Technical University of Denmark

DTU

Reconstructing Dynamic Promoter Activity Profiles from Reporter Gene Data

Kannan, Soumya; Sams, Thomas; Maury, Jerome; Workman, Christopher

[Link back to DTU Orbit](#)

DTU Library
Technical Information Center of Denmark

ACS SyntheticBiology

Research Article

Reconstructing Dynamic Promoter Activity Profiles from Reporter Gene Data

Soumya Kannan,[†] Thomas Sams,[‡] Jérôme Maury,[§] and Christopher T. Workman*,[†]

[†]Department of Biotechnology and Biomedicine, Technical University of Denmark, 2800 Lyngby, Denmark
[‡]Department of Electrical Engineering, Technical University of Denmark, 2800 Lyngby, Denmark
[§]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Lyngby, Denmark

Ⓢ Supporting Information

ABSTRACT: Accurate characterization of promoter activity is important when designing expression systems for systems biology and metabolic engineering applications. Promoters that respond to changes in the environment enable the dynamic control of gene expression without the necessity of inducer compounds, for example. However, the dynamic nature of these processes poses challenges for estimating promoter activity. Most experimental approaches utilize reporter gene expression to estimate promoter activity. Typically the reporter gene encodes a fluorescent protein that is used to infer a constant promoter activity despite the fact that the observed output may be dynamic and is a number of steps away from the transcription process. In fact, some promoters that are often thought of as constitutive can show changes in activity when growth conditions change. For these reasons, we have developed a system of ordinary differential equations for estimating dynamic promoter activity for promoters that change their activity in response to the environment that is robust to noise and changes in growth rate. Our approach, inference of dynamic promoter activity (PromAct), improves on existing methods by more accurately inferring known promoter activity profiles. This method is also capable of estimating the correct scale of promoter activity and can be applied to quantitative data sets to estimate quantitative rates.

KEYWORDS: *promoter activity, dynamic model, ordinary differential equations, fluorescent proteins, growth profiling*
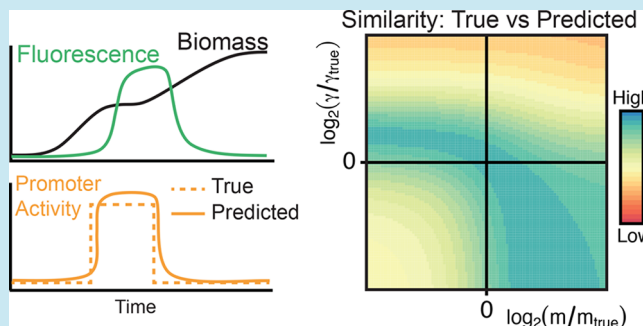
Recent developments in the fields of systems and synthetic biology have greatly expanded our ability to use engineering principles to model and design cellular pathways. One important example is the use of modular genetic elements to control the expression of protein products or enzymes governing fluxes through metabolic pathways.[1,2] This is most often achieved through the engineering of promoters that provide control of transcription rates, resulting in the desired level of proteins in the cell.

Promoter activity is typically assessed by measuring reporter gene expression, for example using fluorescent proteins, under the control of the promoter in question. In 1996, Subramanian and Srienc developed a mathematical model describing green fluorescent protein (GFP) accumulation in mammalian cells as a means of analyzing gene expression, but took transcription rate to be a constant value.[3] In 2001, Leveau and Lindow modified this earlier method, presenting a model of promoter activity, which they defined as the combined rate of transcription and translation, based on measurement of GFP fluorescence driven by the promoter of interest.[4] However, this model relies on three assumptions that are not necessarily valid in conditions that are relevant particularly to industrial bioprocesses, which are the ultimate targets of most metabolic engineering efforts: first, that the growth rate of the culture is

constant, second, that the culture is in the exponential growth phase, and third, that protein levels are at steady state. As a result, this model also assigns to a given promoter a single value corresponding to its "activity" or "strength". The Leveau and Lindow model has been employed in various studies with a wide range of applications.[5−7] Additional studies similarly characterize promoter strength as a single intrinsic quality of the promoter, though not using the Leveau and Lindow model.[8−11]

However, gene expression is commonly measured as a variable rather than a constant quantity over time, since transcriptional regulation changes as organisms grow and adapt to changing environments over the course of batch fermentations. Time-series gene expression data is therefore important for characterizing genetic functions and changes in gene regulation.[12] For example, model organisms such as *Escherichia coli* and *Saccharomyces cerevisiae* will metabolize different carbon substrates by order of preference,[13−16] undergoing a diauxic shift as the organism switches from utilizing one carbon source to another in a phenomenon first

described in bacteria by Jacques Monod in 1941.[17] As a result, cultures may have several exponential growth phases with different rates of growth, separated by lag phases of little to no growth. It is known that gene expression profiles differ in these different phases[13,15,18−21] and that many active transcriptional regulatory events occur during the diauxic shift. Additionally, simply measuring reporter protein expression or computing reporter protein synthesis rate over time does not accurately capture the dynamics of the promoter itself, which can respond much faster to activation signals.[22] Promoter activity in low-growth or static phases is of particular interest for applications in industrial biotechnology where product formation is typically desired only after biomass accumulation. In particular, it is a more efficient use of resources to induce product formation after biomass accumulation where growth can be limited and resources can be converted to products as opposed to more biomass.[23] During this stage of a culture's growth, more commonly called stationary phase, promoters that are active or changing activity when the culture is not growing are necessary to optimize production of desired secondary metabolites.

The ability to construct profiles of dynamic promoter activity over time using simple fluorescence and biomass measurements in order to understand these dynamic regulatory effects is an area of active research.[9,24−29,22] However, current methods either depend directly on growth rate, thus introducing numerical difficulty in accurately capturing activity in lag and stationary phases when growth rate is close to 0, or do not explicitly consider this situation. Here, we present a model and approach for inference of dynamic promoter activity (PromAct) that does not directly depend on growth rate, and show that it is able to capture activity regardless of whether or not biomass accumulation is occurring. This model includes no underlying assumptions about the organism, experiment, time resolution of sampling, or mode of regulation of the promoter of interest. PromAct takes only protein expression and biomass data as inputs and is implemented in the statistical programming language R,[30] which is easily accessible and widely used.

RESULTS AND DISCUSSION

Development of a Model for Dynamic Promoter Activity in Cell Culture. We have modeled fluorescent protein (FP) accumulation using simple mass-action kinetics based on the schematic shown in Figure 1A, ultimately resulting in a single expression for transcription rate dependent only on measured reporter protein expression and biomass (eq 13). We can use first-order kinetics for degradation[3,24] by assuming that the FP accumulation never reaches levels that saturate the cellular proteases.

