Naama Barkai

Miri Carmi
Sven Bergmann
Andreas Doncic
Avigdor Eldar
Gilgi Friedlander
Jan Ihmels
Amnon Koren
Ofer Sarig
Daphna Strauss
Uri Shahar
Ron Chen
Sagi Levy
Roy Mizrachi

Design principles of biological networks

Department of Molecular Genetics; Physics of Complex Systems

Tel. 972 8 934 **4429** **Fax.** 972 8 934 **4108**
E-mail: naama.barkai@weizmann.ac.il
Web page: http://barkai-serv.weizmann.ac.il/GroupPage

In recent years, molecular biology is undergoing a transition from a discipline that focuses on the identification of individual components to the study of networks composed of many interacting components. Although the importance of applying quantitative approaches for analyzing network function is widely realized, theoretical and experimental methods still need to be developed. Moreover, the scope of biological issues that can be rigorously addressed by such "systematic" studies is not well established, and we are still far from understanding even simple bio-molecular networks.

Our aim is to study biochemical and genetics networks, using both theoretical and experimental approaches, to elucidate basic principles governing their design and function. We focus on two types of networks. First, we study relatively small sub-systems whose components are well characterized experimentally. Second, we study cellular networks using large-scale data generated by various genomic techniques.

The main challenge in studying limited-sized and relatively isolated networks is to elucidate core mechanism underlying network function. This is done by analyzing quantitatively the kinetics defined by different molecular models, and differentiating between them based on existing experimental evidence or biological constraints. It is important to emphasize that experimental evidence alone are usually not sufficient for pinpointing the underlying mechanism, since typically the *in-vivo* quantitative parameters such as rate constants or enzyme concentrations are not available, and functional experiments generate only a low resolution data due to the difficulty of performing precise, quantitative measurements. As a consequence, different models can be identified that are consistent with the observed data and additional, biologically-driven constraints are required in order to differentiate between them. An example for a biological constraint which we found to be highly useful is the robustness principle, namely the capacity to buffer quantitative perturbations in gene dosage. This ability to differentiate between models is also crucial for generating testable predictions. An important aspect of our work is to follow the theoretical analysis with quantitative experiments to verify and extent theoretical predictions.

A framework similar to the one described was applied previously for studying bacterial networks, in particular the *e. coli* chemotactic network. We extended this work to patterning networks that establish morphogen gradients in multi-cellular organisms. By focusing on specific networks employed during the development of the fruit-fly *Drosophila*, we address several general challenges:

Generating sharp activation peaks from shallow sources. This was addressed in the context of the network that generates the BMP gradient which patterns the dorsal ectoderm of the *Drosophila* embryo.

Generating long-range gradient while preserving the capacity to buffer fluctuations in morphogen production rate. This was addressed in the context of wing imaginal disc patterning in *Drosophila*.

Reducing patterning error by decoding the transient, rather than the steady-state morphogen profile during rapid patterning. This principle was elucidated in the context of the Bicoid system which defines the anterior-posterior polarity of the *Drosophila* embryo.

Generating thresholds for gene induction from gradual activation profiles. We address this issue in the context of patterning the ventral ectoderm in the early *Drosophila* embryo.

As a second approach, we study cellular networks using large-scale data generated by various

genomic techniques. The amount of data, and the size of networks considered, comes at the expense of data quality and the resolution of the mechanisms that can be resolved. The challenges we address range from developing bioinformatics tools for organizing large-scale information, to using this information for elucidating common design features. In addition to analyzing existing data, we are generating new genomic data for validating and extending computational results. We address the following issues:

Recurrent Signature': a novel bioinformatics approach for analyzing large-scale expression data. This method overcomes several well-recognizes drawback of commonly used clustering methods; Genes can be assigned to several overlapping clusters, co-regulation is defined in a context-specific manner and *a-priori* information can be easily included in the analysis.

The Recurrent Signature method extends the SVD approach to situations of noisy data. This was established theoretically and validated numerically using computer-generated data, as well as expression data from *S. cerevisiae*.

Systematic application of the Recurrent Signature method to large scale expression data generated testable function predictions. We applied the method to study and compare gene expression data from six diverse organisms (*E. coli, S. cerevisae, C. elegans, Drosophila, Arabidopsis* and human).

Software package for applying the Recurrent Signature algorithm and for analyzing the resulting clusters. Several packages were developed and are available on our website

We utilize those bioinformatics tools to identify common features that may present design principles: *Principles of Metabolic gene regulation.* We characterized in details the co-regulation of metabolic genes in *S. cerevisiae.* This study revealed several design features of metabolic gene regulation. For example, we found that co-regulation reduces metabolites dissipation at junction points by biasing metabolic flux toward linearity. This is achieved since typically only few of the branches at junction points are co-regulated. In addition, we found that in many cases, Individual members of isozyme families are differentially regulated with distinct pathways utilizing the associated reaction. This indicates that rather than being redundant, isozymes are primarily dedicated to specific processes, thus reducing crosstalk between separate pathways

A novel unified response to genetic perturbations. We identified a global transcription program, termed the Internal Stress Response (ISR) that is triggered in response to numerous unrelated genetic alterations. The ISR is distinct from the genomic response to general environmental stress. We suggested, and verified experimentally using genome-wide expression profiling, that adaptation of yeast to recognized perturbations, such as changes in temperature or carbon source, is fined-tuned with external conditions but is not linked directly with cellular growth capacity. In contrast, the transcriptional response to unrecognized perturbations, including gene deletion or the addition of specific drugs, is mediated through a distinct mechanism that is linked to internal growth efficiency.

Selected Publications

Barkai, N., and Leibler, S. (1997) Robustness in Simple Biochemical Networks. Nature, 387, 913-916

Barkai, N., and Leibler, S. (2000) Circadian clocks limited by noise Nature. 403, 267-688.

Imhels, J., Friedlander, G., Bergmann, S., Sarig, O., and Barkai, N. (2002) Defining Modular Organization in Yeast Transcriptional Network, Nature Genetics 31, 370-377.

Eldar, A., Dorfman, R., Weiss, D., Ashe. A., Shilo, B-Z., and Barkai, N. (2002) Robustness of Morphogen Gradients in Early Embryonic Patterning, Nature, 419 304-308.

Eldar, A., Rosin, D., Shilo, B., and Barkai, N. (2003) Self-enhanced morphogen degradation underlies robustness of morphogen gradients, Developmental Cell, 5 635-646

Imhels, J., Levy, R., and Barkai, N. (2004) Principles of transcription control in the metabolic network of *s. cerevisiae* Nature Biotechnology, 22 86-92.

Bergmann, S., Ihmels, J., and Barkai N. (2003) Similarities and differences in the genome-wide expression data of six organisms. PLoS, 2(1) 85-93.

Acknowledgements:

Supported by ISF, Israeli Ministry of Science (Tashtiot project), Minerva, NIH and HFSP; incumbent of the Soretta and Henry Shapiro Career Development Chair.