UNIVERSITY OF AMSTERDAM

Understanding Gene Expression Variability in its Biological Context Using Theoretical and Experimental Analyses of Single Cells.

H. Kempe

Summary

Understanding gene expression variability in its biological context using theoretical and experimental analyses of single cells

Traditional gene expression studies have largely ignored cell-to-cell variability in transcription. Current methods allow for single cell analyses and have shown considerable variability in gene expression, even in populations of isogenic cells being exposed to the same growth environment. In this thesis, we assess the impact of various parameters of gene expression variability using experimental systems that enable quantification of the gene expression status of single cells. Based on the obtained data we parameterized mathematical models of gene expression.

In the first two chapters we focused on the effects of cellular volume growth on gene expression. Cell growth largely relates to an increase in cellular volume and the amount of nuclear DNA, i.e. in the interphase of the cell cycle cells double their cellular volume and DNA, and in mitosis they divide into two daughter cells. In general, the concentrations of reactants dictate the reaction rates of the chemical reactions inside cells. This means that to keep protein production constant in a cell, the number of mRNA molecules (and other reactants) need to increase at the same rate as cellular volume. However, gene expression variability is traditionally quantified in absolute numbers by quantifying mRNA copy numbers, which ignores the impact of cellular growth.

We combined single-cell mRNA expression levels with cell volume measurements and observed that the number of mRNA molecules scales proportionally with cell volume. This means that biologically relevant transcription variability (based on mRNA concentrations) is much lower than generally determined by single-molecule mRNA analysis that only takes transcription counts into consideration. The proportionality between mRNA numbers and cell volume implies that the mRNA concentration remains homeostatic over the course of the cell cycle. This indicates that regulatory mechanisms are in place to counteract both the dilution of DNA concentration when cells grow and the doubling of DNA during DNA replication. Our data show the importance of measuring cell growth in combination with gene expression cell-to-cell variability. Hence, it is crucial to understand how cells grow and to measure how growth affects gene expression. We combined experimental data and a theoretical framework of bacterial cell cell growth to construct a model to analyze gene expression in a growing population of cells. The obtained simulations were validated with an experimental setup that combined cell growth measurements with protein measurements in bacterial cells. This algorithm is implemented in StochPy (a python simulation package), which can now be used for simulations and analyses of gene expression networks in growing cells.

In chapter 2 we observed the effects of local chromatin structure in which genes are embedded on gene expression variability. This cell-to-cell variability can only be explained by changes

in the dynamics of the expression model. In the fourth chapter we describe how we modified the chromatin environment of a reporter gene and measured synthesized transcripts in real time when inducing transcription inactivation of the activated reporter gene. We observed that reporter gene inactivation is preceded by a delayed response and that targeting a chromatin regulatory protein (epigenetic reader protein methyl-CpG binding protein-2) to the reporter gene accelerates the response to signals suppressing active transcription.

In chapter 5 we demonstrate the effects of UV-damage on gene expression. We quantified nuclear single-molecule mRNA numbers to determine the UV-damage-induced transcription changes at the single gene level. Our approach enabled us to determine the relationship between UV-dose, gene size, transcription recovery, and DNA repair at the single gene level. Since the probability on gene damage is proportional to its length, gene length is an important parameter for UV damage-induced transcription stalling. By measuring the recovery of transcription after UV-exposure we were able to estimate the half-life of DNA damages. Intriguingly, we observed two distinct half-lives of DNA damage depending on the UV-dose supporting recent observations that transcription coupled and global DNA repair pathways can be discriminated by exposing cells to a defined UV-dose.

In chapter 6 we studied the effect of gene expression variability in breast cancer cells. Although relapse to treatment is observed in up to 30-40% of patients, the exact mechanism involved is unknown. It is expected that changes in the expression level of aromatase induces resistance and stimulates tumor progression. We demonstrate that in breast cancer cells treated with aromatase inhibitors, a subpopulation of cells overexpresses aromatase. Our data suggest that treatment resistance may be initiated by a subset of cells with an altered expression status. Overall, we analyzed gene expression variability in dynamic cells and environments and demonstrate how single cell techniques can be utilized to understand gene expression variability.