At the single-cell level, we have formulated the model as follows:

$$\frac{\partial r}{\partial t} = P(t) - (\gamma_r + \mu(t))r(t) \tag{1}$$

$$\frac{\partial n}{\partial t} = \beta r(t) - (\gamma_n + \mu(t) + m)n(t) \tag{2}$$

$$\frac{\partial f}{\partial t} = mn(t) - (\gamma_f + \mu(t))f(t) \tag{3}$$

where r is the number of mRNAs, n is the number of nonfluorescent proteins, f is the number of mature fluorescent proteins, $P(t)$ is the promoter activity, defined as transcription
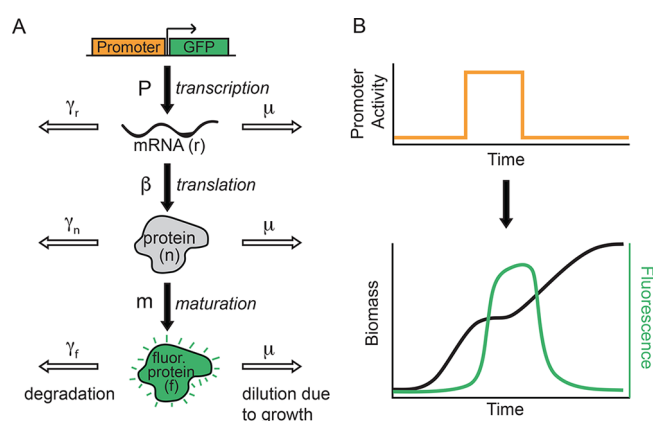

Figure 1. Schematic of reporter gene expression and associated measurements. (A) Schematic of fluorescent protein accumulation. A fluorescent protein (FP) is transcribed at a time-varying rate $P(t)$. The mRNA is translated into an immature nonfluorescent protein with rate β, which folds into the mature FP with rate m. (B) Schematic illustration of promoter activity during low growth.

rate in number of mRNAs per unit time, β is the translation rate in number of proteins per mRNA per unit time, m is the maturation rate of the fluorescent protein in units of inverse time, $\gamma_{r,n,f}$ are the degradation rates of the mRNA, non-fluorescent and fluorescent proteins, respectively, and $\mu(t)$ is the growth rate as a function of time. This general formulation, or similar formulations, are commonly used to describe fluorescent protein expression and accumulation.[3,4,7,24,26,27,29,22,31]

It is currently difficult to measure the number of proteins in a single cell at a high temporal resolution for many strains or replicates, though technologies are being developed to address this problem.[32] However, currently it is more common to make bulk measurements of the fluorescence or gene expression and biomass of an entire cell culture, and most experimentalists have access to instrumentation for this purpose. Using such data, the profile of the average cell in the population can be obtained, which is valid as the measured fluorescence F (often measured in relative fluorescence units [RFU]) is proportional to the number of FP molecules within the culture.[33] Additionally, we assume measured biomass (often measured in units of optical density [OD] or light scattering units [LSU]) is proportional to cell count and thus, by extension, template DNA concentration, given an instrument that is properly calibrated and whose accurate detection ranges are known. While it is true that at the single cell level stochastic fluctuations in gene expression occur,[34,35] the relationship between cell-level stochastic events and events within the population as a whole is not well understood, so we seek here to model an average cell and the whole population.[22] As a result, we can define the following equations based on measurements from a cell culture:[29]

$$\frac{\partial b}{\partial t} = \mu(t)b(t) \tag{4}$$

$$\frac{\partial R}{\partial t} = P(t)b(t) - \gamma_r R(t) \tag{5}$$

$$\frac{\partial N}{\partial t} = \beta R(t) - (\gamma_n + m)N(t) \tag{6}$$

$$\frac{\partial F}{\partial t} = mN(t) - \gamma_f F(t) \tag{7}$$

where $R(t)$ is the amount of mRNA in the culture over time in units corresponding to measured units for protein amounts (*i.e.*, RFU), $N(t)$ is the amount of nonfluorescent protein in measured units, $F(t)$ is the amount of fluorescent protein in measured units, usually RFU, and $b(t)$ is the measured biomass, usually measured in units of OD or, more rarely, LSU or cell dry weight (CDW). All other variables are as before but redefined in appropriate units, *i.e.*, $P(t)$ is in units of RFU per unit biomass per unit time, *etc.* Note that this model is linear in mapping $F(t)$ to $P(t)b(t)$. By considering the entire population of cells in a fixed volume, dilution is no longer relevant.[22] Solving eq 5 for $P(t)$ gives the following expression:

$$P(t) = \frac{1}{b(t)}\left(\frac{\partial R}{\partial t} + \gamma_r R(t)\right) \tag{8}$$

However, $R(t)$ is not measured—$F(t)$ is. Thus, we need to find an expression for $R(t)$ in terms of $F(t)$ and substitute this into eq 8. First, solve for $N(t)$ in eq 7 to get

$$N(t) = \frac{1}{m}\left(\frac{\partial F}{\partial t} + \gamma_f F(t)\right) \tag{9}$$

Taking the derivative with respect to time gives

$$\frac{\partial N}{\partial t} = \frac{1}{m}\frac{\partial^2 F}{\partial t^2} + \frac{\gamma_f}{m}\frac{\partial F}{\partial t} \tag{10}$$

Substituting eqs 9 and 10 into eq 7 and solving for $R(t)$ gives

$$R(t) = \frac{1}{\beta}\left(\frac{1}{m}\frac{\partial^2 F}{\partial t^2} + \frac{1}{m}(\gamma_n + \gamma_f + m)\frac{\partial F}{\partial t} + \frac{\gamma_f}{m}(\gamma_n + m) F(t)\right) \tag{11}$$

Taking the derivative of eq 11 with respect to time yields

$$\frac{\partial R}{\partial t} = \frac{1}{\beta m}\left(\frac{\partial^3 F}{\partial t^3} + (\gamma_n + \gamma_f + m)\frac{\partial^2 F}{\partial t^2} + (\gamma_n + m)\gamma_f\frac{\partial F}{\partial t}\right) \tag{12}$$

Finally, substituting eq 11 and 12 into eq 8 gives transcription rate in terms of only measurable quantities: measured fluorescence $F(t)$ and its derivatives, and measured biomass $b(t)$.

$$P(t) = \frac{1}{\beta m b(t)}\left(\frac{\partial^3 F}{\partial t^3} + C_2\frac{\partial^2 F}{\partial t^2} + C_1\frac{\partial F}{\partial t} + C_0 F(t)\right) \tag{13}$$

where

$$C_0 = \gamma_f \gamma_r (\gamma_n + m) \tag{14}$$

$$C_1 = (\gamma_f + \gamma_r)(\gamma_n + m) + \gamma_f \gamma_r \tag{15}$$

$$C_2 = \gamma_n + \gamma_f + m + \gamma_r \tag{16}$$

This can be calculated if we have measured fluorescence and biomass $F(t)$ and $b(t)$, all degradation rates (for fluorescent proteins, we assume $\gamma_n = \gamma_f$,[4] the translation rate β and the maturation rate m.

Model Implementation. In practice, the promoter activity quantity is computed by fitting spline models to fluorescence and biomass data, from which derivatives are numerically computed.[36] PromAct has been implemented as an R function which takes vectors of fluorescence and biomass data and associated times as inputs and returns a profile of transcription rates in units of input fluorescence units per unit of input biomass per unit time. PromAct runs fast enough that on all data sets tested, no significant scaling effects have been observed. The smooth.Pspline function runs in $O(n)$ time, *i.e.*, is proportional to the size of the input data set, and this is the limiting factor in analyzing the time complexity of our algorithm. The spline model for biomass is constrained to be monotonically increasing, as biomass, defined as intact cells, tends to only increase in early growth phases where nutrients are in relative excess although the measurement of biomass may fluctuate to a large proportion due to low signal-to-noise. Although the number of intact cells can decrease after the primary growth phases, the measurement of biomass at stationary phase is less subject to noise, and as a result, fluctuations have less influence on promoter activity estimates at high biomass levels. The monotone spline models were fit using functions from the fda package in R.[37] In instances where the basis for the monotone spline cannot be constructed, generally due to the noise structure of the time series, a standard cubic spline using the built-in smooth.spline function from the R stats package is used.[30] In this case, the smoothing parameter is automatically determined by the smoothing algorithm *via* generalized cross-validation, but constrained to be ≤0.1 greater than the smoothing parameter calculated *via* eq 17. Use of the monotone function however plays a key role in reducing the noise that is usually captured by normal smoothing splines at low signal when signal-to-noise ratio is also low. The spline model for fluorescence is a fourth-order spline that is fit using the pspline package in R.[38] Fourth-order splines are employed here as a spline order of minimum $n + 1$ is generally required in order to reliably compute an nth order derivative, and calculation of third-order derivatives is required here. We use a smoothing parameter computed from a formula dependent on the spacing between data points where time is in hours. This formula was determined through empirical observation of how varying the smoothing parameter relative to the time resolution of data is able to balance accurately capturing fluctuations in fluorescence signal while minimizing noise at early times.

