to polypharmacy and concomitant use of medications (Zitnik *et al.* 2018). Furthermore, deciphering drug–cell connectivity data, indispensable for patient-specific drug repositioning, gains momentum by embedding PPI networks using tensor completion algorithms (Bumin *et al.* 2022).

The role of medical imaging in precision medicine. In addition to -omics data, medical images have emerged as an important new data modality that can facilitate precision medicine, including disease detection, diagnosis, and therapeutic interventions (Comaniciu *et al.* 2016, Lambin *et al.* 2017). Often, medical images encompass distinct topological patterns of target entities that can serve as diagnostic signatures or biomarkers, such as the dendritic structure of the trachea or clustering behaviors of immune cells. Combining these topological signatures with deep learning algorithms offers a substantial advantage in various medical image analysis endeavors, including segmentation, classification, registration, and tracking, and can help with the interpretability of deep learning models. Building tools to compute topological and deep learning representations of imaging data inaugurates new avenues for nuanced analysis, unveiling hidden patterns and intricate correlations within multifaceted datasets (Edelsbrunner *et al.* 2002). These developments have catalyzed the birth of topology-infused deep learning techniques for myriad applications, spanning from segmenting retinal vessels (Hu *et al.* 2019, Shit *et al.* 2021) to discerning retinal arteries/veins (Mishra *et al.* 2021) and forecasting protein semantic similarities (Wang *et al.* 2023b).

An important application of network-based precision medicine lies in brain disorders, where medical image analysis intertwines with network and -omics data (Fig. 6D). Specifically, procuring multimodal neuroimaging, neural network configurations, genetic markers, and other biomolecular signatures could allow for gaining insights into the neural architectures of the human brain, the modulation of its functionalities by network topographies, and the genetic interplays that correspond to disease-specific cerebral patterns. An emergent discipline, dubbed connectome genetics, heralds the meticulous delineation of human neural connectivity, unraveling its ties to cognition, behavior, and the genetic underpinnings of individual neural circuit variances (Arnatkeviciute *et al.* 2021a). Graph mining techniques combined with data science methods have been devised, geared towards personalizing diagnosis and therapy by leveraging the multifaceted data from connectome genetics (Jahanshad *et al.* 2013, Arnatkeviciute *et al.* 2021b, Sha *et al.* 2023). The recent advent of GNN-driven deep learning models further deepens our grasp on the intricate shifts within this data, advancing our understanding of neurological diseases and their heterogeneity across patient populations (Zhang and Huang 2019, Zhang *et al.* 2021, Zhao *et al.* 2022).

The role of social and contact networks in healthcare. Biological networks hold significant promise for advancing personalized medicine. In tandem, social, support, and contact networks correlate with individual health outcomes (Fig. 6E), providing valuable insights into patient behaviors and sentiments (Smith and Christakis 2008). Such networks offer real-time perspectives on patient inclinations, such as therapy adherence preferences. Moreover, they can model patient behaviors associated with medication consumption, enabling the formulation of individualized intervention strategies (Guiñazú *et al.* 2020). The confluence of health and social networks has been harnessed to forecast individual health outcomes, including mental health parameters like anxiety and depression. These predictions emerge from a rich tapestry of data sources, including combinations of heterogeneous social network data and wearable health measures (Liu *et al.* 2021a), and dynamic social network interactions (Liu *et al.* 2020).

In global health emergencies, networks detailing interpersonal contacts have been pivotal in predicting disease transmission. The COVID-19 pandemic spurred the creation of composite models that integrate contact information with individual patient attributes (Guzzi *et al.* 2022). Within such models, nodes signify individuals, while links—static or temporal/dynamic—depict interindividual interactions. Distinct individual features, such as health status (e.g. healthy or recovered), are encapsulated as node-associated feature vectors. Grounded in theoretical foundations of susceptible-infectious-recovered models (Guzzi *et al.* 2022), these approaches are nuanced and can account for real-world contact patterns. They allow for simulation and evaluation of public health response strategies, from containment measures to vaccination campaigns (Stegehuis *et al.* 2016, Bryant and Elofsson 2020, Alguliyev *et al.* 2021). For example, designing a vaccination strategy targeting individuals based on contact behaviors could preempt outbreaks. Since the design of a tailored vaccination strategy may save lives and control the epidemic spreading, we believe that more work should be done to improve these models by designing novel simulation algorithms which require less computational power. Actually, many simulation models require the inspection of all the nodes and edges for each simulation run, making them difficult to run on very large graphs (Fortunato 2010, Guzzi *et al.* 2022).

Open questions for network-based precision medicine. Despite notable advancements in network methods for precision medicine, several challenges remain. These include model benchmarking and comparison, integration of multimodal data from individual patients, and strategies to achieve the intricate equilibrium between preserving patient confidentiality and maximizing the utility of these approaches. Evaluating new methods is complex because establishing

ground-truth, i.e. gold-standard or "correct," benchmarks against which various network strategies can be compared (Guo *et al.* 2022) remains challenging. Evaluating precision therapeutics *in vivo* presents even greater challenges, given the impossibility of retroactively altering treatment modalities for the same individual at a specific temporal junction. Garnering multimodal data about a single patient presents its own difficulties, as diverse data types vary in quality and completeness (Wang *et al.* 2014, Zitnik *et al.* 2019a). In light of these complexities, there is a need for graph learning algorithms tailored for data-intensive multimodal networks. Importantly, new network embedding methodologies may provide simplification of these complexities into new modeling paradigms that are easier to comprehend and compute on (Xenos *et al.* 2021, Doria-Belenguer *et al.* 2023, 2024). Furthermore, it is imperative to foster computational paradigms adept at handling patient data in a manner that safeguards privacy while not compromising on scientific robustness and safety (Hunter *et al.* 2012).

Precision medicine stands poised to enable transformative shifts in disease diagnosis, therapeutic interventions, and overall patient care. Network methods and multimodal data integration are instrumental to these ambitions. Addressing intrinsic challenges related to small-sample datasets that lack statistical power and magnifying methods' susceptibility to misinterpretation and unstable performance is paramount for furthering its nascent triumphs. Surmounting these obstacles requires interdisciplinary research involving network biology scientists, clinicians, and healthcare policymakers to ensure that precision medicine evolves as a paradigm for disease diagnosis, prevention, and treatment that works equally well for all patients by taking into account individual differences in lifestyle, socioeconomic factors, environment, and biological characteristics (All of Us Research Program Investigators 2019).

7. Research discussion and future outlook

Even the well-established network biology research topics/problems, such as network inference (Section 2), have many known limitations and thus open questions associated with them. The emerging research problems, such as network-of-networks analysis (Section 3) or determining how the explosion of large language models (LLMs) can benefit network biology, will have even more challenges associated with them, as expected, given that these problems have started to receive attention only recently; such challenges are discussed below. The emerging problems also bring exciting new opportunities. In the following sections, we build upon the discussion about limitations and open questions from the previous sections, link together common themes from the earlier sections, and complement the previous sections by introducing additional open problems and opportunities.

On methodological paradigms and empirical evaluation

The need to compare different categories of approaches designed for the same purpose. For several topics discussed thus far, a common theme has been that it remains unclear how specific categories of approaches for a given purpose compare to each other in terms of methodological (dis)advantages, as well as in which network analysis tasks or biological/biomedical applications they might be (in)appropriate to use. For example, with network alignment, methods from

biological and other (e.g. social) network domains are rarely evaluated against each other (as discussed more below); with network-of-networks analysis, the existing approaches were proposed for different network analysis tasks or biological/biomedical applications and have not yet been compared to each other (Section 3); with hypergraph versus pairwise graph analyses, it remains unclear to what extent different tasks actually benefit from hypergraph-based methods (Section 4).

Focusing more on network alignment, methods for this purpose introduced for biological networks have typically been thoroughly compared to each other (Section 3), including fair comparison of different approach categories, such as global versus local network alignment (Meng *et al.* 2016, Guzzi and Milenković 2017), pairwise versus multiple network alignment (Vijayan *et al.* 2020), or alignment of static versus dynamic networks (Vijayan *et al.* 2017). On the other hand, network alignment methods introduced in network biology have rarely been compared to those introduced in other domains such as social networks, and vice versa, despite having similar if not the same goals—mapping related nodes or network regions across compared networks. This could be because biological networks have significantly fewer nodes and are likely noisier than other (e.g. social) networks (Eyuboglu *et al.* 2023). This could also be because networks in different domains contain different types of data, which makes the methods customized to their specific data types, rendering their comparison challenging or requiring methodological extensions and new developments. Or, it could be because developers of methods in different domains are from different scientific communities and may thus be unaware of each other's scientific discoveries (Section 8). In either case, it is critical to understand the methodological (dis)advantages of approaches from different domains. Their comprehensive and fair comparison could be a step in this direction, guiding the development of more powerful and possibly more generalizable network alignment approaches.

Network biology has traditionally relied on approaches that work directly on graph topology. In contrast, in recent years, the field has seen an increasing interest in network embedding—be it via earlier spectral-based or diffusion/propagation/random-walk-based methods or more recent deep learning methods—which first transform graph topology into compact numerical representation vectors, i.e. embeddings, and then work on these graph representations (Section 5). A comparative study of nonembedding approaches that work directly on graph topology against network embedding methods was performed in a broad set of contexts: network alignment, graph clustering (i.e. community detection), protein function prediction, network denoising, and pharmacogenomics (Nelson *et al.* 2019). The finding was that in terms of accuracy, depending on the context and evaluation measures used, sometimes direct, graph-based methods outperformed network embedding ones and other times, results were reversed; regarding computational complexity/running time, embedding methods outperformed direct, graph-based methods most of the time (Nelson *et al.* 2019). These indicate the need for a deeper combination of these approaches.

Also, network biology has traditionally relied on combinatorial or graph-theoretic techniques, i.e. on manually engineered or user-predefined topological features of nodes or graphs (the field has also relied on additional method types, e.g. those from the physics community within the field of network science, but these are not the focus of discussion here).

For example, a prominent research problem of the graph-theoretic type that has revolutionized the field of network biology is counting graphlets/subgraphs in a graph; various node-, edge-, or network-level features based on these counts are then applicable to many downstream computational tasks and biological/biomedical applications, as discussed in Section 4. More recently, network biology has benefited from the boom in deep learning (e.g. GNNs), which can automatically generate relevant network topological features prominently via graph representation learning (Section 5). It remains unclear which of graph-theoretic versus deep learning approaches (i.e. manually engineered versus automatically generated network topological features) are better and in which contexts. In other words, both approach categories seem to have merits depending on the context. Again, the question is how to combine them for improved performance.

As an example, graphlet-based and GNN-based analyses of protein structure networks were shown to outperform traditional nonnetwork-based analyses of protein sequences and 3D structures in the tasks of protein structure comparison/classification and protein function prediction, respectively (Faisal *et al.* 2017, Newaz *et al.* 2020, Gligorijević *et al.* 2021). Only recently, the graphlet and GNN approaches were evaluated against each other when comparing protein structures, by the authors who proposed using GNNs for studying 3D structures (Gligorijević *et al.* 2021). They found that graphlet-based analyses greatly outperformed GNN-based analyses in accuracy, although they found the latter to scale better to denser protein structure networks (Berenberg *et al.* 2021).

The relatively inferior performance of GNNs compared to graphlet-based approaches in that particular network-based protein structure comparison (Berenberg *et al.* 2021) can potentially be elucidated as follows. Given that network comparison represents an NP-hard undertaking, a viable computational strategy that balances feasibility and efficacy involves the comparison of network substructures. Graphlets, by design, embody such an approach. Early GNNs were initially not designed for modeling subgraphs. So, it might not be surprising that popular GNN architectures cannot count graphlets and subgraphs and thus might not be the right methodological choice for specific scientific problems (Chen *et al.* 2020). Nevertheless, recent advancements in the field have yielded a spectrum of novel GNN methodologies tailored to subgraph modeling and enumeration. Theoretical underpinnings have emerged that show the expressive capacity of GNNs, delineating which classes of GNN architectures are proficient or deficient in quantifying specific subgraph structures (Chen *et al.* 2020, Tahmasebi *et al.* 2020, 2023, Bouritsas *et al.* 2022, Yu *et al.* 2023). For example, while message-passing GNNs have been popular architectures for learning on graphs, recent research has revealed important shortcomings in their expressive power. In response, higher-order GNNs have been developed that substantially increase the expressive power, although at a high computational cost (Tahmasebi *et al.* 2020). These techniques demonstrate the potential to enumerate subgraphs, thus circumventing the established limitations of low-order (message-passing) GNNs while exploiting sparsity to reduce the computational complexity relative to higher-order GNNs (Tahmasebi *et al.* 2020). Further, recent recursive pooling methods centered on local neighborhoods and dynamically rewired message-passing techniques (Gutteridge *et al.* 2023) improve performance for tasks relying on long-range interactions. Finally,

innovative methods based on graph transformers (Ying *et al.* 2021, Zhang *et al.* 2022c) afford a spectrum of trade-offs between expressive capability and efficiency of machine learning models.

Related to the above discussion, recent developments have highlighted the emergence of state-of-the-art geometric deep learning models trained on protein 3D structures (Baek *et al.* 2021, Abramson *et al.* 2024). Many models focus on proteins' structural surfaces and some explicitly incorporate the underlying protein sequence or structural fold information (Dauparas *et al.* 2022, Zhang *et al.* 2023b). Notably, these models have enhanced performance in various tasks associated with predicting interactions between proteins and other biomolecules (Gainza *et al.* 2020, 2023, Baek *et al.* 2024). These tasks encompass critical areas such as protein pocket–ligand prediction, prediction of PPI residues, ultrafast scanning of protein surfaces to forecast protein complexes, and the design of novel protein binders (Gainza *et al.* 2020, 2023). Geometric deep learning methods that model protein 3D structures *as networks* are promising. Such approaches were shown to outperform existing scientific methods traditionally used in a variety of tasks related to structure-based modeling and prediction of protein properties; the existing methods included network approaches that are not based on geometric deep learning (Stärk *et al.* 2022, Wang *et al.* 2022c, Zhang *et al.* 2023c). The tasks in question included drug binding, PPI prediction, and protein fold, function, or reaction prediction/classification (Stärk *et al.* 2022, Wang *et al.* 2022c, Zhang *et al.* 2023c).

A potential avenue to handling different approach categories/paradigms, such as those discussed above, each with its own merits depending on the context, is to propose algorithmic improvements toward reconciling them. Another is to carry out empirical evaluation of different approaches in a variety of different contexts: at various levels of graph structure (e.g. node, edge, subgraph, or entire network), for diverse types of graphs (e.g. heterogeneous, dynamic, spatial), in different computational tasks (e.g. node classification, graph classification, link prediction), and different biological/biomedical applications (e.g. protein function prediction, cancer, aging, drug repurposing). The following sections discuss these two avenues in more detail.

Algorithmic improvements towards reconciling diverse methodological paradigms. An algorithmic solution to handling different approach categories for the same purpose is to design hybrid methods that employ techniques from all associated disciplines. For example, deep learning methods can be combined with a network propagation approach to improve the embedding of multiple networks (Nasser and Sharan 2023). Alternatively, a theory that would unify different approach categories could be proposed. For instance, the field of neural algorithmic reasoning focuses on developing deep learning models that emulate combinatorial algorithms (Veličković and Blundell 2021). As a case in point, a transformer neural architecture, which was initially devised for natural language processing, has been repurposed to tackle the combinatorial traveling salesperson network problems (Bresson and Laurent 2021) and graph-structured datasets (Yun *et al.* 2019). A primary objective of this discipline is to investigate the capacity of (graph) neural networks to learn novel combinatorial algorithms, particularly for NP-hard challenges that necessitate heuristic approaches. Put differently, the aim is to ascertain if deep learning can extract

heuristics from data more effectively, potentially superseding human-crafted heuristic methods that could demand years of dedicated research to formulate for NP-hard problems (Bresson and Laurent 2021).

Another potential solution on the methodological level relies on the fact that current GNN approaches mainly adopt deep learning from other domains outside of network biology. As such, it is necessary to understand the correct inductive biases within a deep learning model that are representative of a biological mechanism under consideration. For example, can and should the hierarchical structures of ontologies, such as the GO or Disease Ontology, be incorporated into the GNN structure used for predicting proteins' functions or disease associations, respectively? Existing work on visible neural networks shows that such an attempt to incorporate a cell's hierarchical structure and function into the architecture of the deep learning model is effective and facilitates interpretability as the model's components naturally correspond to biological entities (Ma *et al.* 2018, Gaudelet *et al.* 2020). Even the hierarchical network-of-networks idea is not only useful as a potent new way to represent and analyze multiscale biological data as discussed in Section 3, but also as a novel graph representation learning methodology for popular network analysis tasks that are not necessarily of the multiscale nature. For example, there exist studies that take multiple networks as input, all at the same scale, and then perform the well-established tasks of graph embedding (Du and Tong 2019) or classification (Wang *et al.* 2022g) via novel hierarchical approaches, e.g. a graph-of-graphs neural network (Wang *et al.* 2022g), or matrix-factorization based data fusion (Malod-Dognin *et al.* 2019).

Another relevant question is how generalizable versus specific an approach should be. One frequent issue is selecting a suitable similarity measure. For instance, this issue arises when deciding which property of a graph should indicate the proximity of its nodes in an embedding produced by a GNN, or when discerning relationships between biomolecules for inferring correlation or regulatory networks by linking nodes with edges. Selecting an optimal similarity measure for a specific task or application often requires extensive empirical assessment, evaluating multiple measures against one another. It remains a challenge to discern whether a universal, principled similarity measure exists. The answer could potentially be specific to individual tasks or applications or broad categories of analogous tasks. The emphasis on generalizability also begs the question of its desirability; sometimes, the focus should be finely tuned to the specific task, application, or audience (Ektefaie *et al.* 2024). Furthermore, in some contexts, dissimilarity (or distance) might be more pertinent than similarity. For example, proteins can have opposing effects on each other despite working on the same functional goal (Weber *et al.* 2020, Badia-i Mompel *et al.* 2023, Szklarczyk *et al.* 2023). As another example, neighboring edges might mean different things, such as up- versus down-regulation of genes. An essential consideration is the selection of distances with theoretical underpinnings that facilitate efficient optimization (Cao *et al.* 2013), including distances that provably uphold the triangle inequality (Ding *et al.* 2006) and distances specified on smooth manifolds that yield symmetric positive semidefinite distance matrices (Wang *et al.* 2018). Moreover, in typically high-dimensional spaces, the compromises entailed when our chosen distances forsake theoretical properties can be significant, potentially distorting

interpretations and downstream analyses (Beyer *et al.* 1999, Radovanović *et al.* 2010).

