Regulatory genomics: A systems level investigation of expression control networks

Molecular genetics

Dr. Yitzhak Pilpel

Dr. Orna Dahan, Dr. Itay Furman, Dr. Maria Rodriguez Martinez, Amir Mitchell, Reut Shalgi, Ilya Venger, Ophir Shalem, Zohar Bloom, Hila Gingold, Asaph Carmi, Avihu Yona

☎ 972 8 934 6058

FAX 972 8 934 4108

@ pilpel@weizmann.ac.il

http://longitude.weizmann.ac.il

A critical challenge for living organisms is to execute intricate cellular programs in the face of environmental variations, genetic changes and the "noise" inherent in molecular processes. Our laboratory works to understand such programs by deciphering the structure, function and evolution of the regulatory networks that control them. Our studies tackle the networks' *structural* level, as well as the *functional* level in more complex processes. At the network structural level, we study transcription, translation, mRNA degradation, and non-coding RNAs. At the functional level, we explore how biological phenomena such as genetic backup, stress response, and evolutionary divergence of species are regulated through the various levels of gene expression. Our research entails a combination of computational and experimental work. Some of the published activities include:

I.1. The regulation and selection of genetic backup circuits (Kafri et al., Nature Genetics, 2005; Kafri et al., PNAS, 2006; Kafri, PNAS, 2008)

Members of certain gene families have long been known to provide functional "backup," to compensate for a mutated family member. Yet the question of whether cells control gene expression so as to provide backup activity when needed, was never before addressed. *We discovered that backup is often a regulated, responsive process, in the sense that when a gene undergoes a mutation, a paralog that in the wild-type cell was expressed under different conditions, changes its expression pattern to compensate for the loss of function of the mutant gene.* Using both single- and double gene deletion phenotypes, coupled with transcription factor binding information, we characterized the promoter architecture that supports controlled backup activity, and *identified recurring principles in many individually studied cases that manifest it.* Consequently, we predicted that backup is not randomly allocated to all the genes within the protein cellular network, but is rather concentrated at the network's "hubs", i.e. the most highly connected proteins in the network. We verified this prediction experimentally measuring epistatic interactions between pairs of paralogs, and concluded that *genetic redundancy is strategically allocated to the network's most vulnerable nodes, in a non-random fashion.*

I.2. Noise in gene expression is regulated for specific networks in different environments (Bar-Even et al., Nature Genetics, 2006; Kafri et al., PNAS, 2006)

In addition to genetic perturbations, cells must cope with non-genetic variations that result from the stochastic nature of the molecular processes governing gene expression. The sources of such "noise;" i.e., cell-to-cell variations among isogenic cells, were analyzed in a few selected types of genes. Yet in order to understand the biological significance of this noise, it was necessary to measure it in multiple genes belonging to diverse regulatory networks, in cells grown under various environmental conditions. In collaboration with our departmental colleague, Naama Barkai, we found that genes involved in different pathways and processes display distinctive noise patterns. Furthermore, we demonstrated that the same gene, examined under varying environmental conditions, often exhibited marked differences in noise levels. One of our more remarkable findings was that stress-related genes display the highest noise levels, suggesting the intriguing possibility that cell-to-cell variation in the expression of these genes may be beneficial, as it may enable non-genetic diversity and selection when cells are stressed. *We concluded that noise is, in fact, a regulated attribute of the genetic network, rather than a mere reflection of uncontrolled stochasticity.* We then proceeded to define, and mathematically quantify, the noise filtering capacities of a variety of genetic control circuits and *provided a first step towards establishing a link between regulatory network architecture and*

