Quantification of Transcriptional Noise at Single Cell Single Molecule Level

Olga Anufrieva[1], Olli Yli-Harja[2] and Meenakshisundaram Kandhavelu[1]*

[1]*Molecular Signalling Lab, Computational Systems Biology Research Group, BioMediTech and Faculty of Biomedical Sciences and Engineering, Tampere University of Technology, Tampere, Finland*
[2]*Computational Systems Biology Research Group, BioMediTech and Faculty of Biomedical Sciences and Engineering, Tampere University of Technology, Tampere, Finland*

*Corresponding Author: Meenakshisundaram Kandhavelu, Molecular Signalling Lab, Computational Systems Biology Research Group, BioMediTech and Faculty of Biomedical Sciences and Engineering, Tampere University of Technology, Tampere, Finland.

Abstract

The stochastic gene expression leads to the substantial cell-to-cell variability. The resulting sub-populations differ by mRNA number per cell. With the aim of measuring the stochastic gene expression at transcription, we use MS2-GFP method for detecting mRNA number per single cell. The mRNA molecules are expressed under the control of *lac* promoter. The *lac* variant promoter is used for the independent validation of the *lac* promoter activity. In the fully induced cells, the cell-to-cell variability is determined by higher extrinsic than intrinsic noise component. Both of the noise components decrease over the phases. Evaluation of such behavior by the degree of population synchrony reveals that totally asynchronous cellular division is responsible for high extrinsic component and for decreasing of both noise components during the bacterial cell phases. In addition, the cell-to-cell variability is determined more by the downstream than the upstream gene activation condition. These findings provide a qualitative basis for developing *in silico* model in which the noise dynamics is tightly controlled during the cell divisions in a heterogeneous population.

Keywords: Gene Expression; Noise; Biosensor; Model; Degree of Synchrony

Introduction

Traditionally, behavior of cells was investigated by the methods that measure the average fluorescent intensity, mRNA or protein numbers in a homogenous population. Such measurements do not describe an individual behavior of the single cells. In the contrast to the previous observations, current study describes a stochastic gene expression in the single *E. coli* cells. In addition, stochastic gene expression is observed over bacterial growth phases, e.g. Lag, Acceleration, Exponential, Retardation and Stationary [1-4]. For these purposes, we use mRNA number per cell measured by the MS2-GFP method, total population experiment.

We describe stochastic gene expression by the cell-to-cell variability (η_{tot}) or total noise [5-7]. At transcription, total noise is approximated by the Fano factor [2,8]. Total noise can be further divided into two components such as extrinsic and intrinsic. An extrinsic component (η_{ext}) originates from the rate of the repressor binding, variability in RNAP copy numbers and gene copy number [9]. An intrinsic component (η_{int}) originates from the frequency of the transcription factor binding and unbinding, mRNA partitioning, mRNA degradation and transcription initiation[9].

Recently, we observed the real-time gene expression noise in live cells over the phase. From the result we shows that noise from the *lac* promoter in *E. coli* is dominated by intrinsic and extrinsic noises in slow and fast dividing cells, respectively [10]. However, population of the cell may exhibits different behaviour than the individual cell given the nature of biased partitioning of mRNA and protein molecules, variation in degradation rate etc. In the present work, we estimate extrinsic component as a difference between mRNA numbers per cell and intrinsic component as a squared difference of total and extrinsic noises, such as $\eta^2_{ext} = \eta^2_{tot} - \eta^2_{int}$ [2]. In such a way, it is possible to

estimate the levels of both noise components in different bacterial growth phases, at population level [1,2]. Such measurements expand our understanding of the stochastic gene expression by observing it over the bacterial cell phases and not in one particular phase, e.g. Exponential phase [11]. To ensure that observed mRNA number distribution is from the single cells of a particular phase, we use MS2-GFP method for tracking mRNA numbers in a real-time, *in vivo* and for observing single cells prior, during or immediately after the cell division [1,2].

In a population, cells may divide synchronously or asynchronously [12]. The degree of synchronously divided cells affects the mRNA partitioning between cells in the bacterial growth phases. The mRNA partitioning and degree of population synchrony certainly have an impact to the stochasticity of the gene expression. For estimating an effect of the (a)synchronous cellular division, we simulate dynamics of the cell-to-cell variability originated from the synchronously divided cells and predict the changes of the cell-to-cell variability calculated from the asynchronously divided cells. Accompanied by the appropriate *in silico* Modeling and statistical validation these observations provide new insights onto cellular adaptive mechanisms operating during the growth phases.

Another important factor that affects the mRNA number per single cell is the balance between upstream and downstream activations of the promoter regions [1]. To understand an effect of distinct activation conditions onto the stochastic gene expression, we use *lac* promoter and *lac* variant promoter induced independently from the upstream or from the downstream [1]. The results of such measurements are important for describing the stochastic gene expression in the bacterial cells, over the cell growth phases.

Overall, these observations allows for predicting how the cell population keeps the stability by modulating the levels of the total noise, intrinsic and extrinsic components over the bacterial growth phases. In perspective, these results could be compared to the stochastic gene expression at translation, over the phases, for revealing the degree of correlation between the noise levels at both processes. As well, the proposed way of noise quantification can be applied for studying the noise levels under the control of other promoters.