Uncertainty quantitation and confidence estimation. Uncertainty quantification presents a unique set of challenges. The inherent structure and complexity of network datasets introduce nuances not observed in other data modalities. The primary challenge lies in distinguishing between aleatoric (data-related) and epistemic (model-related) uncertainties while effectively mitigating potential biases that can distort predictive performance (Zhao *et al.* 2020b, Hüllermeier and Waegeman 2021). Aleatoric uncertainty, stemming from inherent biological variation and limitations of experimental technology, encompasses variability arising from naturally random effects and natural variation intrinsic to the data (Hüllermeier and Waegeman 2021). For instance, in PPI networks, inherent biological variability can lead to uncertainties in node or edge properties. On the other hand, epistemic uncertainty is engendered by a lack of knowledge or limited modeling assumptions. This type of uncertainty is particularly pronounced in graph-based tasks due to the myriad ways graphs can be represented, processed, and interpreted. For instance, different choices in GNN model architectures or graph pooling strategies can introduce varying degrees of epistemic uncertainty (Hüllermeier and Waegeman 2021). Effectively quantifying and addressing these uncertainties is paramount for ensuring reliable and robust findings, especially when making critical decisions based on such models.

Additional considerations for proper empirical method evaluation: benchmark data, performance measures, code and data sharing, best practices. Establishing appropriate benchmark data (including ground-truth data for training and testing/evaluating a predictive model), evaluation measures, and benchmark frameworks is critical to allow for systematic, fair, and unbiased method comparison. Valuable efforts already exist (Table 1). Nonetheless, notably, such frameworks must allow for continuous evaluation as new methods and algorithms will continue to appear. Best practices and guidelines on assessment in network biology are needed.

Lessons learned from challenges in biomedicine such as Critical Assessment of protein Structure Prediction (CASP) (Moult *et al.* 1995, Kryshtafovych *et al.* 2021, 2023), Dialogue on Reverse Engineering Assessment and Methods (DREAM) (Stolovitzky *et al.* 2007, Saez-Rodriguez *et al.* 2016, Meyer and Saez-Rodriguez 2021), and Critical Assessment of protein Function Annotation (CAFA) (Radivojac *et al.* 2013, Jiang *et al.* 2016, Zhou *et al.* 2019) can perhaps help guide the development of best evaluation practices specific to network biology. Such challenges are a paradigm for unbiased and robust evaluation of algorithms for analysis of biological and biomedical data, which crowdsources data analysis to large communities of expert volunteers (Costello and Stolovitzky 2013, Saez-Rodriguez *et al.* 2016). Challenges are done in the form of collaborative scientific competitions. Through these, rigorous validation and reproducibility of methods are promoted, open innovation is encouraged, collaborative communities are fostered to solve diverse and critical biomedical problems and accelerate scientific discovery, the creation and dissemination of well-curated data repositories are enabled, and the integration of predictions from different methods submitted by challenge participants provides a robust solution that often outperforms the best individual solution (Saez-Rodriguez *et al.* 2016).

CASP is the earliest formal method assessment initiative in computational biology (Moult *et al.* 1995). While network biology approaches can be used for CASP's protein structure prediction and CAFA's protein function prediction problems, DREAM was explicitly initiated in response to a network biology need—to reverse-engineer biological networks from high-throughput data (Stolovitzky *et al.* 2007). Since then, numerous DREAM Challenges have been conducted spanning a variety of additional computational (not necessarily network) biology topics, including TF binding, gene regulation, signaling networks, dynamical network models, disease module identification, scRNA-seq and scATAC-seq data analysis, single-cell transcriptomics, and drug combinations (https://dreamchallenges.org/) (Meyer and Saez-Rodriguez 2021). Note that in addition to these initiatives focused solely on computational biology tasks, there exist community benchmark frameworks for general graph-based machine learning that also handle some computational biology tasks, which could thus also serve as significant assets. An example is Open Graph Benchmark (Hu *et al.* 2020b, 2021) (Section 5), which includes the task of predicting protein function from PPI network data with fully reproducible results and directly comparable approaches using the same datasets (https://ogb.stanford.edu/docs/leader_nodeprop/\#ogbn-proteins). Other examples are shown in Table 1.

Interestingly, some of the common themes that emerged from the original 2006 DREAM initiative (Stolovitzky *et al.* 2007) still hold to this date. The current biological network data may not be mechanistically accurate, yet they can still help understand cellular functioning. Exploring condition-specific biological networks is important because network properties can differ in different conditions. While there exist some highly trusted biological data (e.g. the reference HURI PPI network for humans; Luck *et al.* 2020) that may serve as ground truth for understanding (dis)advantages of network algorithms, synthetic network data that are much easier to generate will continue to be necessary for evaluating algorithm performance. However, experimentalists are unlikely to trust any scientific findings from synthetic data or computational approaches evaluated only on such data. Further, regarding ground-truth data for training and testing/evaluating a predictive model, it is critical to have available knowledge on both positive and negative instances in ground-truth data. Examples of the latter are PPIs or protein-functional associations that do not exist in cells. However, such negative instance data are hard to obtain in biology.

To add to the discussion about ground-truth data, using the aging process as an example, ground-truth data about human aging have been obtained in one of two ways: via sequence-based homology from model species (de Magalhães *et al.* 2009) or via differential gene expression analyses in humans (Berchtold *et al.* 2008, Jia *et al.* 2018). In a recent study (Li *et al.* 2021), only 17 genes were shared between the 185 sequence-based and 347 expression-based human aging-related genes. This poses several questions. How do we resolve such discrepancies with datasets on the same biological process resulting from different modalities/technologies, which likely exist in other applications as well? Given their high complementarity, perhaps integrating the different data types could yield more comprehensive insights into the biological process under consideration. However, if any of the other datasets are noisy, or if the different data types have different "signatures" (i.e. features) in a biological network,

their integration could decrease the chances of detecting meaningful biological signals from the network compared to analyzing the different data types individually. Moreover, because different types of biological data collected via biotechnologies (e.g. genomic sequence data versus transcriptomic gene expression data versus interactomic PPI data) are likely to capture complementary functional slices of the given biological process, is it appropriate to use some of these datasets as the ground-truth data to validate predictions obtained via computational analyses of the other datasets? In our example of the aging process, is it appropriate to use sequence-based or expression-based aging-related knowledge to validate network-based aging-related gene predictions? Is this appropriate, especially because sequence-based and expression-based "knowledge" are also computational predictions, i.e. the result of sequence alignment and differential gene expression analysis, respectively? Also, is this appropriate because sequence-based knowledge about human aging are sequence orthologs of aging-related genes in model species? So, would any aspects of the aging process that are unique to humans be missed by the knowledge originally collected in the model species?

Another challenge with empirical evaluation is accurately estimating the absolute and relative performance of machine learning models and quantifying the uncertainty of performance estimates. Network data is inherently relational, thus inevitably violating the assumptions of independent and identically distributed data (Neville *et al.* 2009, 2012). Even further, the problems with long-tailed degree distribution in biological networks and homology between nodes require careful selection of training and test data when evaluating performance accuracy (Park and Marcotte 2012, Hamp and Rost 2015, Lugo-Martinez *et al.* 2021).

Also, to allow for proper method evaluation, the authors of original methods must publicly release complete and easy-to-use code and data from their papers to allow for reproducing the initial studies and applying and evaluating a given method on new data (Heil *et al.* 2021). Journals and other publication venues should and typically do establish requirements for data and code sharing. Consequently, scientific communities have shown remarkable improvements regarding releasing open-source software and data. Yet, ensuring compliance remains an issue. For example, while code or data might be released, they are sometimes incomplete or not easy to use. Or, there are instances when there might be a link (e.g. to GitHub) provided in the corresponding publication to meet the publication venue requirements, but the link might point to a page that says "under construction," to an empty directory, or to a directory containing some files but without a transparent readme file on how to use the information provided. Who should ensure compliance with publication venue requirements, i.e. that complete and easy-to-use code and data are provided to ensure easy reproducibility? The editors of a venue publishing a given paper? The reviewers already volunteering their virtually nonexistent "free" time to evaluate the paper's scientific merits for publication should thus probably not be expected to invest even more effort to verify that the code and data can be run correctly. The authors? The future readers of the article who might be interested in using the method? If the latter two, what should be the repercussions if it is found that the code or data do not exist or are not possible or easy to use? On a related note, how long after publication should the authors be required to maintain the project code and data and

respond to related email inquiries? Hosting of the code and data is not an issue for authors due to availability of archival data repositories such as Zenodo. However, actively maintaining the code and data is an issue, and this is directly related to whether and how long after the project completion the funding by the federal agencies and others might be available for this purpose.

Complete transparency in all decisions (from graph construction to analysis) is crucial. Workflow management systems, such as Nextflow (Di Tommaso *et al.* 2017) and Snakemake (Köster and Rahmann 2012), can enable rapid prototyping and deployment of computational workflows by combining software packages and various tools. Clear documentation, open-source sharing of code and algorithms, and making raw and processed data available can ensure that results are not just a one-off finding but can be consistently reproduced and built upon by the broader scientific community.

On missing data

Network completeness and interaction causality. Much of network biology relies on aging technologies with notable limitations. Focusing on physical PPIs, biotechnologies such as yeast two-hybrid systems (Fields and Song 1989), cross-linking mass-spectrometry (Piersimoni *et al.* 2022), and structural determination of protein complexes (Jacobsen 2007, Rhodes 2010, Saibil 2022) have collectively generated systems-level data that have led to critical methodological advances in network biology. Of course, these efforts to obtain the physical interactome have been complemented by valuable data collection and network inference efforts related to systems-level correlation networks. However, as computational methods are now maturing, the data are starting to lag. High-resolution, high-throughput data-generating technologies, capable of directly identifying pathways and order of molecular events in various experimental and clinical contexts, are the next frontier for deeper understanding of molecular systems.

There is a need to expand from physical and correlation networks toward causal relationships (Belyaeva *et al.* 2021) or simulatable kinetic models (Karr *et al.* 2012). For this, biotechnologies for data collection need to be improved to allow for higher-quality data to build better causal networks and more complete networks. This will also require the development of new (categories of) approaches that can handle the captured causality. Even if/when we have high-quality causal networks and efficient and accurate methods for their analysis, will this suffice to understand biochemical mechanisms? When one knows biochemical mechanisms, one can infer causality. However, causality might not necessarily allow for fully understanding biochemical mechanisms.

Algorithmic research to guide data generation efforts. It will likely be beneficial to integrate multi-omic network data with BKGs to offer precise and targeted treatments for rare diseases (Alsentzer *et al.* 2022). Such network data with richer semantics will more directly help suggest biological hypotheses (Sanghvi *et al.* 2013, Wang *et al.* 2023a) or support iterative data generation and analyses through active learning (Sverchkov and Craven 2017, Zhang *et al.* 2023a). Informing laboratory experiments using predictions from computational studies could be a path forward to build more complete and accurate data, which could lead to developing new, more advanced network analysis methods to further inform and improve laboratory experiments.

How network biology (primarily algorithmic research) can best support the collection and analysis of multimodal data is quite an important question, especially when collecting multimodal data for the same individuals, including building personalized (i.e. individual-specific) networks. An answer here could be to first figure out what question will be asked in which task/application and then design a data collection strategy. One might want to define optimal datasets. Or, one might want to find unifying factors within data modalities; this is precisely why there is a need for multimodal data for the same individuals, at least some of the data/individuals. This might require systematic, comprehensive, and well-funded consortia efforts. Perhaps algorithmic approaches such as active learning can help prioritize what data should be collected, e.g. from specific populations or about particular biological functions. As success in experimentally collecting or computationally inferring various types of biological networks continues to improve, research efforts likely should shift towards obtaining a predictive understanding of personalized networks. Moreover, even within a single individual, molecular networks vary across tissues and cell types, posing additional challenges in defining an individual-specific network.

Network dynamics. Another data component that is currently missing or is very scarce is network dynamics. Various types of time-dependent perturbation data could help infer dynamic biological networks. Examples of tasks/applications that have benefited from dynamic network analysis in biology are as follows.

One example is the task of network alignment: unlike traditional network alignment that has compared static networks (Section 3), recently, the problem of aligning dynamic networks has been defined, and several algorithms have been proposed for solving the newly defined problem (Vijayan *et al.* 2017, Vijayan and Milenković 2018b, Aparicio *et al.* 2019). The challenge here is the lack of experimentally obtained dynamic biological network data, which is why such methods have been evaluated on synthetic networks, computationally inferred dynamic biological networks, or dynamic networks from other domains (Vijayan *et al.* 2017, Vijayan and Milenković 2018b, Aparicio *et al.* 2019).

Another example is a recent network-based study of the dynamics of the protein folding process (Newaz *et al.* 2022). A key challenge is the lack of large-scale data on protein folding intermediates, i.e. 3D conformations of a protein as it undergoes folding to attain its native structure. Experimental data of this type are lacking even on the small scale (Newaz *et al.* 2022). Traditional computational, simulation-based studies, as well as the recent network-based effort (Newaz *et al.* 2022), all approximate the folding intermediates of a protein from the protein's final (or native) 3D structure. Obtaining the actual protein folding intermediates experimentally is unlikely to happen any time soon, especially at a large scale, so computational efforts will be needed. With recent breakthroughs in protein structure prediction, e.g. AlphaFold (Jumper *et al.* 2021), this need represents an excellent opportunity for computational research to help obtain, model, and analyze the resulting dynamic data.

A further example is a dynamic network analysis of the aging process, i.e. predicting new aging-related genes from a dynamic aging-specific PPI network (Section 2). Here, a key challenge is that shockingly, using *newer* aging-related gene expression and PPI network data obtained via newer and thus higher-quality biotechnologies to infer a dynamic aging-

specific network does not yield more accurate aging-related gene predictions than using *older* data of the same type from over a decade ago when dynamic network analyses of aging were pioneered (Li *et al.* 2022c). It was also observed in a different study on active module identification that using newer network data typically did not lead to more biologically meaningful results (Lazareva *et al.* 2021). Going back to aging, it remains unclear whether the issue is with gene expression data, PPI network data, methods for integrating the two to computationally infer a dynamic aging-specific network, network methods used for feature extraction from the aging-specific network, ground-truth data on which genes are aging- versus nonaging-related, or something else entirely (Li *et al.* 2022c).

As our final example, we discuss quantitative and qualitative mathematical modeling of network dynamics from the systems biology perspective (Kestler *et al.* 2008, Le Novere 2015). Quantitative formalisms provide a precise description of the evolution of the system, including its temporal aspects; they are strongly dependent on the availability and precision of the required parameters. At the other end of the spectrum, qualitative (logic) frameworks have the advantage to be simpler, with no requirement for quantitative parameters, allowing analytical analyses. Logical models allow coarse-grained descriptions of the properties of the biological network and bring out key actors and mechanisms controlling the dynamics of the system (Maheshwari and Albert 2017). Recent efforts use -omics data, including single-cell transcriptomes, to construct or contextualize Boolean models (Schwab *et al.* 2021, Montagud *et al.* 2022, Hérault *et al.* 2023).

Towards inclusive and equitable precision medicine. Progress in computational (including network) biology and biomedicine has been hindered by a lack of -omics data encompassing vast human diversity (Cruz *et al.* 2023). Underrepresentation of human genetic diversity has drastically weakened the biological discoveries that would benefit all populations, leading to health disparities. The traditional one-size-fits-all healthcare model meant for a "typical" patient may not work well for everyone. In response, the National Institutes of Health has aimed to invite one million people across the USA to help build one of the most diverse health databases in history, welcoming participants from all backgrounds through the "All of Us" program (https://allofus.nih.gov/). Inclusivity is at the core of the program: participants are diverse in terms of their races, ethnicities, age groups, regions of the country, gender identity, sexual orientation, socioeconomic status, education, disability, and health status. The data collected through the program is expected to lead to discoveries on how our biology, environment, and lifestyle affect our health. Unlike traditional research that has focused on a particular disease or group of people, this program aims to build a diverse database that can inform thousands of studies on a variety of health conditions. Availability of inclusive and diverse -omics data, design of research studies that intentionally and carefully account for such data, and development of computational methods and evaluation frameworks that handle such data in a fair and unbiased manner will be critical for advancing computational biology and biomedicine for all populations and reaching health equity.

Beyond the issue of underrepresentation, certain populations are intrinsically limited in size, such as rare diseases, which are inherently limited in clinical cases (Banerjee *et al.* 2023). Studying a substantial fraction of a small population may still result in data that do not yield health outcomes comparable to those from larger populations. In such scenarios, amassing more data may not be feasible, leading to small-sample datasets that can lack statistical power and magnify the susceptibility of computational models to misinterpretation and unstable performance. Network analysis techniques can play a pivotal role in addressing this challenge. Techniques such as few-shot machine learning (Alsentzer *et al.* 2022) and domain adaptation (He *et al.* 2023) for network methods are instrumental in enabling computational models to learn patterns from small datasets and generalize to newly acquired data. Such models can adapt and generalize across diverse populations, thereby enhancing the robustness and applicability of health outcomes derived from datasets with small numbers of samples.

Other major future research advancements

The interface between network biology and LLMs. LLMs, such as ChatGPT and GPT-4, create opportunities to unify natural language processing and knowledge graph reasoning (Fatemi *et al.* 2023, Pan *et al.* 2024), owing to their wide-ranging applicability. Nevertheless, LLMs often serve as black-box models, presenting limitations in comprehensively capturing and accessing factual knowledge. In contrast, BKGs are structured knowledge models that systematically store extensive factual information. BKGs have the potential to enhance LLMs by providing external knowledge that aids in inference and bolstering interpretability. However, constructing BKGs is intricate and dynamic, posing challenges to existing methods in generating novel facts and representing previously unseen knowledge. Thus, an approach integrating LLMs and BKGs could emerge as a valuable strategy, harnessing their strengths in tandem (Pan *et al.* 2024).

