BIOINFORMATICS ORIGINAL PAPER Vol. 27 no. 19 2011, pages 2714–2720
doi:10.1093/bioinformatics/btr471

Systems biology Advance Access publication August 10, 2011

Automatic detection of changes in the dynamics of delayed stochastic gene networks and *in vivo* production of RNA molecules in *Escherichia coli*

Jarno Mäkelä[1], Heikki Huttunen[1], Meenakshisundaram Kandhavelu[1], Olli Yli-Harja[1,2] and Andre S. Ribeiro[1,*]

[1]Computational Systems Biology Research Group, Department of Signal Processing, Tampere University of Technology, FI-33101 Tampere, Finland and [2]Institute for Systems Biology, Seattle, WA 98103-8904, USA

Associate Editor: Ivo Hofacker

ABSTRACT

Motivation: Production and degradation of RNA and proteins are stochastic processes, difficulting the distinction between spurious fluctuations in their numbers and changes in the dynamics of a genetic circuit. An accurate method of change detection is key to analyze plasticity and robustness of stochastic genetic circuits.

Results: We use automatic change point detection methods to detect non-spurious changes in the dynamics of delayed stochastic models of gene networks at run time. We test the methods in detecting changes in mean and noise of protein numbers, and in the switching frequency of a genetic switch. We also detect changes, following genes' silencing, in the dynamics of a model of the core gene regulatory network of *Saccharomyces cerevisiae* with 328 genes. Finally, from images, we determine when RNA molecules tagged with fluorescent proteins are first produced in *Escherichia coli*. Provided prior knowledge on the time scale of the changes, the methods detect them accurately and are robust to fluctuations in protein and RNA levels.

Availability: Simulator: www.cs.tut.fi/∼sanchesr/SGN/SGNSim.html

Contact: andre.ribeiro@tut.fi

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on May 27, 2011; revised on August 3, 2011; accepted on August 5, 2011

1 INTRODUCTION

Gene regulatory networks (GRNs) are stochastic. However, their behavior is, to some extent, robust, e.g. when responding to environmental changes. The behavior is determined by the structure of the genetic circuits. Thus, when structural changes occur, in many cases there are changes in the dynamics of RNA and protein numbers of some genes. Some such structural changes (e.g. a mutation) can be rare, occurring once in a cell's lifetime. It is thus important to develop robust methods for detecting permanent changes in the dynamics of genetic circuits, and distinguish these from spurious fluctuations in RNA and protein numbers.

We apply automatic change detection methods to simulated and real gene expression data, to recognize candidate change points in RNA and protein numbers' dynamics. Automatic detection of change points is the discovering of points in time where the properties of the time series change. Earliest approaches were based on the Behrens–Fisher problem, a statistical hypothesis test of equal means (Belloni and Didier, 2008; Fisher, 1939). A widely used approximation to solve this problem is the Welch's *t*-test. However, these approaches assume normal distributions, typically causing them to be too sensitive for heavy-tailed distributions. The dynamics of RNA and protein production are usually not normal like, especially if a structural change occurs in the GRN during the observations.

We use two recent change point detection methods, namely, the density ratio method and the kernel change point analysis (Harchaoui *et al.*, 2009; Kawahara and Sugiyama, 2009). Our choice is based on their reported good performance compared with alternative methods. These methods use different approaches: the density ratio method has its roots in statistics and density estimation, while the kernel change point analysis is based on the theory of kernel machines and classification. In our understanding, they represent state-of-the-art approaches to the problem.

To assess the accuracy of the methods, knowing the ground truth signal is needed. For this, we require realistic simulations of RNA and protein expression dynamics. The dynamics of the models ought to be as realistic as possible so as to mimic accurately the temporal dynamics of RNA and protein numbers in real cells.

Recently, a delayed stochastic modeling strategy of gene expression and GRNs was proposed (Ribeiro *et al.*, 2006). It is based on the delayed stochastic simulation algorithm (delayed SSA) (Zhu *et al.*, 2007), and thus it accounts for the key dynamical features of real GRNs, namely, the stochasticity of the chemical kinetics (Arkin *et al.*, 1998), and the duration of events such as the promoter complex formation (McClure, 1980) and transcription elongation (Zhu *et al.*, 2007). This modeling strategy was shown to match the dynamics of RNA and protein production at the single molecule level (Yu *et al.*, 2006; Zhu *et al.*, 2007). Delayed stochastic models of GRNs can be simulated by SGNSim (Ribeiro and Lloyd-Price, 2007), which also allows introducing changes in the structure of the GRN at run time, needed to test the change point detection methods.

We apply and test the accuracy of the automatic detection of change points methods to model GRNs subject to a permanent

*To whom correspondence should be addressed.

change at run time in mean level of a protein, noise strength of a protein's time series and in the frequency of switching of a genetic switch. Further, to verify the applicability of the methods to large-scale clusters of interconnected genes, we test the ability to detect a change in the dynamics of a model GRN with 328 genes, subject to the silencing of a randomly selected gene at run time. Finally, we apply the methods to determine when new RNA molecules are produced, from our temporal measurements by confocal microscopy of RNA tagged with MS2d-GFP in *Escherichia coli*.