$$\text{spar} = \frac{0.189}{\Delta t^2} + 0.07, \quad \text{spar} \le 0.6 \tag{17}$$

Intuitive physical limitations are placed on fitted functions—fluorescence cannot be less than 0 and biomass cannot be less than 0.001 measured units. Output promoter activity is also constrained to non-negative values. The promact package containing this function, as well as the simulated and real data presented here can be downloaded from the Supporting Information.

Model Reconstructs Both Shape and Magnitude of True Promoter Activity. Hypothetical promoter activity functions were used to simulate fluorescence data by numerical integration (see Methods section). Three promoter activity functions were chosen that produce gene expression patterns previously observed in cell cultures over time.[19,28,39] Assuming that the degradation, translation and maturation rates are known, this model is, by visual inspection, able to both quantitatively and qualitatively reconstruct the input promoter activity function (Figure 2). It must be noted that promoter
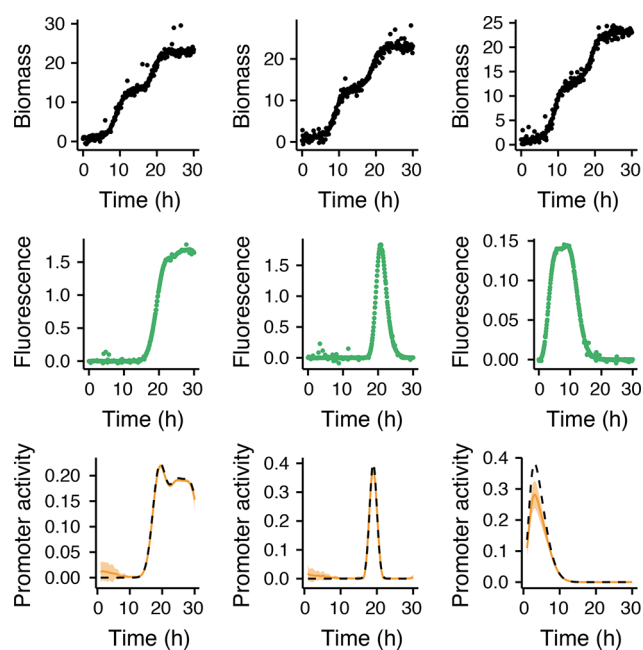
Figure 2. Promoter activity reconstruction from simulated data. Simulated biomass and fluorescence data and promoter activity as reconstructed by our model compared to "true" promoter activity. "True" promoter activity is shown by dashed black line, with model reconstruction in orange. Bottom row shows mean of 100 simulations with shaded area ± one standard deviation. Units of promoter activity are dependent on units in which fluorescence and biomass are measured. Promoter activity functions were chosen to produce gene expression patterns that have been previously observed (Column 1,[39] Column 2,[28] Column 3[19]). Promoter activity is given in units of fluorescence unit per biomass unit per unit of time.

activity is defined in terms of measured units, so the magnitude of the function is reconstructed assuming the input function is in measured units. However, the biological meaning of a single RFU or a single OD unit, for example, is not immediately obvious. Thus, we might more accurately consider that the promoter activity function is reconstructed up to some multiplicative constant that scales the function to biologically relevant units. We also observe that the reconstructed promoter activity is noisier when biomass is low, which results from the fact that dividing by a small number amplifies small fluctuations in the numerator (see eq 13).

It is known that quantitative characterization of promoters or other biological parts is dependent on many external factors, notably instrumentation and data analysis choices[7,40] experimental conditions[7,41] and genetic context.[41,42] In particular, it has been shown that fluorescent protein expression can bias measurements of optical density, a common means of inferring cell abundance.[43] Furthermore, these factors can be intertwined and hard to decouple, adding difficulty to attempts at creating standardized registries of biological parts. Kelly et al.[7] have cleverly proposed circumventing this problem through use of a standard control, as they have found that, while absolute promoter activity is variable when measured for the same promoters by different laboratories, promoter activity relative to that of a reference promoter measured in the same lab varied only within an acceptable range of measurement error. The 2014/2015 iGEM interlab study[40] similarly found a high degree of precision in the ratio of observed fluorescence between various promoters in experiments conducted by 88 institutions worldwide, indicating the promise of such a strategy. Nielsen et al. similarly make use of "Relative Promoter Units" in Cello, a program that can be used for automation of genetic circuit design;[44] in general, this practice has been adopted as relatively standard in sythetic biology. However, both Kelly et al. and
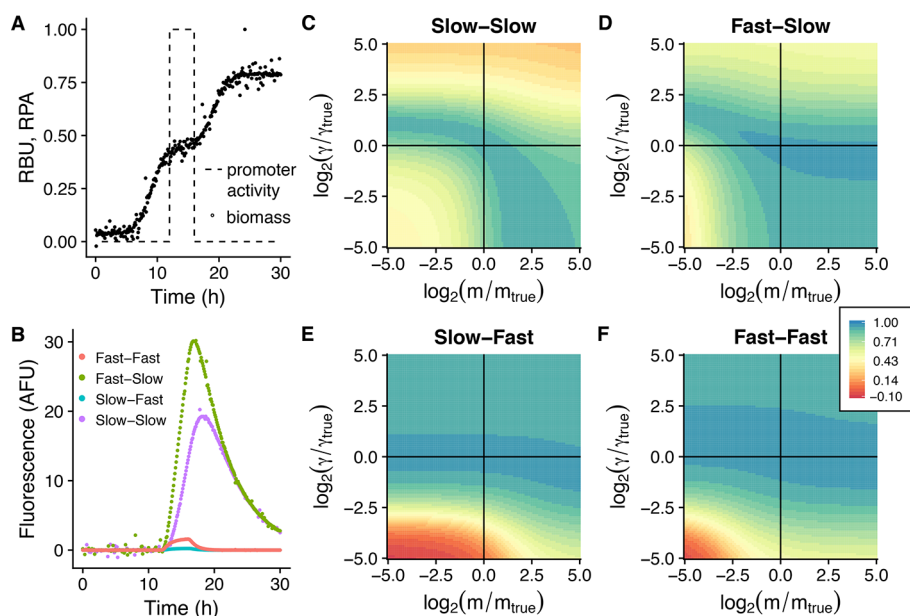

