Many experiments show that the numbers of mRNA and protein are proportional to the cell volume in growing cells. However, models of stochastic gene expression often assume constant transcription rate per gene and constant translation rate per mRNA, which are incompatible with these experiments. Here, we construct a minimal gene expression model to fill this gap. Assuming ribosomes and RNA polymerases are limiting in gene expression, we show that the numbers of proteins and mRNAs both grow exponentially during the cell cycle and that the concentrations of all mRNAs and proteins achieve cellular homeostasis; the competition between genes for the RNA polymerases makes the transcription rate independent of the genome number. Furthermore, by extending the model to situations in which DNA (mRNA) can be saturated by RNA polymerases (ribosomes) and becomes limiting, we predict a transition from exponential to linear growth of cell volume as the protein-to-DNA ratio increases.
Despite the noisy nature of gene expression\textsuperscript{1–6}, various aspects of single cell dynamics, such as volume growth, are effectively deterministic. Recent single-cell measurements show that the growth of cell volume is often exponential. These include bacteria\textsuperscript{7–9}, archaea\textsuperscript{10}, budding yeast\textsuperscript{10,12–15} and mammalian cells\textsuperscript{10,16}. Moreover, the mRNA and protein numbers are often proportional to the cell volume throughout the cell cycle: the homeostasis of mRNA concentration and protein concentration is maintained in an exponentially growing cell volume with variable genome copy number\textsuperscript{17–22}. The exponential growths of mRNA and protein number indicate dynamical transcription and translation rates proportional to the cell volume, rather than the genome copy number. However, current gene expression models often assume constant transcription rate per gene and constant translation rate per mRNA (constant rate model)\textsuperscript{1,5,23}. Assuming a finite degradation rate of mRNAs and non-degradable proteins, these models lead to a constant mRNA number proportional to the gene copy number and linear growth of protein number\textsuperscript{26–28}, incompatible with the proportionality of mRNA and protein number to the exponentially growing cell volume.

Since the cell volume, protein copy number and mRNA copy number grow exponentially throughout the cell cycle, one may expect a sufficient condition to achieve a constant concentration is to let them grow with the same exponential growth rate. However, mathematical analysis suggests this is insufficient. Let us consider the logarithm of protein concentration $c$, which can be written as $\ln(c) = \ln(p) - \ln(V)$. Here $p$ is the protein number and $V$ is the cell volume. If one assumes the protein number and the cell volume grow exponentially but independently, with time-dependent exponential growth rates $\lambda_p(t)$ and $\lambda_V(t)$ respectively, the time derivative of the logarithm of concentration then obeys $d/\ln(c)/dt = \lambda_p(t) - \lambda_V(t)$. Even when the time-averaged growth rates of protein number and cell volume are equal, $\langle \lambda_p(t) \rangle = \langle \lambda_V(t) \rangle$, any fluctuations in the difference between them will accumulate and lead to a random walk behavior of the logarithm of concentration. The homeostasis of protein and mRNA concentrations implies that there must be a regulatory mechanism in place to prevent the accumulation of noise over time.

The main goal of this work is to identify such a mechanism by developing a coarse-grained model taking into account cell volume growth explicitly. Specially, we only consider continuously proliferating cells and do not take account of non-growing cells, e.g., bacterial cells in stationary phase\textsuperscript{29}. The ubiquity of homeostasis suggests that the global machinery of gene expression, RNA polymerases (RNAPs) and ribosomes, should play a central role within the model. Based on the assumption that the number of ribosomes is the limiting factor in translation, we find that the exponential growth of cell volume and protein number originates from the auto-catalytic nature of ribosomes\textsuperscript{30–33}. The fact that ribosomes make all proteins ensures that the protein concentrations do not diverge. Based on the assumption that the number of RNAP is the limiting factor in transcription, we find that the mRNA number also grows exponentially and the mRNA concentration is independent of the genome copy number because of the competition between genes for this global resource\textsuperscript{18–20}. We also study the effects of genome replication. Due to the heterogeneous timing of gene replication, the transcription rate of one gene has a cell cycle dependence. Within our model, it doubles immediately after the gene is replicated and decreases gradually as other genes are replicated. Nevertheless, we find that this leads to a small effect on protein levels. Finally, we extend our model to more general situations in which an excess of RNAP (ribosome) leads to the saturation of DNA (mRNA). We propose a phase diagram of gene expression and cellular growth controlled by the protein-to-DNA ratio. We predict a transition from exponential growth to linear growth of cell volume as the protein-to-DNA ratio passes a threshold.