2 METHODS

2.1 *In vivo* detection of RNA molecules in *E.coli*

RNA detection and quantification *in vivo* in *E. coli* cells DH5α-PRO uses the ability of the coat protein of bacteriophage MS2 to tightly bind specific RNA sequences (Peabody and Lim, 1996). High-resolution detection of single RNA transcripts with 96 tandem repeats of MS2 binding sites in *E.coli* is possible by using dimeric MS2 fused to GFP (MS2d-GFP fusion protein) as a detection tag (Golding *et al.*, 2005). The method uses two genetic constructs. The first is a medium-copy vector expressing the MS2d-GFP fusion protein, whose promoter (P_{tetO}) is regulated by tetracycline repressor. The second is a single copy F-based vector, with a $P_{lac/ara}$ promoter controlling production of the transcript target, specifically mRFP1 followed by a 96 MS2 binding site array. Constructs were generously provided by I. Golding (University of Illinois). Experimental procedures of induction of the target RNA, confocal microscopy and cell and RNA spots segmentation from images are described in Supplementary Material.

2.2 Models of GRNs

We follow the modeling strategy of delayed stochastic GRNs proposed in Ribeiro *et al.* (2006). The models are implemented in the simulator SGNSim (Ribeiro and Lloyd-Price, 2007), and the dynamics is based on the delayed SSA (Zhu *et al.*, 2007), that unlike the SSA (Gillespie, 1977), uses a waiting list to store delayed output events. The algorithm of delayed SSA is presented in Supplementary Material. Delayed reactions are represented as: $A \rightarrow B + C(\tau_1) + D(\tau_2)$. In this reaction, B is instantaneously produced, while C and D are placed on the waitlist until they are released, after τ_1 and τ_2 seconds, respectively. This strategy accounts for the stochastic nature of chemical reactions and for the fact that transcription and translation are multistep processes that take non-negligible time to be completed once initiated. The strategy was validated in Zhu *et al.* (2007) by matching temporal measurements of expression of individual proteins (Yu *et al.*, 2006).

We implement four model GRNs, named models 1, 2, 3 and 4. Models 1 and 2 are identical, and consist of a two-gene network, where Gene 1 represses Gene 2. These models differ in the change at run time. In Model 1, mean protein levels change at run time, while in Model 2 it is the strength of fluctuations in protein levels that changes. Model 3 is a genetic switch whose switching frequency is changed at run time.

We also test if changes in the dynamics of larger GRNs can be detected. Gene networks consist of hundreds to thousands of genes, usually organized in clusters of dozens to hundreds, that are involved in specific tasks in development, metabolism, etc. Changes known to occur in the dynamics of these networks may be caused by mutations, deletions or duplications, or as a response to external signals or stress. Usually, such events cause one to a few genes, along with several of its neighbor genes, to alter the expression level (e.g. from high to low). The models and how the changes at run time are implemented are described in Supplementary Material.

To test if the algorithms of change point detection are successful for large genetic circuits, we apply them to a model of the core gene network of *Saccharomyces cerevisiae* inferred from microarray measurements following gene deletions and overexpressions (Chowdhury *et al.*, 2010). This network contains 328 genes. Inferred connections were verified by gene enrichment. The perturbations consist of selecting genes randomly (see Supplementary Material) and subject them to silencing at run time, one per simulation.

2.3 Methods of change point detection

Formally, the problem of change point detection can be stated as follows. Given a multidimensional time series $\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_N \in \mathrm{R}^n$, which time points K represent change in some sense, given the data samples in the M-point backward window $\boldsymbol{X}_B = (\boldsymbol{x}_{K-M}, \dots, \boldsymbol{x}_{K-1})$ and the M-point forward window $\boldsymbol{X}_F = (\boldsymbol{x}_{K+1}, \dots, \boldsymbol{x}_{K+M})$. To define the dissimilarity of the two windows one can pose the question as a hypothesis testing problem:

$$\begin{cases} H_0 : p_{\boldsymbol{X_F}}(\boldsymbol{x}) = p_{\boldsymbol{X_B}}(\boldsymbol{x}) \\ H_1 : p_{\boldsymbol{X_F}}(\boldsymbol{x}) \neq p_{\boldsymbol{X_B}}(\boldsymbol{x}) \end{cases} \quad (1)$$

where $p_{\boldsymbol{X_F}}(\boldsymbol{x})$ and $p_{\boldsymbol{X_B}}(\boldsymbol{x})$ denote the probability density functions of the forward and backward windows, respectively.

Detection in non-parametric cases is still, in general, an open problem. We apply two recent change point detection methods proposed for the non-parametric case. Namely, we apply a direct density ratio test (uLSIF) (Kawahara and Sugiyama, 2009) and a kernel change point analysis method (KCpA) (Harchaoui *et al.*, 2009), described in Supplementary Material.

3 RESULTS AND DISCUSSION

3.1 Selecting a proper window size

Most parameters of change point detection algorithms can be inferred from the data by cross-validation. However, the detection window length cannot be determined from training data, since changes appear at multiple scales. Thus, the algorithm cannot decide which time scale is biologically relevant. Below, we choose the window size from knowledge of the scale of the biological phenomena studied. Before, we study the performance of the detectors for a wide range of window sizes to search for differences in robustness and to compare performances with larger number of test cases.