Figure 3. Analysis of parameter space. (A) Simulated biomass and promoter activity function used to generate simulated fluorescence data. (B) Fluorescence data simulated from above biomass and promoter activity given four combinations of maturation and degradation rates. (C−F) Correlation coefficient of calculated promoter activity to true promoter activity for 10 000 maturation and protein degradation rate combinations (100 maturation rates × 100 degradation rates) for each of 4 combinations of maturation and degradation rate. Correlation values are binned into 1 of 24 equally spaced intervals within the range shown by the color bar. Higher correlation values are shown in green to blue coloring indicating a greater similarity. Fast maturation = $8 \, h^{-1}$, slow maturation = $0.5 \, h^{-1}$, fast degradation = $5 \, h^{-1}$, slow degradation = $0.2 \, h^{-1}$.
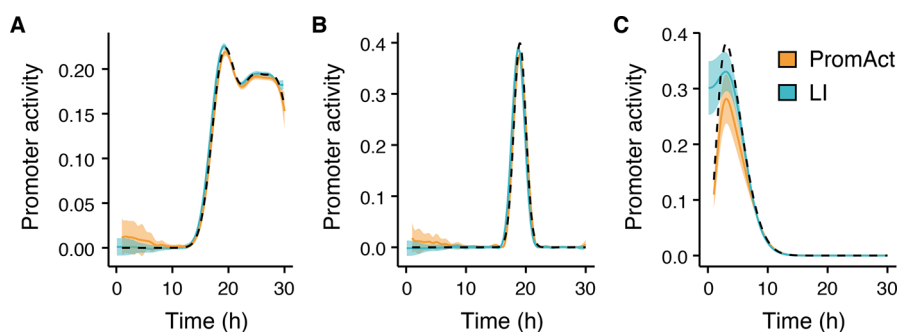
Figure 4. Comparison of promoter activity calculation methods given simulated data. Promoter activity inferred *via* the PromAct method is shown in orange and *via* the LI method[29] is shown in blue. Lines are means of promoter activity calculated for 100 data simulations and shaded area shows ± one standard deviation.

Nielsen *et al.* only consider promoter activity as a single intrinsic value of the promoter, as opposed to a dynamic value that changes as the cell population grows, as we have here. These two ideas are not incompatible; it is possible that a time series of promoter activity could be normalized to the dynamic profile of a reference promoter as is done in a static context. The use of relative promoter activity is compatible with our approach given that the maturation rate of fluorscent protein, the degradation rate of both nonmature and mature protein, and the degradation rate of the mRNA are the same as in the control.

An Evaluation of Robustness to Inaccuracies in Maturation and Degradation Rates. To account for the fact that degradation and maturation rates of a particular FP may not be measured with perfect accuracy, we have evaluated the ability of our model to reconstruct a promoter activity input function across a range of parameters that deviate from the "true" values for four combinations of high and low degradation and maturation rates[45] (Figure 3). This is also relevant because it is known that even the same FPs can exhibit different values of these parameters when expressed in different organisms,[46] and in general, there are no comprehensive databases or standardized methods of measuring these parameters. Additionally, these are user-input parameters, so it is relevant to explore how the user's choices may affect model performance. We examined the correlation coefficient, an indicator of qualitative similarity, between calculated and "true" promoter activity over time as in some cases, even with the wrong parameters, the model is able to reproduce the "true" promoter activity qualitatively but not quantitatively. In general, we see that in all cases, calculated and "true" promoter activity are highly correlated at zero fold change, as expected. Examples of promoter activity profile comparisons where maturation and degradation parameters differ can be seen in Figure S1. Notably, for fast degrading proteins, as long as the degradation rate is within a 1−2 fold difference from the "true" rate, the maturation rate does not appear to affect correlation (Figure 3D,E). For slow-degrading proteins, it appears that decreasing degradation rate from the "true" value can be compensated for by increasing maturation rate (Figure 3B,C). A similar analysis based on the Euclidean distance, as a dissimilarity measure, can be seen in Figure S2.

Comparison to Existing Models. Other models exist that attempt to capture the phenomenon of dynamic variation in transcription rate.[8,9,24−26,29] Of these, the linear inversion (LI) method described by Zulkower *et al.*[29] is the most generally applicable, is also based on the same underlying differential eqs

5−7, and has an easily accessible computational implementation. We will compare this to the model we have developed here. Promoter activity functions of different shapes, magnitudes and with activity occurring at different phases of growth were examined and compared across all promoter activity calculation methods (Figure 4).

The LI method has relatively low noise at early times compared to our method. It is also able to capture various promoter activity input functions from noisy simulations, given that changes in activity occur after a sufficient amount of biomass has accumulated (Figure 4). PromAct is additionally able to capture the particular features of activity functions at low biomass (initial lag phase), while the LI method tends to be biased in this regime (Figure 4C), likely due to the regularization, which otherwise serves to effectively reduce noise. However, as it has been shown that gene expression occurs in this growth phase,[19] it is important to be able to accurately capture promoter activity dynamics in this regime.

Additionally, though there is a general trend toward high time resolution measurement systems, many experimenters still make manual fluorescence measurements with spectrophotometers, which can limit time resolution to the scale of hours in some cases. Thus, we compared the performance of these models for a variety of different time-steps between measurements (Figure 5), as it is generally desirable to have a model that will perform for a large range of step sizes. We assumed parameter values set to replicate common GFP variants (see Simulation of Biomass and Fluorescence Data in the Methods section) and that the same parameter values as used to simulate data were used to infer promoter activity. In general, and as expected, both methods become worse as step size increased as a higher time resolution allows for more detailed capture of rapid changes in fluorescence levels and thus in the promoter activity function. In general, promoter activity reconstructions by our model are consistently better correlated to "true" promoter activity than those by the LI method; at the 1.5 h step size, our method has a median correlation coefficient 3.40% greater than that of the LI method over the 100 data simulations analyzed in Figure 5. Additionally, when increasing step size from 0.1 to 1.5 h, PromAct only has a decrease in median correlation coefficient to "true" promoter activity of 2.51% as compared to 4.16% for the LI method. However, it should be noted that both have a median correlation coefficient over 0.9 for all step sizes examined. For longer step sizes, the LI method begins to fail (data not shown). Additionally, output step size of the LI method is very restricted as compared to our
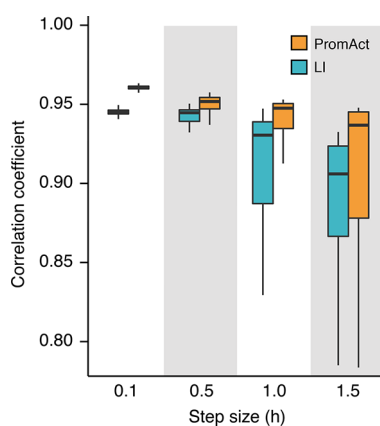
Figure 5. Comparison of models over a range of data temporal resolutions. Promoter activity for a step function occurring during the diauxic shift lag phase was calculated using data simulated with different rates of sampling ranging from one measurement taken every 6 min to one every 1.5 h, and the correlation coefficient between "true" promoter activity and each calculated promoter activity was computed. Our model is shown in orange and LI method shown in blue. Boxes show median with lower and upper hinges corresponding to first and third quartiles and whiskers extending from the lowest value within 1.5 times the interquartile range to the highest such value, over 100 simulations.

method, which allows for construction of a promoter activity time series with arbitrary precision in step size.

Case Study 1: GFP Driven by a Xylose-Inducible Promoter. To test PromAct on experimental data of a promoter driving FP expression, we examined the xylose-inducible pXylA promoter driving expression of GFPmut3b in *E. coli* (Figure 6C). Using PromAct, we found that transcription turned on at about 1.6 h, reached a maximum at about 3.32 h and shut off at about 5.5 h after induction (Figure 6D). This matches what is observed in the fluorescence and biomass data. Biomass starts rapidly increasing around 1.5 h. The ratio of rate of increase in fluorescence to rate of increase reaches a

maximum in biomass at about 3.84 h, reflecting the time required for translation and maturation processes to occur following the occurrence of the maximum transcription rate. Fluorescence begins to decrease around 5.72 h, allowing for the translation and maturation of existing transcripts after the promoter has been shut off.

