Minireview

Genome-wide analysis of the context-dependence of regulatory networks

Balázs Papp and Stephen Oliver

Address: Faculty of Life Sciences, University of Manchester, Michael Smith Building, Oxford Road, Manchester M13 9PT, UK.

Correspondence: Stephen Oliver. E-mail: steve.oliver@manchester.ac.uk

Published: 27 January 2005

Genome ***Biology*** 2005, **6:**206

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2005/6/2/206

Abstract

Genome-wide analytical tools are now allowing the discovery of the design rules that govern regulatory networks. Two recent studies in yeast have helped reveal the relatively small number of transcription-factor control strategies that cells employ to maximize their regulatory options using only a small number of components.

One of the earliest benefits of the complete genome sequences of major model organisms was the development of hybridization-array technology - DNA microarrays, or chips - which has enabled the mRNA levels for every gene in a genome to be monitored simultaneously [1]. This gives a picture of the transcriptome, the complete set of genes being expressed in a given cell or organism under a particular set of conditions. It should be possible to exploit such transcriptome data together with information on regulatory interactions to determine how cells regulate their gene-expression programs. But most efforts to map genome-scale transcription regulatory networks either have produced a network relevant only to one growth condition [2] or have included all previously described regulatory interactions, thus representing the total regulatory potential of the genome [3]. These static representations miss the importance of environmental transitions and ignore the time-dependence of regulatory interactions. In other words, the context-dependence that is intrinsic to functional genomics studies [4] has been lost or ignored.

A complete and dynamic description of gene regulation should enable us to answer a number of fundamental questions. What is the mechanistic basis of context-dependent regulatory interactions? How can a relatively small number of regulators respond to a huge variety of conditions? Can we identify 'design principles' in the architecture of transcriptional regulatory networks? What are the main functional differences between the underlying regulatory networks of the endogenous (developmental) and exogenous (sensory) gene-expression programs?

Context-dependence of regulatory interactions

Two approaches have recently been applied to mapping the gene regulatory networks of the budding yeast *Saccharomyces cerevisiae* in different physiological contexts. In the first, Harbison *et al.* [5] determined which sites on yeast chromosomes were occupied by which transcription factors under a number of environmental conditions. This analysis was performed for almost all of the yeast transcription factors and used chromatin immunoprecipitation array technology (ChIP-chip). In this method [6], living yeast cells are treated with a chemical cross-linking agent to 'freeze' protein-DNA interactions; chromatin fragments bearing specific transcription factors are then isolated by immunoprecipitation using antibodies against those factors. The DNA sites bound by the factors are then identified by hybridizing the DNA to a microarray. In this way, the genome occupancy of each transcription factor was examined in yeast grown in a rich medium; the occupancy of many of the regulators was also analyzed in at least one of 12

other environmental conditions [5]. In the second, purely computational, approach, Luscombe *et al*. [7] inferred the active part of the yeast regulatory network under five conditions by integrating gene-expression data with a static transcriptional network assembled from previously described regulatory interactions.

The first approach [5] should help us understand the specific functions of transcriptional regulators in terms of their binding behavior. Four general regulatory strategies emerged. In the first, termed 'condition invariant', the transcription factor binds the same set of promoters under different environmental conditions, but its activity depends on some additional requirements, such as ligand binding [8,9]. In the second, 'condition enabled', the transcription factor does not bind promoters under one set of conditions but binds a number of them in other conditions where it is present. In the third, 'condition expanded', the factor binds a core set of promoters under one condition but binds a larger set in a different condition where its level increases. In the fourth, 'condition altered', the factor binds different sets of promoters under different conditions. In fact, more than 40% of the transcriptional regulators investigated were found to alter their set of target genes in an environment-specific way.

If such a large proportion of transcriptional regulators display context-dependent activity, it is obviously important to determine the mechanisms by which their specificity is changed. This can occur both through direct modifications to the protein, such as phosphorylation, and through interactions with other regulator proteins [10]. Thus, the regulation of gene expression in a context-dependent manner may rely, to a large extent, on the combinatorial action of transcriptional regulators. Combinatorial regulation is not only an economic way to express a large number of regulatory states using only a limited number of regulators [11], but it also enables the transcription machinery to perform complex logical computations on the input signals [12,13]. The generality of combinatorial regulation in yeast is highlighted by the results of Luscombe *et al*. [7]: although many individual regulators are used in more than one condition, only a minor proportion of pairs of regulators participate in multiple transcriptional programs.

Design principles of gene regulatory networks

Systems biology can be regarded as the application of engineering principles to the understanding of biological 'machines'. In this context, there have been attempts to uncover the design principles of transcriptional networks [3,14], although it should always be remembered that these networks are the products of evolution, rather than design. So far, it is mainly the functions of local structures, such as network motifs (recurring network patterns) and regulatory cascades (a set of transcription factors that regulate each other sequentially), that have been investigated in detail. There appear to be significant differences between regulatory networks that are exogenous (that is, responsive to external stimuli such as stress) and those that are endogenous (that is, internal to the cell itself, such as the regulators of the cell cycle or meiosis). For instance, feed-forward loops, in which transcription factor X regulates transcription factor Y, with X and Y together regulating gene Z [15], represent a device to provide a rapid response in one direction - for example, ON to OFF - but a delayed response in the opposite direction - OFF to ON - thus enabling the circuit to be sensitive to sustained rather than transient signals. Feed-forward loops are found to be prevalent in, but not exclusive to, endogenous expression programs [7].

