

Frank J. Bruggeman, Jorrit J. Hornberg, Fred C. Boogerd and Hans V. Westerhoff
Introduction to systems biology
Summary

The developments in the molecular biosciences have made possible a shift to combined molecular and system-level approaches to biological research under the name of Systems Biology. It integrates many types of molecular knowledge, which can best be achieved by the synergistic use of models and experimental data. Many different types of modeling approaches are useful depending on the amount and quality of the molecular data available and the purpose of the model. Analysis of such models and the structure of molecular networks have lead to the discovery of principles of cell functioning overarching single species. Two main approaches of systems biology can be distinguished. Top-down systems biology is a method to characterize cells using system-wide data originating from the Omics in combination with modeling. Those models are often phenomenological but serve to discover new insights into the molecular network under study. Bottom-up systems biology does not start with data but with a detailed model of a molecular network on the basis of its molecular properties. In this approach, molecular networks can be quantitatively studied leading to predictive models that can be applied in drug design and optimization of product formation in bioengineering. In this chapter we introduce analysis of molecular network by use of models, the two approaches to systems biology, and we shall discuss a number of examples of recent successes in systems biology.

Christophe Rothan and Mathilde Causse
Natural and artificially induced genetic variability in crop and model plant species for plant systems biology
Summary

The sequencing of plant genomes which was completed a few years ago for *Arabidopsis thaliana* and *Oryza sativa* is currently underway for numerous crop plants of commercial value such as maize, poplar, tomato grape or tobacco. In addition, hundreds of thousands of expressed sequence tags (ESTs) are publicly available that may well represent 40-60% of the genes present in plant genomes. Despite its importance for life sciences, genome information is only an initial step towards understanding gene function (functional genomics) and deciphering the complex relationships between individual genes in the framework of gene networks. In this chapter we introduce and discuss means of generating and identifying genetic diversity, i.e. means to genetically perturb a biological system and to subsequently analyze the systems response, e.g. the changes in plant morphology and chemical composition. Generating and identifying genetic diversity is in its own right a highly powerful resource of information and is established as an invaluable tool for systems biology.

Christine H. Foyer, Guy Kiddle and Paul Verrier

Transcriptional profiling approaches to understanding how plants regulate growth and defence: a case study illustrated by analysis of the role of Vitamin C

Summary

In this chapter, basic technical aspects concerning the design of DNA microarray experiments are discussed including sample preparation, hybridization conditions and statistical significance of the acquired data are detailed. Given that microarrays are perhaps the most used tool in plant systems biology there is much experience in the pitfalls in using them. Herein important considerations are presented for both the experimental biologists and data analyst in order to maximize the utility of these resources. Finally a case study using the analysis of vitamin C deficient plants is presented to illustrate the power of this approach in enhancing comprehension of important and complex biological functions.

Lars Hennig and Claudia Köhler

Case studies for transcriptional profiling

Summary

DNA microarrays are frequently used to study transcriptome regulation in a wide variety of organisms. Although they are an invaluable tool for the acquisition of large scale dataset in plant systems biology, a number of surprising results and unanticipated complications are often encountered that illustrate the limitations and potential pitfalls of this technology. In this chapter, we will present examples of real world studies from two classes of microarray experiments that were designed to (i.) identify target genes for transcriptional regulators and (ii.) to characterize complex expression patterns to reveal unexpected dependencies within transcriptional networks.

Cameron Johnson and Venkatesan Sundaresan

Regulatory small RNAs in plants

Summary

The discovery of microRNAs in the last decade altered the paradigm that protein coding genes are the only significant components for the regulation of gene networks. Within a short period of time small RNA systems within regulatory networks of eukaryotic cells have been uncovered that will ultimately change the way we infer gene regulation networks from transcriptional profiling data. Small RNAs are involved in the regulation of global activities of genic regions via chromatin states, as inhibitors of “selfish” sequences (transposons, retroviruses), in establishment or maintenance of tissue/organ identity, and as modulators of the activity of transcription factor as well as ‘house keeping’ genes. With this chapter we provide an overview of the central aspects of small RNA function in plants and the features that distinguish the different small RNAs. We furthermore highlight the use of computational prediction methods for identification of plant miRNAs/precursors and their targets and provide examples for the experimental

validation of small RNA candidates that could represent trans-regulators of downstream genes. Lastly, the emerging concepts of small RNAs as modulators of gene expression constituting systems networks within different cells in a multicellular organism are discussed.