The potential synergies between traditional text and structured knowledge graphs are becoming increasingly evident. Language model pretraining has proven invaluable in extracting knowledge from text corpora to bolster various downstream tasks. Yet, these models predominantly focus on single documents, often overlooking interdocument dependencies or broader knowledge scopes. Recent advances (Yasunaga *et al.* 2022b, McDermott *et al.* 2023) address this limitation by conceptualizing text corpora as interconnected document graphs. By placing linked documents in shared contexts and adopting self-supervised objectives combining masked language modeling and document relation prediction, such methods can achieve considerable progress in tasks like multi-hop reasoning and few-shot question answering. On a parallel front, while text-based language models have garnered substantial attention, knowledge graphs can complement text data, offering structured background knowledge that provides a useful scaffold for reasoning. In an emerging line of inquiry, studies (Yasunaga *et al.* 2022a) explore self-supervised paradigms to construct a unified foundation model, intertwining text and knowledge graphs. These approaches pretrain models by unifying two self-supervised reasoning tasks, masked language modeling, and link prediction, marking an exciting direction for future advancements in network biology.

LLMs, traditionally associated with the processing of natural language, possess a flexibility that extends their utility beyond text data (Luo *et al.* 2022a). The underlying architectures, especially transformer-based designs like BERT and GPT variants, can be adapted to learn from any sequential data. In biology, this adaptability implies that LLMs can be

trained on biological sequences, such as DNA, RNA, and proteins (Rao *et al.* 2019, Xu *et al.* 2022, Lin *et al.* 2023). Rather than processing words or sentences, these models can assimilate nucleotide or amino acid sequences, thereby capturing intricate patterns and dependencies in genomic and proteomic data (Meier *et al.* 2021, Dauparas *et al.* 2022, Lin *et al.* 2023, McDermott *et al.* 2023). These cross-disciplinary advances in LLMs highlight their potential to advance the frontiers of computational biology. In addition to large sequence-based pretrained models like LLMs, an emerging area of structure-based pretrained models is concerned with generating new network structures, such as protein and small molecule networks (Townshend *et al.* 2021, Rodrigues and Ascher 2022, Wang *et al.* 2022b, Bennett *et al.* 2023, Gainza *et al.* 2023).

Interpretabilty. Interpretability in network biology involves elucidating mechanisms of disease and health, such as tumor growth and immune responses. However, deep graph learning models are black-box systems with limited immediate interpretability as they produce outputs through a series of complex, nonlinear transformations of input data points. This poses challenges in domains where clear insights are imperative. For instance, while dimensionality reduction techniques and graph representation learning algorithms produce compact latent feature representations of high-dimensional data and graphs, they often sacrifice the interpretability of the features they produce. Conversely, graph-theoretic signatures, which capture network motifs, graphlets, or other substructures, can amplify understanding of networks by identifying relevant structural patterns.

Future research directions in interpretability must focus on integrating domain-specific knowledge into model training and evaluation. By directly incorporating biological constraints and prior knowledge into model architectures, we can enhance interpretability without compromising predictive performance. Additionally, developing explainable techniques tailored explicitly for network biology is crucial. Exploring hybrid models combining interpretable statistical models with deep learning approaches is another promising avenue. Such models can leverage the strengths of both types to produce interpretable and accurate predictions. Likewise, creating advanced visualization tools that effectively convey complex model outputs and biological insights to researchers and clinicians is essential. These tools should be intuitive and enable interactive exploration of model predictions and features.

Reproducibility. Reproducibility in network biology research is a multifaceted challenge due to several reasons. (i) Graph construction: How a graph is constructed can drastically impact the insights drawn from it. For example, consider the problem of inferring an association PPI network. The decision to include only direct interactions versus both direct and indirect interactions can lead to vastly different network topologies. Choosing a threshold to determine an edge (e.g. a particular strength of interaction or confidence level) can also significantly alter the graph. (ii) Edge definitions: What constitutes an edge can be subjective and is often based on the specific context. In a gene co-expression network, for instance, the definition of an edge might be based on a particular correlation coefficient threshold. A slight variation in this threshold can lead to including or excluding numerous interactions, thus changing the network's structure and potentially its inferred properties. (iii) Latent embeddings: Graph-based machine learning methods used to compute embeddings can have a significant effect on the results.

Different embedding techniques capture different types of structural and feature-based information, leading to variations in tasks like node classification or link prediction. (iv) Dynamic nature of biological networks: Biological systems are inherently dynamic. A PPI network at one point in time or under one set of conditions might differ from the network under another state. Thus, reproducing results requires the same methodology and the same or equivalent biological conditions. (v) Finally, graph sampling: In many cases, a subgraph or sample is taken due to the massive size of networks or computational constraints. The method and randomness inherent in this sampling can lead to nonreproducible results if not carefully controlled.

Towards wide adoption and translation of algorithmic innovation into practical and societal impact. The recommended method evaluation and data generation improvements discussed above are needed not just for method developers—typically, computational scientists—to be able to properly evaluate their new approaches against existing ones, but even more importantly, for adoption by end users—experimental scientists and in the long run, clinicians, healthcare workers, and patients (Section 8 comments more on this topic, including training needed for noncomputational folks to use network approaches). The disconnect between computational and experimental scientists, even those dedicated to the common scientific goals (Ramola *et al.* 2022), suggests that efforts are necessary to overcome both technical and social challenges in interdisciplinary research fields. Computational scientists might need to consider not only traditionally algorithmic evaluation measures, such as precision, recall, and other performance criteria, but also measures that evaluate the utility and feasibility of integrating methods into scientific and clinical workflows (Huang *et al.* 2022, 2023a). Additionally, computational scientists are primarily incentivized to develop new algorithms and prototype software. In contrast, experimental and clinical scientists expect tools that are robust, trustworthy, and exhibit few glitches in practice. Authoritative evaluations, carried out by independent and interdisciplinary researchers on tasks directly relevant to downstream applications, are essential (Marbach *et al.* 2012a, Choobdar *et al.* 2019). Rapid and broad dissemination of these evaluations, recommendations, and guidelines for best practices should be prioritized in network biology.

Major milestones in network biology. The pinnacle of success for network biology would likely be a comprehensive and dynamic understanding of the entire cellular or organismal interactome across different conditions and life stages. This would include PPIs, gene regulation, metabolic pathways, cell signaling, and more. We can imagine a complete map of every biological interaction in an organism, from the level of genes and molecules up to tissues and organs, with the ability to zoom in on details and see dynamic changes over time or under different conditions. Another significant milestone would be the seamless integration of network biology with other disciplines to provide a holistic understanding of life. This means connecting the molecular interactome with tissue-level networks, organ systems, and interorganismal interactions, such as those seen in symbiosis or ecosystems. From a practical standpoint, a significant success measure would be the application of network biology insights to develop novel and more effective therapeutic interventions. This could mean identifying critical network nodes or interactions to target diseases, leading to innovative treatments.

Drawing parallels from the reference human genome, the equivalent for network biology could be a reference interactome—a standardized and comprehensive map of all known biological interactions within a human cell. This would serve as a baseline for studying disease, development, aging, and other biological processes. Any deviations from this reference in specific cell types, conditions, or diseases could be studied in detail.

Just as AlphaFold (Jumper *et al.* 2021) has made waves in predicting protein structures, a comparable success in network biology might be the development of tools that can accurately predict the emergent properties of a biological system from its underlying network. Given a set of interactions, this would mean the tool could foresee the system's response to a drug, its behavior under certain conditions, or its evolution over time.

8. Additional discussion on scientific communities, education, and diversity

The question of who are network biologists or computational biologists is hard. Ideally, a computational biologist would have the interest and knowledge to both develop core computational methods and understand fundamental biological mechanisms. That raises the question of how to properly train more of such researchers to advance computational biology, including its subarea of network biology that models and analyzes biological systems as networks. For example, based on the personal experience of some of the authors of this article, in a network biology course, computationally focused students might enjoy computational but not biological aspects (e.g. in a general network science course, students typically choose a nonbiology domain to work on, such as technological or social networks). In contrast, biology students might enjoy biological but not computational aspects. So, efforts might be needed to convince students to be genuinely excited about both developing computational approaches and understanding biological mechanisms. Systematically identifying and addressing gaps in current computational biology training programs or starting new interdisciplinary training programs might be needed, along with appropriate support and resources from funding agencies.

Some of these gaps are as follows. An essential part of efficient training would be to have robust, well-known, and trustworthy software tools that are readily available and easy to use, especially by those who are not proficient in computing; clearly, both developing and sustaining such software requires resources. Similar holds for building and making available datasets easily accessible by people who are not proficient in biology to help them get involved easily. Another important part would be exposing students to interdisciplinary collaborative teams to train them to work together on the same research questions with scientists from different disciplines.

Another vital part of training relates to hiring and promoting computational biology faculty who would offer the training. A challenge here, based on the personal experience of some of the authors of this article, seems to be as follows. When hiring a computational biologist in a traditional computationally focused department (e.g. computer science, applied mathematics, statistics, or physics), someone who is more trained in biology may be viewed as not enough of a computational scientist, even when they are proficient in using existing computational methods to uncover new

biological knowledge and possibly also at least occasionally develop new computational methods for studying biological systems. Similarly, in a traditional biology-focused department, a more computationally trained person may be viewed as not enough of a biological scientist, even when they evaluate their new computational methods on biological data and possibly at least occasionally yield new knowledge about biological systems. Yet, both kinds of candidates can be great for both department types. Hence, hiring and promotion groups might need to think differently about interdisciplinary computational biology research. This is especially true in departments where these groups do not have computational biologists or where there are no specific, interdisciplinary departments like biomedical data science or computational biology.

There exists an additional challenge even when focusing on computationally oriented researchers within computational biology. Scientific communities that could benefit (from) the field of network biology include graph theory, network science, data mining, machine learning, and artificial intelligence. These communities often use different terminology for the same concepts (e.g. network alignment versus graph matching or graph clustering versus network community detection). Distinct scientific communities may all analyze biological network data, or address identical computational challenges across various application domains, such as biological versus social networks. However, they often do not attend the same research forums. For instance, attendees of the prominent computational biology conference, Intelligent Systems for Molecular Biology (ISMB), might not necessarily participate in data mining conferences like Knowledge Discovery and Data Mining (KDD) or artificial intelligence conferences such as Neural Information Processing Systems (NeurIPS), and vice versa. Consequently, advancements in one domain might remain obscure in another. Organizing scientific symposia to convene computational scientists from traditionally distinct network biology communities, focusing on universally relevant topics, could help bridge this gap.

The above discussion items can be seen as diversity-focused, be it diversity in one's training and skills or scientific communities they belong to (Nielsen *et al.* 2018). Many other aspects of diversity exist in science, and we focus on some of them here. The International Society for Computational Biology (ISCB) is a globally recognized entity advocating for and advancing scholarship, research, training, outreach, and inclusive community building in computational biology and its professions. This is why we rely on ISCB's demographic statistics to represent the current state in the computational biology field. According to a demographic survey of the ISCB membership, whose results are publicly available in the 2022/2023 ISCB Equity, Diversity, and Inclusion (EDI) report (https://www.iscb.org/edi-resources), among those who responded, 32.8% indicated "female," 60% indicated "male," 0.4% indicated "non-binary," and 6.8% indicated "prefer not to declare." Regarding ethnic origin, in the same report, 53% of those who responded with anything but "prefer not to declare" indicated a non-European descent. Some additional EDI statistics are as follows. At the time of the 2020/2021 ISCB EDI report (the latest report that offered this type of information), 41% of the ISCB Board of Directors were female, and 57% of the Executive Committee (elected officers) were female; 61% of selected keynote speakers at the Intelligent Systems for Molecular Biology (ISMB),

ISCB's flagship and most prestigious conference, were female since 2016. Regarding ISCB awards, fellows election, and other honors, the final selection shows a good gender balance that reflects the membership. However, during the nomination stage, in 2022/2023, for the innovator award, senior scientist award, and fellows election, 22%, 28%, and 25% of the nominees were female, respectively, compared to 32.8% of the entire ISCB membership being female. ISCB does not have such data yet on ethnicity.

Enhancing awareness and mitigating biases when nominating candidates for honors or inviting candidates as conference speakers is a pathway to improving diversity in the computational biology field. Another more ambitious goal is to achieve diversity statistics in the field that mirror those of the general population. This should be accomplished for all undergraduate students, graduate students, postdoctoral fellows, and faculty (across various ranks), not only by addressing the "leaky pipeline" issue (Alper 1993, Sarraju *et al.* 2023), but also by identifying and eliminating institutional barriers to establish an inclusive support infrastructure (Stevens *et al.* 2021). This might only be achievable over a longer period. Also, biology-focused subfields of computational biology are currently more gender-diverse than its computationally focused subfields. Thus, diversity in computational biology might be more readily achieved by recruiting trainees from biology-focused subfields and equipping them with the requisite computational skills rather than the reverse. However, sourcing from computational subfields remains essential. Yet, disciplines like computer science, mathematics, and physics can act as gatekeepers and entering these fields without the appropriate background can be challenging (Torbey *et al.* 2020, Mervis 2022). Because innovative concepts can emerge from diverse sources and all individuals, it is imperative to eliminate gatekeeping barriers.

Additional diversity-related challenges include the need to recognize and mitigate potential implicit biases; limited access to registration and travel funds to conferences based on their locations, especially for those in middle and low-income countries; current lack of ethnicity data to evaluate diversity efforts of computational biology conferences and communities, including ISCB; empirical research into equity in science, etc. Systematic and properly funded initiatives by universities and professional societies are necessary to achieve this. And so are individual efforts by the members of the scientific community. Everyone should be responsible for contributing to joint diversity efforts for the field to make significant and sufficient progress.

Acknowledgements

Author contributions

Marinka Zitnik (Conceptualization [lead], Investigation [lead], Supervision [lead], Visualization [lead], Writing—original draft [lead], Writing—review & editing [lead]), Michelle M. Li (Conceptualization [equal], Investigation [equal], Visualization [lead], Writing—original draft [equal], Writing—review & editing [equal]), Aydin Wells (Conceptualization [equal], Investigation [equal], Visualization [lead], Writing—original draft [equal], Writing—review & editing [equal]), Kimberly Glass (Conceptualization [equal], Investigation [equal], Supervision [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), Deisy Morselli Gysi (Conceptualization [equal], Investigation [equal], Supervision [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), Arjun Krishnan (Conceptualization [equal], Investigation [equal], Supervision [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), T.M. Murali (Conceptualization [equal], Investigation [equal], Supervision [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), Predrag Radivojac (Conceptualization [equal], Investigation [equal], Supervision [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), Sushmita Roy (Conceptualization [equal], Investigation [equal], Supervision [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), Anaïs Baudot (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Serdar Bozdag (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Danny Z. Chen (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Lenore Cowen (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Kapil Devkota (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Anthony Gitter (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft

[supporting], Writing—review & editing [supporting]), Sara Gosline (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Pengfei Gu (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Pietro H. Guzzi (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Heng Huang (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Meng Jiang (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Ziynet Nesibe Kesimoglu (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Mehmet Koyuturk (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Jian Ma (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Alexander R. Pico (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Nataša Pržulj (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Teresa M. Przytycka (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Benjamin J. Raphael (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Anna Ritz (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Roded Sharan (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Yang Shen (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Mona Singh (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Donna K. Slonim (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Hanghang Tong (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Xinan Holly Yang (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Byung-Jun Yoon (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft

[supporting], Writing—review & editing [supporting]), Haiyuan Yu (Conceptualization [supporting], Investigation [supporting], Visualization [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), and Tijana Milenković (Conceptualization [lead], Funding acquisition [lead], Investigation [lead], Project administration [lead], Supervision [lead], Visualization [lead], Writing—original draft [lead], Writing—review & editing [lead]).

Conflict of interest

In Section 8, we rely on ISCB's diversity statistics. These statistics are publicly available, and so there is no conflict of interest. Yet, to remedy any potential perceived conflict of interest, we declare that Predrag Radivojac is the President of ISCB and currently serves on the Board of Directors of ISCB. In addition, Tijana Milenković currently serves on the ISCB Board of Directors and the ISCB EDI Committee. The remaining authors have no conflicts of interest to declare.

Funding

Data availability

Not applicable.

References

Abdar M, Pourpanah F, Hussain S *et al.* A review of uncertainty quantification in deep learning: techniques, applications and challenges. *Inf Fusion* 2021;**76**:243–97.

Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;**630**:493–500.

Agarwal C, Queen O, Lakkaraju H *et al.* Evaluating explainability for graph neural networks. *Sci Data* 2023;**10**:144.

Agarwal S, Branson K, Belongie S. Higher order learning with graphs. In: *Proceedings of the International Conference on Machine Learning.* p. 17–24. New York, NY: Association for Computing Machinery, 2006.

Agrawal M, Zitnik M, Leskovec J. Large-scale analysis of disease pathways in the human interactome. In: *Proceedings of the Pacific Symposium on Biocomputing.* p. 111–22. 2018.

Alguliyev R, Aliguliyev R, Yusifov F. Graph modelling for tracking the COVID-19 pandemic spread. *Infect Dis Model* 2021;**6**:112–22.

All of Us Research Program Investigators. The "All of Us" research program. *N Engl J Med* 2019;**381**:668–76.

Alper J. The pipeline is leaking women all the way along. *Science* 1993;**260**:409–11.

Alsentzer E, Finlayson S, Li M *et al.* Subgraph neural networks. In: *Proceedings of the Advances in Neural Information Processing Systems.* Vol. 33, p. 8017–29. Red Hook, NY: Curran Associates, Inc., 2020.

Alsentzer E, Li MM, Kobren SN *et al.* Few shot learning for phenotype-driven diagnosis of patients with rare genetic diseases. medRxiv, 2022.12.07.22283238, 2022, preprint: not peer reviewed.

Antelmi A, Cordasco G, Polato M *et al.* A survey on hypergraph representation learning. *ACM Comput Surv* 2023;**56**:1–38.

Aparicio D, Ribeiro P, Milenković T *et al.* Temporal network alignment via GoT-WAVE. *Bioinformatics* 2019;**35**:3527–9.

Arici KM, Tuncbag N. Unveiling hidden connections in omics data via pyPARAGON: an integrative hybrid approach for disease network construction. bioRxiv, 2023.07.13.547583, 2023, preprint: not peer reviewed.

Arnatkeviciute A, Fulcher BD, Bellgrove MA *et al.* Where the genome meets the connectome: understanding how genes shape human brain connectivity. *Neuroimage* 2021a;**244**:118570.

Arnatkeviciute A, Fulcher BD, Oldham S *et al.* Genetic influences on hub connectivity of the human connectome. *Nat Commun* 2021b;**12**:4237.

Aronson SJ, Rehm HL. Building the foundation for genomics in precision medicine. *Nature* 2015;**526**:336–42.