Materials and Methods

Strains and media

An *Escherichia coli* Dh5α-PRO strain, K12 variant, is used as a host system to study mRNA expression at a single cell level, over the bacterial growth phases. One system is co-transformed with the two plasmids. The first plasmid is a single copy bacterial artificial chromosome (BAC), pTRUEBLUE-BAC2 (ChlR), tagged with 96 target binding sites (96xbs) for MS2 protein coding genes. This gene cassette is located under the control of wild type *lac* promoter with the Catabolic Activator Protein (CAP) site in the upstream and Operator 1 site in the downstream for binding of the CAP and *lac* repressor proteins, respectively [1,2,13]. This promoter is activated by cyclic AMP (cAMP) and Isopropyl-D-1-thiogalactopyranoside (IPTG). The second plasmid is a medium-copy reporter, pPROTET.E (KanR), expressing MS2 coat protein tagged with green fluorescent protein (GFP) coding gene (GFP-mut3) that is located under the control of promoter LtetO-1 [1,2,13]. Another system, also, consists of two plasmids constructed as described above but still possesses a number of modifications. The BAC carries mRFP1 coding gene and 96xbs under the control of a *lac* variant promoter, *lac/ara1*, with araI1-I2 binding site for araC and *lac* Operator 1 binding site for *lac* repressor in the upstream and downstream, respectively [1,2,14]. All the strains are grown in LB medium, in the presence of the required antibiotics. Used medium is composed of 10 g/L of tryptone (T7293-Sigma Aldrich-USA), 5 g/L of yeast extract (MC 001-LabM-UK) and 10 g/L of NaCl (S3014-LabM-UK).

Cell phase determination and induction

The generation time of each phase is calculated from the OD_{600} measurements. Multiple technical and biological repeats are used to calculate the generation times to ensure that the majority of the cells in a population are of specific cell phase (Figure 1A). To monitor the divided cells transient from phase to phase, image acquisition is initiated within 5 - 10 minutes after the generation time. In such a way we are certain that the majority of the cells in a population are of specific cell phase [1,2]. Based on generation times of the cells such phases as Lag, Acceleration, Exponential, Retardation and Stationary were determined [3]. To study transcriptional noise, *lac* and *lac* vari-

ant promoters are activated by corresponding inducers for 15 minutes. To detect an individual mRNA molecule, PLtetO-1-MS2-GFPmut3 is activated by 50 ng/ml of anhydrotetracycline (aTc) (lot number 2-0401-001, IBA GmbH, Germany) 20 minutes prior to the detection of mRNAs. The P_{lac}DRS and P_{lac}URS inductions are achieved by 1 mM IPTG (L6758-Sigma Aldrich-USA) and endogenous cAMP, respectively [13]. The $P_{lac/ara1}$DRS and $P_{lac/ara1}$URS inductions are achieved by 0.1% L-arabinose (A3256-Sigma Aldrich-USA) and 1 mM IPTG, respectively. As a control, cells without 96 binding sites were used. In such cells no fluorescent spots are detected and, consequently, no mRNAs are produced (Supplementary figure 1). Image acquisition is performed using the microscope immediately after the division time of each phase. Multiple biological and technical repeats are performed to reproduce the results.
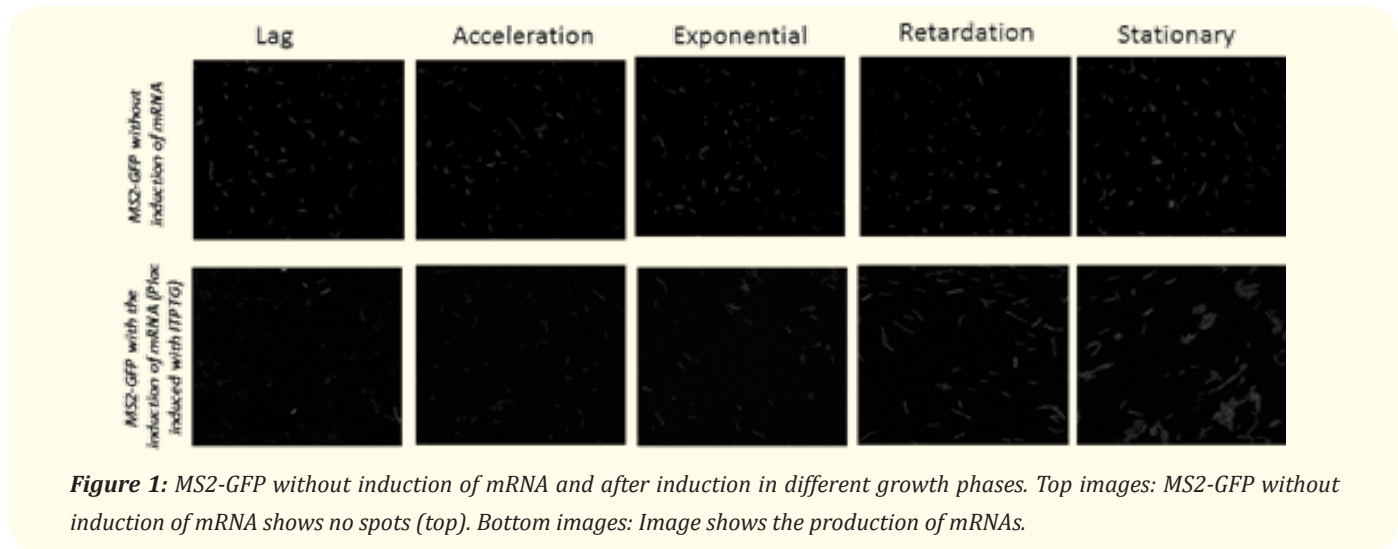

Figure 1: MS2-GFP without induction of mRNA and after induction in different growth phases. Top images: MS2-GFP without induction of mRNA shows no spots (top). Bottom images: Image shows the production of mRNAs.