We first apply KCpA and uLSIF algorithms to three models. We change at run time mean protein levels in Model 1, noise in protein numbers in Model 2 and frequency of switching between noisy attractors in Model 3 (Section 2.2). Examples of the time series of protein numbers in these models, prior and after a change are shown in Supplementary Material. The performance is assessed by the area under the receiver operating characteristics (ROC) curve or the AUC criterion (Kay, 1998). The ROC curve and the corresponding AUC value were based on the specificity–sensitivity coordinates obtained by varying the detection threshold.

To improve the detection performance, the data are appended by auxiliary variables. This attempts to cast the change in, e.g. variance into a change in the mean of the auxiliary variable. For example, in Model 2, the change is in the degree of the fluctuations in protein numbers. One can convert this into a change in 'local variance' (that is, variance within a small time interval). Appending the local variance estimates for both proteins to mean levels significantly improves the detection of changes in noise in protein numbers.

We also include auxiliary variables in Model 3. The added feature is the average absolute difference of consecutive samples. This improves performance because of the nature of the change (in switching frequency) and because the levels of the two protein are dynamically coupled. The change in switching frequency from low to high is reflected in the difference between consecutive samples of $|P_1 - P_2|$ (similarly one could have used the sum of P_1 and P_2).

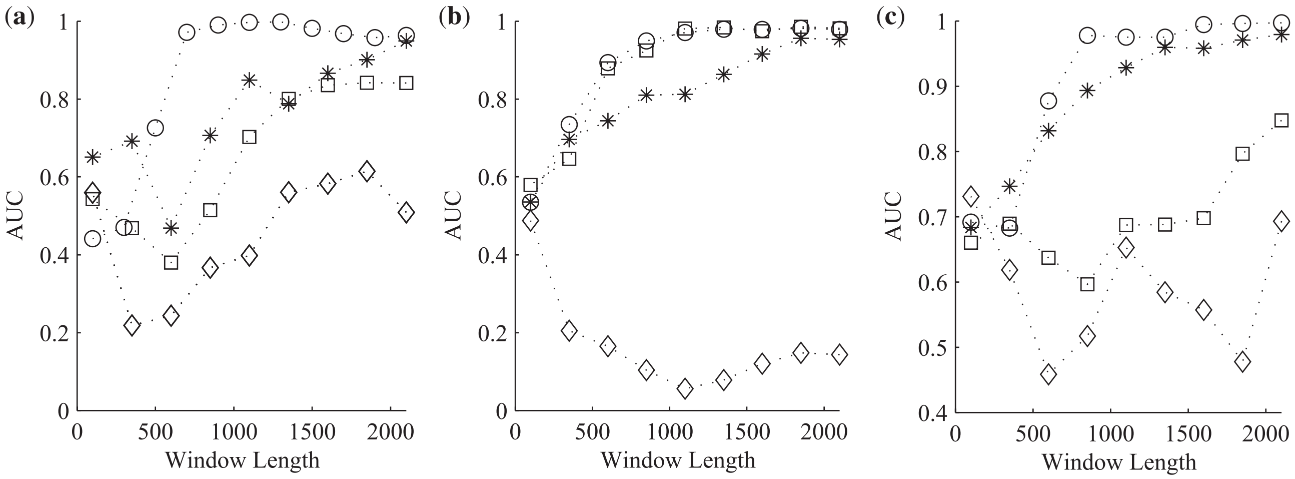

Fig. 1. Effect of window length on change point detection methods. KCpA (linear kernel) (open circle), KCpA (polynomial kernel) (open square), KCpA (RBF kernel) (asterisk) and uLSIF (diamond) detection results for each model. Horizontal axis is the size of forward and backward windows (window length size 1000 uses 1000 + 1000 = 2000 samples for detection). Vertical axis is the area under the ROC curves (AUC) for different window lengths when applied to Models 1 (**a**), 2 (**b**) and 3 (**c**).

The performance of detection of change points for various window sizes is summarized in Figure 1. For all models, KCpA seems more effective (higher AUC values for most window lengths). Also, uLSIF is more sensitive to the choice of parameters, but their manual adjustment can improve the performance to comparable levels. However, the cross-validated automatic parameter setup does not provide good results. We acknowledge that the *ad hoc* addition of auxiliary variables may favor KCpA over uLSIF. Without it, the two detectors exhibited similar (poor) performances, although for Model 2, uLSIF performed slightly better. This is likely because uLSIF models more extensively the characteristics of the distribution, while KCpA assumes Gaussian densities with equal covariances. However, for our practical purposes, neither method was robust enough, so we decided to improve the detection by adding auxiliary variables.

In general, widening the window improves performance. The best robustness of detection is obtained by KCpA with the linear kernel. This is not surprising, since simple linear models typically exhibit small variance (e.g. repeated experiments tend to have similar results). The drawback is a high bias if the model is not complex enough. However, according to Figure 1, the linear model with KCpA seems sufficient for our data.