Artzy-Randrup Y, Fleishman SJ, Ben-Tal N *et al.* Comment on "network motifs: simple building blocks of complex networks" and "superfamilies of evolved and designed networks". *Science* 2004;**305**:1107.

Atz K, Grisoni F, Schneider G. Geometric deep learning on molecular representations. *Nat Mach Intell* 2021;**3**:1023–32.

Ausiello G, Laura L. Directed hypergraphs: introduction and fundamental algorithms—a survey. *Theor Comput Sci* 2017;**658**:293–306.

Bader JS, Chaudhuri A, Rothberg JM *et al.* Gaining confidence in high-throughput protein interaction networks. *Nat Biotechnol* 2004;**22**:78–85.

Badia-I Mompel P, Wessels L, Müller-Dott S *et al.* Gene regulatory network inference in the era of single-cell multi-omics. *Nat Rev Genet* 2023;**24**:739–54.

Baek M, DiMaio F, Anishchenko I *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 2021;**373**:871–6.

Baek M, McHugh R, Anishchenko I *et al.* Accurate prediction of protein–nucleic acid complexes using rosettafoldna. *Nat Methods* 2024;**21**:117–21.

Bagal V, Aggarwal R, Vinod P *et al.* MolGPT: molecular generation using a transformer-decoder model. *J Chem Inf Model* 2021;**62**:2064–76.

Bajpai AK, Davuluri S, Tiwary K *et al.* Systematic comparison of the protein-protein interaction databases from a user's perspective. *J Biomed Inform* 2020;**103**:103380.

Banerjee J, Taroni JN, Allaway RJ *et al.* Machine learning in rare disease. *Nat Methods* 2023;**20**:803–14.

Barabási AL. *Network Science*. Cambridge: Cambridge University Press, 2016.

Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet* 2011;**12**:56–68.

Baryshnikova A, Costanzo M, Myers CL *et al.* Genetic interaction networks: toward an understanding of heritability. *Annu Rev Genomics Hum Genet* 2013;**14**:111–33.

Basha O, Argov C, Artzy R *et al.* Differential network analysis of multiple human tissue interactomes highlights tissue-selective processes and genetic disorder genes. *Bioinformatics* 2020;**36**:2821–8.

Basha O, Shpringer R, Argov CM *et al.* The DifferentialNet database of differential protein–protein interactions in human tissues. *Nucleic Acids Res* 2018;**46**:D522–6.

Battiston F, Cencetti G, Iacopini I *et al.* Networks beyond pairwise interactions: structure and dynamics. *Phys Rep* 2020;**874**:1–92.

Batzner S, Musaelian A, Sun L *et al.* E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat Commun* 2022;**13**:2453.

Baur B, Shin J, Zhang S *et al.* Data integration for inferring context-specific gene regulatory networks. *Curr Opin Syst Biol* 2020;**23**:38–46.

Belyaeva A, Cammarata L, Radhakrishnan A *et al.* Causal network models of SARS-CoV-2 expression and aging to identify candidates for drug repurposing. *Nat Commun* 2021;**12**:1024.

Bennett NR, Coventry B, Goreshnik I *et al.* Improving de novo protein binder design with deep learning. *Nat Commun* 2023;**14**:2625.

Bepler T, Berger B. Learning the protein language: evolution, structure, and function. *Cell Syst* 2021;**12**:654–69.e3.

Berchtold NC, Cribbs DH, Coleman PD *et al.* Gene expression changes in the course of normal brain aging are sexually dimorphic. *Proc Natl Acad Sci USA* 2008;**105**:15605–10.

Berenberg D, Gligorijević V, Bonneau R. *Talk in the 3DSIG Track at the Intelligent Systems For Molecular Biology and European Conference on Computational Biology*. https://www.youtube.com/watch?v=1SuojEkR6ZA. Youtube, 2021.

Berge C. *Graphs and Hypergraphs*. Oxford: Elsevier Science, 1985.

Beyer K, Goldstein J, Ramakrishnan R *et al.* When is "Nearest Neighbor" Meaningful? In: *Proceedings of the International Conference on Database Theory*. p. 217–35. Springer Berlin Heidelberg, 1999.

Bilodeau C, Jin W, Jaakkola T *et al.* Generative models for molecular discovery: recent advances and challenges. *Wiley Interdiscip Rev Comput Mol Sci* 2022;**12**:e1608.

Boluki S, Esfahani MS, Qian X *et al.* Incorporating biological prior knowledge for Bayesian learning via maximal knowledge-driven information priors. *BMC Bioinformatics* 2017;**18**:552–80.

Bondy JA, Hemminger RL. Graph reconstruction—a survey. *J Graph Theory* 1977;**1**:227–68.

Bonneau R, Reiss DJ, Shannon P *et al.* The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets de novo. *Genome Biol* 2006;**7**:R36–16.

Borgwardt KM, Ong CS, Schonauer S *et al.* Protein function prediction via graph kernels. *Bioinformatics* 2005;**21**(Suppl 1):i47–56.

Bouritsas G, Frasca F, Zafeiriou S *et al.* Improving graph neural network expressivity via subgraph isomorphism counting. *IEEE Trans Pattern Anal Mach Intell* 2022;**45**:657–68.

Bresson X, Laurent T. The transformer network for the traveling salesman problem. arXiv, 2103.03012, 2021, preprint: not peer reviewed.

Bryant P, Elofsson A. Modelling the dispersion of SARS-CoV-2 on a dynamic network graph. medRxiv, 2020.10.19.20215046, 2020, preprint: not peer reviewed.

Bumin A, Ritz A, Slonim DK *et al.* FiT: fiber-based tensor completion for drug repurposing. In: *Proceedings of the ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*. p. 1–10. New York, NY: Association for Computing Machinery, 2022.

Butler A, Hoffman P, Smibert P *et al.* Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;**36**:411–20.

Cai Y, Jiang X, Li Y *et al.* Resolving power equipment data inconsistency via heterogeneous network alignment. *IEEE Access* 2023;**11**:23980–8.

Callahan TJ, Tripodi IJ, Stefanski AL *et al.* An open source knowledge graph ecosystem for the life sciences. *Sci Data* 2024;**11**:363.

Cambini R, Gallo G, Scutellà MG. Flows on hypergraphs. *Math Program* 1997;**78**:195–217.

Cannistraci CV, Valsecchi MG, Capua I. Age-sex population adjusted analysis of disease severity in epidemics as a tool to devise public health policies for COVID-19. *Sci Rep* 2021;**11**:11787.

Cao M, Zhang H, Park J *et al.* Going the distance for protein function prediction: a new distance metric for protein interaction networks. *PLoS One* 2013;**8**:e76339.

Cao W, Yan Z, He Z *et al.* A comprehensive survey on geometric deep learning. *IEEE Access* 2020;**8**:35929–49.

Carlin DE, Fong SH, Qin Y *et al.* A fast and flexible framework for network-assisted genomic association. *iScience* 2019;**16**:155–61.

Chandak P, Huang K, Zitnik M. Building a knowledge graph to enable precision medicine. *Sci Data* 2023;**10**:67.

Chandrasekaran S, Price ND. Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* 2010;**107**:17845–50.

Chasman D, Gancarz B, Hao L et al. Inferring host gene subnetworks involved in viral replication. *PLoS Comput Biol* 2014;**10**: e1003626.

Chen C, Tong H, Xie L et al. FASCINATE: fast cross-layer dependency inference on multi-layered networks. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. p. 765–74. New York, NY: Association for Computing Machinery, 2016.

Chen M, Ju CJT, Zhou G et al. Multifaceted protein–protein interaction prediction based on Siamese residual RCNN. *Bioinformatics* 2019;**35**:i305–14.

Chen S, Witten DM, Shojaie A. Selection and estimation for mixed graphical models. *Biometrika* 2014;**102**:47–64.

Chen Z, Chen L, Villar S et al. Can graph neural networks count substructures? In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 33, p. 10383–95. Red Hook, NY: Curran Associates, Inc., 2020.

Cheng F, Desai RJ, Handy DE et al. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat Commun* 2018;**9**:2691.

Cheng F, Zhao J, Wang Y et al. Comprehensive characterization of protein–protein interactions perturbed by disease mutations. *Nat Genet* 2021;**53**:342–53.

Chitra U, Park TY, Raphael BJ. NetMix2: a principled network propagation algorithm for identifying altered subnetworks. *J Comput Biol* 2022;**29**:1305–23.

Chitra U, Raphael B. Random walks on hypergraphs with edge-dependent vertex weights. In: *Proceedings of the International Conference on Machine Learning*. p. 1172–81. PMLR, 2019.

Choobdar S, Ahsen ME, Crawford J et al.; DREAM Module Identification Challenge Consortium. Assessment of network module identification across complex diseases. *Nat Methods* 2019; **16**:843–52.

Christensen RG, Enuameh MS, Noyes MB et al. Recognition models to predict DNA-binding specificities of homeodomain proteins. *Bioinformatics* 2012;**28**:i84–9.

Chu Z, Huang F, Fu H et al. Hierarchical graph representation learning for the prediction of drug-target binding affinity. *Inf Sci* 2022; **613**:507–23.

Chuang HY, Lee E, Liu YT et al. Network-based classification of breast cancer metastasis. *Mol Syst Biol* 2007;**3**:140.

Comaniciu D, Engel K, Georgescu B et al. Shaping the future through innovations: from medical imaging to precision medicine. *Med Image Anal* 2016;**33**:19–26.

Cong J, Hagen L, Kahng A. Random walks for circuit clustering. In: *Proceedings of the IEEE International ASIC Conference and Exhibit*. p. P14–2. 1991.

Corso G, Stärk H, Jing B et al. DiffDock: diffusion steps, twists, and turns for molecular docking. In: *Proceedings of the International Conference on Learning Representations*. 2023.

Coşkun M, Koyutürk M. Node similarity-based graph convolution for link prediction in biological networks. *Bioinformatics* 2021; **37**:4501–8.

Costanzo M, VanderSluis B, Koch EN et al. A global genetic interaction network maps a wiring diagram of cellular function. *Science* 2016; **353**:aaf1420.

Costello JC, Stolovitzky G. Seeking the wisdom of crowds through challenge-based competitions in biomedical research. *Clin Pharmacol Ther* 2013;**93**:396–8.

Cowen L, Ideker T, Raphael BJ et al. Network propagation: a universal amplifier of genetic associations. *Nat Rev Genet* 2017;**18**:551–62.

Crawford J, Milenković T. ClueNet: clustering a temporal network based on topological similarity rather than denseness. *PLoS One* 2018;**13**:e0195993.

Cruz LA, Cooke Bailey JN, Crawford DC. Importance of diversity in precision medicine: generalizability of genetic associations across ancestry groups toward better identification of disease susceptibility variants. *Annu Rev Biomed Data Sci* 2023;**6**:339–56.

Dao P, Kim YA, Wojtowicz D et al. BeWith: a between-within method to discover relationships between cancer modules via integrated analysis of mutual exclusivity, co-occurrence and functional interactions. *PLoS Comput Biol* 2017;**13**:e1005695.

Dauparas J, Anishchenko IV, Bennett NR et al. Robust deep learning–based protein sequence design using ProteinMPNN. *Science* 2022; **378**:49–56.

De Cao N, Kipf T. MolGAN: an implicit generative model for small molecular graphs. In: *Proceedings of the International Conference on Machine Learning Workshop on Theoretical Foundations and Applications of Deep Generative Models*. 2018.

De Domenico M. More is different in real-world multilayer networks. *Nat Phys* 2023;**19**:1247–62.

De Domenico M, Solé-Ribalta A, Cozzo E et al. Mathematical formulation of multilayer networks. *Phys Rev X* 2013;**3**:041022.

de Magalhães JP, Budovsky A, Lehmann G et al. The human ageing genomic resources: online databases and tools for biogerontologists. *Aging Cell* 2009;**8**:65–72.

Dehghannasiri R, Yoon BJ, Dougherty ER. Optimal experimental design for gene regulatory networks in the presence of uncertainty. *IEEE/ACM Trans Comput Biol Bioinform* 2015a;**12**:938–50.

Dehghannasiri R, Yoon BJ, Dougherty ER. Efficient experimental design for uncertainty reduction in gene regulatory networks. In: *Proceedings of the MidSouth Computational Biology and Bioinformatics Society Conference*. Vol. 16, p. 1–18. Springer, 2015b.

Demetci P, Santorella R, Sandstede B et al. SCOT: single-cell multi-omics alignment with optimal transport. *J Comput Biol* 2022; **29**:3–18.

Devkota K, Murphy JM, Cowen LJ. GLIDE: combining local methods and diffusion state embeddings to predict missing interactions in biological networks. *Bioinformatics* 2020;**36**:i464–73.

Di Tommaso P, Chatzou M, Floden EW et al. Nextflow enables reproducible computational workflows. *Nat Biotechnol* 2017;**35**:316–9.

Ding C, He X, Xiong H et al. Transitive closure and metric inequality of weighted graphs: detecting protein interaction modules using cliques. *Int J Data Min Bioinform* 2006;**1**:162–77.

Ding K, Wang S, Luo Y. Supervised biological network alignment with graph neural networks. *Bioinformatics* 2023;**39**:i465–74.

Dong Y, Chawla NV, Swami A. metapath2vec: scalable representation learning for heterogeneous networks. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. p. 135–44. New York, NY: Association for Computing Machinery, 2017.

Doria-Belenguer S, Xenos A, Ceddia G et al. A functional analysis of omic network embedding spaces reveals key altered functions in cancer. *Bioinformatics* 2023;**39**:281.

Doria-Belenguer S, Xenos A, Ceddia G et al. The axes of biology: a novel axes-based network embedding paradigm to decipher the functional mechanisms of the cell. *Bioinform Adv* 2024;**4**:vbae075.

Du B, Tong H. MrMine: multi-resolution multi-network embedding. In: *Proceedings of the ACM International Conference on Information and Knowledge Management*. p. 479–88. New York, NY: Association for Computing Machinery, 2019.

Ducournau A, Bretto A. Random walks in directed hypergraphs and application to semi-supervised image segmentation. *Comput Vis Image Underst* 2014;**120**:91–102.

Dwivedi VP, Joshi CK, Luu AT et al. Benchmarking graph neural networks. *J Mach Learn Res* 2022a.

Dwivedi VP, Rampášek L, Galkin M et al. Long range graph benchmark. *Adv Neural Inf Process Syst* 2022b;**35**:22326–40.

Edelsbrunner H, Letscher D, Zomorodian A. Topological persistence and simplification. *Discrete Comput Geometry* 2002;**28**:511–33.

Ektefaie Y, Dasoulas G, Noori A et al. Multimodal learning with graphs. *Nat Mach Intell* 2023;**5**:340–50.

Ektefaie Y, Shen A, Bykova D et al. Evaluating generalizability of artificial intelligence models for molecular datasets. bioRxiv, 2024.02.25.581982, 2024, preprint: not peer reviewed.

Elmarakeby HA, Hwang JH, Arafeh R et al. Biologically informed deep neural network for prostate cancer discovery. *Nature* 2021; **598**:348–52.

Emmert-Streib F, Dehmer M, Shi Y. Fifty years of graph matching, network alignment and network comparison. *Inf Sci* 2016; **346–347**:180–97.

Esteva A, Robicquet A, Ramsundar B *et al.* A guide to deep learning in healthcare. *Nat Med* 2019;**25**:24–9.

Evans R, O'Neill M, Pritzel A *et al.* Protein complex prediction with AlphaFold-Multimer. bioRxiv, 2021.10.04.463034, 2021, preprint: not peer reviewed.

Eyuboglu S, Zitnik M, Leskovec J. Mutual interactors as a principle for phenotype discovery in molecular interaction networks. *Pac Symp Biocomput* 2023;**28**:61–72.

Faisal F, Newaz K, Chaney J *et al.* GRAFENE: graphlet-based alignment-free network approach integrates 3D structural and sequence (residue order) data to improve protein structural comparison. *Sci Rep* 2017;**7**:14890.

Faisal FE, Meng L, Crawford J *et al.* The post-genomic era of biological network alignment. *EURASIP J Bioinform Syst Biol* 2015a;**2015**:3.

Faisal FE, Milenković T. Dynamic networks reveal key players in aging. *Bioinformatics* 2014;**30**:1721–9.

Faisal FE, Zhao H, Milenković T. Global network alignment in the context of aging. *IEEE/ACM Trans Comput Biol Bioinform* 2015b; **12**:40–52.

Faith JJ, Hayete B, Thaden JT *et al.* Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol* 2007;**5**:e8.

Fatemi B, Halcrow J, Perozzi B. Talk like a graph: encoding graphs for large language models. arXiv, 2310.04560, 2023, preprint: not peer reviewed.

Fecho K, Thessen A, Baranzini SE *et al.*; Biomedical Data Translator Consortium. Progress toward a universal biomedical data translator. *Clin Transl Sci* 2022;**15**:1838–47.

Feinberg EN, Sur D, Wu Z *et al.* PotentialNet for molecular property prediction. *ACS Cent Sci* 2018;**4**:1520–30.

Feng S, Heath E, Jefferson B *et al.* Hypergraph models of biological networks to identify genes critical to pathogenic viral response. *BMC Bioinformatics* 2021;**22**:287–21.

Fields S, Song O. A novel genetic system to detect protein-protein interactions. *Nature* 1989;**340**:245–6.

Forster DT, Li SC, Yashiroda Y *et al.* BIONIC: biological network integration using convolutions. *Nat Methods* 2022;**19**:1250–61.

Fortunato S. Community detection in graphs. *Phys Rep* 2010; **486**:75–174.

Franzese N, Groce A, Murali T *et al.* Hypergraph-based connectivity measures for signaling pathway topologies. *PLoS Comput Biol* 2019;**15**:e1007384.

Friedman N, Nachman I, Peér D. Learning Bayesian network structure from massive datasets: the "sparse candidate" algorithm. In: *Proceedings of the Conference on Uncertainty in Artificial Intelligence*. p. 206–15. San Francisco, CA: Morgan Kaufmann Publishers Inc., 1999.

Fu H, Huang F, Liu X *et al.* MVGCN: data integration through multi-view graph convolutional network for predicting links in biomedical bipartite networks. *Bioinformatics* 2022;**38**:426–34.

Gainza P, Sverrisson F, Monti F *et al.* Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nat Methods* 2020;**17**:184–92.

Gainza P, Wehrle S, Hall-Beauvais AV *et al.* De novo design of protein interactions with learned surface fingerprints. *Nature* 2023; **617**:176–84.

