ACS Synthetic Biology

Research Article

pubs.acs.org/synthbio

Characterization of Intrinsic Properties of Promoters

Timothy J. Rudge,[†,⊥] James R. Brown,[†,§,⊥] Fernan Federici,[†,‡,⊥] Neil Dalchau,[§] Andrew Phillips,[§] James W. Ajioka,[∥] and Jim Haseloff[*,†]

[†]Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, United Kingdom

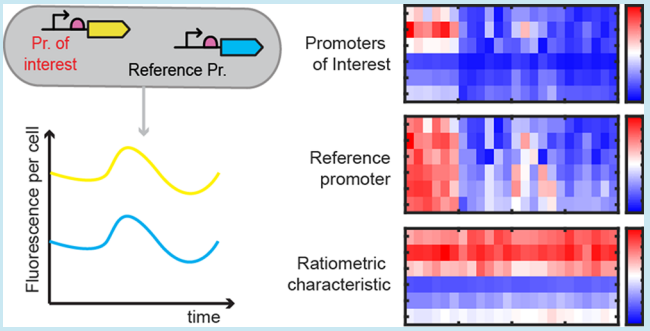
[‡]Departamento de Genética Molecular y Microbiologia, Faculty de Ciencias Biológicas, Pontificia Universidad Católica de Chile, Santiago 8331150, Chile

[§]Biological Computation Group, Microsoft Research, Cambridge CB1 2FB, United Kingdom

[∥]Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, United Kingdom

Supporting Information

ABSTRACT: Accurate characterization of promoter behavior is essential for the rational design of functional synthetic transcription networks such as logic gates and oscillators. However, transcription rates observed from promoters can vary significantly depending on the growth rate of host cells and the experimental and genetic contexts of the measurement. Furthermore, *in vivo* measurement methods must accommodate variation in translation, protein folding, and maturation rates of reporter proteins, as well as metabolic load. The external factors affecting transcription activity may be considered to be extrinsic, and the goal of characterization should be to obtain quantitative measures of the intrinsic characteristics of promoters. We have developed a promoter characterization method that is based on a mathematical model for cell growth and reporter gene expression and exploits multiple *in vivo* measurements to compensate for variation due to extrinsic factors. First, we used optical density and fluorescent reporter gene measurements to account for the effect of differing cell growth rates. Second, we compared the output of reporter genes to that of a control promoter using concurrent dual-channel fluorescence measurements. This allowed us to derive a quantitative promoter characteristic (ρ) that provides a robust measure of the intrinsic properties of a promoter, relative to the control. We imposed different extrinsic factors on growing cells, altering carbon source and adding bacteriostatic agents, and demonstrated that the use of ρ values reduced the fraction of variance due to extrinsic factors from 78% to less than 4%. This is a simple and reliable method to quantitatively describe promoter properties.



KEYWORDS: *promoter, transcription, characterization, ratiometric, fluorescence, quantification, design, modeling*

A major aim of synthetic biology is the rational design and construction of DNAs composed of genetic parts from various sources to achieve defined novel functions. Promoters, regulatory proteins, operators, and translation initiation elements from bacteria and bacteriophages have been combined to build transcription networks or circuits encoding toggle switches,[1] oscillators,[2] logic,[3] and simple computation[4] in *Escherichia coli*. While these studies demonstrate the potential of the synthetic biology approach, the scope of designed genetic systems remains limited due to lack of reliable data on the behavior of genetic parts. There are usually many candidate parts that could be used to build a given genetic network topology, and the designer must select those likely to achieve the desired function. As the number of available parts increases, exhaustive testing or trial-and-error become infeasible, and quantitative modeling becomes essential. Part characterization is then required to parametrize these models in such a way that they are predictive of practical genetic network operation. However, characterization of promoters is problematic because their behavior can vary unpredictably in different contexts.

This variation is partly due to dependence on the host cell. Genetic parts and host strains are usually chosen to minimize specific regulatory interactions between native and synthetic circuits, for example, by avoiding CAP binding sites in promoters.[5] However, broad utilization of host resources cannot be avoided. For example, initiation of transcription from a promoter sequence generally depends on the availability of the cell's native RNA polymerase (RNAP), associated sigma factors, and RNA nucleotides. Translation of transcribed mRNAs into proteins, including fluorescent reporters, requires host ribosomes, tRNAs, and amino acids. Empirical correlations among chromosome and plasmid copy number, ribosome and RNAP levels, and growth rate have been identified, which lead to fluctuation in gene expression.[6−8] DNA sequence local to genetic parts can also significantly affect their behavior. In particular, the activity of promoters has been shown to depend

Received: August 24, 2015
Published: October 5, 2015

DOI: 10.1021/acssynbio.5b00116
ACS Synth. Biol. 2016, 5, 89−98

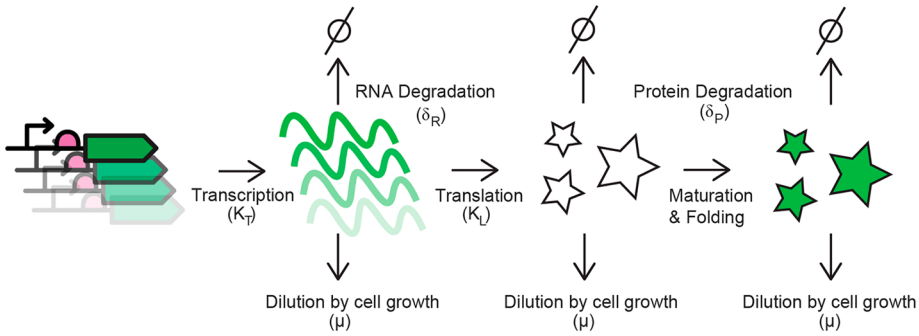


Figure 1. Steps required for fluorescent reporter synthesis. (Left−right) Transcription is initiated at promoter sequences in each copy of a reporter gene at rate K_T, the resulting mRNA is degraded at rate δ_R and diluted by cell growth (μ), and translation of mRNAs into immature fluorescent proteins occurs at rate K_L, followed by folding and maturation into active fluorescent reporters. Proteins are also degraded (δ_P) and diluted by cell growth (μ).

on the adjacent transcript sequence.[9] The mechanistic details of these relationships are not fully understood. For the goal of characterization, this means that both the promoter of interest and any indirect (fluorescence, luminescence, colorimetric) reporter measurements of its activity are subject to unpredictable variation. Moreover, expression of a reporter causes metabolic load that will likely affect the operation of the promoter under study.

Relative measurement is a common approach to reducing variation in measurements and has been applied to promoter characterization. In higher organisms, our laboratory has applied relative measurement of promoters to account for differences between cell types and the accessibility of different tissues to measurement.[10] In bacteria, Kelly et al.[11] measured the activity of a set of promoters in *E. coli* growing under different conditions and hosted in different strains. Each promoter was measured individually in separate experiments under each growth condition and strain. One of these promoters was chosen as a reference, and its mean activity was used to normalize the other promoters in the set. The result was a relative measure of promoter activity with lower variance than absolute promoter activity. More recently, Keren et al.[12] screened a library of around 1800 *E. coli* promoters expressing GFP in 10 different growth media. Supporting the results of Kelly et al., they found that the activity of 70−90% of the promoters was scaled by a constant factor when changing growth conditions. Furthermore, they showed evidence that promoters deviating from this global scaling were those specifically regulated by the change in conditions, e.g., metabolic operons affected by choice of carbon source.