### Results

**Model of stochastic gene expression.** In constant rate models, the transcription rate per gene and the translation rate per mRNA are constant\textsuperscript{1,5,24} (Fig. 1a). Constant rate models predict a constant mRNA number proportional to the gene copy number and independent of the cell volume. However, experimental observations on plant and mammalian cells have revealed a proportionality between mRNA number and cell volume for cells with a constant genome copy number\textsuperscript{18–20}. Moreover, even

![Fig. 1](image-url) The growing cell model of stochastic gene expression in comparison with constant rate models. a In the constant rate model, the transcription rate is proportional to the gene copy number, and the translation rate is proportional to the mRNA number. These assumptions imply that the gene number and mRNA number are the limiting factors in gene expression. b In Phase 1 of the growing cell model, we introduce as limiting factors RNA polymerases (RNAPs) and ribosomes. Genes with different colors are transcribed with different rates. Here $k_0$ is a constant and the gene regulation is coarse-grained into the gene allocation fraction $\phi_i = g_i / \sum_j g_j$, $g_i$ is the effective copy number of gene $i$ (also accounting for the promoter strength). $n$ is the total number of RNAPs. Translation rates of mRNA depend on the number of active ribosomes ($f_r$), the translation rate $k_0$, and the fraction of mRNA $i$ in the total pool of mRNA. In a later section (A unified phase diagram of gene expression and cellular growth), we will relax our assumptions and consider situations in which the limiting factors of gene expression become the gene number and the mRNA number.
comparing the cells before and after the genome replication (S phase), the proportionality coefficient between mRNA and cell volume does not exhibit any obvious change. In contrast, a constant transcription rate per gene would predict a doubled transcription rate after the replication of the whole genome, leading to a higher mRNA concentration. In one class of constant rate models, a deterministic exponential growth of cell volume is explicitly considered. The resulting perturbation on the concentrations due to genome replication is suppressed in the long lifetime limit, but still significant for short lifetime molecules, e.g., mRNA (see Fig. 1 in ref.25).

Considering translation, various experiments have shown that the number of ribosomes is the limiting factor rather than the number of mRNAs. The most direct evidence is the growth law: the growth rate of cells is proportional to the fraction of ribosomal proteins in the total proteome (with a constant factor depending on the growth condition)35 both for bacterial cells36,37, budding yeast cells38, and mammalian cells39,40. This means a constant fraction of ribosomes are actively translating mRNAs. These results suggest that in general cells are below the saturation limit in which there are too many ribosomes that the mRNAs can bind. We will therefore assume the biological situation in which mRNAs in the cell compete for the limiting resource of actively translating ribosomes, therefore the translation rate of one type of mRNA is proportional to the number of active ribosomes times its fraction in the total pool of mRNAs.

Considering transcription, experiments have shown that mutants of fission yeasts altered in cell size regulated global transcription to maintain similar transcription rates per cell volume regardless of the cellular DNA content. The changes in total transcription correlated with coordinated changes in gene occupancy by RNA polymerases37. These results suggest that the number of RNAPs may be the limiting factor in transcription rather than the gene number, and similar evidence has been shown for bacterial cells38 and mammalian cells39. However, in the same experiments on fission yeast37, it has also been found that in cell-cycle-arrested mutants, total transcription rates stopped increasing as the cell volume exceeded a certain value, which suggested DNA became limiting for transcription at low DNA concentration. This result suggests that an excess of RNAPs may lead the gene number to become the limiting factor in certain conditions. In this section, we will focus on the scenario that both RNAP and ribosome are limiting in gene expression, which we denote as Phase 1. In this phase, we will show that the mRNA number and the protein number are proportional to the cell volume and grow exponentially. In a later section (A unified phase diagram of gene expression and cellular growth), we will consider a more general model in which the limiting nature of RNAPs and ribosomes may break down and the dynamics of mRNA and protein number is different.