Gao J, Zhao Q, Ren W *et al.* Dynamic shortest path algorithms for hypergraphs. *IEEE/ACM Trans Netw* 2015a;**23**:1805–17.

Gao M, Zhou H, Skolnick J. Insights into disease-associated mutations in the human proteome through protein structural analysis. *Structure* 2015b;**23**:1362–9.

Gao Z, Jiang C, Zhang J *et al.* Hierarchical graph learning for protein–protein interaction. *Nat Commun* 2023;**14**:1093.

Garrido-Rodriguez M, Zirngibl K, Ivanova O *et al.* Integrating knowledge and omics to decipher mechanisms via large-scale models of signaling networks. *Mol Syst Biol* 2022;**18**:e11036.

Gärtner T, Flach P, Wrobel S. On graph kernels: hardness results and efficient alternatives. In: *Proceedings of the Conference on Learning Theory*. p. 129–43. Springer Berlin Heidelberg, 2003.

Gaudelet T, Malod-Dognin N, Lugo-Martinez J *et al.* Hypergraphlets give insight into multi-scale organisation of molecular networks. In: *Proceedings of the International Conference on Complex Networks and Their Applications*. p. 41. 2017.

Gaudelet T, Malod-Dognin N, Pržulj N. Integrative data analytic framework to enhance cancer precision medicine. *Netw Syst Med* 2021;**4**:60–73.

Gaudelet T, Malod-Dognin N, Sánchez-Valle J *et al.* Unveiling new disease, pathway, and gene associations via multi-scale neural network. *PLoS One* 2020;**15**:e0231059.

Ghersi D, Singh M. Interaction-based discovery of functionally important genes in cancers. *Nucleic Acids Res* 2014;**42**:e18.

Gillespie ME, Jassal B, Stephan R *et al.* The Reactome Pathway Knowledgebase 2022. *Nucleic Acids Res* 2022;**50**:D687–92.

Gilmer J, Schoenholz SS, Riley PF *et al.* Neural message passing for quantum chemistry. In: *International Conference on Machine Learning*. p. 1263–72. PMLR, 2017.

Glass K, Huttenhower C, Quackenbush J *et al.* Passing messages between biological networks to refine predicted interactions. *PLoS One* 2013;**8**:e64832.

Glass K, Quackenbush J, Spentzos D *et al.* A network model for angiogenesis in ovarian cancer. *BMC Bioinformatics* 2015;**16**:115–7.

Gligorijevic V, Malod-Dognin N, Pržulj N. Integrative methods for analyzing big data in precision medicine. *Proteomics* 2016a;**16**:741–58.

Gligorijevic V, Malod-Dognin N, Pržulj N. Patient-specific data fusion for cancer stratification and personalised treatment. In: *Proceedings of the Pacific Symposium on Biocomputing*. 2016b.

Gligorijević V, Pržulj N. Methods for biological data integration: perspectives and challenges. *J R Soc Interface* 2015;**12**:20150571.

Gligorijević V, Renfrew PD, Kosciólek T *et al.* Structure-based protein function prediction using graph convolutional networks. *Nat Commun* 2021;**12**:3168.

Gong X, Li H, Zou N *et al.* General framework for E(3)-equivariant neural network representation of density functional theory Hamiltonian. *Nat Commun* 2023;**14**:2848.

Gosline SJC, Spencer SJ, Ursu O *et al.* SAMNet: a network-based approach to integrate multi-dimensional high throughput datasets. *Integr Biol (Camb)* 2012;**4**:1415–27. ISSN 1757–9708.

Greene CS, Krishnan A, Wong AK *et al.* Understanding multicellular function and disease with human tissue-specific networks. *Nat Genet* 2015;**47**:569–76.

Greenfield A, Hafemeister C, Bonneau R. Robust data-driven incorporation of prior knowledge into the inference of dynamic regulatory networks. *Bioinformatics* 2013;**29**:1060–7.

Grover A, Leskovec J. node2vec: scalable feature learning for networks. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. p. 855–64. New York, NY: Association for Computing Machinery, 2016.

Gu S, Jiang M, Guzzi PH *et al.* Modeling multi-scale data via a network of networks. *Bioinformatics* 2022;**38**:2544–53.

Gu S, Johnson J, Faisal F *et al.* From homogeneous to heterogeneous network alignment via colored graphlets. *Sci Rep* 2018;**8**:12524.

Gu S, Milenković T. Data-driven network alignment. *PLoS One* 2020; **15**:e0234978.

Gu S, Milenković T. Data-driven biological network alignment that uses topological, sequence, and functional information. *BMC Bioinformatics* 2021;**22**:34–24.

Gui H, Liu J, Tao F *et al.* Large-scale embedding learning in heterogeneous event data. In: *Proceedings of the IEEE International Conference on Data Mining*. p. 907–12. Institute of Electrical and Electronics Engineers Inc., 2016.

Guiñazú MF, Cortés V, Ibáñez CF *et al.* Employing online social networks in precision-medicine approach using information fusion predictive model to improve substance use surveillance: a lesson from twitter and marijuana consumption. *Inf Fusion* 2020;**55**:150–63.

Guney E, Menche J, Vidal M *et al.* Network-based in silico drug efficacy screening. *Nat Commun* 2016;**7**:10331.

Guo MG, Sosa DN, Altman RB. Challenges and opportunities in network-based solutions for biological questions. *Brief Bioinform* 2022;**23**:bbab437.

Gutteridge B, Dong X, Bronstein MM *et al*. DRew: dynamically rewired message passing with delay. In: *Proceedings of the International Conference on Machine Learning*. p. 12252–67. PMLR, 2023.

Guzzi HP, Petrizzelli F, Mazza T. Disease spreading modeling and analysis: a survey. *Brief Bioinform* 2022;**23**:bbac230.

Guzzi PH, Milenković T. Survey of local and global biological network alignment: the need for reconciling the two sides of the same coin. *Brief Bioinform* 2017;**19**:472–81.

Gysi DM, do Valle ÌF, Zitnik M *et al*. Network medicine framework for identifying drug-repurposing opportunities for COVID-19. *Proc Natl Acad Sci USA* 2021;**118**:e2025581118.

Gysi DM, Nowick K. Construction, comparison and evolution of networks in life sciences and other disciplines. *J R Soc Interface* 2020; **17**:20190610.

Gysi DM, Voigt A, Fragoso T *et al*. wTO: an R package for computing weighted topological overlap and a consensus network with integrated visualization tool. *BMC Bioinformatics* 2018;**19**:392–16.

Halu A, De Domenico M, Arenas A *et al*. The multiplex network of human diseases. *NPJ Syst Biol Appl* 2019;**5**:15.

Hamilton W, Bajaj P, Zitnik M *et al*. Embedding logical queries on knowledge graphs. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 31. Red Hook, NY: Curran Associates, Inc., 2018.

Hamilton W, Ying Z, Leskovec J. Inductive representation learning on large graphs. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 30. Red Hook, NY: Curran Associates, Inc.,2017a.

Hamilton WL. *Graph Representation Learning*. Morgan & Claypool Publishers, 2020.

Hamilton WL, Ying R, Leskovec J. Representation learning on graphs: methods and applications. *IEEE Data Eng Bull* 2017b;**40**:52–74.

Hamp T, Rost B. More challenges for machine-learning protein interactions. *Bioinformatics* 2015;**31**:1521–5.

Haque A, Engel J, Teichmann SA *et al*. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med* 2017;**9**:75–12.

Hashemifar S, Neyshabur B, Khan AA *et al*. Predicting protein–protein interactions through sequence-based deep learning. *Bioinformatics* 2018;**34**:i802–10.

Hassani K, Khasahmadi AH. Contrastive multi-view representation learning on graphs. In: *Proceedings of the International Conference on Machine Learning*. p. 4116–26. PMLR, 2020.

Hawe JS, Theis FJ, Heinig M. Inferring interaction networks from multi-omics data. *Front Genet* 2019;**10**:535.

He H, Queen O, Koker T *et al*. Domain adaptation for time series under feature and label shifts. In: *Proceedings of the International Conference in Machine Learning*. PMLR, 2023.

Heckerman D, Chickering DM, Meek C *et al*. Dependency networks for inference, collaborative filtering, and data visualization. *J Mach Learn Res* 2000;**1**:49–75.

Heil BJ, Hoffman MM, Markowetz F *et al*. Reproducibility standards for machine learning in the life sciences. *Nat Methods* 2021;**18**:1132–5.

Hein M, Setzer S, Jost L *et al*. The total variation on hypergraphs–learning on hypergraphs revisited. In: *Proceedings of the Advances in Neural Information Processing Systems*. p. 2427–35. Red Hook, NY: Curran Associates, Inc., 2013.

Hérault L, Poplineau M, Duprez E *et al*. A novel boolean network inference strategy to model early hematopoiesis aging. *Comput Struct Biotechnol J* 2023;**21**:21–33.

Hetzel L, Fischer DS, Günnemann S *et al*. Graph representation learning for single-cell biology. *Curr Opin Syst Biol* 2021;**28**:100347.

Heumos L, Schaar AC, Lance C *et al*.; Single-cell Best Practices Consortium. Best practices for single-cell analysis across modalities. *Nat Rev Genet* 2023;**24**:550–72.

Himmelstein DS, Baranzini SE. Heterogeneous network edge prediction: a data integration approach to prioritize disease-associated genes. *PLoS Comput Biol* 2015;**11**:e1004259.

Himmelstein DS, Lizee A, Hessler C *et al*. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *Elife* 2017; **6**:e26726.

Hu W, Fey M, Ren H *et al*. OGB-LSC: a large-scale challenge for machine learning on graphs. In: *Proceedings of the Conference on Neural Information Processing Systems*. Vol. 35. 2021.

Hu W, Fey M, Zitnik M *et al*. Open graph benchmark: datasets for machine learning on graphs. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 33, p. 22118–33. Red Hook, NY: Curran Associates, Inc., 2020a.

Hu W, Liu B, Gomes J *et al*. Strategies for pre-training graph neural networks. In: *Proceedings of the International Conference on Learning Representations*. 2020b.

Hu X, Li F, Samaras D *et al*. Topology-preserving deep image segmentation. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 32. Red Hook, NY: Curran Associates, Inc., 2019.

Huang K, Chandak P, Wang Q *et al*. Zero-shot prediction of therapeutic use with geometric deep learning and clinician centered design. medRxiv, 2023.03.19.23287458, 2023a, preprint: not peer reviewed.

Huang K, Fu T, Gao W *et al*. Therapeutics data commons: machine learning datasets and tasks for drug discovery and development. In: *Proceedings of the Conference on Neural Information Processing Systems*. 2021.

Huang K, Fu T, Gao W *et al*. Artificial intelligence foundation for therapeutic science. *Nat Chem Biol* 2022;**18**:1033–6.

Huang K, Jin Y, Candes E *et al*. Uncertainty quantification over graph with conformalized graph neural networks. *Adv Neural Inf Process Syst* 2023b;**36**:26699–721.

Huang K, Xiao C, Glass LM *et al*. SkipGNN: predicting molecular interactions with skip-graph networks. *Sci Rep* 2020;**10**:21092–16.

Huang S, Poursafaei F, Danovitch J *et al*. Temporal graph benchmark for machine learning on temporal graphs. *Adv Neural Inf Process Syst* 2024;**36**.

Huang T, Glass K, Zeleznik OA *et al*. A network analysis of biomarkers for type 2 diabetes. *Diabetes* 2019;**68**:281–90.

Hüllermeier E, Waegeman W. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Mach Learn* 2021;**110**:457–506.

Hulovatyy Y, Chen H, Milenković T. Exploring the structure and function of temporal networks with dynamic graphlets. *Bioinformatics* 2015;**31**:i171–80.

Hulovatyy Y, Milenković T. SCOUT: simultaneous time segmentation and community detection in dynamic networks. *Sci Rep* 2016; **6**:37557.

Hulovatyy Y, Solava R, Milenković T. Revealing missing parts of the interactome via link prediction. *PLoS One* 2014;**9**:e90073.

Hunter LE, Hopfer C, Terry SF *et al*. Reporting actionable research results: shared secrets can save lives. *Sci Transl Med* 2012; **4**:143cm8.

Huynh-Thu VA, Irrthum A, Wehenkel L *et al*. Inferring regulatory networks from expression data using tree-based methods. *PLoS One* 2010;**5**:e12776.

Hwa Chu J, Hersh CP, Castaldi PJ *et al*. Analyzing networks of phenotypes in complex diseases: methodology and applications in COPD. *BMC Syst Biol* 2014;**8**:78–12.

Ietswaart R, Gyori BM, Bachman JA *et al*. GeneWalk identifies relevant gene functions for a biological context using network representation learning. *Genome Biol* 2021;**22**:55–35.

Ihler E, Wagner D, Wagner F. Modeling hypergraphs by graphs with the same mincut properties. *Inf Process Lett* 1993;**45**:171–5.

Ingraham J, Garg V, Barzilay R *et al*. Generative models for graph-based protein design. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 32. Red Hook, NY: Curran Associates, Inc.,2019.

Jacobsen NE. *NMR Spectroscopy Explained: Simplified Theory, Applications and Examples for Organic Chemistry and Structural Biology*. John Wiley & Sons, 2007.

Jahanshad N, Rajagopalan P, Hua X *et al.*; Alzheimer's Disease Neuroimaging Initiative. Genome-wide scan of healthy human connectome discovers spon1 gene variant influencing dementia severity. *Proc Natl Acad Sci USA* 2013;**110**:4768–73.

Jia K, Cui C, Gao Y *et al.* An analysis of aging-related genes derived from the genotype-tissue expression project (GTEx). *Cell Death Discov* 2018;**4**:26.

Jiang T, Zeng Q, Zhao T *et al.* Biomedical knowledge graphs construction from conditional statements. *IEEE/ACM Trans Comput Biol Bioinform* 2021;**18**:823–35.

Jiang Y, Oron T, Clark WT *et al.* An expanded evaluation of protein function prediction methods shows an improvement in accuracy. *Genome Biol* 2016;**17**:184–19.

Jiménez J, Doerr S, Martínez-Rosell G *et al.* DeepSite: protein-binding site predictor using 3D-convolutional neural networks. *Bioinformatics* 2017;**33**:3036–42.

Jin W, Barzilay R, Jaakkola T. Junction tree variational autoencoder for molecular graph generation. In: *Proceedings of the International Conference on Machine Learning.* p. 2323–32. PMLR, 2018.

Jumper JM, Evans R, Pritzel A *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;**596**:583–9.

Kaiser J. NIH plots million-person megastudy. *Science* 2015;**347**:817.

Karr JR, Sanghvi JC, Macklin DN *et al.* A whole-cell computational model predicts phenotype from genotype. *Cell* 2012;**150**:389–401.

Kesimoglu ZN, Bozdag S. SUPREME: multiomics data integration using graph convolutional networks. *NAR Genom Bioinform* 2023a;**5**:lqad063.

Kesimoglu ZN, Bozdag S. GRAF: graph attention-aware fusion networks. arXiv, 2303.16781, 2023b, preprint: not peer reviewed.

Kestler HA, Wawra C, Kracher B *et al.* Network modeling of signal transduction: establishing the global view. *Bioessays* 2008;**30**:1110–25.

Kim CY, Baek S, Cha J *et al.* HumanNet v3: an improved database of human gene networks for disease research. *Nucleic Acids Res* 2021;**50**:D632–9.

Kim S, Shin SY, Lee IH *et al.* PIE: an online prediction system for protein–protein interactions from text. *Nucleic Acids Res* 2008;**36**:W411–5.

Kim S, Sung J, Foo M *et al.* Uncovering the nutritional landscape of food. *PLoS One* 2015;**10**:e0118697.

Kim YA, Basso RS, Wojtowicz D *et al.* Identifying drug sensitivity subnetworks with NETPHIX. *iScience* 2020;**23**:101619.

Kinsley AC, Rossi G, Silk MJ *et al.* Multilayer and multiplex networks: an introduction to their use in veterinary epidemiology. *Front Vet Sci* 2020;**7**:596.

Kipf TN, Welling M. Variational graph auto-encoders. In: *Proceedings of the Neural Information Processing Systems Workshop on Bayesian Deep Learning.* 2016.

Kivelä M, Arenas A, Barthelemy M *et al.* Multilayer networks. *J Complex Netw* 2014;**2**:203–71.

Klamt S, Haus UU, Theis F. Hypergraphs and cellular networks. *PLoS Comput Biol* 2009;**5**:e1000385.

Kobren SN, Chazelle B, Singh M. PertInInt: an integrative, analytical approach to rapidly uncover cancer driver genes with perturbed interactions and functionalities. *Cell Syst* 2020;**11**:63–74.e7.

Kobren SN, Singh M. Systematic domain-based aggregation of protein structures highlights DNA-, RNA- and other ligand-binding positions. *Nucleic Acids Res* 2019;**47**:582–93.

Koller D, Friedman N. *Probabilistic Graphical Models: Principles and Techniques.* MIT Press, 2009.

Kong J, Ha D, Lee J *et al.* Network-based machine learning approach to predict immunotherapy response in cancer patients. *Nat Commun* 2022;**13**:3703.

Köster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* 2012;**28**:2520–2.

Kovács IA, Luck K, Spirohn K *et al.* Network-based prediction of protein interactions. *Nat Commun* 2019;**10**:1240.

Krieger S, Kececioglu J. Computing optimal factories in metabolic networks with negative regulation. *Bioinformatics* 2022a;**38**:i369–77.

Krieger S, Kececioglu J. Heuristic shortest hyperpaths in cell signaling hypergraphs. *Algorithms Mol Biol* 2022b;**17**:12.

Krieger S, Kececioglu J. Computing shortest hyperpaths for pathway inference in cellular reaction networks. In: *Proceedings of the International Conference on Research in Computational Molecular Biology.* p. 155–73. Switzerland: Springer Nature, 2023.

Kryshtafovych A, Antczak M, Szachniuk M *et al.* New prediction categories in CASP15. *Proteins* 2023;**91**:1550–7.

Kryshtafovych A, Schwede T, Topf M *et al.* Critical assessment of methods of protein structure prediction (CASP)-round XIV. *Proteins* 2021;**89**:1607–17.

Ku WL, Duggal G, Li Y *et al.* Interpreting patterns of gene expression: signatures of coregulation, the data processing inequality, and triplet motifs. *PLoS One* 2012;**7**:e31969.

