The causes and consequences of gene expression flexibility

Allison Heath[1], Lydia Kavraki[1] and Gábor Balázsi[2]

***Short Abstract* — Cell populations can employ responsive or preventive strategies to survive in stress (such as drug treatment). Here we find that the same set of yeast genes with "flexible" (noisy and responsive) expression are used in both preventive and responsive survival strategies. We identify additional properties of these genes, and conclude on the possible causes and consequences of gene expression flexibility.**

***Keywords* — Gene expression, noise, survival, evolution.**

I. Introduction

CELL populations can employ responsive or preventive survival strategies to cope with environmental stress (such as drug treatment) [1,2]. For example, noise and random phenotype switching represent preventive, "bet hedging" survival strategies that keep subpopulations "on guard" before the stress arrives [2,3], while in responsive strategies defense is triggered by and follows the stress.

If cell defense is based on gene expression change, it is interesting to ask how the genes of model organisms are assigned to various defensive tasks. In this work, we use genome-scale data from *S. cerevisiae* to answer the questions: (i) are all yeast genes equally responsive? (ii) are all yeast genes equally noisy? (iii) are noisy genes also responsive? (iv) what other properties distinguish highly responsive or noisy genes from other genes?

II. Methods and Results

If proteins were synthesized and degraded at a uniform rate, the intrinsic noise in protein expression should have an inverse dependence on protein abundance [4]. However, the rate of protein synthesis in yeast is not uniform in time [3], which leads to dramatic deviations from the noise-abundance relationship expected in Poisson processes. In this study, we took the "extra" gene expression noise (in excess of the Poisson expectation) [5] and compared it to the responsiveness of the same genes in ~1000 microarray experiments [6]. Importantly, while noise measures the variation of gene expression among single cells in a constant environment, the responsiveness measures the amplitude of gene expression change in millions of cells across ~1,000 different conditions.

We found a highly significant correlation between the noise and responsiveness of yeast genes, indicating that the same set of genes with "flexible" expression are employed in both preventive and responsive survival strategies. The correlation between gene responsiveness (a macroscopic property of cell populations) and expression noise (a microscopic property of single cells) is reminiscent of the fluctuation-dissipation theorem in Physics.

We used data from several recent genome-scale studies to identify other gene properties that can be the cause or consequence of gene expression flexibility. We found that genes with flexible expression also tend to (i) have a TATA box in their core promoter [7], (ii) be highly regulated, (iii) have high mutational variance [8], (iv) be dispensable (non-essential), (v) have few protein interaction partners, (vi) be induced by stress, and (vii) be frequently duplicated/deleted during evolution [9].

Considering the time scales involved, it is likely that of the properties listed above, (i-ii) are the causes, while (iii-vii) are evolutionary consequences of gene expression flexibility.

III. Conclusion

In summary, we found an unexpected correlation between gene expression noise and gene responsiveness to diverse conditions. This implies that the same set of genes is used in defensive or responsive survival strategies. The molecular basis of this correlation is unclear, but other properties associated with these genes might provide a clue.

These findings suggest an evolutionary scenario in which a subset of genes first explores the gene expression space to improve population survival, followed by the stabilization of beneficial expression changes via deletions or duplications. As a consequence, genes with flexible expression are also expected to be highly evolvable.

Acknowledgements: This work was supported by M. D. Anderson start-up funds (to GB) and by NSF (to LK).

[1]Department of Computer Science, Rice University, 6100 Main Street, Houston, TX 77005. E-mails: aheath@cs.rice.edu, kavraki@cs.rice.edu.

[2]Department of Systems Biology - U950, UT M. D. Anderson Cancer Center, Houston, TX, 77054. E-mail: gbalazsi@mdanderson.org

References

[1] Thattai M, van Oudenaarden A (2004) *Genetics* **167**, 523-530.
[2] Wolf DM, Vazirani VV, Arkin AP (2005) *J Theor Biol* **234**, 255-262.
[3] Blake WJ, et al. (2006) *Mol Cell* **24**, 853-865.
[4] Paulsson J (2004) *Nature* **427**, 415-418.
[5] Newman JR et al. (2006) *Nature* **441**, 840-846.
[6] Tirosh I, et al. (2006) *Nat Genet.* **38**, 830-834.
[7] Basehoar AD, Zanton SJ, Pugh BF (2004) *Cell* **116**, 699-709.
[8] Landry CR, et al. (2007) *Science* **317**, 118-121.
[9] Wapinski I, et al. (2007) *Nature* **449**, 54-61.