To address the limiting nature of RNAP, we define an effective gene copy number $g_i$ for each gene to account for its copy number and the binding strength of its promoter, which determines its ability to compete for RNAPs. The transcription rate for one specific gene $i$ is proportional to the fraction of RNAPs that are working on its gene(s), $\phi_i = g_i / \sum_j g_j$, which we denote as the gene allocation fraction. Gene regulation is thus coarse-grained into the gene allocation fraction $\phi_i$. The transcription rate is independent of the genome copy number since a change in the genome number leaves the allocation fraction of one gene invariant, a conclusion which is consistent with a number of experimental results on various organisms18–20,37.

In fact, explicit gene regulation can also be included in our model (Methods), with a time-dependent $g_i$. In such scenarios, $g_i$ may be a function of protein concentrations (for instance, the action of transcription factors modifies the transcription rate).

Such models will lead to more complex dynamics of mRNA and protein concentrations. However, since we are interested in the global behavior of gene expression and cell volume growth, we do not focus on these complex regulations in this manuscript. Our conclusions regarding the exponential growth of mRNA and protein number for constitutively expressed genes and the exponential growth of cell volume on the global level are not affected by the dynamics of gene expression of particular genes.

In the following, $m_i, p_i, r_i$ represents the numbers of mRNAs, proteins, ribosomes and RNA polymerases, respectively. Proteins ($p$) also include RNAPs ($n$) and ribosomes ($r$)40. We consider the degradation of mRNA with degradation time $r$ for all genes. The protein number decreases only through cell divisions (though adding a finite degradation rate for proteins does not affect our results). The stochastic dynamics of gene expression within Phase 1 of our model are summarized in the following sets of equations and Fig. 1b, 

$$m_i \frac{d}{dt} m_i = m_i/r_i \rightarrow m_i - 1$$

$$p_i \frac{d}{dt} p_i = p_i + 1$$

$$k_i \left( \frac{m_i}{\sum_j m_j} \right) \phi_i$$

Here $k_0, k_r$ are constants, characterizing the transcription (translation) rate of a single RNAP (ribosome). $\phi_i$ is the fraction of active ribosomes, which we assume to be constant in a given nutrient environment30,32. We note that nonspecifically bound RNAPs have been reported in bacteria40,41. We will discuss their effect later. For simplicity, we first assume the values of $\phi_i$ do not change in time. This can be formally thought of as corresponding to an instantaneous replication of the genome. In reality, a finite duration of DNA replication and the varying time of replication initiation for different genes lead to $\phi_i$'s that change during the DNA replication. We later analyze a more complete version of the model which includes these gene dosage effects, but we first consider the simplified scenario of constant $\phi_i$ that will capture the essential features of the problem.

We assume the cell volume is approximately proportional to the total protein mass, i.e., $V \propto M = \sum_j p_j$, which is a good approximation for bacteria42,43 and mammalian cells17. To simplify the following formulas, we consider each protein has the same mass and set the cell density as 1.

Due to the fast degradation of mRNA compared with the cell cycle duration44,45, the mRNA number can be well approximated as being in steady state. We can express the ensemble-averaged number of mRNA of gene $i$ as

$$\langle m_i(t) \rangle = k_0 \phi_i \langle n(t) \rangle \tau.$$
averages of two protein numbers in the steady state is set by the ratio of their production rate, therefore \( \frac{p_i}{p_j} = \phi_i/\phi_j \). The average number of mRNA traces the number of RNA polymerases according to Eq. (4), and therefore also grows exponentially. Throughout the cell cycle we have

\[
\langle m_i(t) \rangle = m_b(i) \exp(\mu t),
\]

(6a)

\[
\langle p_i(t) \rangle = p_b(i) \exp(\mu t),
\]

(6b)

where \( m_b(i) (p_b(i)) \) is the number of mRNA (protein) of gene \( i \) at cell birth.