In summary, several studies have suggested that variation in the activity of constitutive promoters is largely due to global sources that preserve their relative levels of activity and that specific regulation of promoters is observed as a change in this relative activity.[10−12] However, both Kelly et al.[11] and Keren et al.[12] measured the activity of promoters individually in separate experiments and computed relative activities between experiments. Such measurements cannot capture global variation due to the metabolic load of an introduced synthetic gene circuit, slight differences in growth conditions, or the initial state of inoculated cells. Previous work in our laboratory developed dual-reporter plasmids to enable concurrent ratiometric measurement of promoter pairs[13] and a computational analysis method[14] that further reduced variance in measures of promoter activity. Note that none of these approaches to characterization explicitly addressed variation during the time period that a promoter is active, but they considered only peak transcription. Hence, characterization of promoters based on existing methods is subject to variation due to extrinsic factors.

Intrinsic Promoter Characteristics. Characterization is the process of estimating quantitative measures of part behavior, which we call characteristics. Promoters drive transcription, so any characteristic of its behavior must relate to transcription rate:

Promoter Characteristic: a quantitative measure of transcription rate obtained from a given promoter sequence.

However, transcription rates are not intrinsic to promoters because they are highly dependent on the context in which they are measured: DNA molecule, host cell, and experimental conditions. Transcription rates are also very difficult to measure directly and nondestructively *in vivo*, meaning that indirect reporters, such as fluorescent proteins, are most often employed. These reporters depend on processes that are subject to variation not specific to the promoter driving transcription or necessarily correlated with transcription rate (Figure 1). Both the context of measurement and the measurement system itself, therefore, introduce extrinsic variation to estimated promoter characteristics. In this study, we seek a measure of promoter activity that is reliable in spite of this extrinsic variation, that is, an intrinsic promoter characteristic:

Intrinsic Promoter Characteristic: a quantitative measure of transcription that is specific to a given promoter and consistent in a range of contexts.

Here, we describe the systematic development of a method for *in vivo* characterization of promoters based on dual-channel measurement of fluorescent reporters. Using this method, we derived a ratiometric promoter characteristic that reduced the fraction of variance due to extrinsic factors to less than 4%, suggesting that it is intrinsic to the promoter.

RESULTS

Simultaneous Measurement of Two Promoters in a Single Replicon. Promoter activity is commonly measured from fusions upstream of a reporter gene.[11,12,15] Fluorescent reporter genes are convenient because they encode a single protein and do not require substrates or precursors to generate a measurable signal. In this approach, the promoter sequence is placed upstream of a ribosome binding site (RBS) and fluorescent protein coding sequence (CDS), either in a plasmid or integrated into a genomic locus. However, the specific sequence to which a promoter is fused can significantly affect its

DOI: 10.1021/acssynbio.5b00116
ACS Synth. Biol. 2016, 5, 89−98

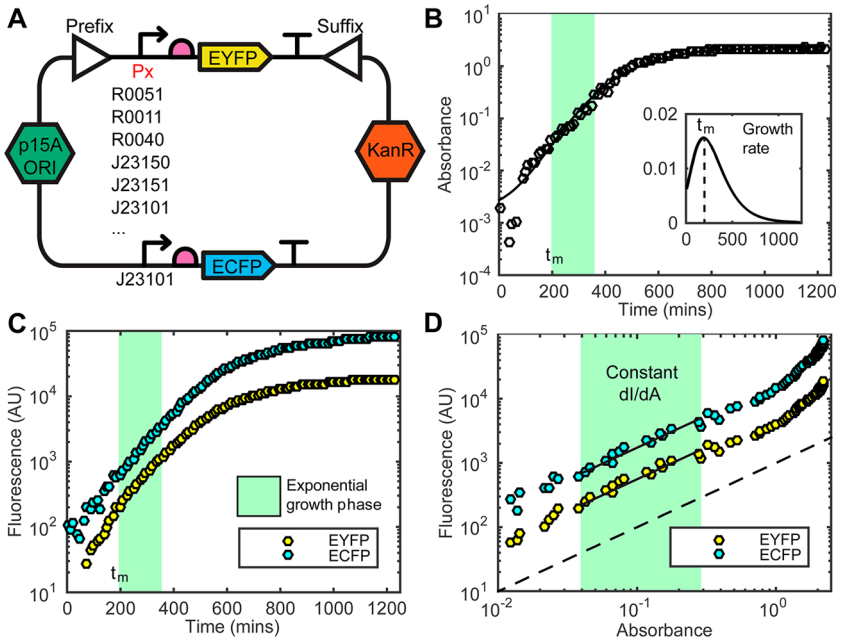


Figure 2. Dual-channel concurrent measurement of promoter activity. (A) Reporter plasmid contains one of six promoters of interest and a reference promoter (J23101). Each promoter is fused to a fluorescent protein gene (EYFP/ECFP) containing a common ribosome binding sequence (purple semicircle) and bidirectional terminator. (B) Absorbance at 600 nm as a measure of culture biomass; black line shows fitted Gompertz model, and the inset is growth rate. (C) Plate-reader fluorescence intensity measurements from pJ23151, with the defined period of exponential phase marked. (D) Plotting fluorescence intensity ($I(t)$) against absorbance ($A(t)$) shows a clear linear relation in exponential phase (green). Log−log plot with dashed line indicating linear relation.

activity[9] and will determine the translation rate of the reporter protein being measured. Synthesis of mature fluorescent protein is required to generate a measurable fluorescence signal. This is a multistep process (Figure 1) beginning with transcription of mRNA, followed by translation, folding, and, finally, maturation. At each step, degradation and dilution by cell growth act to reduce cellular concentrations. Promoter activity itself and the fluorescence measurements used to characterize it are, thus, dependent on the context of local DNA sequence and host cell.

We designed promoter−reporter fusions to provide similar local DNA sequence context by incorporating a common RBS (BBa_R0034)[16] immediately adjacent to the transcription start site of the promoter of interest. Six promoters of interest were chosen following Kelly et al.[11] These promoters are "synthetic" derivatives of Lambda phage promoter PL (R0051, R0011, R0040) and *de novo* synthetic promoters (J23101, J23150, J23151) and were taken from the Registry of Standard Biological Parts.[16] To test the effect of fluorescent protein choice, we measured the activity of these promoters upstream of GFPmut3, mCherry, EYFP, and ECFP (Figure S1A, Supporting Information, and analysis method below). We found that EYFP, ECFP, and mCherry preserved rank order and that EYFP and ECFP also preserved relative magnitudes of promoter activities. GFPmut3 deviated from the other reporters, especially for apparently strong promoters R0051 and R0011. EYFP and ECFP are reported to be similar in maturation half-lives (39 ± 7 and 49 ± 9 min, respectively),[17] stability (half-lives > 24 h),[17] and transcript sequence (19 nucleotide substitutions). Combined with their good spectral separation (ex./em. 514/527 nm for EYFP and 434/477 nm for ECFP), these findings suggested that EYFP/ECFP would make a good pair for comparative analysis.