Kuchaiev O, Milenković T, Memišević V *et al.* Topological network alignment uncovers biological function and phylogeny. *J R Soc Interface* 2010;**7**:1341–54.

Kuijjer ML, Hsieh PH, Quackenbush J *et al.* lionessR: single sample network inference in R. *BMC Cancer* 2019a;**19**:1003–6.

Kuijjer ML, Tung MG, Yuan G *et al.* Estimating sample-specific regulatory networks. *iScience* 2019b;**14**:226–40.

Lambin P, Leijenaar RTH, Deist TM *et al.* Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;**14**:749–62.

Langhauser F, Casas AI, Vi Dao VT *et al.* A diseasome cluster-based drug repurposing of soluble guanylate cyclase activators from smooth muscle relaxation to direct neuroprotection. *NPJ Syst Biol Appl* 2018;**4**:8.

Larsen SJ, Röttger R, Schmidt HHHW *et al.* E. coli gene regulatory networks are inconsistent with gene expression data. *Nucleic Acids Res* 2019;**47**:85–92.

Lazareva O, Baumbach J, List M *et al.* On the limits of active module identification. *Brief Bioinform* 2021;**22**:bbab066.

Le Novere N. Quantitative and logic modelling of molecular and gene networks. *Nat Rev Genet* 2015;**16**:146–58.

Lee B, Zhang S, Poleksic A *et al.* Heterogeneous multi-layered network model for omics data integration and analysis. *Front Genet* 2019;**10**:1381.

Lee Y, Kim AH, Kim E *et al.* Changes in the gut microbiome influence the hypoglycemic effect of metformin through the altered metabolism of branched-chain and nonessential amino acids. *Diabetes Res Clin Pract* 2021;**178**:108985.

Leiserson M, Vandin F, Wu H *et al.* Pan-Cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat Genet* 2015;**47**:106–14.

Lelong S, Zhou X, Afrasiabi C *et al.* BioThings SDK: a toolkit for building high-performance data APIs in biomedical research. *Bioinformatics* 2022;**38**:2077–9.

Leordeanu M, Sminchisescu C. Efficient hypergraph clustering. In: *Proceedings of the International Conference on Artificial Intelligence and Statistics.* p. 676–84. PMLR, 2012.

Li J, Huang Y, Chang H *et al.* Semi-supervised hierarchical graph classification. *IEEE Trans Pattern Anal Mach Intell* 2022a;**45**:6265–76.

Li MM, Huang K, Zitnik M. Graph representation learning in biomedicine and healthcare. *Nat Biomed Eng* 2022b;**6**:1353–69.

Li MM, Huang Y, Sumathipala M *et al.* Contextual AI models for single-cell protein biology. *Nat Methods* 2024.

Li MM, Zitnik M. Deep contextual learners for protein networks. In: *Proceedings of the International Conference on Machine Learning Workshop on Computational Biology.* 2021.

Li Q, Button-Simons KA, Sievert MA *et al.* Enhancing gene co-expression network inference for the malaria parasite plasmodium falciparum. bioRxiv, 2023.05.31.543171, 2023b, preprint: not peer reviewed.

Li Q, Milenković T. Supervised prediction of aging-related genes from a context-specific protein interaction subnetwork. *IEEE/ACM Trans Comput Biol Bioinform* 2022;**19**:2484–98.

Li Q, Newaz K, Milenković T. Improved supervised prediction of aging-related genes via weighted dynamic network analysis. *BMC Bioinformatics* 2021;**22**:520–6.

Li Q, Newaz K, Milenković T. Towards future directions in data-integrative supervised prediction of human aging-related genes. *Bioinform Adv* 2022c;**2**:vbac081.

Li Y, Vinyals O, Dyer C *et al.* Learning deep generative models of graphs. arXiv preprint arXiv: 1803.03324, 2018a, preprint: not peer reviewed.

Li Y, Yu R, Shahabi C *et al.* Diffusion convolutional recurrent neural network: data-driven traffic forecasting. In: *Proceedings of the International Conference on Learning Representations*. 2018b.

Lichtblau Y, Zimmermann K, Haldemann B *et al.* Comparative assessment of differential network analysis methods. *Brief Bioinform* 2017;**18**:837–50.

Lin Z, Akin H, Rao R *et al.* Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 2023; **379**:1123–30.

Liu R, Krishnan A. Open biomedical network benchmark: a Python toolkit for benchmarking datasets with biomedical networks. In: *Proceedings of the Machine Learning in Computational Biology*. p. 23–59. PMLR, 2024.

Liu S, Hachen D, Lizardo O *et al.* The power of dynamic social networks to predict individuals' mental health. In: *Proceedings of the Pacific Symposium on Biocomputing*. Vol. 25, p. 635–46. 2020.

Liu S, Vahedian F, Hachen D *et al.* Heterogeneous network approach to predict individuals' mental health. *ACM Trans Knowl Discov Data* 2021a;**15**:1–26.

Liu X, Wang Y, Ji H *et al.* Personalized characterization of diseases using sample-specific networks. *Nucleic Acids Res* 2016;**44**:e164.

Liu X, Zhang F, Hou Z *et al.* Self-supervised learning: generative or contrastive. *IEEE Trans Knowl Data Eng* 2021b;**35**:1–876.

Liu Y, Jin M, Pan S *et al.* Graph self-supervised learning: a survey. *IEEE Trans Knowl Data Eng* 2022;**35**:1–5900.

Lopes-Ramos CM, Chen CY, Kuijjer ML *et al.* Sex differences in gene expression and regulatory networks across 29 human tissues. *Cell Rep* 2020;**31**:107795.

Lopes-Ramos CM, Kuijjer ML, Ogino S *et al.* Gene regulatory network analysis identifies sex-linked differences in colon cancer drug metabolism. *Cancer Res* 2018;**78**:5538–47.

Luck K, Kim DK, Lambourne L *et al.* A reference map of the human binary protein interactome. *Nature* 2020;**580**:402–8.

Lugo-Martinez J, Pejaver V, Pagel KA *et al.* The loss and gain of functional amino acid residues is a common mechanism causing human inherited disease. *PLoS Comput Biol* 2016;**12**:e1005091.

Lugo-Martinez J, Radivojac P. Generalized graphlet kernels for probabilistic inference in sparse graphs. *Netw Sci* 2014;**2**:254–76.

Lugo-Martinez J, Zeiberg D, Gaudelet T *et al.* Classification in biological networks with hypergraphlet kernels. *Bioinformatics* 2021; **37**:1000–7.

Luo R, Sun L, Xia Y *et al.* BioGPT: generative pre-trained transformer for biomedical text generation and mining. *Brief Bioinform* 2022a; **23**:bbac409.

Luo X, Ju W, Qu M *et al.* CLEAR: cluster-enhanced contrast for self-supervised graph representation learning. *IEEE Trans Neural Netw Learn Syst* 2022b;**35**:899–912.

Ma J, Yu MK, Fong S *et al.* Using deep learning to model the hierarchical structure and function of a cell. *Nat Methods* 2018;**15**:290–8.

Ma L, Shao Z, Li L *et al.* Heuristics and metaheuristics for biological network alignment: a review. *Neurocomputing* 2022;**491**:426–41.

Magnano CS, Gitter A. Automating parameter selection to avoid implausible biological pathway models. *NPJ Syst Biol Appl* 2021; **7**:12.

Maheshwari P, Albert R. A framework to find the logic backbone of a biological network. *BMC Syst Biol* 2017;**11**:122–18.

Maleki S, Saless D, Wall DP *et al.* HyperNetVec: fast and scalable hierarchical embedding for hypergraphs. In: *Proceedings of the International Conference on Network Science*. p. 169–83. Springer International Publishing, 2022.

Malod-Dognin N, Ceddia G, Gvozdenov M *et al.* A phenotype driven integrative framework uncovers molecular mechanisms of a rare hereditary thrombophilia. *PLoS One* 2023;**18**:e0284084.

Malod-Dognin N, Pancaldi V, Valencia A *et al.* Chromatin network markers of leukemia. *Bioinformatics* 2020;**36**:i455–63.

Malod-Dognin N, Petschnigg J, Pržulj N. Precision medicine—a promising, yet challenging road lies ahead. *Curr Opin Syst Biol* 2018; **7**:1–7.

Malod-Dognin N, Petschnigg J, Windels S *et al.* Towards a data-integrated cell. *Nat Commun* 2019;**10**:805.

Malod-Dognin N, Pržulj N. GR-Align: fast and flexible alignment of protein 3D structures using graphlet degree similarity. *Bioinformatics* 2014;**30**:1259–65.

Mamano N, Hayes W. SANA: simulated annealing far outperforms many other search algorithms for biological network alignment. *Bioinformatics* 2017;**33**:2156–64.

Manessi F, Rozza A, Manzo M. Dynamic graph convolutional networks. *Pattern Recognit* 2020;**97**:107000.

Manske M, Böhme U, Püthe C *et al.* GeneDB and Wikidata. *Wellcome Open Res* 2019;**4**:114.

Marbach D, Costello JC, Küffner R *et al.*; DREAM5 Consortium. Wisdom of crowds for robust gene network inference. *Nat Methods* 2012a;**9**:796–804.

Marbach D, Roy S, Ay F *et al.* Predictive regulatory models in drosophila melanogaster by integrative inference of transcriptional networks. *Genome Res* 2012b;**22**:1334–49.

Margolin AA, Nemenman I, Basso K *et al.* ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* 2006;**7**(Suppl 1):S7–15.

Mateo J, McKay RR, Abida W *et al.* Accelerating precision medicine in metastatic prostate cancer. *Nat Cancer* 2020;**1**:1041–53.

McDermott MB, Yap B, Szolovits P *et al.* Structure-inducing pre-training. *Nat Mach Intell* 2023;**5**:612–21.

Meier J, Rao R, Verkuil R *et al.* Language models enable zero-shot prediction of the effects of mutations on protein function. *Proc Adv Neural Inf Process Syst* 2021;**34**:29287–303.

Menche J, Sharma A, Kitsak M *et al.* Uncovering disease-disease relationships through the incomplete interactome. *Science* 2015; **347**:1257601.

Meng L, Striegel A, Milenković T. Local versus global biological network alignment. *Bioinformatics* 2016;**32**:3155–64.

Mervis J. Fix the system, not the students. *Science* 2022;**375**:956–9.

Meyer MJ, Beltrán JF, Liang S *et al.* Interactome INSIDER: a structural interactome browser for genomic studies. *Nat Methods* 2018; **15**:107–14.

Meyer P, Saez-Rodriguez J. Advances in systems biology modeling: 10 years of crowdsourcing DREAM challenges. *Cell Syst* 2021; **12**:636–53.

Milano M, Milenković T, Cannataro M *et al.* L-HetNetAligner: a novel algorithm for local alignment of heterogeneous biological networks. *Sci Rep* 2020;**10**:3901–2322.

Milenković T, Pržulj N. Uncovering biological network function via graphlet degree signatures. *Cancer Inform* 2008;**6**:257–73.

Milo R, Itzkovitz S, Kashtan N *et al.* Superfamilies of evolved and designed networks. *Science* 2004;**303**:1538–42.

Miraldi ER, Pokrovskii MV, Watters A *et al.* Leveraging chromatin accessibility for transcriptional regulatory network inference in T helper 17 cells. *Genome Res* 2019;**29**:449–63.

Mishra S, Wang YX, Wei CC *et al.* VTG-Net: a CNN based vessel topology graph network for retinal artery/vein classification. *Front Med (Lausanne)* 2021;**8**:750396.

Mitra K, Carvunis AR, Ramesh SK *et al.* Integrative approaches for finding modular structure in biological networks. *Nat Rev Genet* 2013;**14**:719–32.

Montagud A, Béal J, Tobalina L *et al.* Patient-specific Boolean models of signalling networks guide personalised treatments. *Elife* 2022; **11**:e72626.

Moore JE, Purcaro MJ, Pratt HE *et al.*; ENCODE Project Consortium. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* 2020;**583**:699–710.

Moris N, Pina C, Arias AM. Transition states and cell fate decisions in epigenetic landscapes. *Nat Rev Genet* 2016;**17**:693–703.

Morris RT, O'Connor TR, Wyrick JJ. Ceres: software for the integrated analysis of transcription factor binding sites and nucleosome positions in saccharomyces cerevisiae. *Bioinformatics* 2010;**26**:168–74.

Morselli Gysi D, Barabási AL. Non-coding RNAs improve the predictive power of network medicine. *Proc Natl Acad Sci USA* 2023;**120**: e2301342120.

Morselli Gysi D, de Miranda Fragoso T, Zebardast F *et al.* Whole transcriptomic network analysis using co-expression differential network analysis (CoDiNA). *PLoS One* 2020;**15**:e0240523.

Mosca R, Céol A, Aloy P. Interactome3D: adding structural details to protein networks. *Nat Methods* 2013;**10**:47–53.

Moult J, Pedersen JT, Judson R *et al.* A large-scale experiment to assess protein structure prediction methods. *Proteins* 1995;**23**:ii–v.

Mucha P, Richardson T, Macon K *et al.* Community structure in time-dependent, multiscale, and multiplex networks. *Science* 2010;**328**:876–8.

Murgas KA, Saucan E, Sandhu R. Hypergraph geometry reflects higher-order dynamics in protein interaction networks. *Sci Rep* 2022;**12**:20879.

Nasser R, Sharan R. BERTwalk for integrating gene networks to predict gene-to pathway-level properties. *Bioinform Adv* 2023;**3**:vbad086.

Natarajan N, Dhillon IS. Inductive matrix completion for predicting gene-disease associations. *Bioinformatics* 2014;**30**:i60–8.

Needham EJ, Parker BL, Burykin T *et al.* Illuminating the dark phosphoproteome. *Sci Signal* 2019;**12**:eaau8645.

Nelson W, Žitnik M, Wang B *et al.* To embed or not: network embedding as a paradigm in computational biology. *Front Genet* 2019; **10**:381.

Neph S, Stergachis AB, Reynolds A *et al.* Circuitry and dynamics of human transcription factor regulatory networks. *Cell* 2012;**150**:1274–86.

Neville J, Gallagher B, Eliassi-Rad T. Evaluating statistical tests for within-network classifiers of relational data. In: *Proceedings of the IEEE International Conference on Data Mining*. p. 397–406. 2009.

Neville J, Gallagher B, Eliassi-Rad T *et al.* Correcting evaluation bias of relational classifiers with network cross validation. *Knowl Inf Syst* 2012;**30**:31–55.

Newaz K, Ghalehnovi M, Rahnama A *et al.* Network-based protein structural classification. *R Soc Open Sci* 2020;**7**:191461.

Newaz K, Milenković T. Graphlets in network science and computational biology. In: *Analyzing Network Data in Biology and Medicine: An Interdisciplinary Textbook for Biological, Medical and Computational Scientists*. p. 193–240. Cambridge: Cambridge University Press, 2019.

Newaz K, Milenković T. Inference of a dynamic aging-related biological subnetwork via network propagation. *IEEE/ACM Trans Comput Biol Bioinformatics* 2022;**19**:974–88.

Newaz K, Piland J, Clark PL *et al.* Multi-layer sequential network analysis improves protein 3D structural classification. *Proteins* 2022;**90**:1721–31.

Newman M. *Networks*. Oxford University Press, 2018.

Newman ME. Modularity and community structure in networks. *Proc Natl Acad Sci USA* 2006;**103**:8577–82.

Ni J, Koyuturk M, Tong H *et al.* Disease gene prioritization by integrating tissue-specific molecular networks using a robust multi-network model. *BMC Bioinformatics* 2016;**17**:453–13.

Nicholson DN, Greene CS. Constructing knowledge graphs and their biomedical applications. *Comput Struct Biotechnol J* 2020; **18**:1414–28.

Nielsen MW, Bloch CW, Schiebinger L. Making gender diversity work for scientific discovery and innovation. *Nat Hum Behav* 2018; **2**:726–34.

Nishihara R, Glass K, Mima K *et al.* Biomarker correlation network in colorectal carcinoma by tumor anatomic location. *BMC Bioinformatics* 2017;**18**:304–14.

Niu P, Soto MJ, Yoon BJ *et al.* TRIMER: transcription regulation integrated with metabolic regulation. *iScience* 2021;**24**:103218.

Nykter M, Price ND, Aldana M *et al.* Gene expression dynamics in the macrophage exhibit criticality. *Proc Natl Acad Sci USA* 2008; **105**:1897–900.

Ourfali O, Shlomi T, Ideker T *et al.* SPINE: a framework for signaling-regulatory pathway inference from cause-effect experiments. *Bioinformatics* 2007;**23**:i359–66.

Padi M, Quackenbush J. Detecting phenotype-driven transitions in regulatory network structure. *NPJ Syst Biol Appl* 2018;**4**:16.

Page RD. Wikidata and the bibliography of life. *PeerJ* 2022;**10**:e13712.

Pai S, Bader GD. Patient similarity networks for precision medicine. *J Mol Biol* 2018;**430**:2924–38.

Pai S, Hui S, Isserlin R *et al.* netDx: interpretable patient classification using integrated patient similarity networks. *Mol Syst Biol* 2019; **15**:e8497.

Pan S, Luo L, Wang Y *et al.* Unifying large language models and knowledge graphs: a roadmap. *IEEE Trans Knowl Data Eng* 2024; **36**:3580–99.

Pandey J, Koyutürk M, Kim Y *et al.* Functional annotation of regulatory pathways. *Bioinformatics* 2007;**23**:i377–86.

Papanikolaou N, Pavlopoulos GA, Theodosiou T *et al.* Protein–protein interaction predictions using text mining methods. *Methods* 2015; **74**:47–53.

Pareja A, Domeniconi G, Chen J *et al.* EvolveGCN: evolving graph convolutional networks for dynamic graphs. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34, p. 5363–70. 2020.

Park J, Hescott BJ, Slonim DK. Pathway centrality in protein interaction networks identifies putative functional mediating pathways in pulmonary disease. *Sci Rep* 2019;**9**:5863.

Park Y, Marcotte EM. Flaws in evaluation schemes for pair-input computational predictions. *Nat Methods* 2012;**9**:1134–6.

Patten JJ, Keiser PT, Morselli-Gysi D *et al.* Identification of potent inhibitors of SARS-CoV-2 infection by combined pharmacological evaluation and cellular network prioritization. *iScience* 2022; **25**:104925.

Paull EO, Carlin DE, Niepel M *et al.* Discovering causal pathways linking genomic events to transcriptional states using tied diffusion through interacting events (TieDIE). *Bioinformatics* 2013; **29**:2757–64.

