Preface

Researchers are all familiar with the spectacular success of reductionist approaches for elucidating biological mechanisms over the last century or so. Nevertheless, recent technological advances, particularly among the leading sub-specialities of molecular biology, present challenges to reductionism on at least two fronts.

First, the explosive growth in the scale of data generated by modern, highly parallel, computerized, and increasingly automated lab methods has become a major impediment. For instance, next-generation sequencing technology is capable of producing terabytes of sequence information per day. It is also now mainstream to report global expression patterns for thousands of genes in multiple strains or cell lines subjected to different conditions. Similarly, phenotypic screens analyzing singly or multiply gene-disrupted cells and organisms have recently become prominent. But how will such data be analyzed, interpreted, and presented when extended to patients where thousands of gene isoforms have unique functions in individual tissues? Secondly, it has become apparent that biological processes do not trace back to individual molecules, and that the discrete pathway diagrams of text books do not accurately reflect the more extensive interaction networks likely present in cells.

Yet, while extremely large datasets describing gene sequences, mRNA transcripts, protein interactions, and metabolite concentrations are increasingly commonplace, these represent only starting "parts lists" that are usually insufficient to unlock mechanistic insights on their own right. Notable examples include a marked increase in organism complexity that does not correspond to aggregate genome size (e.g. yeast versus fly versus human), the formation of individual proteins into macromolecular assemblies (i.e. protein complexes) in order to perform intricate biochemical tasks, and the combinatorial action of sequence-specific DNA-binding transcription factors, microRNAs, and chromatin regulators on transcriptional responses.

Fortunately, despite the clumsiness of our observations, nature herself has the clear objectives of survival and propagation in mind, so concepts emerging from the study of biological entities, such as networks (e.g. functional interactions linking genes, proteins, metabolites, etc.) suggest that order rather than chaos prevails. These principles include modular and hierarchical organization, reactive information-driven causal-response behaviours, systems robustness, co-evolution, and self-organization. Moreover, the experience of the last decade argues that continued success in the field of network biology will rely increasingly on accurate high-throughput methods for generating quantitative "omic" datasets in combination with innovative integrative analytical tools designed to discover organizing patterns and principles underlying the myriad data.

This book attempts to bring both approaches together in a single volume. It includes practical descriptions of the experimental and computational approaches currently prevalent in network biology.

Among the 10 experimental methods, three of the major protein interaction mapping approaches are described over three chapters. Seesandra Rajagopala and Peter Uetz describe how different versions of the yeast two hybrid assay serve as complementary protein interaction mapping tools (Chapter 1) while Zuyao Ni and co-workers provide detailed instructions on mapping mammalian protein interactions using a lentiviral affinity tagging and expression system (Chapter 2). In a description of the protein arrays methodology, Mikael Bauer and co-workers highlight the flexibility and power of this technique (Chapter 3). Next, Oliver Schilling and co-workers (Chapter 4) describe a quantitative mass spectrometry-based method for mapping proteome networks through C-termini; Mihaela Sardiu and Michael Washburn describe a set of procedures for quantifying probabilistic protein interaction networks using label-free proteomics (Chapter 5); Haiming Huang and Sachdev Sidhu describe approaches for mapping interactions between peptide recognition modules and short linear motifs to build protein interaction networks (Chapter 6). Mohan Babu and colleagues present methodology for carrying out large-scale genetic interaction (epistasis) screens in *E. coli* (Chapter 7); Yifat Cohen and Maya Schuldiner show how integrated use of high-content microscopy and gene disruption libraries provide unprecedented visual descriptions of dynamic cell processes (Chapter 8); Kim Blakely and co-workers describe the application of related gene knockdown approaches to mammalian cells based on shRNA libraries (Chapter 9). Finally, Susana Neves and Laura Forrest provide an in-depth overview of DNA sequence analysis technologies for molecular phylogenetics (Chapter 10).

A variety of cutting-edge computational approaches currently used in network biology are then surveyed over the next 11 chapters. Gregory Clark and co-workers describe the use of co-evolutionary methods for predicting protein interactions that rival alternative approaches, such as abundance based methods (Chapter 11); Daniele Merico and colleagues describe a method for finding functionally coherent sets of genes using a Plugin for the Cytoscape network visualization application applicable for many different data types (Chapter 12); Alexei Vazquez and co-workers outline statistical methods for determining error rates in protein interaction networks (Chapter 13) while a complementing method by Gabriel Musso and co-workers reports on methods for evaluating the quality of published datasets (Chapter 14); David Fung and co-workers tackle the widespread problem of classifying network data from clinical samples based on pathway and network topology analysis (Chapter 15); Jennifer Smith and co-workers describe methods for mapping transcriptional regulatory networks from chromatin immunoprecipitation data using models of gene regulation with combinatorial DNA binding activity (Chapter 16); Colm Ryan and colleagues describe a method for predicting epistatic interactions based on nearest neighbour imputation (Chapter 17); Iain Wallace, Corey Nislow and coworkers highlight an application of the widely used Cytoscape application for displaying chemogenomic protein networks (Chapter 18); One of the aims of studying protein networks is to gain a better understanding of the coordinated role of protein assemblies, or complexes, in the biology of the cell; Benjamin Webb, Andrej Sali, and co-workers (Chapter 19) show how combining such network data with constraints for biophysical experiments, sophisticated integrated models of protein complexes may be assembled. Sara Mostafavi and co-workers describe new algorithms for characterizing node attributes and assigning function based on annotated data as a core set of tools applicable to many networks (Chapter 20). In a key final chapter, Alexander Ratushny and colleagues discuss building models of biomolecular network processes based on reaction kinetics and other parameters (Chapter 21).

This book emphasizes the practical application of these approaches – protocols along with troubleshooting guides and benchmarking criteria – and allows for a more comprehensive description than is typically encountered in primary research papers. We are indebted to all the many contributing authors who have strived to make the procedures in this volume concise and accessible, and hope the ensemble will be of value to a broad assortment of readers, ranging from graduate students new to the art to seasoned professionals looking to polish their skill sets. We also encourage the emerging systems biology community to apply these methods regularly in the lab or at the computer. After all, the limitation of past reductionist approaches demand that we cast the net more widely.

Dublin, Ireland ***Gerard Cagney, Ph.D.***
Toronto, ON, Canada ***Andrew Emili, Ph.D.***

http://www.springer.com/978-1-61779-275-5

Network Biology
Methods and Applications
Cagney, G.; Emili, A. (Eds.)
2011, XV, 437 p., Hardcover
ISBN: 978-1-61779-275-5
A product of Humana Press