Luscombe *et al*. [7] report that not only does the frequency of certain motifs differ between endogenous and exogenous regulatory networks, but also the length of regulatory cascades varies between these two contexts. It has been shown theoretically [16] that cascades optimized for both rapid turn-on and turn-off kinetics have a response time proportional to the number of steps in the pathway, resulting in slow responses for multi-step cascades. As expected, cascades with short path lengths prevail in exogenous regulatory networks, presumably reflecting the need to achieve rapid and reversible responses [16]. In contrast, endogenous networks with long cascades regulate multi-step processes that proceed at a slower rate and for which fast response times may be less important. Moreover, many endogenous programs (for example, developmental pathways) are irreversible and need not be optimized for fast reversible changes [16].

Even if all transcription-factor-promoter interactions were mapped with high precision under a large number of conditions, we would still be far from having a complete model of gene regulation. First, information on the type (positive or negative) and kinetics of regulatory interactions is generally lacking; thus in order to understand the dynamic behavior of a transcriptional network it should be parameterized so as to add this kind of information [17]. Second, the functional activity of transcription factors is not necessarily regulated at the transcriptional level or through interactions with other transcription factors. Ligand binding [8,9] and post-translational modifications [10] could explain how certain regulators change their activity or specificity in a context-dependent manner. Third, the availability of promoters can also be regulated by chromatin structure, which in turn is modulated by proteins without sequence-specific DNA-recognition properties. Although a recent study investigated the genome-wide occupancy of certain chromatin regulators [18], it is clear that we need to learn more about how these are recruited to specific genomic regions with the help of transcription factors. Finally, in most cases, the ultimate signal to start a gene-expression program must come from the environment (in the widest sense of the term) and not from the transcriptional

network itself. Thus, it is essential to integrate the outputs of signaling networks with the inputs of gene regulatory networks to build a more complete representation of the cell's information processing machinery.

References

1. Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray.** *Science* 1995, **270:**467-470.
2. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I, *et al.*: **Transcriptional regulatory networks in *Saccharomyces cerevisiae*.** *Science* 2002, **298:**799-804.
3. Shen-Orr SS, Milo R, Mangan S, Alon U: **Network motifs in the transcriptional regulation network of *Escherichia coli*.** *Nat Genet* 2002, **31:**64-68.
4. Oliver SG: **Guilt-by-association goes global.** *Nature* 2000, **403:**601-603.
5. Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, Hannett NM, Tagne JB, Reynolds DB, Yoo J, *et al.*: **Transcriptional regulatory code of a eukaryotic genome.** *Nature* 2004, **431:**99-104.
6. Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, *et al.*: **Genome-wide location and function of DNA binding proteins.** *Science* 2000, **290:**2306-2309.
7. Luscombe NM, Babu MM, Yu H, Snyder M, Teichmann SA, Gerstein M: **Genomic analysis of regulatory network dynamics reveals large topological changes.** *Nature* 2004, **431:**308-312.
8. Kirkpatrick CR, Schimmel P: **Detection of leucine-independent DNA site occupancy of the yeast Leu3p transcriptional activator *in vivo*.** *Mol Cell Biol* 1995, **15:**4021-4030.
9. Sellick CA, Reece RJ: **Modulation of transcription factor function by an amino acid: activation of Put3p by proline.** *EMBO J* 2003, **22:**5147-5153.
10. Zeitlinger J, Simon I, Harbison CT, Hannett NM, Volkert TL, Fink GR, Young RA: **Program-specific distribution of a transcription factor dependent on partner transcription factor and MAPK signaling.** *Cell* 2003, **113:**395-404.
11. Pilpel Y, Sudarsanam P, Church GM: **Identifying regulatory networks by combinatorial analysis of promoter elements.** *Nat Genet* 2001, **29:**153-159.
12. Gerhart J, Kirschner M: *Cells, Embryos and Evolution*. Oxford: Blackwell Science; 1997.
13. Buchler NE, Gerland U, Hwa T: **On schemes of combinatorial transcription logic.** *Proc Natl Acad Sci USA* 2003, **100:**5136-5141.
14. Zaslaver A, Mayo AE, Rosenberg R, Bashkin P, Sberro H, Tsalyuk M, Surette MG, Alon U: **Just-in-time transcription program in metabolic pathways.** *Nat Genet* 2004, **36:**486-491.
15. Mangan S, Zaslaver A, Alon U: **The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks.** *J Mol Biol* 2003, **334:**197-204.
16. Rosenfeld N, Alon U: **Response delays and the structure of transcription networks.** *J Mol Biol* 2003, **329:**645-654.
17. Ronen M, Rosenberg R, Shraiman BI, Alon U: **Assigning numbers to the arrows: parameterizing a gene regulation network by using accurate expression kinetics.** *Proc Natl Acad Sci USA* 2002, **99:**10555-10560.
18. Robert F, Pokholok DK, Hannett NM, Rinaldi NJ, Chandy M, Rolfe A, Workman JL, Gifford DK, Young RA: **Global position and recruitment of HATs and HDACs in the yeast genome.** *Mol Cell* 2004, **16:**199-209.