Systems Biology and Network Analysis: From Multi-omics Integration to Causal and Mechanistic Understanding

Gurkan Bebek
Center for Proteomics and Bioinformatics, Case Western Reserve University, Cleveland, OH 44106, USA
E-mail: gurkan@case.edu

Rachael Hageman Blair
Department of Biostatistics, University at Buffalo, Buffalo, NY, USA
E-mail: hageman@buffalo.edu

Filipa Blasco Tavares Pereira Lopes
Mass Spectrometry and Proteomics Core, University of Utah, Salt Lake City, UT, USA
E-mail: u6071085@utah.edu

Joshua Welch
Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA
E-mail: welch@umich.edu

Onur Mutlu
Information Technology and Electrical Engineering Department, ETH Zürich, Zurich, 8092, Switzerland
E-mail: aonur.mutlu@inf.ethz.ch

Serguei Pakhomov
Department of Pharmaceutical Care & Health Systems, University of Minnesota, Minneapolis, MN 55455, USA
E-mail: pakh0002@umn.edu

1. Session Introduction: Systems Biology and Network Analysis

Advances in high-throughput molecular profiling continue to generate increasingly complex datasets spanning multiple biological scales. These developments have created new opportunities to investigate biological systems through integrative computational approaches that combine molecular measurements with network, pathway, and knowledge-based representations. At the same time, extracting mechanistic insight from these data remains a fundamen-

© 2026 The Authors. Open Access chapter published by World Scientific Publishing Company and distributed under the terms of the Creative Commons Attribution Non-Commercial (CC BY-NC) 4.0 License.

tal challenge. The papers presented in this session address this challenge through advances in network analysis, representation learning, causal knowledge modeling, perturbation analysis, and multi-modal integration. Collectively, they illustrate how systems-level approaches can facilitate biological discovery, generate mechanistic hypotheses, and support translational applications.

2. Session Summary

This session includes an invited talk and twelve peer-reviewed papers selected for oral and poster presentations. The contributions span multiple areas of systems biology, including single-cell modeling, multi-omics integration, network propagation, causal knowledge representation, and translational network biology. Although the methodological approaches vary, a common theme is developing computational frameworks that connect large-scale molecular data to biological mechanisms.

2.1. *Representation Learning and Foundation Models*

Several contributions focus on the opportunities and challenges associated with large-scale representation learning. Whalley examines gene embedding structures across several widely used single-cell foundation models and demonstrates that embedding outliers are not consistently associated with biologically important genes, highlighting limitations in the biological interpretation of learned representations. Kim and Hwang address computational scalability by introducing a cell-adaptive token allocation strategy that substantially reduces the computational cost of single-cell foundation model inference while preserving downstream performance. Together, these studies illustrate the growing role of foundation models in computational biology while emphasizing the need for both interpretability and efficiency.

2.2. *Multi-Omics Integration and Dynamic Modeling*

A second group of papers focuses on integrating heterogeneous molecular data. Mo and Olshen present iClusterPlus2, an extension of the iCluster framework that enables scalable integration of bulk and single-cell multi-omics datasets through shared latent representations. Zhang et al. introduce scMultiNODE, which combines neural ordinary differential equations with optimal transport methods to model cellular trajectories from unaligned single-cell modalities. Zheng et al. investigate relationships between polygenic and proteomic risk scores and show that network-based approaches can reveal complementary biological signals that are not apparent when each data source is considered independently. These studies highlight continuing efforts to bridge multiple molecular layers within unified analytical frameworks.

2.3. *Perturbation Analysis and Network Propagation*

Understanding how perturbations propagate through biological systems remains a central challenge in systems biology. Liang and Singh introduce COMPASS, a framework that decomposes perturbation responses into shared and perturbation-specific components in large-scale Perturb-seq experiments. By separating common cellular responses from gene-specific effects,

the method improves interpretation of regulatory perturbation data. Bobak et al. examine signal propagation in co-expression networks by adapting concepts from network interference and spillover analysis. Their work shows that interactions between network modules can reveal biologically relevant signals that direct measurements alone may not detect.

2.4. *Knowledge Graphs and Mechanistic Representation*

Several papers emphasize the importance of structured biological knowledge for mechanistic interpretation. Hoehndorf et al. introduce PhysioMap, a causal knowledge graph that represents physiological processes and supports analyses based on mechanistic relationships and interventions. Caufield et al. present KGBioportal, a platform that transforms biomedical ontologies into graph-based resources that can be incorporated into knowledge graph applications, network analyses, and machine learning workflows. Together, these contributions reflect increasing interest in combining data-driven approaches with formal biological knowledge representations.

2.5. *Translational Applications and Methodological Evaluation*

The remaining papers demonstrate how network-based approaches can support translational research while also highlighting the importance of methodological rigor. Karunakaran et al. reconstruct a GPX4-centered interaction network to investigate disease mechanisms and identify potential therapeutic opportunities in Sedaghatian-type spondylometaphyseal dysplasia. Nemecio and Brubaker integrate microbial, host, and pharmacological information to study microbial drug metabolism across human-associated microbiomes. Debnath and Ghosh examine benchmark datasets used for drug-synergy prediction and show that commonly used validation strategies can substantially overestimate performance because of information leakage. Collectively, these studies illustrate both the promise of network-based discovery and the importance of careful model evaluation.

3. Summary

Several themes emerge across the papers in this session. Biological networks continue to provide a common framework for integrating heterogeneous datasets, interpreting machine-learning models, and connecting molecular measurements to higher-level phenotypes. At the same time, there is increasing emphasis on approaches that support mechanistic interpretation rather than prediction alone, whether through perturbation modeling, causal knowledge representation, or multi-modal data integration. The contributions presented in this session illustrate how advances in computational methods are helping bridge the gap between large-scale molecular data and biological mechanisms.