3.2 Detection of change points in genetic circuits

Examples of the application of KCpA with linear kernel to one realization of the time series of each model are shown in Figure 2a–c, and the results for uLSIF in Figure 2d–f. The inputs for change point detection are the time series in Supplementary Figure S3a–S3c. In all cases, the true change point occurs at midpoint of the time series. KCpA outperforms uLSIF, as it is less sensitive to spurious, transient changes. We used in all cases a window length of 1100 samples as it enhances detection when compared with smaller lengths, and further increases in length did not improve the detection significantly (Fig. 1). This length corresponds to 10^5 s simulation time, allowing spurious transient fluctuations to be recognized as such. The length is realistic given the time scales of transient fluctuations in protein numbers in bacteria and eukaryotic cells. In the models, the effects of fluctuations in protein numbers last for long periods of time ($10^3 - 10^4$ s). Therefore, it is expected the need of using a window size longer by one order of magnitude. In cells, since proteins have lifetimes of the order of tenths of minutes, fluctuations last by a similar order of magnitude. For example, oscillations in protein P53 numbers in Human cells have a period of 5 h (Geva-Zatorsky *et al.*, 2006) and the period of oscillation of a repressilator engineered in *E.coli* is $\sim 10^4$ s (Elowitz and Leibler, 2000).

For the KCpA method (Fig. 2a–c), we performed multiple tests on independent time series, all of which with identical initial conditions. In the figures, we show the results of three of such independent runs for the three models. The aim is to test if KCpA is robust to the stochasticity of the time series, which will cause different spurious, transient fluctuations in each independent simulation. Visibly, the algorithm is highly robust in the first two models (Fig. 2a and b), in the sense that the change is accurately detected in all independent simulations and at the moment following occurrence of the structural change.

The results for Model 3 are more complex. Namely, while in all cases the change is detected (as depicted in Fig. 2c), the detection takes place at different points in time following the change. This is explained by the nature of the change. What changes is the frequency of switching between noisy attractors. For such a change to be detected (even by a human observer), switches between the two protein levels must take place (so that the algorithm can 'measure' the frequency). In some simulations, switches will take place shortly after the change, while in others it takes longer time. The duration of switches follows approximately an exponential distribution (Ribeiro and Lloyd-Price, 2007) causing the interval between switches to vary widely. Due to that, it is expected that the algorithm, from different runs, will detect the change in the dynamics in different moments following the structural change in the genetic circuit. Relevantly,

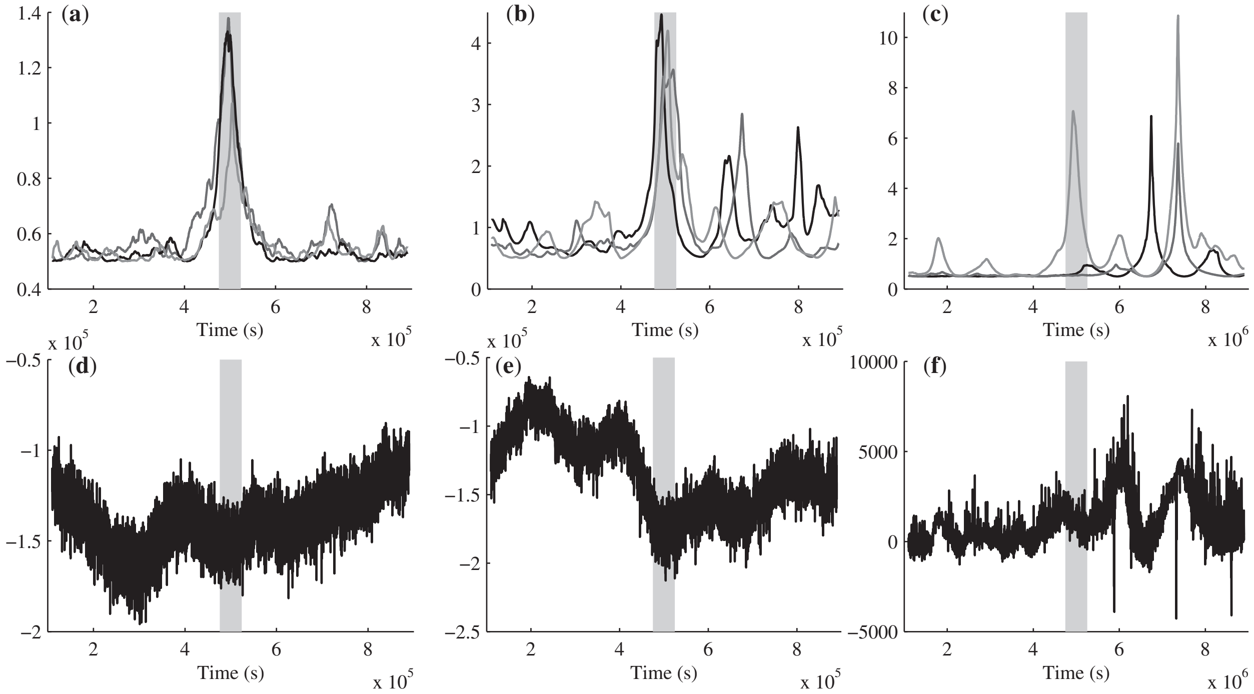

Fig. 2. Results of change point detection for KCpA (linear kernel) in models 1 (**a**), 2 (**b**) and 3 (**c**) and uLSIF in models 1 (**d**), 2 (**e**) and 3 (**f**). Vertical axis are the KCpA and uLSIF indicator outputs. In all cases, the window length is 1100 samples. Ground truth (in gray) used to compute the ROC curves. uLSIF parameters were selected by 10-fold cross-validation. In (a), (b) and (c) the results of applying KCpA to the time series from three independent runs are shown for assessing robustness to the stochastic fluctuations in protein numbers.

even for this case, in all runs KCpA detected the change, while at different moments.