Live cells imaging

Induced cells are pelleted by the centrifugation at 6000 RPM. Then, the cells are suspended in the required volume of LB and placed between 1% LB-agarose gel pad and a cover slip on a microscopic slide. We used Nikon, Eclipse Ti-E confocal microscope for the image acquisition. Specifically, we use 100x N.A. 1.49 oil immersion objective. This microscope is equipped with a hardware autofocus module, motorized z-drive and Nikon's Perfect Focus System (PFS) to maintain the cells in focus during the image acquisition. The PFS function automatically corrects small changes in focus to capture the cell. Nikon NIS-Elements-C was used to acquire the images. Fluorescence of individual cells are measured using a 488 nm argon ion laser (Melles-Griot) and a 515/30 nm emission filter.

Cell segmentation and image processing

From the microscopy images, we detect cells by a semi-automated method consisting of manual masking of the region of the images displaying cells having fluorescent spots **(Supplementary movie).** Cells with spots are analysed by principal-component analysis (PCA) and segmented by sub-cell object detection methods such as kernel density estimation (Gaussian kernel) along with Otsu's threshold method [15]. These methods allow for extracting the information about locations, orientations and dimensions of the masked cells [2]. Unwanted noise in the intensity signals of the spots is reduced by detection of precise location of the segmented fluorescent spots that reduces false negatives or missed spots. Besides, accurate selection of the threshold value minimizes intra-class variance [16]. Extracted fluorescent intensity of the spots per cell is used for measurement of the mRNA number per cell by the scaling procedure described by the Sala., *et al* [2]. In such a way sufficient number of the cells are analysed in each condition [1]. The total number of analysed cells in each experimental conditions over the phases are listed in the supplementary document (Table S1).

Cell phases	Lag	Acceleration	Exponential	Retardation	Stationary
Conditions					
P_{lac}URS-DRS	547	432	508	955	440
$P_{lac-ara1}$URS-DRS	135	549	667	516	465
$P_{lac-ara1}$URS	182	260	472	302	55
$P_{lac-ara1}$DRS	385	467	314	459	455

Table S1: *The number of analyzed cells in each experimental condition, over the phases.*

Noise quantification

In a homogenous population, we approximate the cell-to-cell variability (η_{tot}) or total noise as the Fano factor (defined as the standard deviation divided by the mean) or burstiness factor [2,8].

$$\eta_{tot} = \frac{\acute{o}^2}{m}$$

In a population, an extrinsic component (ŋext) is calculated as the variances of the mRNA number per cell from the arbitrary pair of strains [17].

$$\eta_{ext} = \frac{|mV_2 - nV_1|}{mn^2 - nm^2}$$

where, V_1 and V_2 are two single cells in the population with a number of produced mRNAs at a particular phase, m and n are arbitrary pair of strains with a different copy number of the genes. For avoiding additional noise origins, we used a single-copy number gene. To be able to observe the contribution of extrinsic and intrinsic components to the total noise, an intrinsic component is approximated as

$$\varsigma_{int} = \frac{|n^2V_1 - m^2V_2|}{mn^2 - nm^2}$$

$$\varsigma_{int} = \sqrt{\left(\frac{\acute{o}^2}{m}\right)^2 - \left(V_2 - V_1\right)^2}$$

In such a way, intrinsic and extrinsic components make an orthogonal contribution to the total noise and can be used for its description. Recently, such estimation of the gene expression noise was performed by the Sala., *et al*. who approximated the levels of intrinsic and extrinsic components originated from the *lac* promoter and *lac* variant promoter, at the Exponential phase in the *E. coli* cells [2].

Approximation of population synchronization

In perfectly synchronized population, cells divide simultaneously at a given time points, while in asynchronous population, cells divide randomly over a range of time [12]. Our cell populations are a priori asynchronous as probability of RNAP binding to the promoter, duration of the open complex formation and time required for the RNA polymerase to bind to the promoter is not controlled simultaneously. However, the time, cells proceed from one phase to another, is tightly regulated as done by Sala., *et al* [1,2]. Thus, we are able to simulate the degree of synchrony of the cell-to-cell variability (η_{tot}) subjected to the cell phases. The *in silico* model for prediction of the cellular behavior is implemented in Matlab 2013a. In addition, the model approximates the behavior of asynchronous divisions during the cell phases. By using sum of squares due to error (SSE) and coefficient of determination (R-square) we estimate the goodness of the performed analysis (Table 2).