We assembled the six constitutive promoters mentioned above into dual-channel reporter plasmids (Figure 2A). Each promoter of interest was fused to the EYFP reporter and common RBS. Following Kelly et al.,[11] promoter J23101 was chosen as a reference and fused to the ECFP reporter and common RBS in reverse orientation in the plasmid backbone. Each reporter gene was transcriptionally isolated with a bidirectional terminator (BBa_B0015). This plasmid design enabled concurrent measurement of two promoter−reporter fusions with the same gene dosage (plasmid copy) and very similar local DNA contexts (RBS and CDS) for each promoter. We refer to these plasmids according to the name of the promoter of interest as p<EYFP promoter>, e.g., pR0051. In the following, we analyze measurements of *E. coli* cultures carrying each of the six plasmids under a range of growth conditions.

Calculating Fluorescent Protein Synthesis Rate from Time-Course Data. The dual-channel reporter plasmids described above allowed us to measure concurrently both culture absorbance ($A(t)$, Figure 2B) and fluorescence channels (EYFP and ECFP, Figure 2C) in a microplate fluorometer. Here, we outline the derivation of estimates of the rate of synthesis of fluorescent proteins based on these measurements. We show how this rate is related to the transcription rate. At each step, we highlight assumptions made in the analysis and techniques to avoid amplification of errors in the required computations.

As described in previous studies,[12,18−20] the multistep process (Figure 1) of mature fluorescent protein synthesis can be summarized by a time-varying synthesis rate for each cell $F_p(t)$. With a stable protein P (half-life approximately 24 h for EYFP and ECFP),[17] we may neglect degradation ($\delta_p = 0$), resulting in the following differential equation

DOI: 10.1021/acssynbio.5b00116
ACS Synth. Biol. 2016, 5, 89−98

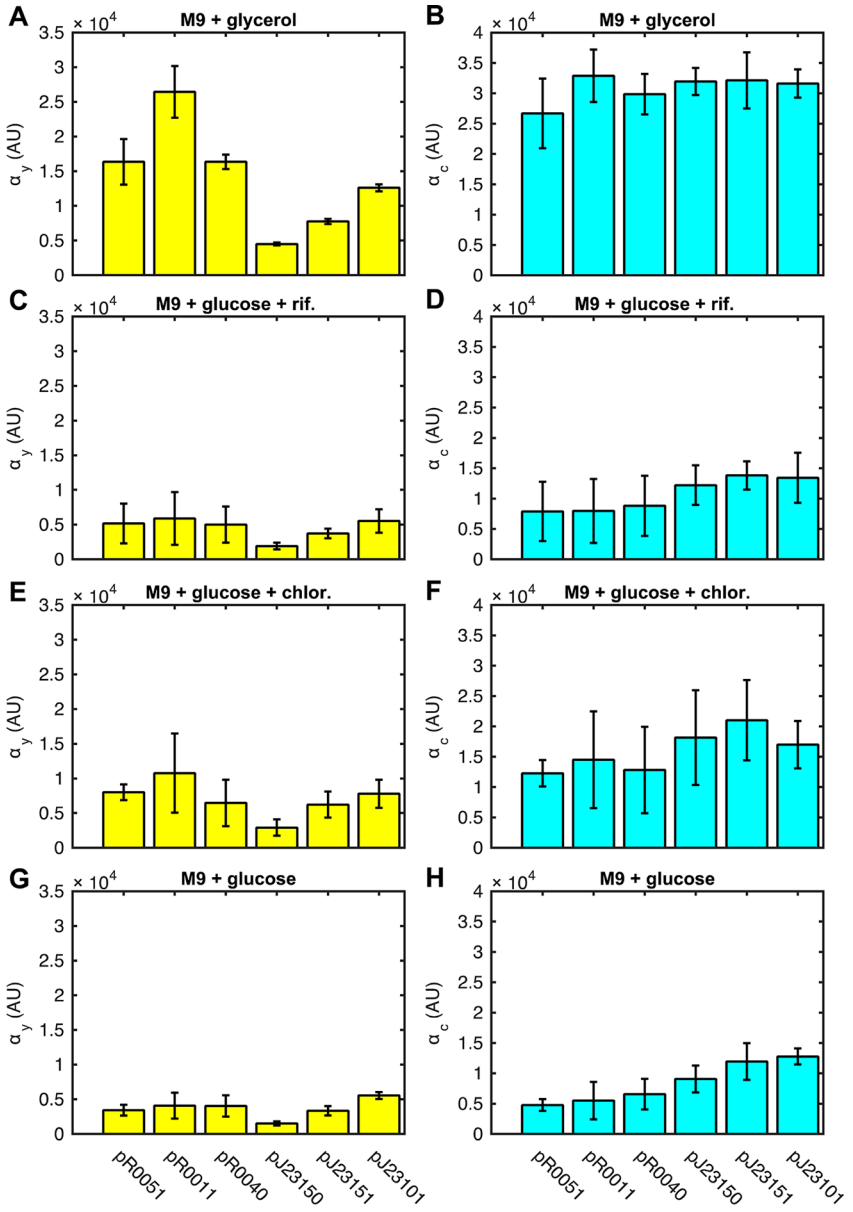


Figure 3. Promoter characteristics α_p measured from dual channel plasmids in four growth conditions. EYFP (α_y) and ECFP (α_c) promoter characteristics measured from cells growing in M9 + 0.2% glycerol (A, B), M9 + 0.4% glucose + 1 μg/mL chloramphenicol (C, D), M9 + 0.4% glucose + 1 μg/mL rifampicin (E, F), and M9 + 0.4% glucose (G, H). Error bars show one standard deviation.

$$\frac{dP}{dt} = F_p(t) - \mu(t)\,P(t)$$

where the second term represents dilution by cell growth at average rate $\mu(t) = 1/A(t)\ dA/dt$, and $A(t)$ is culture density measured by absorbance (Figure 2B). Assuming that fluorescence is detected linearly, and that absorbance is a good measure of cell number, then measured fluorescence intensity (Figure 2C) is given by

$$I_p(t) = A(t)\,P(t)$$

From this simple model, the usual expression for the protein synthesis rate of each cell can be derived[21,22]

$$F_p(t) = \frac{1}{A(t)}\frac{dI_p}{dt} \quad (1)$$

However, calculation of $F_p(t)$ presents several technical problems. Early in time-course experiments, culture density $A(t)$ is very low and thus subject to significant background noise (see Supporting Information), which is amplified in computing $1/A(t)$. Furthermore, fluorescence signal during this period is also low, and computation of dI_p/dt amplifies measurement noise. This period of low culture density corresponds to lag and exponential growth phases. σ^{70} promoters, such as those examined here, are known to be most active during exponential growth,[6] and characterization of their behavior must focus on this time period. It is, therefore, important to accurately identify the period of exponential growth phase and to quantify promoter activity during this time.