We now compare the results of KCpA and uLSIF in each model. Figure 2a and d illustrate the change point detection results for Model 1. KCpA largely outperforms uLSIF. As mentioned, KCpA determines accurately the exact moment when the protein levels start changing for this model. After, the two protein levels only fluctuate around a mean level, and KCpA does not detect any significant changes. We conclude that KCpA is appropriate to detect changes in mean expression levels in highly stochastic time series, since fluctuations due to noise in the chemical kinetics are not confused with the change in mean expression levels.

Next, we test the ability of detecting changes in the degree of fluctuations in a protein's level (Model 2). In our example, the noise strength in protein numbers changes from 0.63 to 0.73, as measured by the square of the coefficient of variation (SD over the mean). The results of the detection are shown in Figure 2b and e. Again KCpA outperforms uLSIF, indicating the true change shortly after the midpoint of the time series. In comparison to the first case, the results are not as clear as there are a few false matches after the true change point. Nevertheless, the moment at which the structure changed was correctly identified. In addition, the highest peak occurs at the true change point. Thus, we conclude that KCpA detects changes in the noise strength of temporal expression levels even from time series of protein numbers that are highly stochastic both before and after the change.

We now compare KCpA and uLSIF in detecting a change in the frequency of switching of a genetic switch (Model 3). The decrease in switching frequency is due to weaker fluctuations in the protein numbers and leads to a slight increase in mean number of proteins since the decrease in fluctuations is not symmetric in relation to the mean protein level (see Supplementary Material). Thus, there are two changes, in mean and in fluctuations, which also have different time scales to be completed once initiated. The results of the detection are shown in Figure 2c and f. In this case, both the uLSIF and the KCpA have a poor performance. The KCpA does detect the true change point from one of the realizations of the data, but this might only be a coincidence. However, both methods are able to detect the strong changes in the mean levels (due to the switching dynamics) that occur at time 6×10^6 and 7.5×10^6 s (see Supplementary Fig. S1c). Thus, we conclude that for networks with switching dynamics, the detection of change of frequency requires complex analysis of the results, namely, there actually was a detection of a change point (but not of the frequency of switching), from which one can infer that the structure of the switch changed at run time. From this point of view, the uLSIF detector outperforms KCpA in this case.

3.3 Detecting change points in a complex genetic circuit

We simulate the dynamics of models from a core genetic network model of *S.cerevisiae* with 328 genes (Chowdhury *et al.*, 2010). Both topology as well as genes' transfer functions were inferred from microarray measurements following deletions and overexpressions of a gene, in optimal environmental conditions. Perhaps due to this, it was observed that unless the inferred model network is perturbed,

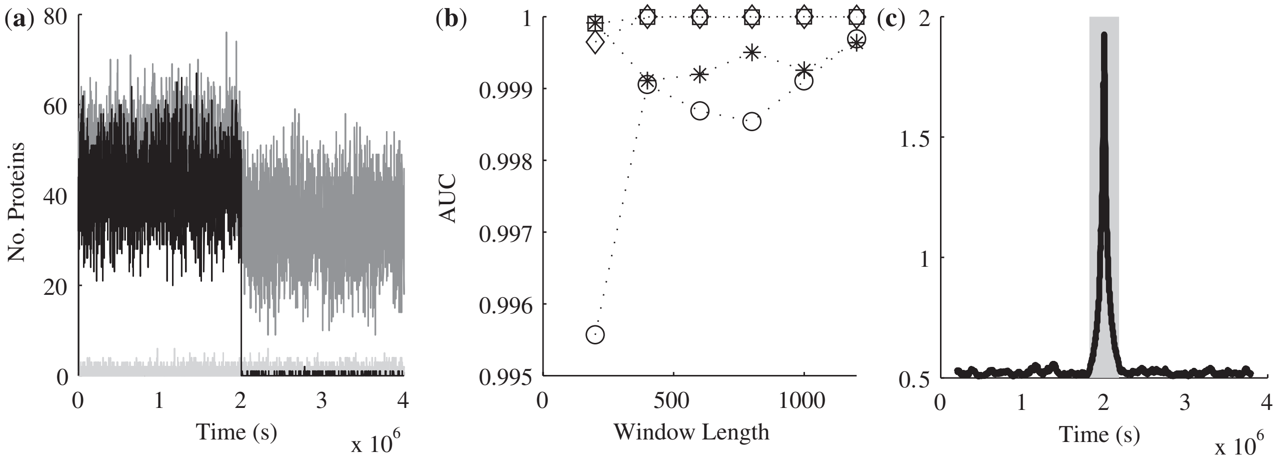

Fig. 3. (**a**) Example of protein numbers of 3 out of the 328 genes from a simulation (two were affected by the perturbation, one was not). (**b**) Results of detection. The vertical axis is the AUCs for all window sizes and all genes (the symbols 'open circle', 'asterisk', 'square' and 'diamond' are results when silencing different genes). Window size is 200 in all cases. (**c**) Example of detection results of KCpA for window size 200. Vertical axis is KCpA indicator output.

its dynamics remain relatively stable (Chowdhury *et al.*, 2010) even when modeled with the stochastic modeling strategy.