Erich Brunner, Bertran Gerrits, Mike Scott and Bernd Roschitzki

Differential display and protein quantification

Summary

High-throughput quantitation of proteins is of essential importance for all systems biology approaches and provides complementary information on steady-state gene expression and perturbation-induced systems responses. This information is necessary because it is, e.g. difficult to predict protein concentrations from the level of mRNAs, since regulatory processes at the posttranscriptional level adjust protein concentrations to prevailing conditions. Despite its importance, quantitative proteomics is still a challenging task because of the high dynamic range of protein concentrations in the cell and the variation in the physical properties of proteins. In this chapter we review the current status of and options for protein quantification in high-throughput experiments and discuss the suitability and limitations of different existing methods.

Sven Schuchardt and Albert Sickmann

Protein identification using mass spectrometry: a method overview

Summary

With the introduction of soft ionization techniques such as matrix assisted laser desorption ionization (MALDI)- and electrospray ionization (ESI), proteins have become accessible to mass spectrometric analyses. Since then, mass spectrometry has become the method of choice for sensitive, reliable and inexpensive protein and peptide identification. With the increasing number of full genome sequences for a variety of organisms and the numerous protein databases constructed thereof, all the tools necessary for the high-throughput protein identification with mass spectrometry are in place. This chapter highlights the different mass spectrometric techniques currently applied in proteome research by giving a brief overview of methods for identification of posttranslational modifications and discussing their suitability of strategies for protein quantification.

Dirk Steinhauser and Joachim Kopka

Methods, applications and concepts of metabolite profiling: primary metabolism

Summary

In the last decade of the 20th century the concept of a comprehensive analysis of the metabolic complement in biological systems, termed metabolomics or alternately metabonomics, was established as the last of four corner stones for phenotypic studies in

the post-genomic era. With genomic, transcriptomic, and proteomic technologies in place and metabolomic phenotyping under rapid development all necessary tools appear to be available today for a full functional assessment of biological phenomena at all major system levels of life. This chapter attempts to describe and discuss crucial steps of establishing and maintaining a gas chromatography - electron impact ionization - mass spectrometry (GC-EI-MS) based metabolite profiling platform. GC-EI-MS can be perceived as the first and exemplary profiling technology aimed at simultaneous and non-biased analysis of primary metabolites from biological samples. The potential and constraints of this profiling technology are among the best understood. Most problems are solved as well as pitfalls identified. Thus GC-EI-MS serves as an ideal example for students and scientists who intend to enter the field of metabolomics. This chapter will be biased towards GC-EI-MS analyses but aims at discussing general topics, such as experimental design, metabolite identification, quantification and data mining.

Lloyd W. Sumner, David V. Huhman, Ewa Urbanczyk-Wochniak and Zhentian Lei
Methods, applications and concepts of metabolite profiling: primary metabolism:
secondary metabolism

Summary

Plants manufacture a vast array of secondary metabolites/natural products for protection against biotic or abiotic environmental challenges. These compounds provide increased fitness due to their anti-microbial, anti-herbivory, and/or alleopathic activities.

Secondary metabolites also serve fundamental roles as key signaling compounds in mutualistic interactions and plant development. Metabolic profiling and integrated functional genomics are advancing the understanding of these intriguing biosynthetic pathways and the response of these pathways to environmental challenges. This chapter provides an overview of the basic methods, select applications, and future directions of metabolic profiling of secondary metabolism. The emphasis of the application section includes the combination of primary and secondary metabolic profiling. The future directions section describes the need for increased chromatographic and mass resolution, as well as the inevitable need and benefit of spatially and temporally resolved metabolic profiling.

Martine Dieuaide-Noubhani, Ana-Paula Alonso, Dominique Rolin, Wolfgang Eisenreich and Philippe Raymond

Metabolic flux analysis. Recent advances in carbon metabolism in plants

Summary

Isotopic tracers are used to both trace metabolic pathways and quantify fluxes through these pathways. The use of different labelling methods recently led to profound changes in our views of plant metabolism. Examples are taken from primary metabolism, with sugar interconversions, carbon partitioning between glycolysis and the pentose phosphate pathway, or metabolite inputs into the TCA cycle, as well as from secondary metabolism with the relative contribution of the plastidial and cytosolic pathways to the biosynthesis

of terpenoids. While labelling methods are often distinguished according to the instruments used for label detection, emphasis is put here on labelling duration. Short time labelling is adequate to study limited areas of the metabolic network. Long term labelling, when designed to obtain metabolic and isotopic steady state, allows to calculate various fluxes in large areas of central metabolism. After longer labelling periods, large amounts of label accumulate in structural or storage compounds: their detailed study through the retrobiosynthetic method gives access to the biosynthetic pathways of otherwise undetectable precursors. This chapter presents the power and limits of the different methods, and illustrates how they can be associated with each other and with other methods of cell biology, to provide the information needed for a rational approach of metabolic engineering.