In order to enable accurate calculation of $F_p(t)$, we note that eq 1 can be rewritten using the chain rule as

DOI: 10.1021/acssynbio.5b00116
ACS Synth. Biol. 2016, 5, 89−98

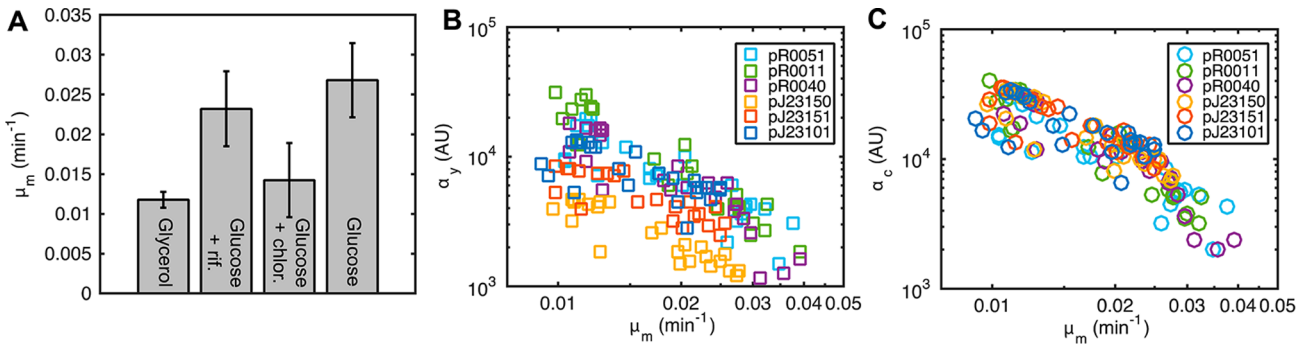


Figure 4. Promoter characteristic α_p and peak growth rate μ_m in different growth media. (A) Mean and standard deviation (error bars) of peak growth rate (μ_m) for cells containing the six dual-channel plasmids growing in four different growth media. (B, C) Promoter characteristics for (B) the promoter of interest (α_y) and (C) the reference promoter (α_c) for each plasmid measured in the four growth media showed a clear negative relation to peak growth rate.

$$F_p(t) = \frac{1}{A(t)}\frac{dA}{dt}\frac{dI_p}{dA} = \mu(t)\frac{dI_p}{dA} \quad (2)$$

where $\mu(t)$ is the growth rate of the culture. The form of eq 2 states an explicit dependence of gene expression on growth rate, which has been indicated by several previous studies in *E. coli*.[6,8,12,18,23] The fluorescent protein synthesis rate $F_p(t)$ is often used as a proxy for transcription rate, assuming that other processes occur at fixed rates,[12,18−20] but as illustrated in Figure 1, it also incorporates extrinsic variation due to gene copy number and translation and maturation rates.

Fluorescent Protein Synthesis Rate Is Proportional to Growth Rate in Exponential Phase. The analysis above shows that the relation between fluorescence intensity $I(t)$ and culture optical density $A(t)$ may in itself be informative of time variation in promoter activity. For all plasmids used in this study, we found that the slopes (dI_c/dA) and (dI_y/dA) remain approximately constant until growth falls off during the transition to stationary phase. In Figure 2D, we illustrate this relation by plotting measurements of fluorescence $I_c(t)$ (ECFP) and $I_y(t)$ (EYFP) against the corresponding $A(t)$ for one of the six plasmids. The shapes of the curves $I_c(A)$ and $I_y(A)$ show that fluorescent protein synthesis is proportional to growth rate during exponential growth phase. The constant of proportionality is given by $\alpha_p = dI_p/dA$, where p indicates the fluorescent reporter EYFP or ECFP. This is the relative rate of fluorescent protein synthesis to biomass synthesis as measured by culture density.

In order to quantify this relation, we needed to identify the time period of exponential growth phase, which is associated with the peak in culture growth rate $\mu(t)$. Accurate estimation of growth rates suffers from the issues with noise amplification described above, due to both differentiation of and dividing by small noisy $A(t)$ measurements. We therefore fit a Gompertz model[14,24] to measurements of $A(t)$ (Figure 2B, black line). The Gompertz model applied to bacterial culture growth is parametrized by lag time (λ), peak growth rate (μ_m), and carrying capacity (K), and the time of peak growth is given directly as $t_m = K/e\mu_m + \lambda$ (where $e = \exp(1)$). To avoid problems with low signal during early culture growth, we considered exponential growth phase as the period from peak growth (t_m) extending for four doubling periods (4 ln $2/\mu_m$). Exponential phase measurements identified in this way are highlighted in green in Figure 2. Despite large variation in growth rate (e.g., see Figure 2B, inset) during exponential phase, the slope of fluorescence against optical density (α_p) remained constant. We confirmed this relation for all experiments shown in Figure 3 (see Supporting Information) by linear regression, giving $R^2 = 0.97 \pm 0.039$ for both EYFP and ECFP.

Hence, the value of α_p quantifies fluorescent protein synthesis in relation to growth rate, parametrizing a simple linear model of fluorescent protein synthesis rate

$$F_p(t) = \alpha_p\mu(t)$$

We computed the values of α_y and α_c for the six dual-reporter plasmids in cells grown in M9 minimal media with glycerol (Figure 3A,B). For each experiment, plates were inoculated from two separate colonies, each with three replicates ($N = 6$ in total). The coefficient of variation (CV) of alpha estimates for each plasmid ranged from 4 to 20% for the promoter of interest (EYFP) and 7 to 21% for the reference (ECFP) promoter (see Supporting Information Table S1). Analysis of covariance (ANCOVA) of the α_y and α_c values estimated by regression for each plasmid showed significant differences between replicates ($p < 0.05$). This means that variance was not solely due to the analysis method and that this method could distinguish between replicates at the 5% significance level. For a given growth condition, then, α_p was characteristic of promoter activity despite large changes in fluorescent protein synthesis over exponential growth phase.

The relative rate of fluorescent protein synthesis to growth thus gives a promoter characteristic α_p that is consistent over exponential phase and that eliminates growth rate as a source of extrinsic variation.

Fluorescent Protein Synthesis Rate Varies Significantly under Different Growth Conditions. We showed above that we could estimate promoter characteristics (α_p) for a given growth condition (M9 with glycerol) with a CV of <22%. We next considered the effect of different growth conditions. If growth rate were the major source of extrinsic variation,[8,12] then we would expect the values of α_y and α_c to be similar under different conditions. This is because they parametrize the linear relation between fluorescent protein synthesis and growth $F_p(t) = \alpha_p\mu(t)$. We repeated the estimation of α_y and α_c for each plasmid in M9 media with glucose as the main carbon source and with the addition of bacteriostatic drugs, rifampicin and chloramphenicol (Figure 3C−H). These growth conditions introduced extrinsic factors to reporter expression, and use of our dual-channel plasmid allowed us to examine concurrently their effect on the promoter of interest and reference promoter. We found dramatic