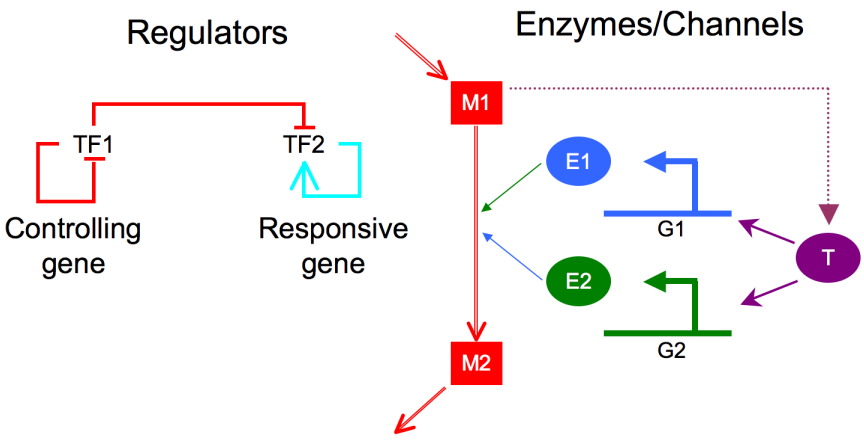

Fig. 1 *Two types of genetic switches, observed in biological systems, which can support responsive backup. On the right is the case of two paralogous enzymes that can catalyze the same reaction (converting between two metabolites, M1 and M2). A shared transcription factor, T, which is induced by M1 may convert the information about mutation in one of the genes to the promoter of the other gene. On the left is a case where the two paralogs are transcription factors that may exert negative regulation on each other. TF2 may respond to the deletion of TF1 and provide backup.*