Results

Measuring mRNAs in live cells

Production of mRNA molecules in single cells is studied by MS2-GFP system. The system consists of RNA-binding protein, MS2 coat protein that is fused to the GFP and is expressed along with the target mRNA tagged with RNA motif. To image mRNAs, multiple copies of RNA motif are introduced into the gene encoding target mRNA. In such a way, mRNA molecules are expressed in single cells along with MS2 coat protein fused to GFP. Consequent binding of multiple copies of fused MS2-GFP to target mRNAs ensures that produced molecules possess considerably higher fluorescent signal than the surrounding cytoplasmic region of the cell [18]. It allows for observing production of mRNAs as a fluorescent spots on the microscopic images. Fluorescent signal of cells with spots and consequent scaling of fluorescent signal into mRNA number is done as described in the Methods.
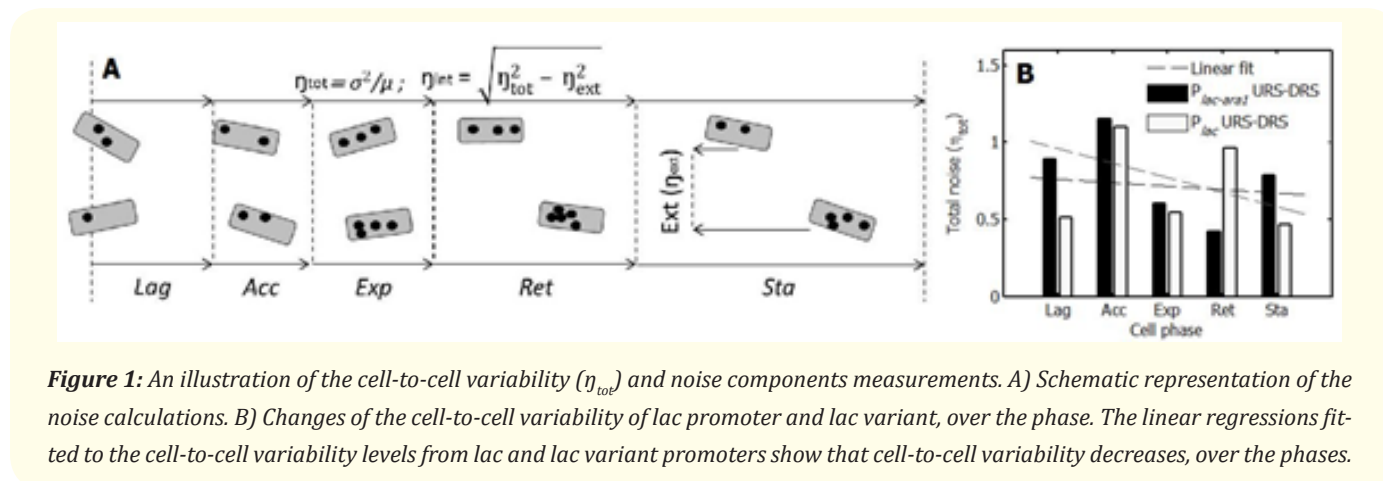
The cellular MS2-GFP level was already reported by Ido golding in 2005, Cell. It was found that Based on fluorescence measurements (6 different experiments, > 2700 cells), each cell contains 104 MS2-GFP molecules (approximately 10 mM) in this induction range. Of these molecules, typically only 3% - 4% are bound to RNA targets, with a maximum fraction of 10% at the highest RNA levels (> 10 transcripts per cell). However, to eliminate the possible error of incomplete occupation of RNA stem loops by MS2 coat protein, we use additional images from time-series experiment with cells under same conditions for tracking appearance of first fluorescent spot in 50 random cells. The fluorescence intensity of each spot of 50 random cells was measured (sala 2015) and cell background correction of spot intensities for calculating mean intensity of single mRNA molecule (Sala., *et al*. 2015a, b). Such way of mRNA measuring avoids misleading detection of single spot as single mRNA molecule in cells with high level of fluorescent spot that may contain several transcripts. Thus, we assure that the number of mRNAs, especially higher than the one, is accurately detected in the single cells.

We use total population experiment, MS-GFP method, for describing the mRNA distribution variation, estimation of stochastic gene expression in different bacterial growth phases. In the present method, we observed mRNA distribution immediate after division as described in method. At the moment after division the cells would have fully produced mRNAs and they are all portioned between the sister cells. Hence we are able to quantify the effect of cell phase. The procedure of measuring the fluorescent signal of single mRNA molecule, described above, ensures that mRNA number from cells with low or high fluorescent levels of spots is accurately scaled to the number of mRNA molecules. And the spots with lower fluorescent intensity represent currently being transcribed mRNA (half mRNA) which are not considered in mRNA population statistics, thus, ensures the accuracy of the results.

Cell-to-cell variability in the bacterial cell phases

In a homogenous population, stochastic gene expression leads to the cell-to-cell variability (η_{tot}) or total noise. It is present not only at a particular gene but, also, exists at the network level [19-22]. Fluctuation of different gene activities as well as the degree of these fluctuations significantly affect the population diversity [23,24]. In the present study, mRNA number distribution is used to measure the cell-to-cell variability and intrinsic/extrinsic components in different bacterial growth phases. At transcription, cell-to-cell variability is approximated by the burstiness factor or Fano factor [2,8]. We find that the relationship between the burstiness factor and the mean mRNA number in the cell growth phases is super-Poissonian (Table 1). It means that examined promoters possess bursty mRNA synthesis. Super-Possonian dynamics of the promoter activities agrees with the previous observations [14,25,26], although, some of the studies observed sub-Poissonian behaviour of the mRNA synthesis [27,28]. Our results may be different from those observed recently due to using of different hosts and promoter systems.

In the fully induced cells, the changes of the cell-to-cell variabilities originated from the *lac* promoter and *lac* variant promoter are similar, over cell growth phases (Figure 1). The similarity is described by the linear regressions. In addition, the linear regressions indicate that the levels of the cell-to-cell variability under control of both promoters increase from Lag to Acceleration phase and decrease from the Exponential to Stationary phase. An increased cell-to-cell variability in the cells at the beginning of the cell cycle may appear due to the fast cellular division rates and the bias in mRNA partitioning between the cells [29]. These findings demonstrate an importance of the cell growth phases for the dynamics of the cell-to-cell variability originated from the fully induced *lac* and *lac* variant promoters.