DOI: 10.1021/acssynbio.5b00116
ACS Synth. Biol. 2016, 5, 89−98

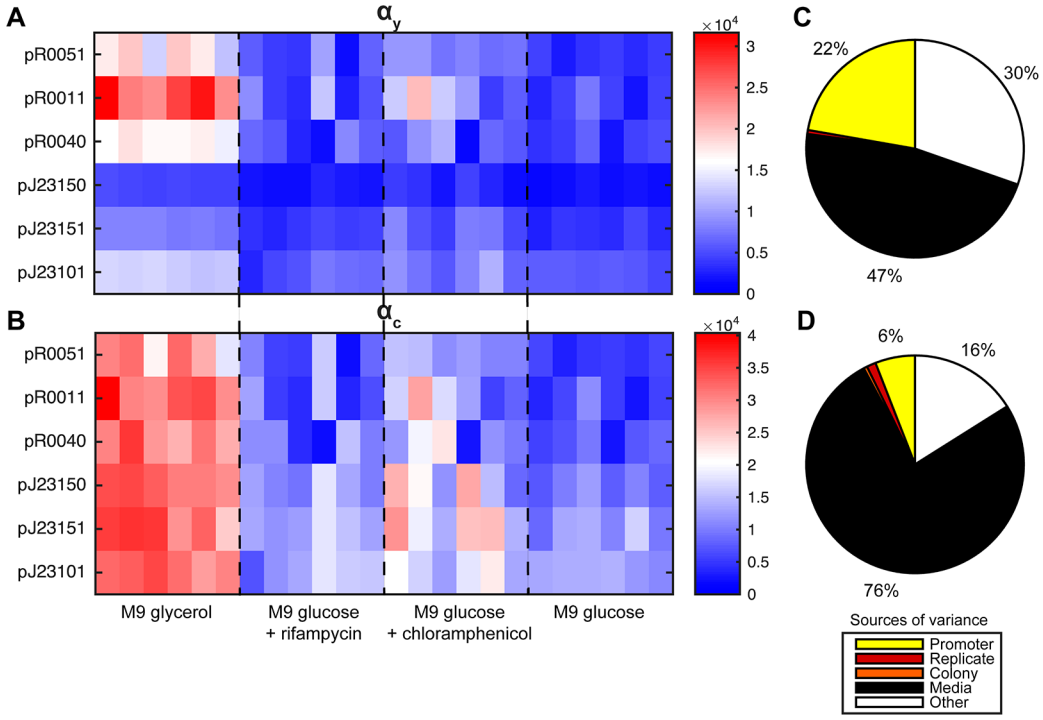


Figure 5. Heat maps of promoter characteristics α_p from six plasmids in four growth conditions. (A, B) Heat maps of promoter characteristics α_y (A) and α_c (B) from each of six plasmids (y-axis) with six measurements (three replicates of two colonies) under each of four growth conditions (x-axis). (C, D) Analysis of variance (four-way ANOVA) shows that single-channel characterization of the promoter of interest (α_y) is not robust to media changes (22% of variance attributed to plasmid or promoter identity). Pie chart sectors not labeled were <1%.

differences in promoter characteristics (both α_y and α_c) under these growth conditions (Figure 3A–H) despite maintenance of the linear relation $F_p(t) = \alpha_p\mu(t)$ (see Supporting Information for regressions).

Clearly, α_p is not an intrinsic characteristic of a promoter even over the limited range of conditions we tested in this study. Peak growth rates (μ_m) in each of the four conditions we tested were significantly different (ANOVA, $p < 10^{-36}$), but the estimated α_y and α_c tended to decrease under conditions leading to faster growth (Figure 4). This result is somewhat surprising given the literature on correlations between growth rate and gene expression, and it highlights the complexity of interactions among growth rate, gene/plasmid copy, and transcription and translation rates.

To quantify the reliability of promoter characteristics, we performed analysis of variance (four-way ANOVA) on the estimated α_y and α_c from the experiments shown in Figure 3. In this analysis, we partitioned the variance in these measures of promoter activity into that due to the identity of the promoter of interest (plasmid), experimental replicate, inoculating colony, and the four growth conditions tested (Figure 5). In each ANOVA test (for α_y and α_c), all factor effects were significant ($p < 0.01$), but their contribution to variance was very different for each measure. Ideally, a robust characteristic of promoters would be subject to variance only due to the identity of the promoter being measured and would be invariant to nonspecific factors.

For α_y (Figure 5A,C), we found that only 22% of variance was attributed to the identity of the promoter being measured ($N = 24$), and the largest contribution to variance was due to changing growth conditions (47%, $N = 36$). A further 30% of the variance in α_y was not attributed to the four factors included in the analysis and represents some combination of technical (due to equipment) and biological (due to cells) variations. For α_c (Figure 5B,D), variance was dominated by growth conditions (76%, $N = 36$), with little effect (6%, $N = 24$) from the identity of the host plasmid, i.e., the EYFP promoter. Hence, in our experiments, the activity of the reference promoter was largely independent of the promoter of interest. Similar to the EYFP channel, α_c was also subject to variance from sources unspecified (16%).

Mathematical Derivation of Promoter Characteristics from Fluorescent Protein Synthesis Rates. We now outline a simple model of the observed variation in promoter characteristic α_p, the rate of fluorescent protein synthesis relative to growth, and in the following section, we use it to derive an intrinsic promoter characteristic.

Previous studies have shown that promoter activity is a saturating function of growth rate[6,8] or is proportional to growth rate.[12,23] These observations are consistent with each other at subsaturating growth rates such as those that might be expected in minimal media. Note that in these reports the growth rate variation considered was between experimental conditions rather than over the population growth cycle. Our results confirm the observed dependence of promoter activity on growth rate during exponential growth phase under the given growth conditions. However, comparing estimated promoter characteristics (α_p) between experimental conditions showed a decreasing trend in relation to peak growth rate (Figure 4B,C).

The empirical relation between promoter activity and growth rate suggests that growth rate is an indicator of limiting factors (or resources) common to both biomass production and promoter activity (e.g., RNA polymerase holoenzyme).[23] Intuitively, as limiting factors increase, growth rate will eventually saturate to some maximum. Denoting limiting

DOI: 10.1021/acssynbio.5b00116
ACS Synth. Biol. 2016, 5, 89–98

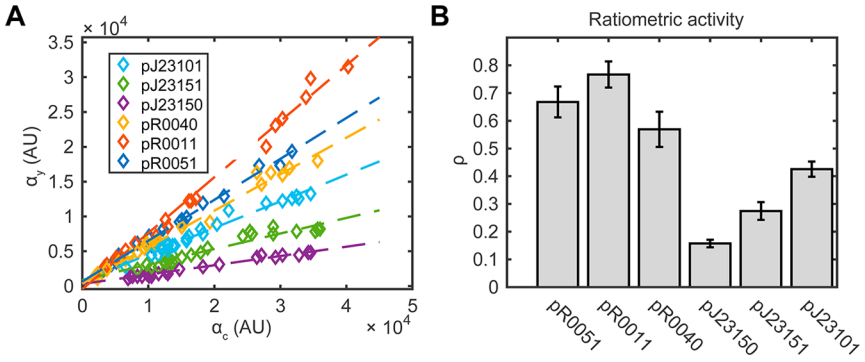


Figure 6. Promoter characteristics from dual-channel plasmids were well-correlated across all growth conditions. (A) Characteristics of promoter of interest (y-axis) and reference promoter (x-axis) for each plasmid (colors) in all four growth conditions. Lines show regression fits ($R > 0.95$ in all cases). (B) The ratios of promoter characteristics ($\rho = \alpha_y/\alpha_c$) were maintained in all growth conditions (CVs < 12%, $N = 24$; error bars show standard deviation).