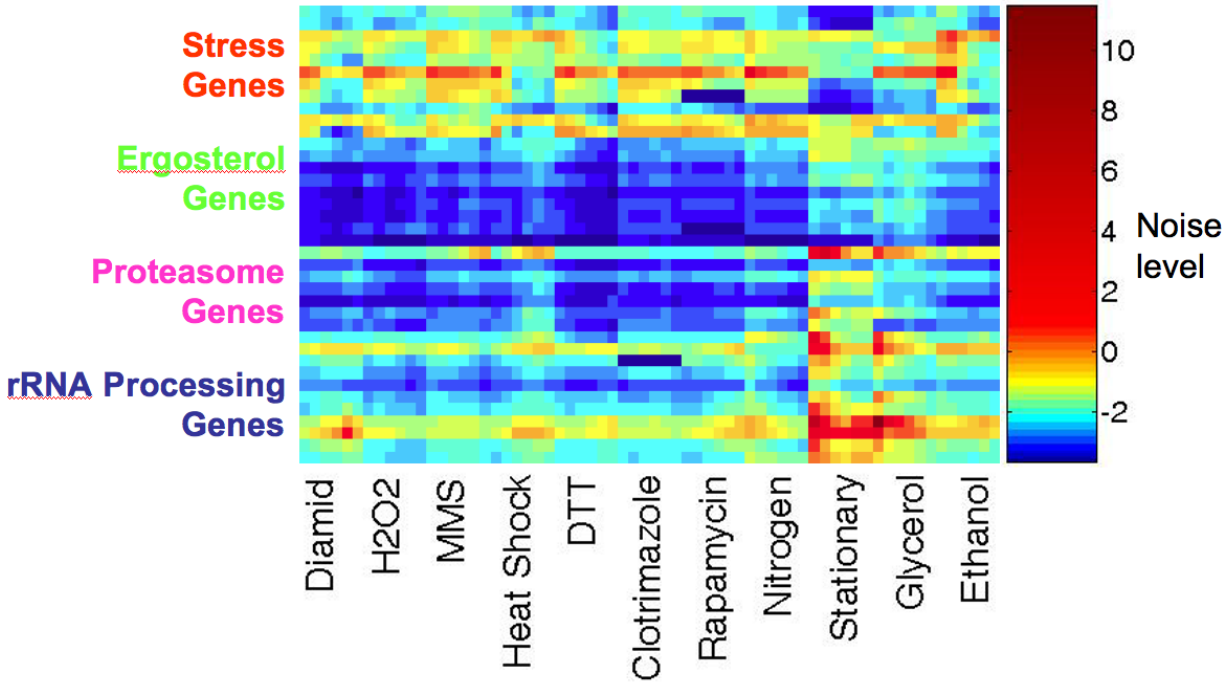

Fig 2. *Noise in gene expression was measured for 43 genes under 11 growth conditions. The noise (i.e. cell-to-cell variation) is depicted with a color code – red means high noise, blue stands for low noise. The genes are organized in rows and are clustered according to their function. The various environmental conditions are depicted below the relevant columns. Clearly the stress genes show the highest level of noise while the proteasomal and ergosterol genes show the lower noise levels.*

gene expression noise.

I.3. The regulation of translation efficiency is involved in phenotypic divergence among species (Man and Pilpel, Nature Genetics, 2007)

Regulation of gene expression not only influences cell-to-cell variation, but also impacts divergence among species during evolution. Seen within the broader evolutionary context, a critical question then arises: how are phenotypic differences between species encoded in their genomes? It has long been known that differential *transcriptional* regulation of orthologous genes in various species plays a role in phenotypic diversity. Yet until now, very little was known about the potential of differential *translational* regulation to affect species divergence. Analyzing ten fully sequenced yeast genomes, *we established the critical role played by translational regulation in the phenotypic divergence of species*. For each gene in the ten yeast genomes we analyzed we predicted its translational efficiency by measuring the adaptation of its codons to the organism's tRNA pool. Mining this dataset, we found that genes belonging to the same pathways and complexes often exhibit similar patterns of translational efficiency in the various species. In many cases, a clear rationale for these patterns lay in the fact that species sharing certain phenotypic characteristics, such as a particular metabolic preference, efficiently regulate the translation of the associated genes. *We therefore concluded that during evolution, extensive pressure acts on codons in hundreds of genes, tuning their translational efficiency and, ultimately, their encoded protein expression levels, according to their functions and to the organisms' needs*.

I.4. mRNA transcription and degradation networks (Tabach et al., Molecular Systems Biology, 2005; Shalgi et al., Genome Biology, 2005; Shalgi et al., PLoS Comp. Biol., 2007; Lapidot et al., PLoS Genetics, 2008)

Current studies of the transcriptome mainly focus on promoters and enhancers that control mRNA production. *Realizing that the transcriptome state reflects a balance between mRNA production and degradation, we established the first catalog of regulatory motifs that appear to control the decay kinetics of all transcripts in yeast*. By combining information on mRNA production and degradation, we succeeded in advancing the modeling of the transcriptome to a new stage. In the mammalian network, we studied the coupling between transcriptional and post-transcriptional regulation mediated through network motifs consisting of both transcription factors (TFs) and micro-RNAs (miRs). At the level of transcriptional control, we undertook collaborations with Institute scientists Varda Rotter (Molecular Cell