Since the topology is highly clustered and the mean connectivity $\simeq$5, and this was inferred from observing how many genes' expression level was affected by the deletion or overexpression of another gene (Chowdhury *et al.*, 2010), we can expect that when perturbing the model network by gene silencing, only a few genes' expression level will be affected, on average.

Due to that, and given our prior knowledge of the topology of this network, it is possible, for simplicity, to provide the algorithm with the data comprising only the perturbed gene and its near neighbors (between 20 and 30 genes' expression levels, selected based on smallest path length to the randomly perturbed node). In general, adding non-informative data can only decrease the performance of any detection algorithm. Therefore, attempting to detect from the expression of all 328 genes, is likely harder than when using a subset of genes. Nevertheless, this decrease ought to be minimal in this case, since even most of the selected genes were found to also be non-informative (no clear change was observed in the time series of protein numbers following the perturbation).

The dynamics is simulated for a period of time and one gene, chosen at random (see Supplementary Material), is silenced at midpoint. Examples of the protein numbers of a few genes of the network are shown in Figure 3a. Note that the mean expression levels (30–50 when gene expression is active, and close to 0 when repressed) are within realistic intervals for *S.cerevisiae* (Bar-Even *et al.*, 2006). As expected, we observed that following a perturbation, only a small fraction of genes was dynamically perturbed.

Figure 3b shows the results of detection by KCpA for varying window sizes. We use only KCpA as it exhibited the best results so far. The detection is highly accurate and robust to varying window size. In all, we ran four experiments. In each case, a different gene from the GRN was silenced. In two cases, the silenced gene was included in the simulated measurement data, while for the two other cases, we only included measurements of non-silenced genes. As one can see, the AUC's are in all case very close to 1, and it seems irrelevant whether the particular silenced gene is included or not. The accuracy in the four cases differs slightly. This is expected, since different genes have different number of outputs and thus its silencing will have differing range of effects in the GRN's dynamics. Figure 3c shows the detection results of KCpA in one case.

3.4 Time series of images of *E.coli* cells expressing RNA target for MS2d-GFP

We now apply the methods to detect changes in the number of tagged RNA molecules in cells over time from series of images taken by confocal microscopy of *E.coli* cells expressing RNA target for MS2d-GFP. The time series of total fluorescence intensity of MS2d-GFP-RNA spots in a cell is shown in Figure 4a (extracted from the images of the cell shown in Supplementary Material). Note that some time points are missing due to the microscope getting out-of-focus. From the images, one can see that in that period the objects become blurred and determining the appearance of RNA molecules becomes impossible even by a human observer. Inclusion of these outliers in the data would only result in detection of the problems of robustness of the images acquisition process.

Instead of compensating this by acquiring new data or by interpolating the existing data, we marked the out-of-focus time points as missing data and apply the detection algorithms to the remaining points. After all, missing data are common in biological measurements, and any practical algorithm should tolerate it. In our case, missing data are treated in a natural manner: both methods define a forward and a backward window, but do not restrict them to be of equal size. Thus, their use with missing data are straightforward, and it is interesting to their response to missing data.

Figure 4b and c show the results of the detection. From Figure 4a, it is visible that the fluorescence intensity of RNA-MS2d-GFP spots over time is a very noisy signal, although once the target RNA is tagged by MS2d-GFP it does not degrade (Golding *et al.*, 2005). This implies that decreases in fluorescence are not due to RNA degradation. One cause is the movements in and out of focus

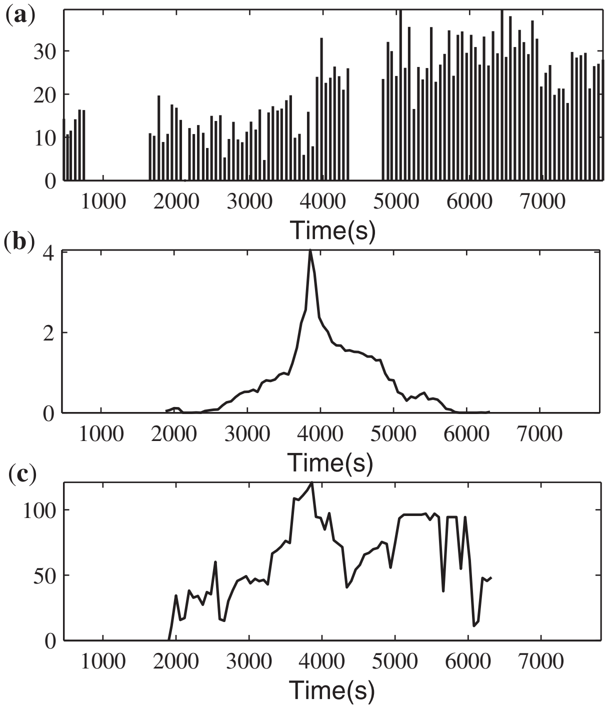

Fig. 4. (**a**) Time series of fluorescence intensity of RNA spots from images of an *E.coli* cell. Images taken for 2 h, one per minute (vertical axis is total spot intensity in arbitrary units). (**b**) Result of KCpA (vertical axis is the KCpA indicator output). (**c**) Result of uLSIF (vertical axis is the uLSIF indicator output). In both cases, window size is 25 (1500 s).