factors as $R(t)$, we represent saturation of growth rate with a Michaelis–Menten equation

$$\mu(t) = \mu^*\left(\frac{kR(t)}{1 + kR(t)}\right)$$

where μ^* is the theoretical maximum growth rate with increasing $R(t)$ and k is the rate of use of limiting factors in growth. In the case of low $R(t)$, that is, $kR(t) \ll 1$, this simplifies to

$$\mu(t) \approx \mu^* kR(t)$$

Solving for $R(t)$ gives

$$R(t) = \frac{1}{k}\frac{\mu(t)}{\mu^*}$$

Similarly, promoter activity is also dependent on limiting factors $R(t)$, and again at low values of $R(t)$, we have the transcription rate (Figure 1)

$$K_T(t) \approx K_T^* mR(t) = K_T^*\frac{m}{k}\frac{\mu(t)}{\mu^*}$$

where m is the rate of use of limiting factors in promoter activity and K_T^* is the theoretical maximum transcription rate at saturation ($mR(t) \gg 1$). Hence, promoter activity is proportional to growth rate. With short mRNA half-lives (typically, ~2 min), we can assume quasi-steady-state, and assuming first-order degradation at fixed rate δ_R, we have mRNA concentration

$$M(t) = \frac{K_T(t)}{\mu(t) + \delta_R}$$

Approximating translation as a first-order process and assuming that maturation is in steady-state, the synthesis rate of mature fluorescent proteins from each cell then follows

$$F_p(t) = \varphi_p n(t)\, K_L(t)\left(\frac{K_T(t)}{\mu(t) + \delta_R}\right)$$

and substituting for transcription rate gives

$$F_p(t) = \varphi_p n(t)\, K_L(t)\left(\frac{1}{\mu(t) + \delta_R}\right)K_T^*\frac{m}{k}\frac{\mu(t)}{\mu^*}$$

where φ_p is the fraction of fluorescent proteins in the mature state and the terms $n(t)$ and $K_L(t)$ are the time-varying plasmid copy number and translation rate, respectively. Relating this model to the equation for fluorescent protein synthesis rate given in eq 2 gives an expression for the promoter characteristic

$$\alpha_p = \varphi_p n(t)\, K_L(t)\left(\frac{1}{\mu(t) + \delta_R}\right)K_T^*\frac{m}{k\mu^*} \quad (3)$$

This model predicts that promoter characteristics (α_y and α_c) are subject to variation from a number of sources. First, gene copy $n(t)$ and translation efficiency $K_L(t)$ can vary in different media and over time[8] due to availability of ribosomes and other limiting factors.[7] Allocation of resources to growth (μ^* and k) would likely depend on growth media and other experimental conditions. Finally, despite accounting for growth rate dependence of promoter activity, α_p is predicted to decrease with increasing growth rate ($\mu(t)$) and maximal growth rate (μ^*), as we found from our results in different media (Figure 4). However, it seems reasonable to assume that the affinity of a promoter for available resources (m) and its maximal attainable activity (K_T^*) are specific or intrinsic characteristics and are not dependent on growth conditions.

Now consider promoter characteristics measured from our dual-channel plasmids. Since both promoters are encoded on the same plasmid, we can assume that gene copy ($n(t)$) is common. Similarly, with concurrent *in vivo* measurement, we have promoters active in the same cells, which exhibit particular allocation of resources to growth (μ^* and k). Furthermore, the close similarity in our fluorescent reporter transcripts (ECFP and EYFP) suggests that translation efficiency $K_L(t)$ and mRNA degradation rate δ_R would be similar. Our model then predicts that the characteristics of the promoter of interest (α_y) and the reference promoter (α_c) across a range of conditions should be correlated due to common sources of variation. Confirming this prediction, across the four growth conditions we tested, with three replicates of two colonies in each condition, variation in the estimated promoter characteristics (Figure 3A–H) for EYFP and ECFP was closely correlated ($R > 0.95$ for each plasmid; Figure 6A). Flow cytometry of cells grown in M9 + 0.4% glucose also indicated correlation at the single-cell level ($R > 0.7$) for all plasmids (see Supporting Information).

Ratiometric Measurement Approximates Intrinsic Characteristics of Promoters. The model derived above shows that the fluorescent reporter-based promoter character-

DOI: 10.1021/acssynbio.5b00116
ACS Synth. Biol. 2016, 5, 89–98

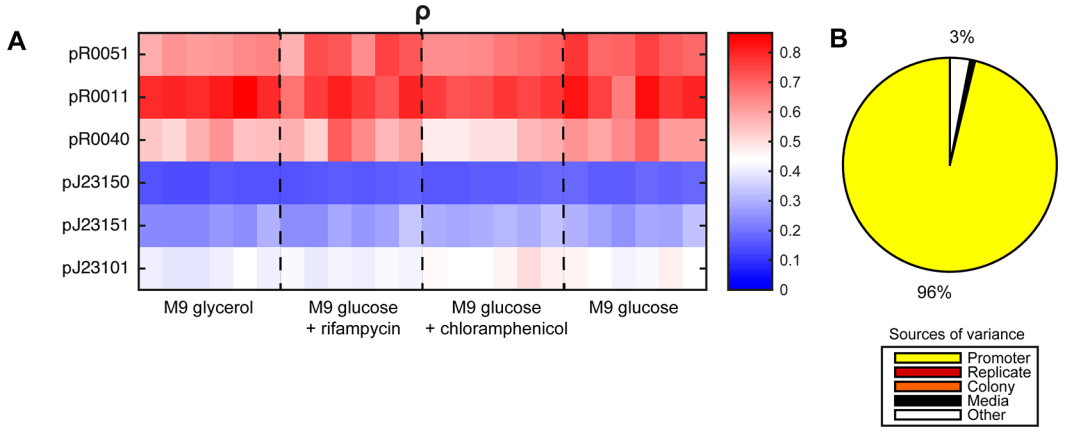


Figure 7. Heat map of ratiometric promoter characteristic ρ from six plasmids under four growth conditions. (A) Heat map of ratiometric promoter characteristics $\rho = \alpha_y/\alpha_c$ from each of six plasmids (y-axis), with six measurements (three replicates of two colonies) under each of four growth conditions (x-axis). (B) Analysis of variance (four-way ANOVA) shows that normalization to an *in vivo* reference promoter provides a reliable characteristic with 96% of variance explained by promoter identity. Pie chart sectors not labeled were <1%.

istic α_p is a product of intrinsic factors (m and K_T^*) and extrinsic factors due to mRNA degradation, translation, fluorescent protein maturation, and cell growth. Without accounting for all extrinsic factors, individual promoter characteristics (α_y and α_c) were not reliable under changing growth conditions. In the absence of direct measurement of most of these extrinsic factors, eq 3 suggests that we can use our dual-channel measurements to extract an intrinsic promoter characteristic via the ratio

$$\rho = \frac{\alpha_y}{\alpha_c} = \frac{\varphi_y K_{T,y}^* m_y}{\varphi_c K_{T,c}^* m_c} \approx \frac{K_{T,y}^* m_y}{K_{T,c}^* m_c}$$

where we have assumed common copy number ($n(t)$), translation rate ($K_L(t)$), and mRNA degradation rate (δ_R), consistent with our dual-channel plasmid design (Figure 2). A further reasonable approximation given their similar maturation half-lives is that the fractions of EYFP and ECFP in the mature state are the same ($\varphi_y \approx \varphi_c$).