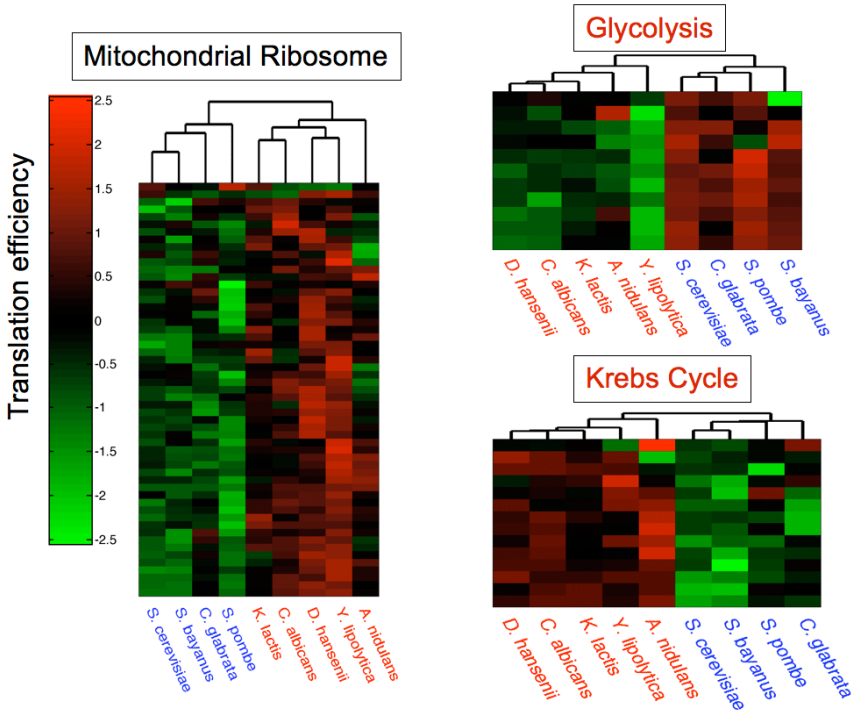

Fig. 3 *Translation efficiency of orthologous yeast genes computed in a variety of yeast species. In each matrix rows correspond to genes and columns to species. The color code (see bar on left) depict the efficiency of translation of each gene in each species, such that red colors correspond to high efficiency and green colors to low efficiency. The species names are given below each matrix, anaerobic species in blue, while aerobic in red. Three sets of genes are shown – mitochondrial ribosomal proteins, glycolytic enzymes and enzymes of the Krebs cycle. While aerobic genes are efficiently translated in the aerobes, the glycolytic ones are optimized in the anaerobic species.*

Biology) and Eytan Domany (Physics of Complex Systems), in order to tackle the challenge of determining how entire mammalian networks are regulated. We predicted a module of transcription factors that regulate the genes constituting the core cell cycle machinery, and then confirmed our predictions experimentally to show the function of the regulatory modules. *We found how these regulators map internal and external proliferative signals onto the expression levels of genes involved in the cell cycle machinery.*

Selected publications

Kafri, R. Bar-Even, A. & Pilpel, Y. (2005) Transcription control reprogramming in genetic backup circuits. *Nature Genetics* 37(33), 295-9

Tabach, Y. Milyavsky, M. Shats, I. Brosh, R. Zuk, O. Yitzhaky, A. Mantovani, R. Domany, E. Rotter, V*. & Pilpel, Y*. (2005) Genome-wide transcription regulatory circuits controlling cellular malignant transformation. *Molecular Systems Biology* doi:10.1038/msb4100030 *corresponding authors

Bar-Even A., Paulsson J., Maheshri N., Carmi M., O'Shea E., Pilpel Y.* & Barkai N.* (2006) Noise in protein expression scales with natural protein abundance *Nature Genetics* 38(6):636-43 *corresponding authors

Kafri R., Levi M. & Pilpel Y. (2006) The regulatory utilization of genetic redundancy through responsive backup circuits. *Proc Natl Acad Sci U S A.* 103(31):11653-8

Shalgi R, Lieber D. Oren M. & Pilpel Y. (2007) Global and local architecture of the mammalian microRNA-transcription factor regulatory network. *PloS. Comput. Biol.* 3(7): e131

Man O & Pilpel Y. (2007) Differential translation efficiency of orthologous genes is involved in phenotypic divergence of yeast species. *Nature Genetics* 39(3):415-21

Kafri R*, Dahan O*, Levy J, & Pilpel Y. (2008) Preferential protection of protein interaction network hubs in yeast: Evolved functionality of genetic redundancy. *Proc Natl Acad Sci U S A* *These authors contributed equally to this work

Acknowledgements

Y.P. is an incumbent of the Aser Rothstein career development chair in genetic diseases. External support: Israel Science Foundation, European Union, European Research Council, Minerva Foundation, European Molecular Biology Organization

INTERNAL support

The Ben May Charitable Trust, The Tauber Foundation, The Clore Center for Biological Physics, Samuel M. and Helene Soref Foundation, A grant from Mr. Walter Strauss