The functional form of promoter activity also matches with the hypothesis of what is occurring in this simple experiment. *E. coli* are able to metabolize xylose and thus, the xylose could serve as both an inducer and carbon source in this experiment. If the cells are consuming the xylose, we would expect that upon reaching steady state, the carbon source has run out, meaning that there is no xylose left. The pXylA promoter would thus be lacking an inducer and would cease initiating transcription of GFP, leading to a decrease in fluorescence signal as existing GFPs decay but no new ones are produced. This would occur slightly delayed, as existing transcripts may still be translated as well, and existing nonfluorescent proteins will still mature and add to the fluorescence signal. PromAct constructs a promoter activity profile in accordance with this prediction—soon after promoter activity reaches 0, the cultures reach steady state and the fluorescence signal begins to decrease (Figure 6C,D).

Case Study 2: GFP Driven by the *E. coli fis* Promoter. As a second test of PromAct on experimental data, we examined a data set of fluorescence and absorbance data of *E. coli* strains containing GFPmut3b driven by the promoter of Fis, a global transcription regulator in *E. coli*, obtained from de Jong *et al.* (see Figure 10a of ref 27).[27] As detailed by the authors, *fis* expression is known to be induced by an upshift in glucose and is subsequently decreased as the culture enters the exponential growth phase.[27,51−53] As seen both here (Figure 6B) and by the authors of this study (Figure 10c of ref 27), this is exactly what is predicted by the kinetic model we have proposed—there is a strong burst of transcription at the beginning, but as the culture enters the exponential phase of growth around 2 h, transcription rate decreases to a basal level in which some transcription still occurs at much lower rates,
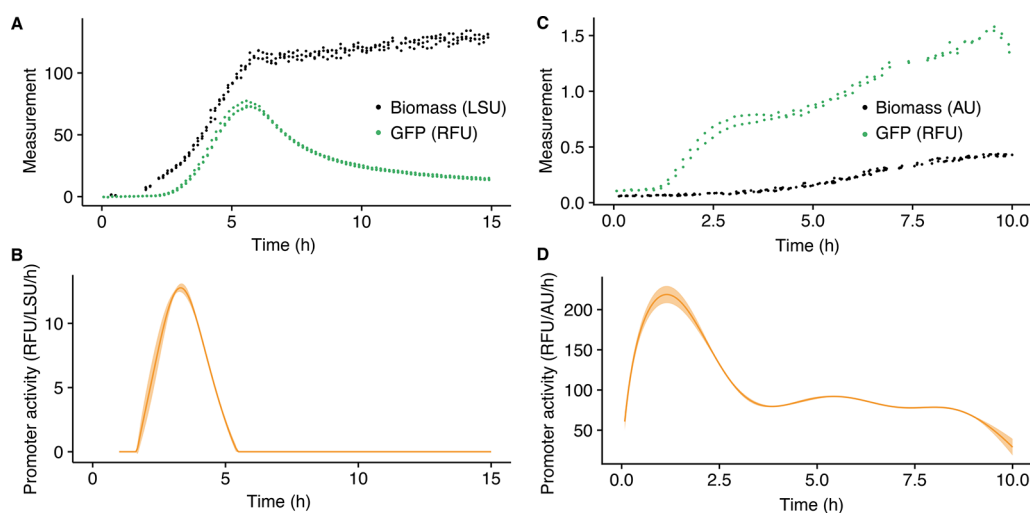

Figure 6. Promoter activity calculated from experimental measurements of fluorescence and biomass. (A) Fluorescence in relative fluorescence units (RFU) and biomass measured in light scattering units (LSU) of *E. coli* expressing GFP under control of pXylA ($n = 3$), and (B) calculated promoter activity given $m = \log(2)/0.125\ \mathrm{h}^{-1}$,[47] $\gamma_n = \gamma_f = \log(2)/24\ \mathrm{h}^{-1}$[48] and $\gamma_r = \log(2)/0.05\ \mathrm{h}^{-1}$.[49,50] (C) Fluorescence and biomass measured in absorption units at 600 nm (AU) of *E. coli* expressing GFP under control of the *fis* promoter ($n = 2$), and (D) calculated promoter activity given $m = \log(2)/0.417\ \mathrm{h}^{-1}$, $\gamma_n = \gamma_f = 0.72\ \mathrm{h}^{-1}$ and $\gamma_r = 18\ \mathrm{h}^{-1}$.[27] Lines are means of n calculated promoter activities and shaded region represents one standard deviation.

leading to the slower accumulation of GFP that is demonstrated by the more slowly increasing fluorescence signal past this point.

The main difference to note between the promoter activity profile obtained by this analysis and that obtained by de Jong *et al.* is quantitative, a result of the units used to express transcription rate. Here, we show promoter activity as fluorescence per absorbance unit per hour; de Jong *et al.* use units of per minute, and further, do not consider activity per biomass unit. Thus, we obtain a different magnitude of activity. We include this, as previously discussed, in order to model the behavior of the "average" cell in the population, as absorbance units could be converted to cell count. In addition, notably, these fluorescence and biomass data were not taken at the same time points, but this does not prove to be a limitation of our method.

Future Considerations: Use of Fluorescent RNA Input. One drawback of our model is that it suffers from more noise in the regime where biomass is close to 0, as small fluctuations in the numerator of eq 13 are amplified when divided by small numbers, and even small differences in the smoothed biomass profile relative to the observed values can create significant noise in this regime. This results from the indirect method of computing transcription rate, as fluorescent protein measurements must be translated backward through several biosynthetic steps in order to estimate the rate of mRNA production.

One potential solution to this problem lies not with the model formulation or numerics, but rather the setup of experiments aimed at measuring dynamic transcription rate profiles. Recent innovations in *in vivo* RNA imaging include a suite of RNA aptamers that bind fluorophores to create fluorescing RNA-fluorophore complexes which can be detected and imaged in the same way as fluorescent proteins.[54,55] In particular, the Spinach aptamer, which mimics the excitation and emission wavelength of common GFP variants has been optimized for *in vivo* use,[56] which is especially useful given that most instrumentation for measuring fluorescence in cells uses filters for GFP wavelengths. Spinach has previously been used in simple experiments for *in vivo* part characterization and time-series measurement of gene expression,[57] which could potentially be extended to use in online measurement instruments. The promact R package, available in Supporting Information, includes a setting for this type of measurement as well. However, use of fluorescent mRNAs is not currently as widespread as use of fluorescent proteins, necessitating models considering protein level measurements for the time being.

In Summary. Using simulated data, our PromAct model appears appreciably robust to values of maturation and degradation rate (Figure 3), time resolution of input data (Figure 5), and acceptable levels of measurement noise. It is easily accessible and installed as an R package (Supporting Information) that is designed to be as simple as possible to use.

PromAct also offers some improvements over another well-developed and easily accessible model, the LI model presented by Zulkower *et al.* in 2015.[29] In general, we find that both models are able to capture the "true" promoter activity function from simulated noisy data when time resolution of data points are sufficiently close in time, and the promoter activity occurs at times where a sufficient amount of biomass is present. However, PromAct is able to capture the promoter activity accurately at relatively low time resolution (one measurement per hour) and in the initial lag phase before biomass accumulation has begun, which are important considerations

in some contexts.[19] Additionally, PromAct has been developed particularly without a dependence on growth rate, although the LI model is also able to accurately reconstruct promoter activity in low growth phases, such as the diauxic shift and stationary phase, so long as sufficient biomass is present.