We computed the ratio ρ for the experiments (Figure 3) in which each plasmid was measured under four growth conditions. Even though single-channel characteristics of promoter activity varied significantly (around 4-fold) among these conditions, the computed ratios showed low coefficients of variation (Figure 6B; CVs 6−12%). Analysis of variance (four-way ANOVA as above, $p < 0.01$ for all factors) showed that 96% ($N = 24$) of variance was due to the identity of the promoter of interest (Figure 7). Variance due to growth conditions was largely eliminated (<1%, $N = 36$), and unspecified sources of variance were also reduced (3%). In separate experiments, we measured the six plasmids in cells growing in LB rich media and found that estimates of ρ agreed closely with those measured in M9 minimal media (see Supporting Information).

Thus, we define ρ as an intrinsic ratiometric promoter characteristic that reliably quantifies the behavior of a promoter with respect to an *in vivo* reference.

DISCUSSION

Synthetic biology aims to create an engineering discipline for the design of functional genetic circuits. Promoters often perform critical functions in such circuits,[3,4] so obtaining reliable quantitative characteristics of their activity is essential to this design process. In this work, we studied fluorescent reporters fused to previously well-characterized constitutive promoters carried on plasmids. Using microplate fluorometer measurements, we highlighted technical issues affecting estimation of promoter activity from time-series data. We developed a simple analytical approach to computing promoter activity that overcame several of these issues to give accurate characteristics. This analysis explicitly revealed the dependence of promoter activity computation on growth rate during exponential phase as a linear model. The slope of this linear model (α_p) gave a promoter characteristic that was largely invariant to approximately 4-fold changes in growth rate over exponential phase.

We then subjected cells to extrinsic factors known to affect promoter activity by growing them with different carbon sources and in the presence of bacteriostatic drugs. While α_p was reliable over exponential growth phase under a given experimental condition, with these additional extrinsic factors we observed significant variation. Furthermore, this variation showed a negative relation to the peak growth rate observed under each test growth condition. Statistical analysis showed that 78% of the variance in characteristics was attributed to extrinsic factors, of which 47% was due to the imposed extrinsic factor and 30% to unidentified sources. Only 22% of variance was attributed to the identity of the promoter.

Our design of dual-channel fluorescent reporter plasmids allowed us to concurrently measure and compute promoter characteristics (α_y, α_c) from a promoter of interest and a reference in very similar genetic contexts. We derived a simple mathematical model of these promoter characteristics. This model confirmed our observation that α_p decreased with increasing peak growth rate. It also suggested that promoter characteristics α_y and α_c would be correlated due to common extrinsic factors. This was supported by our data in which promoter characteristics measured from the same plasmid were closely correlated across all conditions tested ($R > 0.95$). This analysis then predicted that the ratio $\alpha = \alpha_y/\alpha_c$ would be consistent when cells were subjected to extrinsic factors. Our data showed that the fraction of variance in promoter characteristics attributed to imposed extrinsic factors was reduced from 47% to less than 1% after computing the ratio $\rho = \alpha_y/\alpha_c$. Other unidentified sources of variance were reduced from 30% to just 3% of variance. Overall, the identity of the

DOI: 10.1021/acssynbio.5b00116
ACS Synth. Biol. 2016, 5, 89−98

promoter accounted for 96% of the variance, and only 4% was due to extrinsic factors. We therefore propose the ratiometric characteristic (ρ) with respect to an *in vivo* reference as an intrinsic promoter characteristic.

We showed that common extrinsic factors dominated the variation in constitutive transcription from promoters hosted on the same plasmid and that ratiometric characteristics were largely unaffected by this variation. The mechanisms by which extrinsic factors affect promoters and fluorescent reporter measurements of them are currently poorly understood and thus cannot be incorporated into quantitative models. Our approach utilized an *in vivo* constitutive reference promoter placed in a genetic context as similar as possible to that of the promoter of interest, with the same RBS and almost identical fluorescent reporter coding sequence. This context-matched reference promoter, therefore, provided a live concurrent readout of common extrinsic effects on promoter activity. This means that, in principle, ratiometric characteristics could be used to estimate the activity of promoters operating under novel conditions from measurements of the reference promoter alone.

While this study focused on constitutive promoters, ratiometric characterization can also enable accurate quantification of specifically regulated promoters (e.g., inducible) by minimizing the effects of common extrinsic factors and revealing specific regulatory effects on promoters measured under different conditions. These conditions might, for example, be concentrations of inducers (e.g., IPTG, aTC) that bind transcription factors associated with promoters (e.g., LacI, TetR). Such approaches will be essential to the accurate design of functional genetic circuits by parametrizing models of transcription regulation. Previous work from our laboratory demonstrated this approach to modeling of a homoserine lactone regulated promoter.[14]

The results presented here were based on microplate fluorometer measurements of bacteria in bulk culture. However, the principle of ratiometric promoter characterization could be applied to other techniques in which concurrent dual-channel fluorescence measurements can be made and even to multicellular organisms. Previous work from our laboratory successfully applied a similar approach to confocal microscopy images of plant tissues.[10] Thus, we have presented an approach to characterization that enables simple and reliable quantification of promoter activity under a range of conditions and organisms and that we hope will further progress toward rational design of synthetic genetic circuits.

METHODS

Protocol. A step-by-step protocol for ratiometric characterization of a promoter of choice using our plasmids is given in Supporting Information.

Microbial Strains and Growth Conditions. *E. coli* strain EC10G (Invitrogen) was used for all experiments. Growth media were based on M9 minimal medium[25] with 0.2% (w/v) casamino acids, kanamycin (50 μg mL^{-1}) for selection, and four different additional supplement combinations:

(A) 0.4% (w/v) glucose
(B) 0.4% (w/v) glucose +1 μg mL^{-1} chloramphenicol
(C) 0.4% (w/v) glucose +1 μg mL^{-1} rifampicin
(D) 0.2% (w/v) glycerol

Plasmids. All constructs used in this study were constructed from BioBricks obtained from the Registry of Biological Parts distribution kit (Registry of Standard Biological Parts, MIT, http://partsregistry.org) and assembled into pSB3K3CY. pSB3K3CY was created by cloning BBa_J23101 promoter, ribosome binging site BBa_B0034, cyan fluorescent protein BBa_E0020, and BBa_B0015 terminator into pSB3K3 vector backbone between the kanamycin resistance gene and p15A origin of replication by Gibson assembly.[26] Each of the promoters used in this study, R0051, R0011, R0040, J23101, J23150, and J23151, was fused to RBS BBa_B0034, EYFP BBa_E0030, and bidirectional terminator BBa_B0015; these cassettes were subsequently cloned between prefix and suffix sequences of pSB3K3CY using BioBrick assembly (sequences listed in Supporting Information).