We denote the concentrations of mRNA and protein of gene \( i \) as \( c_i^m = m_b/V \) and \( c_i = p_b/V \) respectively. According to Eqs. (1)–(3), the deterministic equations of the above variables become (see details in Methods)

\[
\frac{dc_i}{dt} \approx \mu (\phi_i - c_i).
\]

(7a)

\[
\frac{dc_i^m}{dt} \approx \frac{1}{\tau} \left( k_i \phi_i \phi_\sigma \tau - c_i^m \right).
\]

(7b)

A fixed point exists for the dynamics of \( c_i \) and \( c_i^m \), namely \( c_i = \phi_i \) and \( c_i^m = k_i \phi_i \phi_\sigma \tau \). This fixed point is stable due to the global nature of RNAPs and ribosomes: any noises arising from the copy number of RNAPs (ribosomes) equally affect all RNAs (proteins), and therefore leave the relative fraction of one type of mRNA (protein) in the total pool of RNAs (proteins) invariant. The average concentrations of mRNA and protein of gene \( i \) become \( \langle c_i \rangle = \phi_i \) and \( \langle c_i^m \rangle = k_i \phi_i \phi_\sigma \tau \). The results are independent of the cell volume and genome copy number agreeing with experimental data on various organisms\(^{18,20,22}\). We take cell division explicitly into account and, for concreteness, use the “adder” model for cell division by considering an initiator protein \( I \). The initiator protein accumulates from cell birth, triggers the cell division once it reaches the division threshold \( I \), and is then destroyed (or “reset”, e.g., after initiation of DNA replication in bacteria, the ATP-bound DnaA is dephosphorylated to the ADP-bound form\(^{46,47}\)). During a division event, we assume proteins and mRNAs are divided between the two daughter cells following a binomial distribution\(^{49}\). The initiator protein sets the scale of absolute protein number, and the average number of proteins produced in one cell cycle is equal to \( \Delta(i) = I_i \phi_i/\phi_i \)\(^{47}\). Since the protein number grows twofold during one cell cycle, the average protein number of gene \( i \) at cell birth is \( p_b(i) = I_i \phi_i/\phi_i \) and the corresponding average mRNA number at cell birth is \( m_b(i) = k_b I_i \tau \phi_i \phi_\sigma/\phi_i \). We remark that the exact molecular mechanism of cell division does not affect our results.

We corroborate the above analytical calculations with numerical simulations. These will also capture the stochastic fluctuations in gene expression levels, which are not included in the previous analysis. Due to the short lifetime of mRNAs, the production of proteins can be approximated by instantaneous bursts\(^{54}\). We introduce the burst size parameter \( b_0 \) as the average number of proteins made per burst, \( b_0 = k_i \phi_i \langle r(t) \rangle / \langle \sum m_i \rangle \times \tau = k_i \phi_i \phi_\sigma/\phi_\delta \), independent of the cell volume. \( \phi_\delta \) for \( N = 200 \) proteins are uniformly sampled in logarithmic space, with the sum over \( \phi_\delta \) (including ribosome and RNAP) constraint to be precisely one. We choose the parameters to be biologically relevant for bacteria: the doubling time \( T = \ln(2)/\mu = 150 \text{ min} \), \( r_b = 10^3 \), \( b_0 = 0.8 \), \( I_i = 20 \), \( \phi_i = 0.2 \), \( \phi_\delta = 0.7 \) and \( \tau = 3.5 \) min, see other numerical details in Methods. Our conclusions are independent of the specific choice of parameters.

In Fig. 2a, we show the typical trajectories from our simulations of cell volume, protein number and mRNA number for the same gene over multiple generations. To verify the exponential growth of protein and mRNA, we average the protein and mRNA numbers given a fixed relative phase in the cell cycle progression, which is normalized by the generation time and changes from 0 to 1. The averaged values of protein and mRNA numbers (circles) are well predicted by exponential growth, Eqs. (6a) and (6b) (black lines) without any fitting parameters, as shown in Fig. 2b with 3 single trajectories in the background. We also simulate a regulated gene with a time-dependent gene copy number and obtain qualitatively similar results (Methods, Supplementary Fig. 1).