Peck Justice SA, McCracken NA, Victorino JF *et al.* Boosting detection of low-abundance proteins in thermal proteome profiling experiments by addition of an isobaric trigger channel to TMT multiplexes. *Anal Chem* 2021;**93**:7000–10.

Peng C, Xia F, Naseriparsa M *et al.* Knowledge graphs: opportunities and challenges. *Artif Intell Rev* 2023;**56**:1–32.

Peng H, Wang H, Du B *et al.* Spatial temporal incidence dynamic graph neural networks for traffic flow forecasting. *Inf Sci* 2020; **521**:277–90.

Perez De Souza L, Alseekh S, Brotman Y *et al.* Network-based strategies in metabolomics data analysis and interpretation: from molecular networking to biological interpretation. *Expert Rev Proteomics* 2020;**17**:243–55.

Perozzi B, Al-Rfou R, Skiena S. DeepWalk: online learning of social representations. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. p. 701–10. New York, NY: Association for Computing Machinery, 2014.

Persikov AV, Singh M. De novo prediction of DNA-binding specificities for Cys_2His_2 zinc finger proteins. *Nucleic Acids Res* 2014; **42**:97–108.

Petti M, Farina L. Network medicine for patients' stratification: from single-layer to multi-omics. *WIREs Mech Dis* 2023;**15**:e1623.

Piersimoni L, Kastritis PL, Arlt C *et al.* Cross-linking mass spectrometry for investigating protein conformations and protein-protein interactions—a method for all seasons. *Chem Rev* 2022;**122**:7500–31.

Pierson E, Koller D, Battle A *et al.*; GTEx Consortium. Sharing and specificity of co-expression networks across 35 human tissues. *PLoS Comput Biol* 2015;**11**:e1004220.

Pinheiro PO, Rackers J, Kleinhenz J *et al.* 3D molecule generation by denoising voxel grids. *Adv Neural Inf Process Syst* 2024;**36**.

Pio-Lopez L, Valdeolivas A, Tichit L *et al.* MultiVERSE: a multiplex and multiplex-heterogeneous network embedding approach. *Sci Rep* 2021;**11**:8794.

Pirhaji L, Milani P, Leidl M *et al.* Revealing disease-associated pathways by network integration of untargeted metabolomics. *Nat Methods* 2016;**13**:770–6.

Pržulj N. Biological network comparison using graphlet degree distribution. *Bioinformatics* 2007;**23**:e177–83.

Pržulj N, Corneil DG, Jurisica I. Modeling interactome: scale-free or geometric? *Bioinformatics* 2004;**20**:3508–15.

Pržulj N, Malod-Dognin N. Network analytics in the age of big data. *Science* 2016;**353**:123–4.

Przytycki PF, Singh M. Differential allele-specific expression uncovers breast cancer genes dysregulated by cis noncoding mutations. *Cell Syst* 2020;**10**:193–203.e4.

Purkait P, Chin TJ, Sadri A *et al.* Clustering with hypergraphs: the case for large hyperedges. *IEEE Trans Pattern Anal Mach Intell* 2017;**39**:1697–711.

Pushpakom SP, Iorio F, Eyers PA *et al.* Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov* 2019;**18**:41–58.

Radivojac P, Clark WT, Oron T *et al.* A large-scale evaluation of computational protein function prediction. *Nat Methods* 2013;**10**:221–7.

Radovanović M, Nanopoulos A, Ivanović M. Hubs in space: popular nearest neighbors in high-dimensional data. *J Mach Learn Res* 2010;**11**:2487–531.

Ramadan E, Tarafdar A, Pothen A. A hypergraph model for the yeast protein complex network. In: *Proceedings of the International Parallel and Distributed Processing Symposium*. p. 189. Los Alamitos, CA: IEEE Computer Society, 2004.

Ramola R, Friedberg I, Radivojac P. The field of protein function prediction as viewed by different domain scientists. *Bioinform Adv* 2022;**2**:vbac057.

Rao R, Bhattacharya N, Thomas N *et al.* Evaluating protein transfer learning with tape. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 32. Red Hook, NY: Curran Associates Inc., 2019.

Reshef DN, Reshef YA, Finucane HK *et al.* Detecting novel associations in large data sets. *Science* 2011;**334**:1518–24.

Reyna MA, Chitra U, Elyanow R *et al.* NetMix: a network-structured mixture model for reduced-bias estimation of altered subnetworks. *J Comput Biol* 2021;**28**:469–84.

Rhodes G. *Crystallography Made Crystal Clear, Third Edition: A Guide for Users of Macromolecular Models*. Elsevier, 2010.

Rider A, Milenković T, Siwo G *et al.* Networks are important for systems biology. *Netw Sci* 2014;**2**:139–61.

Ritchie ME, Phipson B, Wu D *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;**43**:e47.

Ritz A, Poirel CL, Tegge AN *et al.* Pathways on demand: automated reconstruction of human signaling networks. *NPJ Syst Biol Appl* 2016;**2**:16002–9.

Ritz A, Tegge AN, Kim H *et al.* Signaling hypergraphs. *Trends Biotechnol* 2014;**32**:356–62.

Rodrigues CH, Ascher DB. CSM-Potential: mapping protein interactions and biological ligands in 3D space using geometric deep learning. *Nucleic Acids Res* 2022;**50**:W204–9.

Rogers JD, Aguado BA, Watts KM *et al.* Network modeling predicts personalized gene expression and drug responses in valve myofibroblasts cultured with patient sera. *Proc Natl Acad Sci USA* 2022;**119**:e2117323119.

Rolland T, Taşan M, Charloteaux B *et al.* A proteome-scale map of the human interactome network. *Cell* 2014;**159**:1212–26.

Roy S, Lagree S, Hou Z *et al.* Integrated module and gene-specific regulatory inference implicates upstream signaling networks. *PLoS Comput Biol* 2013;**9**:e1003252.

Ruan D, Young A, Montana G. Differential analysis of biological networks. *BMC Bioinformatics* 2015;**16**:327.

Saelens W, Cannoodt R, Saeys Y. A comprehensive evaluation of module detection methods for gene expression data. *Nat Commun* 2018;**9**:1090.

Saez-Rodriguez J, Costello JC, Friend SH *et al.* Crowdsourcing biomedical research: leveraging communities as innovation engines. *Nat Rev Genet* 2016;**17**:470–86.

Sahni N, Yi SS, Taipale M *et al.* Widespread macromolecular interaction perturbations in human genetic disorders. *Cell* 2015;**161**:647–60.

Saibil HR. Cryo-EM in molecular and cellular biology. *Mol Cell* 2022;**82**:274–84.

Said A, Bayrak R, Derr T *et al.* NeuroGraph: benchmarks for graph machine learning in brain connectomics. *Adv Neural Inf Process Syst* 2023;**36**:6509–31.

Salazar D, Valencia C, Pržulj N. Multi-project and multi-profile joint non-negative matrix factorization for cancer omic datasets. *Bioinformatics* 2021;**37**:4801–9.

Samieri C, Sonawane AR, Lefèvre-Arbogast S *et al.* Using network science tools to identify novel diet patterns in prodromal dementia. *Neurology* 2020;**94**:e2014–25.

Sanghvi JC, Regot S, Carrasco S *et al.* Accelerated discovery via a whole-cell model. *Nat Methods* 2013;**10**:1192–5.

Sarajlić A, Malod-Dognin N, Yaveroğlu Ö *et al.* Graphlet-based characterization of directed networks. *Sci Rep* 2016;**6**:35098.

Saraph V, Milenković T. MAGNA: maximizing accuracy in global network alignment. *Bioinformatics* 2014;**30**:2931–40.

Sarraju A, Ngo S, Rodriguez F. The leaky pipeline of diverse race and ethnicity representation in academic science and technology training in the United States, 2003–2019. *PLoS One* 2023;**18**:e0284945.

Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature* 2009;**461**:218–23.

Schaefer M, Serrano L, Andrade-Navarro M. Correcting for the study bias associated with protein–protein interaction measurements reveals differences between protein degree distributions from different cancer types. *Front Genet* 2015;**6**:260.

Schiebinger G, Shu J, Tabaka M *et al.* Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell* 2019;**176**:928–43.e22.

Schlichtkrull M, Kipf TN, Bloem P *et al.* Modeling relational data with graph convolutional networks. In: *Proceedings of the International Semantic Web Conference*. p. 593–607. Berlin: Springer-Verlag, 2018.

Schmidt F, Gasparoni N, Gasparoni G *et al.* Combining transcription factor binding affinities with open-chromatin data for accurate gene expression prediction. *Nucleic Acids Res* 2017;**45**:54–66.

Schmidt F, Kern F, Ebert P *et al.* TEPIC 2—an extended framework for transcription factor binding prediction and integrative epigenomic analysis. *Bioinformatics* 2019;**35**:1608–9.

Schmidt M, Niculescu-Mizil A, Murphy K. Learning graphical model structure using L1-regularization paths. In: *Proceedings of the National Conference on Artificial Intelligence*. p. 1278–83. AAAI Press, 2007.

Schwab JD, Ikonomi N, Werle SD *et al.* Reconstructing Boolean network ensembles from single-cell data for unraveling dynamics in the aging of human hematopoietic stem cells. *Comput Struct Biotechnol J* 2021;**19**:5321–32.

Sedgewick AJ, Buschur K, Shi IW *et al.* Mixed graphical models for integrative causal analysis with application to chronic lung disease diagnosis and prognosis. *Bioinformatics* 2018;**35**:1204–12.

Segal E, Pe'er D, Regev A *et al.* Learning module networks. *J Mach Learn Res* 2005;**6**:557–88.

Sha Y, Wang S, Zhou P *et al.* Inference and multiscale model of epithelial-to-mesenchymal transition via single-cell transcriptomic data. *Nucleic Acids Res* 2020;**48**:9505–20.

Sha Z, Schijven D, Fisher SE *et al.* Genetic architecture of the white matter connectome of the human brain. *Sci Adv* 2023;**9**:eadd2870.

Sharan R, Ideker T. Modeling cellular machinery through biological network comparison. *Nat Biotechnol* 2006;**24**:427–33.

Shawe-Taylor J, Cristianini N. *Kernel Methods for Pattern Analysis*. Cambridge: Cambridge University Press, 2004.

Shchur O, Mumme M, Bojchevski A *et al.* Pitfalls of graph neural network evaluation. In: *Proceedings of the Relational Representation Learning Workshop*. 2018.

Shervashidze N, Vishwanathan S, Petri T *et al.* Efficient graphlet kernels for large graph comparison. In: *Proceedings of the Artificial Intelligence and Statistics*. p. 488–95. 2009.

Shit S, Paetzold JC, Sekuboyina A *et al.* clDice-a novel topology-preserving loss function for tubular structure segmentation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. p. 16560–9. 2021.

Siahpirani AF, Roy S. A prior-based integrative framework for functional transcriptional regulatory network inference. *Nucleic Acids Res* 2017;**45**:e21.

Silverbush D, Cristea S, Yanovich-Arad G *et al.* Simultaneous integration of multi-omics data improves the identification of cancer driver modules. *Cell Syst* 2019;**8**:456–66.e5.

Simonovsky M, Komodakis N. GraphVAE: towards generation of small graphs using variational autoencoders. In: *Proceedings of the Artificial Neural Networks and Machine Learning*. p. 412–22. Switzerland: Springer Nature, 2018.

Singh R, Devkota K, Sledzieski S *et al.* Topsy-Turvy: integrating a global view into sequence-based PPI prediction. *Bioinformatics* 2022;**38**:i264–72.

Singh R, Xu J, Berger B. Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proc Natl Acad Sci USA* 2008;**105**:12763–8.

Sledzieski S, Devkota K, Singh R *et al.* TT3D: leveraging precomputed protein 3D sequence models to predict protein–protein interactions. *Bioinformatics* 2023;**39**:btad663.

Sledzieski S, Singh R, Cowen L *et al.* D-script translates genome to phenome with sequence-based, structure-aware, genome-scale predictions of protein-protein interactions. *Cell Syst* 2021;**12**:969–82.6.

Smith KP, Christakis NA. Social networks and health. *Annu Rev Sociol* 2008;**34**:405–29.

Solava R, Michaels R, Milenković T. Graphlet-based edge clustering reveals pathogen-interacting proteins. *Bioinformatics* 2012; **28**:i480–6.

Sonawane AR, DeMeo DL, Quackenbush J *et al.* Constructing gene regulatory networks using epigenetic data. *NPJ Syst Biol Appl* 2021; **7**:45.

Sonawane AR, Platig J, Fagny M *et al.* Understanding tissue-specific gene regulation. *Cell Rep* 2017;**21**:1077–88.

Sonawane AR, Weiss ST, Glass K *et al.* Network medicine in the age of biomedical big data. *Front Genet* 2019;**10**:294.

Sprinzak E, Sattath S, Margalit H. How reliable are experimental protein–protein interaction data? *J Mol Biol* 2003;**327**:919–23.

Stark C, Breitkreutz BJ, Reguly T *et al.* BioGRID: a general repository for interaction datasets. *Nucleic Acids Res* 2006;**34**:D535–9.

Stärk H, Ganea O, Pattanaik L *et al.* EquiBind: geometric deep learning for drug binding structure prediction. In: *Proceedings of the International Conference on Machine Learning*. p. 20503–21. PMLR, 2022.

Stegehuis C, Van Der Hofstad R, Van Leeuwaarden JS. Epidemic spreading on complex networks with community structures. *Sci Rep* 2016;**6**:29748–7.

Stevens KR, Masters K, Imoukhuede PI *et al.* Fund black scientists. *Cell* 2021;**184**:561–5.

Stokes JM, Yang K, Swanson K *et al.* A deep learning approach to antibiotic discovery. *Cell* 2020;**180**:688–702.e13.

Stolovitzky G, Monroe D, Califano A. Dialogue on reverse-engineering assessment and methods: the DREAM of high-throughput pathway inference. *Ann N Y Acad Sci* 2007;**1115**:1–22.

Sun Y, Crawford J, Tang J *et al.* Simultaneous optimization of both node and edge conservation in network alignment via WAVE. In: *Proceedings of the Workshop on Algorithms in Bioinformatics*. p. 16–39. Berlin: Springer-Verlag, 2015.

Sun Y, Han J, Yan X *et al.* PathSim: meta path-based top-k similarity search in heterogeneous information networks. *Proc VLDB Endow* 2011;**4**:992–1003.

Sverchkov Y, Craven M. A review of active learning approaches to experimental design for uncovering biological networks. *PLoS Comput Biol* 2017;**13**:e1005466.

Sychev ZE, Hu A, DiMaio TA *et al.* Integrated systems biology analysis of KSHV latent infection reveals viral induction and reliance on peroxisome mediated lipid metabolism. *PLoS Pathog* 2017; **13**:e1006256.

Szklarczyk D, Kirsch R, Koutrouli M *et al.* The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* 2023;**51**:D638–46.

Tahmasebi B, Lim D, Jegelka S. Counting substructures with higher-order graph neural networks: possibility and impossibility results. arXiv, 2012.03174, 2020, preprint: not peer reviewed.

Tahmasebi B, Lim D, Jegelka S. The power of recursion in graph neural networks for counting substructures. In: *Proceedings of the International Conference on Artificial Intelligence and Statistics*. p. 11023–42. PMLR, 2023.

Tang F, Xu D, Wang S *et al.* Chromatin profiles classify castration-resistant prostate cancers suggesting therapeutic targets. *Science* 2022;**376**:eabe1505.

Tang J, Qu M, Wang M *et al.* LINE: large-scale information network embedding. In: *Proceedings of the International Conference on World Wide Web*. p. 1067–77. International World Wide Web Conferences Steering Committee, 2015.

Teschendorff AE, Feinberg AP. Statistical mechanics meets single-cell biology. *Nat Rev Genet* 2021;**22**:459–76.

Theodoris CV, Xiao L, Chopra A *et al.* Transfer learning enables predictions in network biology. *Nature* 2023;**618**:616–24.

Torbey R, Martin ND, Warner JR *et al.* Algebra I before high school as a gatekeeper to computer science participation. In: *Proceedings of the ACM Technical Symposium on Computer Science Education*. p. 839–44. New York, NY: Association for Computing Machinery, 2020.

Townshend RJ, Eismann S, Watkins AM *et al.* Geometric deep learning of RNA structure. *Science* 2021;**373**:1047–51.

Tseng PT, Cheng YS, Yen CF *et al.* Peripheral iron levels in children with attention-deficit hyperactivity disorder: a systematic review and meta-analysis. *Sci Rep* 2018;**8**:788–11.

Tsuruta H, Yamazaki H, Maeda R *et al.* AVIDa-hIL6: a large-scale VHH dataset produced from an immunized alpaca for predicting antigen-antibody interactions. *Adv Neural Inf Process Syst* 2024; **36**:42077–96.

Tu JJ, Ou-Yang L, Zhu Y *et al.* Differential network analysis by simultaneously considering changes in gene interactions and gene expression. *Bioinformatics* 2021;**37**:4414–23.

Tu K, Cui P, Wang X *et al.* Structural deep embedding for hyper-networks. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. AAAI Press, 2018.

Tuncbag N, Braunstein A, Pagnani A *et al.* Simultaneous reconstruction of multiple signaling pathways via the prize-collecting Steiner Forest problem. *J Comput Biol* 2013;**20**:124–36.

Tuncbag N, Gosline SJ, Kedaigle A *et al.* Network-based interpretation of diverse high-throughput datasets through the Omics Integrator software package. *PLoS Comput Biol* 2016;**12**:e1004879.

Ünsal Ü, Cüvitoğlu A, Turhan K *et al.* NMSDR: drug repurposing approach based on transcriptome data and network module similarity. *Mol Inform* 2023;**42**:e2200077.

Vacic V, Iakoucheva LM, Lonardi S *et al.* Graphlet kernels for prediction of functional residues in protein structures. *J Comput Biol* 2010;**17**:55–72.

Van Der Wijst MG, de Vries DH, Brugge H *et al.* An integrative approach for building personalized gene regulatory networks for precision medicine. *Genome Med* 2018;**10**:96–15.

van Haagen HHHBM, 't Hoen PAC, Botelho Bovo A *et al.* Novel protein-protein interactions inferred from literature context. *PLoS One* 2009;**4**:e7894.

Vanunu O, Magger O, Ruppin E *et al.* Associating genes and protein complexes with disease via network propagation. *PLoS Comput Biol* 2010;**6**:e1000641.

