How the Number of Alleles Influences Gene Expression

Beata Hat,1,2 Pawel Paszek,2 Marek Kimmel,2,3 Kazimierz Piechor1 and Tomasz Lipniacki1,2

Received January 28, 2006; accepted September 20, 2006
Published Online March 7, 2007

The higher organisms, eukaryotes, are diploid and most of their genes have two homological copies (alleles). However, the number of alleles in a cell is not constant. In the S phase of the cell cycle all the genome is duplicated and then in the G2 phase and mitosis, which together last for several hours, most of the genes have four copies instead of two. Cancer development is, in many cases, associated with a change in allele number. Several genetic diseases are caused by haploinsufficiency: Lack of one of the alleles or its improper functioning. In the paper we consider the stochastic expression of a gene having a variable number of copies. We applied our previously developed method in which the reaction channels are split into slow (connected with change of gene state) and fast (connected with mRNA/protein synthesis/decay), the later being approximated by deterministic reaction rate equations. As a result we represent gene expression as a piecewise deterministic time-continuous Markov process, which is further related with a system of partial differential hyperbolic equations for probability density functions (pdfs) of protein distribution. The stationary pdfs are calculated analytically for haploidal gene or numerically for diploidal and tetraploidal ones. We distinguished nine classes of simultaneous activation of haploid, diploid and tetraploid genes. This allows for analysis of potential consequences of gene duplication or allele loss. We show that when gene activity is autoregulated by a positive feedback, the change in number of gene alleles may have dramatic consequences for its regulation and may not be compensated by the change of efficiency of mRNA synthesis per allele.

KEY WORDS: stochastic gene expression, feedback regulation, diploid genes, haploinsufficiency, piecewise deterministic time-continuous Markov process

1 Institute of Fundamental Technological Research, Swietokrzyska 21, 00-049 Warsaw, Poland; e-mail: tomek@rice.edu and tlipnia@ippt.gov.pl.
2 Department of Statistics, Rice University, 6100 Main St. MS-138, Houston, TX 77005.
3 Institute of Automation, Silesian Technical University, 44-100 Gliwice, Poland.
1. INTRODUCTION

Stochasticity in gene expression arises from fluctuation in gene activity, mRNA transcription, protein translation and oligomerization (Refs. 16, 11, 28, 29, recently reviewed in Refs. 9 and 17). Figure 1 illustrates the main steps in gene expression. Control of gene activity is mediated by transcription factors which may bind a specific promoter regions and switch the allele on or off. When the gene is active, RNA polymerase may bind the gene promoter and initiate mRNA transcription. Next, mRNA becomes edited and exported from the nucleus to the cytoplasm, where the protein translation occurs. Accordingly, a single event of allele activation results (if the activation period is sufficiently long), in a burst of mRNA molecules, which is then translated into an even larger burst of proteins. Stochasticity in gene expression causes the population of cells to exhibit a large cell-to-cell variability as observed for example by Takasuka et al. and Stirland et al. for mammalian cells, Raser and O’Shea for budding yeast (Saccharomyces cerevisiae) or Elowitz et al. for bacteria (Escherichia coli).

Prokaryotes are haploid, i.e., most of their genes have only one copy. The higher organisms eukaryotes are diploid, most of their genes have two homologous copies (alleles), which can be independently activated and inactivated. However, the number of gene copies is not constant in cell evolution. In the S phase of cell cycle whole genome is duplicated and then in the G2 phase and mitosis, which together last for several hours, most genes have four copies instead of two. On the contrary, in meiosis 4 daughter haploid cells (gametes) are produced (in two subsequent divisions) out of one diploid cell. The genetic defect of loss of one allele or its transcriptional inactivity can result in haploinsufficiency, which is a hallmark of some diseases. Cancer cells may have gene or chromosomal duplications resulting in a larger number of alleles which may substantially alter cell function and make the disease more aggressive. Transfected cells used in various experiments may have arbitrary and difficult to control number of homologous copies.

Fig. 1. Simplified schematic diagram of gene expression.