The corresponding trajectories of protein and mRNA concentrations are shown in Fig. 2c, with bounded fluctuations around the predicted averaged values (black lines). In contrast, if the protein number and cell volume grow exponentially but independently, the ratio between them will diverge as the effects of noise accumulate, exhibiting a random walk behavior (Fig. 2d).

During a division event, we assume proteins and mRNAs are divided between the two daughter cells following a binomial distribution\(^{49}\). The initiator protein sets the scale of absolute protein number, and the average number of proteins produced in one cell cycle is equal to \( \Delta(i) = I_i \phi_i/\phi_i \)\(^{47}\). Since the protein number grows twofold during one cell cycle, the average protein number of gene \( i \) at cell birth is \( p_b(i) = I_i \phi_i/\phi_i \) and the corresponding average mRNA number at cell birth is \( m_b(i) = k_b I_i \tau \phi_i \phi_\sigma/\phi_i \). We remark that the exact molecular mechanism of cell division does not affect our results.

We corroborate the above analytical calculations with numerical simulations. These will also capture the stochastic fluctuations in gene expression levels, which are not included in the previous analysis. Due to the short lifetime of mRNAs, the production of proteins can be approximated by instantaneous bursts\(^{54}\). We introduce the burst size parameter \( b_0 \) as the average number of proteins made per burst, \( b_0 = k_i \phi_i \langle r(t) \rangle / \langle \sum m_i \rangle \times \tau = k_i \phi_i \phi_\sigma/\phi_\delta \), independent of the cell volume. \( \phi_\delta \) for \( N = 200 \) proteins are uniformly sampled in logarithmic space, with the sum over \( \phi_\delta \) (including ribosome and RNAP) constraint to be precisely one. We choose the parameters to be biologically relevant for bacteria: the doubling time \( T = \ln(2)/\mu = 150 \text{ min} \), \( r_b = 10^3 \), \( b_0 = 0.8 \), \( I_i = 20 \), \( \phi_i = 0.2 \), \( \phi_\delta = 0.7 \) and \( \tau = 3.5 \) min, see other numerical details in Methods. Our conclusions are independent of the specific choice of parameters.

Effects of finite duration of gene replication. So far, we considered a constant \( \phi_i \) throughout the cell cycle assuming an instantaneous replication of the genome. In this section, we relax this condition and study the effects of finite DNA replication time. We consider the bacterial model of DNA replication, specifically, E. coli, for which the mechanism of DNA replication is well characterized\(^{57}\). The duration of DNA replication is constant, and defined as the C period. The corresponding cell division follows after an approximately constant duration known as the D period. Details of the DNA replication model are in the Methods. In Fig. 3a, we show the time trajectories of the gene allocation fraction, mRNA concentration and protein concentration of one gene for a doubling time of \( T = 30 \text{ min} \) with \( C + D = 70 \text{ min} \). The DNA replication introduces a cell cycle dependent modulation of \( \phi_i \). The abrupt increase of \( \phi_i \) corresponds to the replication of the specific gene \( i \) (Fig. 3a) \( \phi_i \rightarrow 2 \phi_i \). However, as other genes are replicated, the relative fraction of gene \( i \) in the total genome decreases. This modulation propagates to the mRNA concentration which essentially tracks the dynamics of \( \phi_i \) due to its short lifetime. The modulation of mRNA concentration affects the protein concentration as well, yet with a much smaller amplitude. These results can be tested experimentally by monitoring the DNA replication process and mRNA concentration simultaneously.
Noise in gene expression can be classified as intrinsic and extrinsic noise\textsuperscript{34}. While intrinsic noise is due to the stochastic nature of the chemical reactions involved in gene expression, extrinsic noise is believed to be due to the fluctuations of external conditions and common to a subset of proteins. Experiments have revealed a global extrinsic noise that affects all protein concentrations in the genome\textsuperscript{50,59,60}. Because all genes are subjected to the finite duration of DNA replication, it is tempting to attribute the finite duration of DNA replication as one of the main sources of global extrinsic noise\textsuperscript{34}. Within our model in the previous section (constant $\phi_i$'s throughout the cell cycle), there is no global extrinsic noise (Supplementary Fig. 3). A global extrinsic noise may emerge after we introduce the time-dependent $\phi_i$ due to DNA replication. However, we find that the coefficient of variation (CV, the ratio between standard deviation and mean) of the most highly expressed proteins is only about 0.02 within the growing cell model (Fig. 3b), much smaller than that found in experiments\textsuperscript{50,59}. We note that a small extrinsic noise due to gene replication is also observed in constant rate models\textsuperscript{26,27}. Moreover, recent experiments and modeling have suggested that a significant part of the extrinsic noise of mRNA expression level can be attributed to the fluctuations of RNAP copy number\textsuperscript{28}. Within our model, RNAP level fluctuations will lead to extrinsic noise in mRNA concentrations.