Especially as systems and synthetic biology strategies take hold for larger scale metabolic engineering and industrial biotechnology, it is important to have efficient and high-throughput methods for quick characterization of genetic elements, particularly in industrially relevant contexts. Synthetic biology tools have been successfully applied to produce a diversity of natural products using cell factories, for example resveratrol,[58] iso- and *n*-butanol,[59] and many more (see Table 1 in ref 1). However, even in well studied organisms such as *E. coli* and *S. cerevisiae*, despite their widespread use as cell factories in industrial biotechnology, only a handful of molecular biology tools exist to fine-tune expression of enzymes in pathways of interest, especially under industrial bioprocess conditions. In particular, there are not many libraries of promoters that are well characterized under such conditions. Additionally, in large scale fermentation processes, it is often a more efficient use of resources to turn on product formation after biomass accumulation is largely complete,[23] creating a necessity for regulatory mechanisms that activate gene expression when growth rate is close to 0. It is this particular situation that violates the assumptions of many existing models for calculating promoter activity. Tools like PromAct that accurately compute dynamic promoter activity under any condition or growth rate offer the possibility of engineering temporal regulation within metabolic pathways of important industrial relevance.[60]

■ METHODS

Simulation of Biomass and Fluorescence Data. Biomass data was simulated using a bilogistic function,[61] which qualitatively reproduces the biomass profile of cultures undergoing two growth phases separated by a diauxic shift lag phase. Simulated biomass data and hypothetical promoter activity functions were used to simulate fluorescence data by numerical integration of eqs 5−7 using the rk function from the deSolve package in R[62] employing a variable step-size Runge−Kutta method using a fourth and fifth ordered pair (more commonly known as the ode45 method). Initial concentrations of mRNA and protein species were set to 0, and the parameter values were set to replicate common GFP variants where possible: the nonfluorescent and fluorescent protein degradation rates (γ_n and γ_f respectively) were set to $\frac{\log(2)}{0.5}$ h^{-1},[4,63] and the maturation rate constant m was set to $\frac{\log(2)}{0.45}$ h^{-1}.[64] The mRNA degradation rate constant γ_r and the translation rate β were set to 1 for convenience and later comparisons to other similar models, as no reliable estimates for these could be found. Similar data simulation procedures are found in the literature.[24,26,65] To simulate homoscedastic noise from measurement, noise vectors of the same length as the vectors of simulated biomass and fluorescence were generated by drawing from a distribution of the residuals of spline fits to 198 measurements of fluorescence and biomass from *S. cerevisiae* strains carrying various promoters driving GFP expression (unpublished data). These noise vectors were added element-wise to the vectors simulated data points and the model was used to reconstruct the initial promoter activity function from

the noisy simulated data. Three promoter activity functions were chosen that qualitatively produce gene expression patterns previously observed in cell cultures over time.[19,28,39]

Analysis of Parameter Space. We evaluated the ability of PromAct to reconstruct a promoter activity function given a different set of maturation and degradation rates than those used to simulate the fluorescence time profiles. We chose four combinations of maturation and degradation rates in the four extremes of hypothetical parameter space (Table 1) adapted

Table 1. Maturation and Degradation Rate Combinations for Parameter Analysis

system type	maturation rate (h^{-1})	degradation rate (h^{-1})
slow–slow	0.5	0.2
fast–slow	8.0	0.2
slow–fast	0.5	5.0
fast–fast	8.0	5.0

from a collection of current estimates of maturation and degradation rates for common FPs.[45] A single biomass and fluorescence profile were simulated for each parameter combination in Table 1 using the procedure described in the Simulation of Biomass and Fluorescence Data section above, given a step function of promoter activity occurring during the diauxic shift of simulated biomass data. Combinations of 100 maturation rates and 100 degradation rates spanning a fold change of −5 to 5 from the parameter used to simulate data were input into the promoter activity model, for a total of 10 000 parameter combinations for each "true" combination given in Table 1. The correlation coefficient and log of Euclidean distance were then computed between the input "true" promoter activity and the reconstructed promoter activity for each of the 10 000 parameter combinations. All analysis was done in R.[30]

Varying Temporal Resolution. We also evaluated the ability of our model to reconstruct promoter activity for a range of sampling resolutions. For rates of one sample every 6 and 30 min and every 1 and 1.5 h, 100 simulated data sets were generated using the procedure described in Simulation of Biomass and Fluorescence Data above given a step function of promoter activity during the diauxic shift of the simulated biomass data. For each of these 100 simulations, correlation coefficient was calculated between the input "true" promoter activity and the promoter activity reconstructed by the model. Boxplot in Figure 5 shows median of correlation coefficients over these 100 simulations, with lower and upper hinges corresponding to the first and third quantiles and whiskers extending from the lowest value within 1.5 times the interquartile range to the highest such value. Again, all analysis was carried out in R[30] excepting any promoter activity calculations by the LI method, which is implemented in Python.

Strain and Growth Conditions. *Escheria coli* were transformed with a pSB1C3 plasmid,[66] a standard high copy number plasmid conferring chloramphenicol resistance, containing the xylose-inducible pXylA promoter[67] driving expression of GFPmut3b.[68] Cultures of transformants were grown overnight (at 37 °C) in LB media containing 6 μg/mL chloramphenicol.

Fluorescence and Biomass Measurement. Optical density at 600 nm (OD600) of overnight cultures was measured, and cultures were diluted in sterile deionized H_2O

to an OD600 of 0.5. In a 48-well BioLector (m2p Laboratories) FlowerPlate, 3 replicates of 300 μL of diluted culture were added to 1.2 mL LB media with 7.5 μg/mL chloramphenicol and 1.25% xylose, giving a final OD600 of 0.1 with final chloramphenicol concentration at 6 μg/mL and 1% xylose. Negative controls were prepared using media without xylose. The BioLector plate was covered with a gas-permeable seal and placed in the BioLector chamber, with an incubation temperature of 37 °C, shaking at 1000 rotations per minute and humidity at 95%. Biomass was measured by absorbance at 620 nm in light scattering units (LSU) and GFP fluorescence was measured by emission at 520 nm following excitation at 488 nm in relative fluorescence units (RFU). Measurements were taken every 10 min for 15 h.

Analysis of pXylA Promoter Activity. Outliers as a result of instrumentation failure were pruned from biomass time courses, and biomass measurements were background corrected by subtracting a factor such that the minimum measurement was corrected to 1 light scattering unit. To background correct fluorescence profiles of induced samples, fluorescence time courses of uninduced samples were combined and fit with a spline model using the smooth.spline function in the base package of R[30] with a smoothing parameter of 0.8. Values predicted by the smoothing spline at each time were subtracted from induced fluorescence profiles to obtain background-corrected profiles. Background-corrected biomass and fluorescence data and corresponding times of measurements for each induced sample were input into the PromAct function using a maturation rate of $\log(2)/0.125$ h^{-1}[47] and degradation rate of $\log(2)/24$ h^{-1}[48] as measured for GFPmut3. mRNA degradation rate was set as $\log(2)/0.05$ based on genome-wide studies of mRNA degradation rates in *E. coli*.[49,50]

Analysis of *fis* Promoter Activity. Fluorescence and absorbance data were extracted from de Jong *et al.*[27] were previously background corrected as a part of the original analysis detailed in that article. These biomass and fluorescence measurements and corresponding times of measurements (converted to hours for consistency with the rest of the data presented here) were input into the PromAct function using a maturation rate of $\log(2)/(0.417)$ h^{-1}, protein degradation rate of 0.72 h^{-1} and mRNA degradation rate of 18 h^{-1} as measured by de Jong *et al.* in their expression system.[27]

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.7b00223.