Figure 1: *An illustration of the cell-to-cell variability (η_{tot}) and noise components measurements. A) Schematic representation of the noise calculations. B) Changes of the cell-to-cell variability of lac promoter and lac variant, over the phase. The linear regressions fitted to the cell-to-cell variability levels from lac and lac variant promoters show that cell-to-cell variability decreases, over the phases.*

Degree of population synchrony

Probabilistic mRNA partitioning between the diving cells is quantified by the degree of synchrony [12]. For estimating the degree of synchrony, we assume that the studied cell populations are a priori asynchronous as they were not imposed to any kind of synchronizing stimulus. Thus, we are able to simulate *in silico* the synchronous behaviour of the cell populations and approximate their asynchronous changes, over the cell phases (Figure 2). The analysis along with the calculated parameters for goodness of fit (R-squared and SSE) reveals that the asynchronic approximation fits better an observed dynamics of the cell-to-cell variability from both *lac* promoter and its variant, than the synchronic simulation (Figure 2A and Table 2). In addition, in the fully induced cells, asynchronous divisions lead to the reduction of cell-to-cell variability compared to the synchronized divisions (Figure 2A). It means that the population with asynchronically dividing cells has higher bias in RNA partitioning than it would be observed in cells dividing synchronically. Nevertheless, decreasing asynchronic approximation indicates that the cell population intends to reduce its diversity over the phases.

Observed synchronous and asynchronous approximations of the cell-to-cell variability originated from the fully induced *lac* and *lac* variant promoters could be determined by the distinct and combinatorial effects of upstream and downstream gene activations. It is known that the downstream activation plays greater role in the dynamics of the cell-to-cell variability and its components over the phases than the upstream [1,2]. To validate an effect of upstream/downstream inductions to stochastic gene expression, we simulate synchronous behavior of the cell-to-cell variability and approximate its asynchronous changes for cells induced from upstream and from downstream. It reveals that cell-to-cell variability is more asynchronous than synchronous, over the bacterial growth phases, in independently induced cells (Figure 2B, Table 2). In addition, asynchronous changes of the cell-to-cell variability decrease over the bacterial cell phases. Decreased asynchronous changes of the cell-to-cell variability indicates reduction of the bias in partitioning of mRNA molecules.

Comparison between the asynchronous approximations of the upstream and downstream gene activations and the asynchronous approximations of the fully induced gene shows that the downstream activation is closer to the full induction condition (Figure 2, Table 2). This information indicates greater role of downstream than upstream activation for the level of the cell-to-cell variability in the fully induced cells.
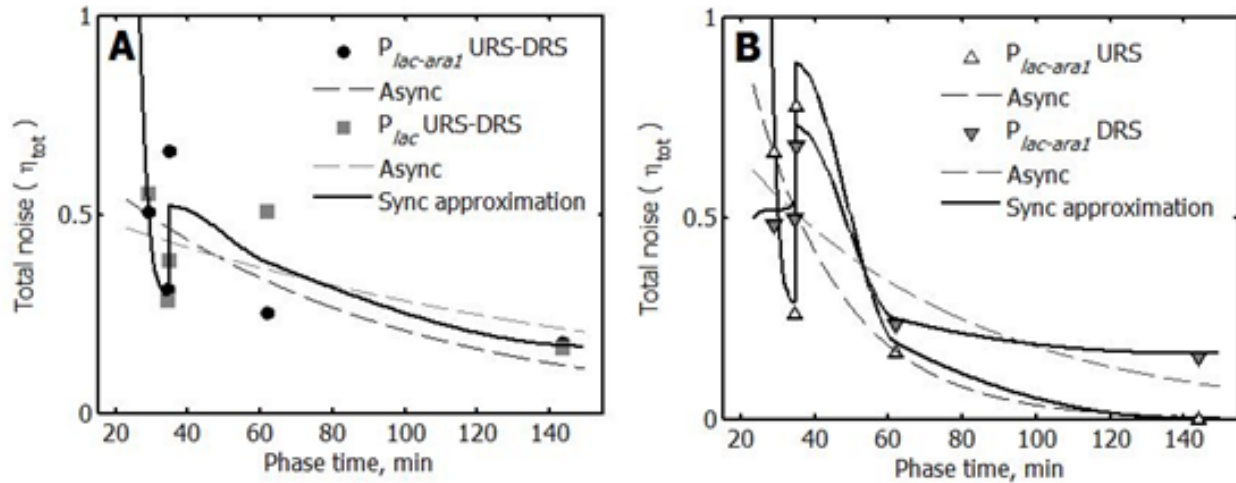
Figure 2: *Approximation of the synchronous and asynchronous changes of the cell to cell diversity. A) Under full induction, lac and lac variant show asynchronous changes of the total noise with respect to the generation time. B) Independent induction reveals an effect of synchrony degree to the upstream and downstream activations of the total noise dynamics with respect to the generation time. Stronger downstream activation changes asynchronically and is closer to the asynchronous behavior of lac and lac variant.*

Parameter	R-square	SSE
Condition		
Async P$_{lac}$ URS/DRS	0.9419	0.0060
Async P$_{lac-ara1}$ URS/DRS	0.8800	0.0187
Async P$_{lac-ara1}$ URS	0.6903	0.1369
Async P$_{lac-ara1}$ DRS	0.7375	0.0478
Sync P$_{lac}$ URS/DRS	0.6093	0.0399
Sync P$_{lac-ara1}$ URS/DRS	0.6042	0.0617
Sync P$_{lac-ara1}$ URS	0.6721	0.1449
Sync P$_{lac-ara1}$ DRS	0.6715	0.1220

Table 2: *Parameters for goodness of fit for the degree of synchrony between cell to cell diversity and cell phases.*

Fluctuation of the noise components under full induction

We find a number of significances, when observing contribution of intrinsic and extrinsic components to the cell-to-cell variability. Firstly, analysis shows that the extrinsic component is generally higher than intrinsic, in the fully induced cells, over the bacterial growth phases. It may be due to the frequency of the repressor proteins binding to the promoter region, fluctuation of the RNAP number or changes in the frequency of magnesium (II) binding to the promoter to form an open complex [2,7,30,31]. Secondly, analysis reveals that both of the noise components are decreasing from the Lag to Stationary phase under the control of *lac* promoter and its variant (Figure 3A). Declining intrinsic component under the control of both promoters indicates decreasing homogeneity of the cell populations over the bacterial growth phases. This cellular behaviour may serve as an adaptive strategy for keeping population stability.