of the RNA spots along the *z*-axis scanned. Another source is endogenous to the tagging method, namely the number of MS2d-GFP molecules bound to the target RNA varies from 40 to 120 over time. Nevertheless, the method was found reliable in detecting, within ≤30 s, the appearance of new RNA molecules in the cell both empirically as well as using semiautomated methods (Golding and Cox, 2004; Golding *et al.*, 2005). From the images (Supplementary Material), a new RNA spot is detected to appear at 4100 s both empirically and by semiautomated methods. This moment should be identified by the point detection method as the one when the most significant change takes place.

The result of KPcA is shown in Figure 4b and of uLSIF in Figure 4c. For the 1D case, all kernels for KCpA are equivalent, giving the same result. In both cases, the window length was selected arbitrarily to 50 frames (i.e. 25 frames in both the backward and forward windows). Again, KPcA performed better, although both methods detect the change in the same location (at ~4000 s). Note that the window size is much smaller than the size used for the models. From the experimental data, we aim to detect the appearance of individual RNA molecules, whose effect on the number of RNAs in the cell is readily observable (given the small number produced by a cell). On the other hand, in the models we detected changes in mean levels of the order of tenths of new proteins, which is a change that requires much longer time to be completed once initiated, and thus wider windows to be detected.

The accuracy of the detection in this case is of relevance for studies of gene expression dynamics from measurements. So far, the MS2d-GFP tagging system is the only method available to detect the appearance of individual RNAs *in vivo*. The analysis of the images is, unfortunately, cumbersome (see movie in Supplementary Material). Our results are promising as they show that these methods can be used to detect in an automated fashion the moments when new RNAs appear, which will provide greater confidence in the results and allow the analysis of much larger samples of cells, making the analysis of the dynamics of gene expression, one molecule at a time, more robust. An automated unbiased analysis will also facilitate comparative studies of transcription activity of different promoters.

4 CONCLUSION

Genetic networks are subject to various structural changes and external signals, which alter their dynamics in various degrees. The detection of changes requires observing the dynamics of gene expression at the single cell, single molecule level. So far, very few direct or indirect methods allow this observation (Fusco *et al.*, 2003; Golding *et al.*, 2005; Yu *et al.*, 2006), and usually the extraction of the data from the measurements is cumbersome. Further, the ground truth signal is commonly unknown, further enhancing the need of using models to develop new methods for detecting changes in the dynamics of genetic circuits.

Changes in gene expression can be complex and diverse, e.g. in time scale. To detect them it is necessary to combine the use of adequate algorithms to particular changes, and provide information of the nature of the change one wishes to detect, which requires prior knowledge of the dynamics of gene expression at the molecular level combined with the development of new data analysis methods to distinguish real changes in signals from spurious fluctuations.

We applied recently developed point change detection methods to this problem. We tested their ability in detecting changes at run time in the dynamics of stochastic models of GRNs. The changes implemented mimic naturally occurring ones. A change in the mean expression level of a gene can occur, e.g. due to gene duplication or to silencing of a gene expressing a repressor. A change in noise in protein levels can occur, e.g. due to changes in the rate of RNA degradation. A change in the switching frequency of a two-gene switch can occur, e.g. due to a change in the number or binding affinity of the repressor proteins. Finally, the silencing of a gene that is part a large gene network can occur as a response to an external signal, and will affect the expression levels of multiple genes in the network.

In most of our test cases, KCpA outperforms uLSIF. This is likely a result of the nature of the changes that we aimed to detect and of the dynamics of protein and RNA levels. uLSIF has problems with cross-validated parameter selection and its results suffer from the sensitivity to the choice of parameter. The best kernel for detection is the linear kernel. This is probably because the changes in our examples are simple changes in mean levels or, for the more complicated cases (Models 2 and 3), the data can be cast to a change in mean level. In theory, the other kernels may detect more complex changes, which is probably the cause for multiple false matches in our case. Future studies may determine which algorithms are more appropriate to which type of change.

When applying KCpA to a model of an inferred core network of 328 genes of *S.cerevisiae* (Chowdhury *et al.*, 2010), we found it to be very accurate in detecting changes in the overall gene expression dynamics of the network, following the silencing of randomly selected genes. The size of the network did not seem to be problematic. This example is of relevance in that it shows that the method can be applied to the analysis of time series of complex gene networks, when affected by a change in either the network's structure or by an external signal or perturbation.