For assembly, a master mix was prepared by combining 100 μL of 5× isothermal reaction buffer, 2 μL of 1 U μL^{-1} T5 exonuclease (Epicenter), 6.25 μL of 2 U μL^{-1} Phusion DNA polymerase (Thermo Scientific), 50 μL of 40 U mL^{-1} Taq DNA ligase (NEB), and water up to a final volume of 375 μL. 15 μL of this reagent−enzyme mix were added to purified DNA fragments totaling 5 μL and incubated for 1 h at 50 °C. 5× isothermal buffer was prepared following Gibson et al.[26] EC10G chemically competent cells were transformed by heat shock and plated on LB agar plates with kanamycin (50 μg mL^{-1}).

Plasmids are available from AddGene.

Plate Fluorometry Assays. Each of the plasmids described above (Figure 2A) was transformed into chemically competent *E. coli* strain EC10G (Invitrogen) and incubated overnight in LB agar plates containing kanamycin (50 μg mL^{-1}) for selection. Next, two colonies of each of these transformations were selected and inoculated into 5 mL of one of the M9 media (A−D, see above) and grown overnight in a shaking incubator at 37 °C for approximately 16 h. Cultures were then diluted 1:100 into fresh identical medium. Then, 200 μL of this diluted culture was added in three replicates to each well of a black 96-well microplate with a clear bottom (Greiner). A BMG Fluostar Omega plate reader was used to measure optical density at 600 nm and fluorescence every ∼12 min. Excitation filter 430/10 nm and emission filter 480/10 nm were used for measuring ECFP, whereas excitation filter 500/10 nm and emission filter 530/10 nm were used for EYFP measurements. The plate was maintained at 37 °C during the measurement assay. Between readings, plates were shaken at 200 rpm.

Data Analysis. Matlab (Mathworks) was used for all data analysis, and a custom python script was used to import data from the BMG spreadsheet format (all code is available from www.github.com/timrudge/platypus). Gompertz models were fitted to OD data using the "nlinfit" matlab function, which implements the Levenberg−Marquardt algorithm.[27] Statistical analysis was carried out with Matlab functions that implement standard methods for linear regression, ANOVA, and ANCOVA ("polyfit", "anovan", and "aoctool").

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.5b00116.

Single-channel promoter characterization plasmid design, promoter activity measured from single-channel plasmids, fluorescence against OD measured during exponential phase growth, statistics of promoter characteristics, promoter characterisation from cells

DOI: 10.1021/acssynbio.5b00116
ACS Synth. Biol. 2016, 5, 89−98

with ratiometric plasmids, flow cytometry of cells with ratiometric plasmids, plasmid maps and sequence information, and step-by-step guide to dual-channel promoter characterization (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: jh295@cam.ac.uk.

Author Contributions

⊥T.J.R., J.R.B., and F.F. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

T.J.R. was supported by a Microsoft Research studentship and by EC FP7 project no. 612146 (PLASWIRES) awarded to J.H., J.R.B. was supported by a Microsoft Research studentship and internship, and F.F. was supported by CONICYT-PAI/Concurso Nacional de Apoyo al Retorno de Investigadores/as desde el Extranjero Folio 82130027 andby EPSRC grant EP/H019162/1 awarded to J.H. J.W.A. acknowledges the EPSRC and the Wellcome Trust for support. The authors would like to thank Anton Kan and Nigel Miller for assistance with flow cytometry. We would like to thank Rodrigo Gutierrez (PUC, Chile) for support and useful comments.

REFERENCES

(1) Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000) Construction of a genetic toggle switch in Escherichia coli. *Nature 403*, 339−342.

(2) Danino, T., Mondragón-Palomino, O., Tsimring, L., and Hasty, J. (2010) A synchronized quorum of genetic clocks. *Nature 463*, 326−330.

(3) Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C., and Voigt, C. A. (2012) Genetic programs constructed from layered logic gates in single cells. *Nature 491*, 249−253.

(4) Daniel, R., Rubens, J. R., Sarpeshkar, R., and Lu, T. K. (2013) Synthetic analog computation in living cells. *Nature 497*, 619−623.

(5) Lutz, R., and Bujard, H. (1997) Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res. 25*, 1203−1210.

(6) Klumpp, S., and Hwa, T. (2008) Growth-rate-dependent partitioning of RNA polymerases in bacteria. *Proc. Natl. Acad. Sci. U. S. A. 105*, 20245−20250.

(7) Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010) Interdependence of cell growth and gene expression: origins and consequences. *Science 330*, 1099−1102.

(8) Klumpp, S., Zhang, Z., and Hwa, T. (2009) Growth rate-dependent global effects on gene expression in bacteria. *Cell 139*, 1366−1375.

(9) Lou, C., Stanton, B., Chen, Y. J., Munsky, B., and Voigt, C. A. (2012) Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol. 30*, 1137−1142.

(10) Federici, F., Dupuy, L., Laplaze, L., Heisler, M., and Haseloff, J. (2012) Integrated genetic and computation methods for in planta cytometry. *Nat. Methods 9*, 483−485.

(11) 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.

(12) 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.

(13) Brown, J. (2013) A design framework for self-organised Turing patterns in microbial populations. Ph.D. Thesis, Department of Plant Sciences, University of Cambridge.

(14) Yordanov, B., Dalchau, N., Grant, P. K., Pedersen, M., Emmott, S., Haseloff, J., and Phillips, A. (2014) A Computational Method for Automated Characterization of Genetic Components. *ACS Synth. Biol. 3*, 578−588.

(15) 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.

(16) *Registry of Standard Biological Parts*, BioBricks Foundation, San Francisco, CA.

(17) Gordon, A., Colman-Lerner, A., Chin, T. E., Benjamin, K. R., Yu, R. C., and Brent, R. (2007) Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nat. Methods 4*, 175−181.

(18) 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.

(19) Berthoumieux, S., de Jong, H., Baptist, G., Pinel, C., Ranquet, C., Ropers, D., and Geiselmann, J. (2013) Shared control of gene expression in bacteria by transcription factors and global physiology of the cell. *Mol. Syst. Biol. 9*, 634.

(20) De Jong, H. (2002) Modeling and simulation of genetic regulatory systems: a literature review. *J. Comput. Biol. 9*, 67−103.

(21) Boyer, F., Besson, B., Baptist, G., Izard, J., Pinel, C., Ropers, D., Geiselmann, J., and De Jong, H. (2010) WellReader: a MATLAB program for the analysis of fluorescence and luminescence reporter gene data. *Bioinformatics 26*, 1262−1263.

(22) 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*, 10555−10560.

(23) Liang, S.-T., Bipatnath, M., Xu, Y.-C., Chen, S.-L., Dennis, P., Ehrenberg, M., and Bremer, H. (1999) Activities of constitutive promoters in Escherichia coli. *J. Mol. Biol. 292*, 19−37.

(24) Zwietering, M. H., Jongenburger, I., Rombouts, F. M., and Van't Riet, K. (1990) Modeling of the bacterial growth curve. *Appl. Environ. Microbiol. 56*, 1875−1881.

(25) (1999) *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., Eds.) John Wiley & Sons, New York.

(26) Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods 6*, 343−345.

(27) (2002) *Numerical Recipes in C++: The Art of Scientific Computing* (Press, W. H., Ed.) Cambridge University Press, Cambridge.

DOI: 10.1021/acssynbio.5b00116
ACS Synth. Biol. 2016, 5, 89−98