> File S1: PromAct: an R package for promoter activity calculation (ZIP)
> Figure S1: Example comparisons of estimated and "true" promoter activities when maturation and degradation vary from the true values; Figure S2: Error landscape relative to values of m and γ differing from the true values using Euclidean distance (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: cwor@dtu.dk.

ORCID ⓞ

Thomas Sams: 0000-0002-6316-2539

Christopher T. Workman: 0000-0002-2210-3743

Author Contributions

JM and CTW proposed the initial model for dynamic promoter activity. SK, TS and CTW further developed the approach presented here. KS performed all analyses presented. All authors contributed to the writing of the manuscript. KS and CTW implemented the PromAct method in R.

Notes

The authors declare no competing financial interest.

The R package PromAct can also be accessed *via* GitHub at https://github.com/soumyakannan/promact.

■ REFERENCES

(1) Jensen, M. K., and Keasling, J. D. (2015) Recent applications of synthetic biology tools for yeast metabolic engineering. *FEMS Yeast Res.*, 1−10.

(2) Borodina, I., Kildegaard, K. R., Jensen, N. B., Blicher, T. H., Maury, J., Sherstyk, S., Schneider, K., Lamosa, P., Herrgård, M. J., Rosenstand, I., Öberg, F., Forster, J., and Nielsen, J. (2015) Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in Saccharomyces cerevisiae via ??-alanine. *Metab. Eng. 27*, 57−64.

(3) Subramanian, S., and Srienc, F. (1996) Quantitative analysis of transient gene expression in mammalian cells using the green fluorescent protein. *J. Biotechnol. 49*, 137−51.

(4) Leveau, J. H. J., and Lindow, S. E. (2001) Predictive and Interpretive Simulation of Green Fluorescent Protein Expression in Reporter Bacteria. *J. Bacteriol. 183*, 6752.

(5) Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005) Tuning genetic control through promoter engineering. *Proc. Natl. Acad. Sci. U. S. A. 102*, 12678.

(6) Qi, L., Haurwitz, R. E., Shao, W., Doudna, J. a., and Arkin, A. P. (2012) RNA processing enables predictable programming of gene expression. *Nat. Biotechnol. 30*, 1002−1006.

(7) Kelly, J. R., Rubin, A. J., Davis, J. H., Ajo-Franklin, C. M., Cumbers, J., Czar, M. J., de Mora, K., Glieberman, A. L., Monie, D. D., and Endy, D. (2009) Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng. 3*, 4.

(8) Ronen, M., Rosenberg, R, Shraiman, B. I., and Alon, U. (2002) Assigning numbers to the arrows: Parameterizing a gene regulation network by using accurate expression kinetics. *Proc. Natl. Acad. Sci. U. S. A. 99*, 10555.

(9) Gerosa, L., Kochanowski, K., Heinemann, M., and Sauer, U. (2013) Dissecting specific and global transcriptional regulation of bacterial gene expression. *Mol. Syst. Biol. 9*, 658.

(10) Keren, L., Zackay, O., Lotan-Pompan, M., Barenholz, U., Dekel, E., Sasson, V., Aidelberg, G., Bren, A., Zeevi, D., Weinberger, A., Alon, U., Milo, R., and Segal, E. (2013) Promoters maintain their relative activity levels under different growth conditions. *Mol. Syst. Biol. 9*, 701.

(11) Rudge, T. J., Brown, J. R., Federici, F., Dalchau, N., Phillips, A., Ajioka, J. W., and Haseloff, J. (2016) Characterization of Intrinsic Properties of Promoters. *ACS Synth. Biol. 5*, 89.

(12) Bar-Joseph, Z., Gitter, A., and Simon, I. (2012) Studying and modelling dynamic biological processes using time-series gene expression data. *Nat. Rev. Genet. 13*, 552.

(13) Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014) Hierarchy of non-glucose sugars in Escherichia coli. *BMC Syst. Biol. 8*, 133.

(14) Görke, B., and Stülke, J. (2008) Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol. 6*, 613−24.

(15) Monod, J. (1966) From enzymatic adaptation to allosteric transitions. *Science 154*, 475−483.

(16) Gancedo, J. M. (1992) Carbon catabolite repression in yeast. *Eur. J. Biochem. 206*, 297−313.

(17) Monod, J. (1941) Recherches sur la croissance des cultures bactériennes, Ph.D. thesis, Université de Paris.

(18) Peng, B., Williams, T. C., Henry, M., Nielsen, L. K., and Vickers, C. E. (2015) Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift: a comparison of yeast promoter activities. *Microb. Cell Fact.*, DOI: 10.1186/s12934-015-0278-5.

(19) Madar, D., Dekel, E., Bren, A., Zimmer, A., Porat, Z., and Alon, U. (2013) Promoter activity dynamics in the lag phase of Escherichia coli. *BMC Syst. Biol. 7*, 136.

(20) Miksch, G., and Dobrowolski, P. (1995) Growth phase-dependent induction of stationary-phase promoters of Escherichia coli in different gram-negative bacteria. *J. Bacteriol. 177*, 5374−5378.

(21) Yu, X., Xu, J., Liu, X., Chu, X., Wang, P., Tian, J., Wu, N., and Fan, Y. (2016) Identification of a highly efficient stationary phase promoter in Bacillus subtilis. *Sci. Rep. 5*, 18405.

(22) Canton, B., Labno, A., and Endy, D. (2008) Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol. 26*, 787−793.

(23) Borodina, I., and Nielsen, J. (2014) Advances in metabolic engineering of yeast Saccharomyces cerevisiae for production of chemicals. *Biotechnol. J. 9*, 609−620.

(24) Bhasi, K., Forrest, A., and Ramanathan, M. (2005) SPLINDID: A semi-parametric, model-based method for obtaining transcription rates and gene regulation parameters from genomic and proteomic expression profiles. *Bioinformatics 21*, 3873−3879.

(25) Finkenstädt, B., Heron, E. A., Komorowski, M., Edwards, K., Tang, S., Harper, C. V., Davis, J. R. E., White, M. R. H., Millar, A. J., and Rand, D. A. (2008) Reconstruction of transcriptional dynamics from gene reporter data using differential equations. *Bioinformatics 24*, 2901−2907.

(26) Brown, D., and Lostroh, C. P. (2008) Inferring gene expression dynamics from reporter protein levels. *Biotechnol. J. 3*, 1437−1448.

(27) De Jong, H., Ranquet, C., Ropers, D., Pinel, C., and Geiselmann, J. (2010) Experimental and computational validation of models of fluorescent and luminescent reporter genes in bacteria. *BMC Syst. Biol. 4*, 55.

(28) Bren, A., Hart, Y., Dekel, E., Koster, D., and Alon, U. (2013) The last generation of bacterial growth in limiting nutrient. *BMC Syst. Biol. 7*, 27.

(29) Zulkower, V., Page, M., Ropers, D., Geiselmann, J., and De Jong, H. (2015) Robust reconstruction of gene expression profiles from reporter gene data using linear inversion. *Bioinformatics 31*, i71−i79.