Notably, analysis shows greater fluctuation of intrinsic and extrinsic components under the control of *lac* promoter than under the control of *lac* variant. It may appear due to the native stochasticity of the former promoter. In addition, analysis reveals higher intrinsic than the extrinsic component in a steady, Exponential phase of the fully induced cells. This result agrees with the recent observations of high level of the intrinsic component, determined by the mRNA distribution in an Exponential bacterial growth phase (Figure 3) [30]. Overall, these observations describe how stochastic gene expression fluctuates over the bacterial cell phases in order to keep the population stability.
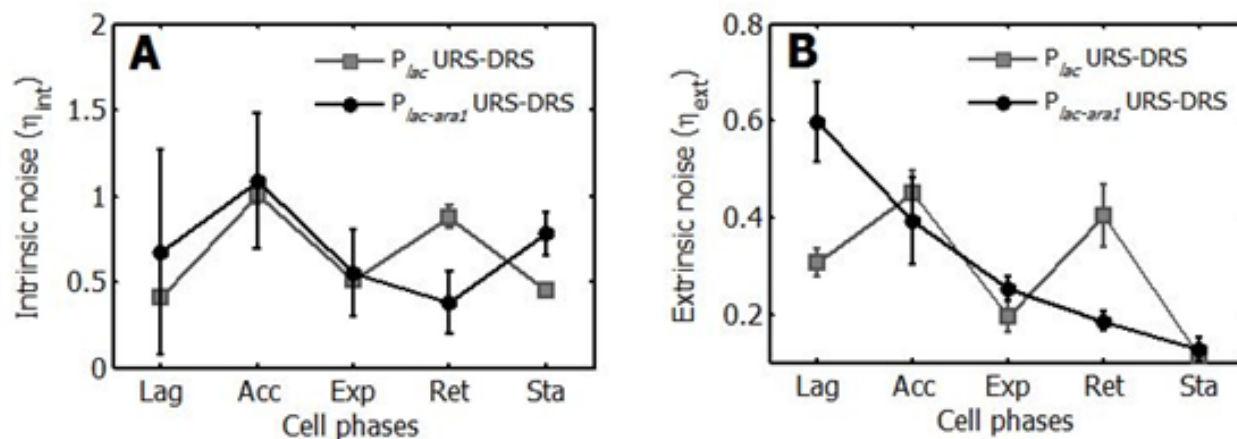

Figure 3: *Phase-wise changes of intrinsic and extrinsic components over the phases under in lac and lac variant. A) Similar behavior of intrinsic component in both constructs shows the level of inherent stochasticity over the phases. B) Decreasing extrinsic component under control of lac and lac variant shows the level of population heterogeneity of population over the phases.*

Role of independent induction in fluctuation of the noise components

We activate *lac* variant promoter independently from the upstream and from the downstream in order to understand the role of the independent inductions in the levels of the intrinsic and extrinsic noise components. As known, upstream promoter activation is faster and has an influence mostly in the Lag and Acceleration phases, whereas downstream activation is slower and is responsible for the greater mRNA distribution in the Retardation and Stationary phases [1,2]. Taking into account these findings, we determine that the upstream gene activation increases intrinsic and extrinsic components in the Lag and Acceleration phases (Figure 4). Yet, observed increased levels of both noise components originated from the upstream activated gene are not significantly higher than the levels of both noise components originated from the downstream activated genes. In the Exponential phase, cells are mainly ruled by the down-regulation as it causes the noise components to increase (Figure 4). In the Retardation and Stationary bacterial growth phases, increased intrinsic and extrinsic components are observed under downstream gene activation (Figure 4). These observations indicate an importance of the cellular phases and types of the gene activation in the levels of the intrinsic/extrinsic noise components.