Victoria J. Nikiforova and Lothar Willmitzer

Network visualization and network analysis

Summary

Network analysis of living systems is an essential component of contemporary systems biology. It is targeted at assemblance of mutual dependences between interacting systems elements into an integrated view of whole-system functioning. In the following chapter we describe the existing classification of what is referred to biological networks and show how complex interdependencies in biological systems can be represented in a simpler form of network graphs. Further structural analysis of the assembled biological network allows getting knowledge on the functioning of the entire biological system. Such aspects of network structure as connectivity of network elements and connectivity degree distribution, degree of node centralities, clustering coefficient, network diameter and average path length are touched. Networks are analysed statically, or the dynamical behaviour of underlying biological systems may be considered. The description of mathematical and computational approaches for determining the dynamics of regulatory networks is provided. Causality as another characteristic feature of a dynamically functioning biosystem can be also accessed in the reconstruction of biological networks, we give the examples of how this integration is accomplished. Further questions about network dynamics and evolution can be approached by means of network comparison. Network analysis gives rise to new global hypotheses on systems functionality and reductionist findings of novel molecular interactions, based on the reliability of network reconstructions, which has to be tested in the subsequent experiments. We provide a collection of useful links to be used for the analysis of biological networks.

Christian H. Ahrens, Ulrich Wagner, Hubert K. Rehrauer, Can Türker and Ralph Schlapbach

Current challenges and approaches for the synergistic use of systems biology data in the scientific community

Summary

Today's rapid development and broad application of high-throughput analytical technologies are transforming biological research and provide an amount of data and analytical opportunities to understand the fundamentals of biological processes undreamt of in past years. To fully exploit the potential of the large amount of data, scientists must be able to understand and interpret the information in an integrative manner. While the sheer data volume and heterogeneity of technical platforms within each discipline already poses a significant challenge, the heterogeneity of platforms and data formats across disciplines makes the integrative management, analysis, and interpretation of data a significantly more difficult task. This challenge thus lies at the heart of systems biology, which aims at a quantitative understanding of biological systems to the extent that systemic features can be predicted. In this chapter, we discuss several key issues that need to be addressed in order to put an integrated systems biology data analysis and mining within reach.

Matthias Steinfath, Dirk Reipsilber, Matthias Scholz, Dirk Walther and Joachim Selbig

Integrated data analysis for genome-wide research

Summary

Integrated data analysis is introduced as the intermediate level of a systems biology approach to analyse different "omics" datasets, i.e. genome-wide measurements of transcripts, protein levels or protein-protein interactions, and metabolite levels aiming at generating a coherent understanding of biological function. In this chapter we focus on different methods of correlation analyses ranging from simple pairwise correlation to kernel canonical correlation which were recently applied in molecular biology. Several examples are presented to illustrate their application. The input data for this analysis frequently originate from different experimental platforms. Therefore, preprocessing steps such as data normalisation and missing value estimation are inherent to this approach. The corresponding procedures, potential pitfalls and biases, and available software solutions are reviewed. The multiplicity of observations obtained in omics-profiling experiments necessitates the application of multiple testing correction techniques. Relevant algorithms to account for multiple testing are also discussed in this chapter.

**Daniel Schönert, Simon Barkow, Stefan Bleuler, Anja Wille, Philipp Zimmermann,
Peter Bühlmann, Wilhelm Gruissen and Eckart Zitzler**

Network analysis of systems elements

Summary

A central goal of postgenomic research is to assign a function to every predicted gene. Because genes often cooperate in order to establish and regulate cellular events, the examination of a gene has also included the search for at least a few interacting genes. This requires a strong hypothesis about possible interactions partners, which has often been derived from what was known about the gene or protein beforehand. Many times, though, this prior knowledge has either been completely lacking, biased towards favored concepts, or only partial due to the theoretically vast interaction space. With the advent of high-throughput technology and robotics in biological research it has become possible to study gene function on a global scale, monitoring entire genomes and proteomes at once. These systematic approaches aim at considering all possible dependencies between genes or their products, thereby exploring the interaction space at a systems scale. This chapter provides an introduction to network analysis and illustrates the corresponding concepts on the basis of gene expression data. First, an overview of existing methods for the identification of co-regulated genes is given. Second, the issue of topology inference is discussed and as an example a specific inference method is presented. And lastly, the application of these techniques is demonstrated for the *Arabidopsis thaliana* isoprenoid pathway.

Plant Systems Biology

Baginsky, S.; Fernie, A.R. (Eds.)

2007, XIV, 358 p. 64 illus., 4 illus. in color., Hardcover

ISBN: 978-3-7643-7261-3

A product of Birkhäuser Basel