Influence of gene copy number on self-regulated gene expression

Jakub Jędrak, Anna Ochab-Marcinek

www.elsevier.com/locate/yjtbi

PII: S0022-5193(16)30252-1
DOI: http://dx.doi.org/10.1016/j.jtbi.2016.08.018
Reference: YJTBI8786

To appear in: *Journal of Theoretical Biology*

Received date: 9 February 2016
Revised date: 15 July 2016
Accepted date: 11 August 2016

Cite this article as: Jakub Jędrak and Anna Ochab-Marcinek, Influence of gen copy number on self-regulated gene expression, *Journal of Theoretical Biology* http://dx.doi.org/10.1016/j.jtbi.2016.08.018

Influence of gene copy number on self-regulated gene expression

Jakub Jędrak and Anna Ochab-Marcinek*
Institute of Physical Chemistry, Polish Academy of Sciences, ul. Kasprzaka 44/52, 01-224 Warsaw, Poland
(Dated: August 11, 2016)

Using an analytically solvable stochastic model, we study the properties of a simple genetic circuit consisting of multiple copies of an self-regulating gene. We analyse how the variation in gene copy number and the mutations changing the auto-regulation strength affect the steady-state distribution of protein concentration.

We predict that one-reporter assay, an experimental method where the extrinsic noise level is inferred from the comparison of expression variance of a single and duplicated reporter gene, may give an incorrect estimation of the extrinsic noise contribution when applied to self-regulating genes.

We also show that an imperfect duplication of an auto-activated gene, changing the regulation strength of one of the copies, may lead to a hybrid, binary+graded response of these genes to external signal.

The analysis of relative changes in mean gene expression before and after duplication suggests that evolutionary accumulation of gene duplications may non-trivially depend on the inherent noisiness of a given gene, quantified by maximal mean frequency of bursts.

Moreover, we find that the dependence of gene expression noise on gene copy number and auto-regulation strength may qualitatively differ, e.g. in monotonicity, depending on whether the noise is measured by Fano factor or coefficient of variation. Thus, experimentally-based hypotheses linking gene expression noise and evolutionary optimisation may be ambiguous as they are dependent on the particular function chosen to quantify noise.

I. INTRODUCTION

Gene copy number variation is an ubiquitous phenomenon that manifests itself in multiplication of gene fragments, single genes, groups of genes up to whole genome. Duplicated genes contribute to gene evolution; subsequent mutations may turn one of gene copies into an inactive pseudo-gene, which may accumulate further mutations without affecting the phenotype [1–3]. Gene copies may be parts of either chromosomal or extra-chromosomal DNA. In bacterial cells, low-copy plasmids appear in the numbers of copies characteristic to plasmid type, and the numbers are conserved during cell division [1]. Bacterial plasmids, as well as the circular molecules of DNA found in mitochondria and chloroplasts may also appear in high numbers of copies (e.g., 20-40 for chloroplasts of higher plants [1]).

Variation in the number of copies of a particular gene in a living cell may strongly affect the concentration of protein encoded for by that gene. This in turn may have a profound impact on the phenotype, and hence on the fitness of the organism. The relationship between copy number variation and phenotype is of great interest in higher eukaryotes such as mammals, including humans, where gene copy number variation is known to be related not only to differences in concentrations of some enzymes (e.g., starch amylase, [4]) but also to several genetic diseases [5, 6] as well as cancer [7]. However, it is usually easier to study experimentally the effects of copy number variation in model unicellular organisms, such as *E. coli* or *S. cerevisiae*; strains of such organisms differing by gene copy number may be relatively easily constructed [8, 9]. Yet, within the existing mathematical models of gene expression [10–17], usually a single gene copy is considered, and the influence of gene copy number on gene expression is neglected. To the best of our knowledge, there are only few papers providing a theoretical description of the influence of copy number variation on gene expression [9, 18–23]

In particular, in Ref. [18], the influence of copy number variation on the gene expression level was studied in the case of four different network motifs, from a simple auto-activated gene (positive feedback) to more complicated, two- and three-gene circuits. This analysis, although thorough and throwing much light on the subject, was nonetheless based on deterministic approach so it neglected the molecular noise, inherent to as small biochemical systems as living cells. In the present paper, we will focus on how the noise produced by self-regulating gene depends on the copy number of that gene.

The dependence of gene expression noise on the strength of negative self-regulation of two gene copies was analysed in Refs. [20, 21]. It was concluded that gene expression noise, measured there by Fano factor, may prevent the evolution of strong negative auto-regulation in diploid cells, and this was proposed as a possible explanation of the observed difference in abundance of negative auto-regulation between *E. coli* (where negative auto-regulation is a frequently appearing network motif) and *S. cerevisiae* and other eukaryotic species (where it is much less frequent). The authors pointed out that it may also account for the fact that duplicated copies of negatively self-regulating genes are relatively rare in *E. coli*, despite the fact that roughly half of all known transcription factors of *E. coli* take part in negative auto-

*electronic address: ochab@ichf.edu.pl