As known, stochastic gene expression is possible only when the increase of the cell-to-cell variability is accompanied by the increase of the mean mRNA number [1,2]. High rate of the mRNA synthesis would result in the increased extrinsic component and decreased intrinsic component. Though, both noise components are decreasing over the bacterial growth phases, we find that extrinsic component is high when promoter is down-regulated. It indicates that the downstream activation type contributes stronger to the fluctuation of the gene expression noise in the fully induced cells, than the upstream. The p-values from the Kolmogorov-Smirnov test support greater down-regulation influence to the cell-to-cell variability in the *lac* and *lac* variant (0.9789 and 0.9996, respectively) than the up-regulation (0.6974 and 0.4000, respectively).
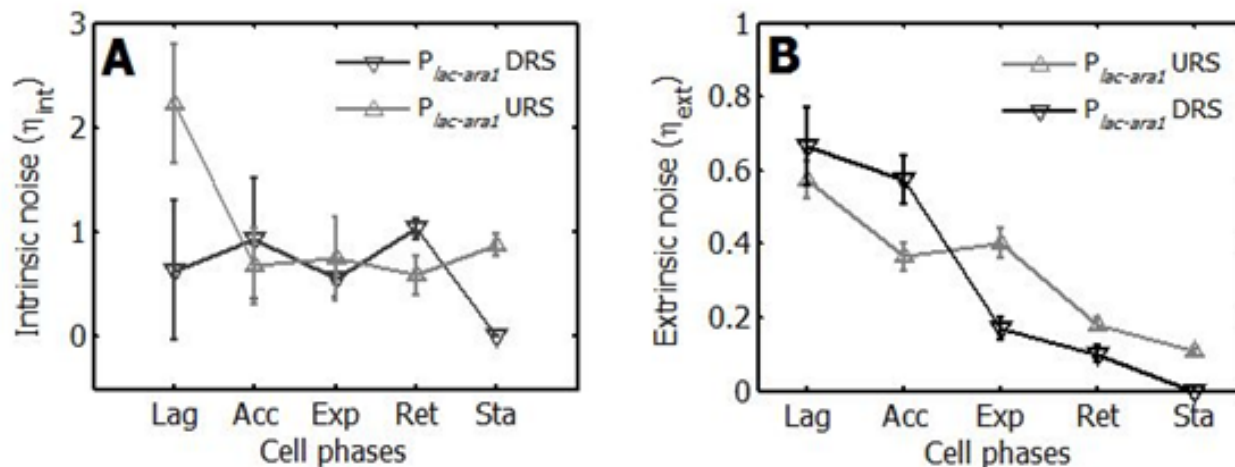
Figure 4: *Fluctuation of intrinsic and extrinsic components under upstream and downstream activations. A) Intrinsic component is generally higher under downstream than upstream activation mode that describes the level of population heterogeneity over the phases. B) Decreasing extrinsic component in bacterial cells indicates reducing of cell to cell diversity over the phases.*

Overall, our results demonstrate how the cell-to-cell variability fluctuates in cells, over the bacterial growth phases. The description of such dynamics is accompanied by the approximation of the degree of synchrony of the cell population, the levels of both noise components and an effect of gene activation conditions. Obtained knowledge shed light on how cellular variability might be profitable for the cells in order to reduce the population diversity in natural, asynchronous conditions. In addition, the results uncover the origins of the noise fluctuations in the fully induced cells at a single cell level and evaluate an effect of the independent activations onto the level of the gene expression noise. The present work serves, also, as an extension of the previously available data about levels of the cell-to-cell variability and both noise components at transcription, in the *E. coli* cells [2]. In perspective, these results can be expanded to construct the circuits with controllable gene expression and to investigate the modulation of the gene expression noise in different cell types, at transcription.

Discussion

The current study presents an approach for evaluating the gene expression noise of the *E. coli* cells, at transcription. An approach uses an experimentally measured mRNA number per cell for the calculation of the cell-to-cell variability and intrinsic/extrinsic components, over the bacterial growth phases. An accuracy of quantitative mRNA measurements is ensured by the known phase generation times of the cells containing fully induced *lac* promoter [1,2]. In addition to the traditional full induction scheme, in which cells are exposed to the inducer from both, upstream and downstream, regions, we also evaluate an effect of the independent, upstream and downstream, inductions.

The results show that population of the fully induced cells displays significant heterogeneity over the cell phases. It is determined mainly by the extrinsic component. High level of the extrinsic component may arise from the degradation of mRNAs, uneven mRNA partitioning between the cells or random births and deaths of individual molecules, small fluctuation of which would create large variation between the cells [32]. Down-regulation of the fully induced gene plays a greater role in the determination of the levels of the noise components than up-regulation. This observation opposes with the previously reported stronger effect of the upstream than the downstream regulation to the mRNA number distribution [33]. Opposite to the previous knowledge results may be due to the different promoter types used in both experiments or different cell phases taken into consideration. Still, our data indicates an importance of the activation modes and cell growth phases in determination of population behaviour.

Overall, we present the study of how an information about the mRNA number distribution can be used to determine stochastic gene expression in the single cells, over bacterial growth phases. Although, an attempts to perform such measurements have been undertaken, they possessed a number of limitations [34]. Firstly, one of the studies hypothesized that extrinsic and intrinsic components could be quantified by smFISH [35]. Our data provides an alternative way of calculating the noise components by using MS2-GFP method. The MS2-GFP allows for quantification of the mRNA number distribution per cell, in a real-time, *in vivo*. Secondly, evaluations of the stochastic gene expression in the bacterial cells was done mainly at a particular cell phase, e.g. Exponential [2,11]. Our analysis expands these observations by describing dynamics of the stochastic gene expression with respect to the bacterial growth phases. In perspective, presented way of quantification of stochastic gene expression may be used to study stochasticity of specific genes at the network level.

Highlights

- Quantification of noise in transcriptional burst during cell cycle.
- The cell-to-cell variability determined by higher extrinsic than intrinsic noise.
- Extrinsic and intrinsic noise decrease over the phases.
- Asynchronous mode is responsible for high extrinsic noise during the cell cycle.

Supporting Information

The supporting file contain table S1, figure 1 and movie 1.

Author Contributions

OA and MK analyzed the data and interpreted the results. MK executed the experiments. OA and MK wrote the paper. All the authors read and agreed the final version of manuscript. MK designed the research work. We do not have any conflict of interest.

Acknowledgement

Funding

Competing Financial Interests

The authors declare no competing financial interests.

Bibliography

1. Sala A., *et al*. "Origins of Transcriptional Transition : Balance between Upstream and Downstream Gene Regulatory Sequences". *MBio* 6.1 (2015): 1-9.

2. Sala A., *et al*. "Transcription closed and open complex dynamics studies reveal balance between genetic determinants and co-factors". *Physical Biology* 12.3 (2015): 036003.