(30) R Core Team (2016) *R: A Language and Environment for Statistical Computing*, https://www.r-project.org/.

(31) Huang, Z., Senocak, F., Jayaraman, A., and Hahn, J. (2008) Integrated modeling and experimental approach for determining transcription factor profiles from fluorescent reporter data. *BMC Syst. Biol. 2*, 64.

(32) Zuleta, I. A., Aranda-Díaz, A., Li, H., and El-Samad, H. (2014) Dynamic characterization of growth and gene expression using high-throughput automated flow cytometry. *Nat. Methods 11*, 443−8.

(33) Wang, X., Wu, M., Li, Z., and Chan, C. (2008) Short time-series microarray analysis: Methods and challenges. *BMC Syst. Biol. 2*, 58.

(34) Elowitz, M., Levine, a., Siggia, E., and Swain, P. (2002) Stochastic gene expression in a single cell. *Science 297*, 1183−6.

(35) Kaern, M., Elston, T. C., Blake, W. J., and Collins, J. J. (2005) Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet. 6*, 451−464.

(36) Zopf, C. J., Quinn, K., Zeidman, J., and Maheshri, N. (2013) Cell-Cycle Dependence of Transcription Dominates Noise in Gene Expression. *PLoS Comput. Biol. 9*, e1003161.

(37) Ramsay, J. O., Wickham, H., Graves, S., and Hooker, G. (2014) *fda: Functional Data Analysis*, https://cran.r-project.org/package=fda.

(38) Ripley, B., and Ramsey, J. (2015) *pspline: Penalized Smoothing Splines*, https://cran.r-project.org/package=pspline.

(39) Picotti, P., Bodenmiller, B., Mueller, L. N., Domon, B., and Aebersold, R. (2009) Full Dynamic Range Proteome Analysis of S. cerevisiae by Targeted Proteomics. *Cell 138*, 795−806.

(40) Beal, J., et al. (2016) Reproducibility of fluorescent expression from engineered biological constructs in E. coli. *PLoS One 11*, 1−22.

(41) Cardinale, S., and Arkin, A. P. (2012) Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. *Biotechnol. J. 7*, 856−866.

(42) Mutalik, V. K., Guimaraes, J. C., Cambray, G., Mai, Q. A., Christoffersen, M. J., Martin, L., Yu, A., Lam, C., Rodriguez, C., Bennett, G., Keasling, J. D., Endy, D., and Arkin, A. P. (2013) Quantitative estimation of activity and quality for collections of functional genetic elements. *Nat. Methods 10*, 347−353.

(43) Hecht, A., Endy, D., Salit, M., and Munson, M. S. (2016) When Wavelengths Collide: Bias in Cell Abundance Measurements Due to Expressed Fluorescent Proteins. *ACS Synth. Biol. 5*, 1024−1027.

(44) Nielsen, A. A. K., Der, B. S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E. A., Ross, D., Densmore, D., and Voigt, C. A. (2016) Genetic circuit design automation. *Science 352*, aac7341.

(45) Milo, R., and Phillips, R. (2015) *Cell Biology by the Numbers*, Garland Science.

(46) Hebisch, E., Knebel, J., Landsberg, J., Frey, E., and Leisner, M. (2013) High Variation of Fluorescence Protein Maturation Times in Closely Related Escherichia coli Strains. *PLoS One 8*, e75991.

(47) Iizuka, R., Yamagishi-Shirasaki, M., and Funatsu, T. (2011) Kinetic study of de novo chromophore maturation of fluorescent proteins. *Anal. Biochem. 414*, 173−178.

(48) Andersen, J. B., Sternberg, C., Poulsen, L. K., Bjørn, S. P., Givskov, M., and Molin, S. (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl. Environ. Microbiol. 64*, 2240.

(49) Bernstein, J. A., Khodursky, A. B., Lin, P.-H., Lin-Chao, S., and Cohen, S. N. (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. U. S. A. 99*, 9697.

(50) Chen, H., Shiroguchi, K., Ge, H., and Sunney Xie, X. (2015) Genome-wide study of mRNA degradation and transcript elongation in Escherichia coli. *Mol. Syst. Biol. 11*, 781.

(51) Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999) Growth Phase-Dependent Variation in Protein Composition of the Escherichia coli Nucleoid. *J. Bacteriol. 181*, 6361−6370.

(52) Mallik, P., Pratt, T. S., Beach, M. B., Bradley, M. D., Undamatla, J., and Osuna, R. (2004) Growth Phase-Dependent Regulation and Stringent Control of fis Are Conserved Processes in Enteric Bacteria and Involve a Single Promoter (fis P) in Escherichia coli. *J. Bacteriol. 186*, 122.

(53) Ninnemann, O., Koch, C., and Kahmann, R. (1992) The E. coli fis promoter is subject to stringent control and autoregulation. *EMBO J. 1*, 1075−1083.

(54) Paige, J. S., Wu, K. Y., and Jaffrey, S. R. (2011) RNA Mimics of Green Fluorescent Protein. *Science 333*, 642−646.

(55) Filonov, G. S., Moon, J. D., Svensen, N., and Jaffrey, S. R. (2014) Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc. 136*, 16299−16308.

(56) Strack, R. L., Disney, M. D., and Jaffrey, S. R. (2013) A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nat. Methods 10*, 1219−24.

(57) Pothoulakis, G., Ceroni, F., Reeve, B., and Ellis, T. (2014) The Spinach RNA aptamer as a characterisation tool for synthetic biology. *ACS Synth. Biol. 3*, 182.

(58) Shin, S. Y., Jung, S. M., Kim, M. D., Han, N. S., and Seo, J. H. (2012) Production of resveratrol from tyrosine in metabolically engineered Saccharomyces cerevisiae. *Enzyme Microb. Technol. 51*, 211−216.

(59) Lan, E. I., and Liao, J. C. (2013) Microbial synthesis of n-butanol, isobutanol, and other higher alcohols from diverse resources. *Bioresour. Technol. 135*, 339−349.

(60) Wanka, F., Arentshorst, M., Cairns, T. C., Jørgensen, T., Ram, A. F. J., and Meyer, V. (2016) Highly active promoters and native secretion signals for protein production during extremely low growth rates in Aspergillus niger. *Microb. Cell Fact. 15*, 145.

(61) Meyer, P. (1994) Bi-logistic growth. *Technol. Forecast Soc. Change 47*, 89−102.

(62) Soetaert, K., Petzoldt, T., and Setzer, R. W. (2010) Solving Differential Equations in R: Package deSolve. *J. Stat. Soft.*, DOI: 10.18637/jss.v033.i09.

(63) Mateus, C., and Avery, S. V. (2000) Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. *Yeast 16*, 1313−1323.

(64) Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995) Improved green fluorescence. *Nature 373*, 663−664.

(65) Bansal, L., Chu, Y., Laird, C., and Hahn, J. (2012) Determining Transcription Factor Profiles from Fluorescent Reporter Systems involving Regularization of Inverse Problems, *Proceedings of 2012 American Control Conference (ACC 2012)*, pp 2725−2730.

(66) pSB1C3. (2017) http://www.snapgene.com/resources/plasmid_files/basic_cloning_vectors/pSB1C3/.

(67) Kreuzer, P., Gartner, D., Allmansberger, R., and Hillen, W. (1989) Identification and sequence analysis of the Bacillus subtilis W23 xylR gene and xyl operator. *J. Bacteriol. 171*, 3840−3845.

(68) Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene 173*, 33.