Finally, we applied the methods to detect, from time series of fluorescence intensity of RNA tagged with MS2d-GFP, single transcription events in live cells. Tagging RNA with MS2d-GFP proteins is, so far, the only method for detecting *in vivo* individual RNA molecules, thus, the correct extraction of information from the measurements, such as when new RNA molecules appear, is of relevance. Also in this case, the best detector appears to be KCpA, which produces a distinctive peak at the true change point.

We note that, to detect changes in the expression level of strongly expressing genes, it may be possible to assume ergodicity of the expression dynamics (similar temporal and ensemble averages), as several strongly expressing genes in, e.g. bacteria and yeast, usually exhibit fast temporal fluctuations. Provided this assumption, the changes can be detected from measurements at several time moments of expression levels across an isogenic cell population, rather than using our method since, while it is also valid, it would be more fastidious. However, generally, this assumption may not be valid. For example, studies *in vivo* and *in vitro* in *E.coli* show that most genes are rarely expressed (Bernstein *et al.*, 2002; Taniguchi *et al.*, 2010). Our measurements in Figure 4a are in agreement, since the promoter has a very slow dynamics, expressing on average only once every 700 s.

We believe that the results are promising. While the signals analyzed were poised with noise from multiple sources, the information of dynamical changes was, to a great extent, successfully extracted, both when detecting changes in the dynamics of model GRNs as well as when detecting when RNA molecules were produced in *E.coli*. Information on the nature of the changes that one wants to detect needs to be provided, to some extent. Particularly, prior knowledge is needed on the expected time length that a change takes to provoke a tangible change in the protein numbers. This is likely to be necessary regardless of the method used, as the dynamics of GRNs is extremely 'rich' in that a variety of mechanisms can affect the system in different ways, and the change may take different time lengths to be completed. In the future, we aim to further develop these methods and use them to analyze fluorescence measurements of expression of genes within genetic circuits in live cells.

Funding: Academy of Finland and Finnish Funding Agency for Technology and Innovation.

Conflict of Interest: none declared.

REFERENCES

Arkin,A. *et al.* (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected Escherichia coli cells. *Genetics*, **149**, 1633–1648.

Bar-Even,A. *et al.* (2006) Noise in protein expression scales with natural protein abundance. *Nat. Genet.*, **38**, 636–643.

Belloni,A. and Didier,G. (2008) On the Behrens-Fisher problem: a globally convergent algorithm and a finite-sample study of the Wald, LR and LM tests. *Ann. Stat.*, **36**, 2377–2408.

Bernstein,J.A. *et al.* (2002) Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent dna microarrays. *Proc. Natl Acad. Sci. USA*, **99**, 9697–9702.

Chowdhury,S. *et al.* (2010) Information propagation within the genetic network of Saccharomyces cerevisiae. *BMC Syst. Biol.*, **4**, 143.

Elowitz,M.B. and Leibler,S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**, 335–338.

Fisher,R. (1939) The comparison of samples with possibly unequal variances. *Ann. Eugenics*, **9**, 174–180.

Fusco,D. *et al.* (2003) Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr. Biol.*, **13**, 161–167.

Geva-Zatorsky,N. *et al.* (2006) Oscillations and variability in the p53 system. *Mol. Syst. Biol.*, **2**, 2006.0033.

Gillespie,D.T. (1977) Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, **81**, 2340–2361.

Golding,I. and Cox,E.C. (2004) RNA dynamics in live Escherichia coli cells. *Proc. Natl Acad. Sci. USA*, **101**, 11310–11315.

Golding,I. *et al.* (2005) Real-time kinetics of gene activity in individual bacteria. *Cell*, **123**, 1025–1036.

Harchaoui,Z. *et al.* (2009) Kernel change-point analysis. *Adv. Neural Inform. Proc. Syst*.

Kawahara,Y. and Sugiyama,M. (2009) Change-point detection in time-series data by direct density-ratio estimation. In *Proceedings of 9th SIAM International Conference on Data Mining*, Vol. 1, SIAM, Nevada, USA, pp. 389–400.

Kay,S. (1998) *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Prentice Hall, PTR, New Jersey, USA.

McClure,W.R. (1980) Rate-limiting steps in rna chain initiation. *Proc. Natl Acad. Sci. USA*, **77**, 5634–5638.

Peabody,D.S. and Lim,F. (1996) Complementation of rna binding site mutations in ms2 coat protein heterodimers. *Nucleic Acids Res.*, **24**, 2352–2359.

Ribeiro,A. *et al.* (2006) A general modeling strategy for gene regulatory networks with stochastic dynamics. *J. Comput. Biol.*, **13**, 1630–1639.

Ribeiro,A.S. and Lloyd-Price,J. (2007) SGN sim, a stochastic genetic networks simulator. *Bioinformatics*, **23**, 777–779.

Taniguchi,Y. *et al.* (2010) Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science*, **329**, 533–538.

Yu,J. *et al.* (2006) Probing gene expression in live cells, one protein molecule at a time. *Science*, **311**, 1600–1603.

Zhu,R. *et al.* (2007) Studying genetic regulatory networks at the molecular level: delayed reaction stochastic models. *J. Theor. Biol.*, **246**, 725–745.