3. J Monod. "The Growth of Bacterial Cultures". *Annual Review of Microbiology* 3 (1949): 371-394.

4. Viswanathan A., *et al*. "Phase-dependent dynamics of the lac promoter under nutrient stress". *Research in Microbiology* 167.6 (2016): 451-461.

5. EM Ozbudak., *et al*. "Regulation of noise in the expression of a single gene". *Nature Genetics* 31.1 (2002): 69-73.

6.   S Ramsey, *et al*. "Transcriptional noise and cellular heterogeneity in mammalian macrophages". *Philosophical Transactions of the Royal Society B: Biological Sciences* 361.1467 (2006): 495-506.

7.   PS Swain., *et al*. "Intrinsic and extrinsic contributions to stochasticity in gene expression". *Proceedings of the National Academy of Sciences of the United States of America* 99.20 (2002): 12795-12800.

8.   LH So., *et al*. "Golding, General properties of transcriptional time series in Escherichia coli". *Nature Genetics* 43.6 (2011): 554-560.

9.   DL Jones., *et al*. "Promoter architecture dictates cell-to-cell variability in gene expression". *Science* 346.6216 (2014): 1533-1536.

10. Anufrieva., *et al*. "Real-time observation of bacterial gene expression noise". *Nano Communication Networks* 8 (2016): 68-75.

11. TC Barnett., *et al*. "Role of mRNA stability in growth phase regulation of gene expression in the group A streptococcus". *Journal of Bacteriology* 189.5 (2007): 1866-1873.

12. Lloyd-Price., *et al*. "Probabilistic RNA partitioning generates transient increases in the normalized variance of RNA numbers in synchronized populations of Escherichia coli". *Molecular BioSystems* 8.2 (2012): 565-571.

13. Golding I and EC Cox. "RNA dynamics in live Escherichia coli cells". *Proceedings of the National Academy of Sciences of the United States of America* 101.31 (2004): 11310-11315.

14. Golding J Paulsson., *et al*. "Real-time kinetics of gene activity in individual bacteria". *Cell* 123.6 (2005): 1025-1036.

15. J Mäkelä., *et al*. "Stochastic sequence-level model of coupled transcription and translation in prokaryotes". *BMC Bioinformatics* 12 (2011): 121.

16. J Mäkelä., *et al*. "In vivo single-molecule kinetics of activation and subsequent activity of the arabinose promoter". *Nucleic Acids Research* 41.13 (2013): 6544-6552.

17. D Volfson., *et al*. "Origins of extrinsic variability in eukaryotic gene expression". *Nature* 439.7078 (2006): 861-864.

18. S Tyagi. "Imaging intracellular RNA distribution and dynamics in living cells". *Nature Methods* 6.5 (2009): 331-338.

19. MB Elowitz., *et al*. "Stochastic gene expression in a single cell". *Science* 297.5584 (2002): 1183-1186.

20. Raj A and van Oudenaarden. "Nature, nurture, or chance: stochastic gene expression and its consequences". *Cell* 135.2 (2008): 216-226.

21. M Acar., *et al*. "Stochastic switching as a survival strategy in fluctuating environments". *Nature Genetics* 40 (2008): 471-475.

22. WJ Blake., *et al*. "Noise in eukaryotic gene expression". *Nature* 422.6932 (2003): 633-637.

23. T Mora and AM Walczak. "Effect of phenotypic selection on stochastic gene expression". *Journal of Physical Chemistry B* 117.42 (2013): 13194-13205.

24. D Zhuravel., *et al*. "Phenotypic impact of regulatory noise in cellular stress-response pathways". *Systems and Synthetic Biology* 4.2 (2010): 105-116.

25. JR Chubb and TB Liverpool. "Bursts and pulses: Insights from single cell studies into transcriptional mechanisms". *Current Opinion in Genetics and Development* 20.5 (2010): 478-484.

26. A Raj., *et al.* "Stochastic mRNA synthesis in mammalian cells". *PLoS Biology* 4.10 (2006): 1707-1719.

27. Sanchez., *et al.* "Regulation of noise in gene expression". *Annual Review of Biophysics* 42 (2013): 469-491.

28. SJ Gandhi., *et al.* "Transcription of functionally related constitutive genes is not coordinated". *Nature Structural and Molecular Biology* 18.1 (2011): 27-34.

29. AB Muthukrishnan., *et al.* "Dynamics of transcription driven by the tetA promoter, one event at a time, in live Escherichia coli cells". *Nucleic Acids Research* 40.17 (2012): 8472-8483.

30. D Madar., *et al.* "Promoter activity dynamics in the lag phase of Escherichia coli". *BMC Systems Biology* 7 (2013): 136.

31. M Kandhavelu., *et al.* "Effects of Mg (2+): on in vivo transcriptional dynamics of the lar promoter". *Biosystems* 107.3 (2012): 129-134.

32. D Huh and J Paulsson. "Non-genetic heterogeneity from stochastic partitioning at cell division". *Nature Genetics* 43 (2011): 95-100.

33. G Fritz., *et al.* "Single cell kinetics of phenotypic switching in the arabinose utilization system of E. coli". *PLoS One* 9.2 (2014): e89532.

34. N Rosenfeld., *et al.* "Gene regulation at the single-cell level". *Science* 307.5717 (2005): 1962-1965.

35. SO Skinner., *et al.* "Measuring mRNA copy number in individual Escherichia coli cells using single-molecule fluorescent in situ hybridization". *Nature Protocols* 8.6 (2013): 1100-1113.