Vasaikar SV, Straub P, Wang J *et al.* LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res* 2018;**46**:D956–63.

Veličković P, Blundell C. Neural algorithmic reasoning. *Patterns* 2021; **2**:100273.

Veličković P, Cucurull G, Casanova A *et al*. Graph attention networks. In: *Proceedings of the International Conference on Learning Representations*. 2018.

Veličković P, Fedus W, Hamilton WL *et al*. Deep graph Infomax. In: *Proceedings of the International Conference on Learning Representations*. 2019.

Verstraete N, Jurman G, Bertagnolli G *et al*. CovMulNet19, integrating proteins, diseases, drugs, and symptoms: a network medicine approach to COVID-19. *Netw Syst Med* 2020;**3**:130–41.

Vijayan V, Critchlow D, Milenković T. Alignment of dynamic networks. *Bioinformatics* 2017;**33**:i180–9.

Vijayan V, Gu S, Krebs ET *et al*. Pairwise versus multiple global network alignment. *IEEE Access* 2020;**8**:41961–74.

Vijayan V, Milenković T. Multiple network alignment via multiMAGNA+. *IEEE/ACM Trans Comput Biol Bioinform* 2018a;**15**:1669–82.

Vijayan V, Milenković T. Aligning dynamic networks with DynaWAVE. *Bioinformatics* 2018b;**34**:1795–8.

Vijayan V, Saraph V, Milenković T. MAGNA++: maximizing accuracy in global network alignment via both node and edge conservation. *Bioinformatics* 2015;**31**:2409–11.

Vishwanathan SVN, Schraudolph NN, Kondor RI *et al*. Graph kernels. *J Mach Learn Res* 2010;**11**:1201–42.

Voitalov I, Zhang L, Kilpatrick C *et al*. The module triad: a novel network biology approach to utilize patients' multi-omics data for target discovery in ulcerative colitis. *Sci Rep* 2022;**12**:21685.

Von Mering C, Krause R, Snel B *et al*. Comparative assessment of large-scale data sets of protein–protein interactions. *Nature* 2002; **417**:399–403.

Waagmeester A, Stupp GS, Burgstaller-Muehlbacher S *et al*. Wikidata as a knowledge graph for the life sciences. *Elife* 2020;**9**:e52614.

Wachman G, Khardon R. Learning from interpretations: a rooted kernel for ordered hypergraphs. In: *Proceedings of the International Conference on Machine Learning*. p. 943–50. New York, NY: Association for Computing Machinery, 2007.

Wang B, Mezlini AM, Demir F *et al*. Similarity network fusion for aggregating data types on a genomic scale. *Nat Methods* 2014; **11**:333–7.

Wang B, Pourshafeie A, Zitnik M *et al*. Network enhancement as a general method to denoise weighted biological networks. *Nat Commun* 2018;**9**:3108.

Wang H, Fu T, Du Y *et al*. Scientific discovery in the age of artificial intelligence. *Nature* 2023a;**620**:47–60.

Wang H, Lian D, Liu W *et al*. Powerful graph of graphs neural network for structured entity analysis. *World Wide Web* 2022a;**25**:609–29.

Wang H, Lian D, Zhang Y *et al*. GoGNN: graph of graphs neural network for predicting structured entity interactions. In: *Proceedings of the International Joint Conference on Artificial Intelligence*. 2021a.

Wang H, Zheng H, Chen DZ. TANGO: a GO-term embedding based method for protein semantic similarity prediction. *IEEE/ACM Trans Comput Biol Bioinformatics* 2023b;**20**:694–706.

Wang J, Lisanza S, Juergens D *et al*. Scaffolding protein functional sites using deep learning. *Science* 2022b;**377**:387–94.

Wang L, Liu H, Liu Y *et al*. Learning hierarchical protein representations via complete 3D graph networks. In: *Proceedings of the International Conference on Learning Representations*. 2022c.

Wang Q, Jiang H, Jiang Y *et al*. Multiplex network infomax: multiplex network embedding via information fusion. *Digit Commun Netw* 2022d;**9**:1157–68.

Wang T, Shao W, Huang Z *et al*. MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. *Nat Commun* 2021b;**12**:3445.

Wang X, Bo D, Shi C *et al*. A survey on heterogeneous graph embedding: methods, techniques, applications and sources. *IEEE Trans Big Data* 2023c;**9**:415–36.

Wang X, Ji H, Shi C *et al*. Heterogeneous graph attention network. In: *Proceedings of the World Wide Web Conference*. p. 2022–32. New York, NY: Association for Computing Machinery, 2019.

Wang X, Madeddu L, Spirohn K *et al*. Assessment of community efforts to advance network-based prediction of protein–protein interactions. *Nat Commun* 2023d;**14**:1582.

Wang Y, Lee H, Fear JM *et al*. NetREX-CF integrates incomplete transcription factor data with gene expression to reconstruct gene regulatory networks. *Commun Biol* 2022e;**5**:1282.

Wang Y, Peng Q, Wang W *et al*. Network alignment enhanced via modeling heterogeneity of anchor nodes. *Knowl Based Syst* 2022f; **250**:109116.

Wang Y, Zhao Y, Shah N *et al*. Imbalanced graph classification via graph-of-graph neural networks. In: *Proceedings of the ACM International Conference on Information and Knowledge Management*. p. 2067–76. New York, NY: Association for Computing Machinery, 2022g.

Watson JL, Juergens D, Bennett NR *et al*. De novo design of protein structure and function with rfdiffusion. *Nature* 2023; **620**:1089–100.

Weber AN, Bittner ZA, Shankar S *et al*. Recent insights into the regulatory networks of NLRP3 inflammasome activation. *J Cell Sci* 2020; **133**:jcs248344.

Weighill D, Ben Guebila M, Glass K *et al*. Gene targeting in disease networks. *Front Genet* 2021;**12**:649942.

Weighill D, Guebila MB, Glass K *et al*. Predicting genotype-specific gene regulatory networks. *Genome Res* 2022;**32**:524–33.

Wen J, Zhang X, Rush EN *et al*. Multimodal representation learning for predicting molecule–disease relations. *Bioinformatics* 2023; **39**:btad085.

Wetzel JL, Zhang K, Singh M. Learning probabilistic protein–DNA recognition codes from DNA-binding specificities using structural mappings. *Genome Res* 2022;**32**:1776–86.

Windels S, Malod-Dognin N, Pržulj N. Identifying cellular cancer mechanisms through pathway-driven data integration. *Bioinformatics* 2022a;**38**:4344–51.

Windels S, Malod-Dognin N, Pržulj N. Graphlet eigencentralities capture novel central roles of genes in pathways. *PLoS One* 2022b; **17**:e0261676.

Winkler S, Winkler I, Figaschewski M *et al*. De novo identification of maximally deregulated subnetworks based on multi-omics data with DeRegNet. *BMC Bioinformatics* 2022;**23**:139–28.

Wright SN, Colton S, Schaffer LV *et al*. State of the interactomes: an evaluation of molecular networks for generating biological insights. bioRxiv, 2024.04.26.587073, 2024, preprint: not peer reviewed.

Wu L, Lin H, Tan C *et al*. Self-supervised learning on graphs: contrastive, generative, or predictive. *IEEE Trans Knowl Data Eng* 2023; **35**:4216–35.

Wu RM, Ding F, Wang R *et al*. High-resolution de novo structure prediction from primary sequence. bioRxiv, 2022.07.21.500999, 2022, preprint: not peer reviewed.

Wu X, Liu Q, Jiang R. Align human interactome with phenome to identify causative genes and networks underlying disease families. *Bioinformatics* 2009;**25**:98–104.

Xenos A, Malod-Dognin N, Milinković S *et al*. Linear functional organization of the omic embedding space. *Bioinformatics* 2021; **37**:3839–47.

Xenos A, Malod-Dognin N, Zambrana C *et al*. Integrated data analysis uncovers new COVID-19 related genes and potential drug repurposing candidates. *Int J Mol Sci* 2023;**24**:1431.

Xie Y, Katariya S, Tang X *et al*. Task-agnostic graph explanations. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 35, p. 12027–39. Red Hook, NY: Curran Associates Inc., 2022a.

Xie Y, Xu Z, Zhang J *et al*. Self-supervised learning of graph neural networks: a unified review. *IEEE Trans Pattern Anal Mach Intell* 2022b;**45**:2412–29.

Xie Y, Zhang Y, Gong M *et al*. MGAT: multi-view graph attention networks. *Neural Netw* 2020;**132**:180–9.

Xiong H, Yan J, Pan L. Contrastive multi-view multiplex network embedding with applications to robust network alignment. In: *Proceedings of the ACM SIGKDD Conference on Knowledge*

Discovery and Data Mining. p. 1913–23. New York, NY: Association for Computing Machinery, 2021.

Xu J, Wickramarathne TL, Chawla NV. Representing higher-order dependencies in networks. *Sci Adv* 2016;**2**:e1600028.

Xu M, Zhang Z, Lu J *et al.* PEER: a comprehensive and multi-task benchmark for protein sequence understanding. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 35, p. 35156–73. 2022.

Yan Y, Zhou Q, Li J *et al.* Dissecting cross-layer dependency inference on multi-layered inter-dependent networks. In: *Proceedings of the International Conference on Information and Knowledge Management*. p. 2341–51. New York, NY: Association for Computing Machinery, 2022.

Yang XH, Goldstein A, Sun Y *et al.* Detecting critical transition signals from single-cell transcriptomes to infer lineage-determining transcription factors. *Nucleic Acids Res* 2022;**50**:e91.

Yasunaga M, Bosselut A, Ren H *et al.* Deep bidirectional language-knowledge graph pretraining. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 35, p. 37309–23. 2022a.

Yasunaga M, Leskovec J, Liang P. *LinkBERT: Pretraining Language Models with Document Links*. Association for Computational Linguistics, 2022b.

Yaveroğlu Ö, Milenković T, Pržulj N. Proper evaluation of alignment-free network comparison methods. *Bioinformatics* 2015; **31**:2697–704.

Yeger-Lotem E, Riva L, Su LJ *et al.* Bridging high-throughput genetic and transcriptional data reveals cellular responses to alpha-synuclein toxicity. *Nat Genet* 2009;**41**:316–23.

Yi K, Zhou B, Shen Y *et al.* Graph denoising diffusion for inverse protein folding. *Adv Neural Inf Process Syst* 2024;**36**.

Yim J, Stärk H, Corso G *et al.* Diffusion models in protein structure and docking. *Wiley Interdiscip Rev Comput Mol Sci* 2024;**14**:e1711.

Yin W, Mendoza L, Monzon-Sandoval J *et al.* Emergence of co-expression in gene regulatory networks. *PLoS One* 2021; **16**:e0247671.

Ying C, Cai T, Luo S *et al.* Do transformers really perform badly for graph representation? In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 34, p. 28877–88. Red Hook, NY: Curran Associates Inc., 2021.

Ying Z, Bourgeois D, You J *et al.* GNNExplainer: generating explanations for graph neural networks. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 32. Red Hook, NY: Curran Associates Inc., 2019.

Yoon BJ, Qian X, Dougherty ER. Quantifying the objective cost of uncertainty in complex dynamical systems. *IEEE Trans Signal Process* 2013;**61**:2256–66.

You J, Du T, Leskovec J. ROLAND: graph learning framework for dynamic graphs. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. p. 2358–66. New York, NY: Association for Computing Machinery, 2022.

You J, Ying R, Ren X *et al.* GraphRNN: generating realistic graphs with deep auto-regressive models. In: *Proceedings of the International Conference on Machine Learning*. p. 5708–17. PMLR, 2018.

You Y, Chen T, Sui Y *et al.* Graph contrastive learning with augmentations. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 33, p. 5812–23. Red Hook, NY: Curran Associates Inc.,2020.

Yu X, Liu Z, Fang Y *et al.* Learning to count isomorphisms with graph neural networks. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI Press, 2023.

Yuan H, Yu H, Wang J *et al.* On explainability of graph neural networks via subgraph explorations. In: *Proceedings of the International Conference on Machine Learning*. p. 12241–52. PMLR, 2021.

Yue X, Wang Z, Huang J *et al.* Graph embedding on biomedical networks: methods, applications and evaluations. *Bioinformatics* 2020;**36**:1241–51.

Yuen HY, Jansson J. Better link prediction for protein-protein interaction networks. In: *Proceedings of the IEEE International Conference on Bioinformatics and Bioengineering*. p. 53–60. 2020.

Yun S, Jeong M, Kim R *et al.* Graph transformer networks. *Adv Neural Inf Process Syst* 2019;**32**.

Zambrana C, Xenos A, Bottcher R *et al.* Network neighbors of viral targets and differentially expressed genes in COVID-19 are drug target candidates. *Sci Rep* 2021;**11**:18985.

Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 2005; **4**:Article17.

Zhang C, Song D, Huang C *et al.* Heterogeneous graph neural network. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. p. 793–803. New York, NY: Association for Computing Machinery, 2019a.

Zhang J, Cammarata L, Squires C *et al.* Active learning for optimal intervention design in causal models. *Nat Mach Intell* 2023a; **5**:1066–75.

Zhang L, Yu G, Guo M *et al.* Predicting protein-protein interactions using high-quality non-interacting pairs. *BMC Bioinformatics* 2018; **19**:525–124.

Zhang R, Ma J. Matcha: probing multi-way chromatin interaction with hypergraph representation learning. *Cell Syst* 2020;**10**:397–407.e5.

Zhang R, Ma J, Ma J. DANGO: predicting higher-order genetic interactions. bioRxiv, 2020.11.26.400739, 2020a, preprint: not peer reviewed.

Zhang R, Zou Y, Ma J. Hyper-SAGNN: a self-attention based graph neural network for hypergraphs. In: *Proceedings of the International Conference on Learning Representations*. 2020b.

Zhang R, Zhou T, Ma J. Multiscale and integrative single-cell Hi-C analysis with higashi. *Nat Biotechnol* 2022a;**40**:254–61.

Zhang S, Tong H, Xia Y *et al.* NetTrans: neural cross-network transformation. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. p. 986–96. New York, NY: Association for Computing Machinery, 2020c.

Zhang S, Tong H, Xu J *et al.* ORIGIN: non-rigid network alignment. In: *Proceedings of the IEEE International Conference on Big Data*. p. 998–1007. Institute of Electrical and Electronics Engineers Inc., 2019b.

Zhang X, Zeman M, Tsiligkaridis T *et al.* Graph-guided network for irregularly sampled multivariate time series. In: *Proceedings of the International Conference on Learning Representations*. 2022b.

Zhang Y, Huang H. Brain connectome based complex brain disorder prediction via novel graph-blind convolutional network. In: *Proceedings of the International Conference on Information Processing in Medical Imaging*. 2019.

Zhang Y, Zhan L, Wu S *et al.* Disentangled and proportional representation learning for multi-view brain connectomes. In: *Proceedings of the Medical Image Computing and Computer Assisted Intervention*. p. 508–18. Berlin, Heidelberg: Springer-Verlag, 2021.

Zhang Z, Liu Q, Hu Q *et al.* Hierarchical graph transformer with adaptive node sampling. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 35, p. 21171–83. Red Hook, NY: Curran Associates Inc., 2022c.

Zhang Z, Lu Z, Zhongkai H *et al.* Full-atom protein pocket design via iterative refinement. *Adv Neural Inf Process Syst* 2023b; **36**:16816–36.

Zhang Z, Xu M, Jamasb AR *et al.* Protein representation learning by geometric structure pretraining. *arXiv preprint arXiv:2203.06125*, 2023c, preprint: not peer reviewed.

Zhao C, Zhan L, Thompson PM *et al.* Revealing continuous brain dynamical organization with multimodal graph transformer. In: *Proceedings of the Medical Image Computing and Computer Assisted Intervention*. p. 346–55. Berlin: Springer-Verlag, 2022.

Zhao J, Wen Q, Sun S *et al.* Multi-view self-supervised heterogeneous graph embedding. In: *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. p. 319–34. Springer International Publishing, 2021.

Zhao L, Song Y, Zhang C *et al.* T-GCN: a temporal graph convolutional network for traffic prediction. *IEEE Trans Intell Transport Syst* 2020a;**21**:3848–58.

Zhao X, Chen F, Hu S *et al.* Uncertainty aware semi-supervised learning on graph data. *Adv Neural Inf Process Syst* 2020b;**33**:12827–36.

Zheng VW, Sha M, Li Y *et al.* Heterogeneous embedding propagation for large-scale e-commerce user alignment. In: *Proceedings of the IEEE International Conference on Data Mining*. p. 1434–9. 2018.

Zhou D, Huang J, Schölkopf B. Learning with hypergraphs: clustering, classification, and embedding. In: *Proceedings of the Advances in Neural Information Processing Systems*. Vol. 19. Cambridge, MA: MIT Press, 2006.

Zhou G, Ewald J, Xia J. OmicsAnalyst: a comprehensive web-based platform for visual analytics of multi-omics data. *Nucleic Acids Res* 2021a;**49**:W476–82.

Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning–based sequence model. *Nat Methods* 2015;**12**:931–4.

Zhou N, Jiang Y, Bergquist T *et al.* The CAFA challenge reports improved protein function prediction and new functional annotations for hundreds of genes through experimental screens. *Genome Biol* 2019;**20**:244–23.

Zhou P, Wang S, Li T *et al.* Dissecting transition cells from single-cell transcriptome data through multiscale stochastic dynamics. *Nat Commun* 2021b;**12**:5609.

Zhu L, Ding Y, Chen CY *et al.* MetaDCN: meta-analysis framework for differential co-expression network detection with an application in breast cancer. *Bioinformatics* 2016;**33**:1121–9.

Zitnik M, Agrawal M, Leskovec J. Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics* 2018;**34**:i457–66.

Zitnik M, Leskovec J. Predicting multicellular function through multilayer tissue networks. *Bioinformatics* 2017;**33**:i190–8.

Zitnik M, Nguyen F, Wang B *et al.* Machine learning for integrating data in biology and medicine: principles, practice, and opportunities. *Inf Fusion* 2019a;**50**:71–91.

Zitnik M, Sosič R, Feldman MW *et al.* Evolution of resilience in protein interactomes across the tree of life. *Proc Natl Acad Sci USA* 2019b;**116**:4426–33.

Zotenko E, Guimarães KS, Jothi R *et al.* Decomposition of overlapping protein complexes: a graph theoretical method for analyzing static and dynamic protein associations. *Algorithms Mol Biol* 2006;**1**:1–11.