A unified phase diagram of gene expression and cellular growth. Experimental observations on \textit{E. coli}\textsuperscript{30} and budding...
yeast support our assumption that ribosomes are limiting for translation. Experimental observations on plant and mammalian cells and fission yeast are also consistent with our assumption that RNA polymerase is limiting for transcription. However, as we discussed in the first section, in the same experiments on fission yeast DNA became limiting for transcription at low DNA concentration. Therefore, we cannot exclude the possibility that in some cases because RNAPs are too abundant, DNA becomes the limiting resource for transcription rather than the number of RNAPs. Similarly, when ribosomes are too abundant relative to the transcript number, the limiting factor for translation becomes the transcript number rather than ribosome number.

In this section, we generalize our model by assuming that each gene has an upper bound on the number of RNAPs (ni) than can simultaneously work on it. A possible extreme case is that the gene is fully loaded with RNAPs, on which RNAPs are only constrained by steric hindrance. The same assumption is made for mRNA with an upper bound of ribosomes (ri) that can work on it simultaneously. We remark that the exact mechanism of DNA and mRNA saturation is beyond our coarse-grained model. If the number of RNAP (ribosome) is above the upper bound, the transcription (translation) rate is limited by the gene (mRNA) number, in a similar fashion to the constant rate models.

We define the protein-to-DNA ratio (PTD ratio) as the sum of protein numbers divided by the sum of effective gene numbers,

\[ \gamma = \sum_i \frac{p_i}{\sum_i g_i}. \]

As the PTD ratio becomes larger, e.g., due to a sufficiently large cell volume with a fixed number of gene, the number of RNAPs (ribosomes) will exceed the maximum load the total genes (mRNAs) can hold. We have discussed thoroughly Phase 1 (neither DNA nor mRNA is saturated) earlier and we summarize our predictions on the transition from Phase 1 to other phases in the following.

Phase 2: In phase 2, the limiting factor in transcription becomes the gene copy number and the transcription rate is proportional to the effective gene copy number, gi. The translation rate is the same as in phase 1. To simplify the formula, we assume all ribosomes are active (to include the effect of an inactive fraction, r should be replaced by f(0). c The gene expression dynamics in phase 3, in which both DNA and mRNA are saturated. The translation rate is proportional to the mRNA number, ri is the upper bound of the number of ribosomes that can work on one mRNA simultaneously.
Here $n_i$ is the upper bound of the number of RNAPs that can work on one gene and $\phi_i$ is the gene allocation fraction of RNAP. Because mRNA is not saturated, the protein number and the cell volume grow exponentially with the same growth rate as Phase 1, Eq. (5), and the homeostasis of protein concentration is still valid. However, because the production rate of mRNA is now proportional to the gene copy number, the mRNA concentration is not constant anymore as the cell volume grows (Methods). In Phase 2, even though the transcription rate doubles after the genome is replicated, the translation rate is proportional to the relative fraction of mRNA in the total pool of mRNAs and therefore still independent of the genome copy number. The average protein concentration is equal to the gene allocation fraction ($(c_i) = \phi_i$). Recent proposed theoretical models of gene expression are consistent with this phase\textsuperscript{65}. In terms of transcription, our model in Phase 2 is equivalent to constant rate models and we have confirmed that for both bacteria and mammalian cells, the typical lifetime of mRNA is short enough compared with the doubling time to distinguish Phase 1 and Phase 2 (Supplementary Fig. 4).

Phase 3: As the cell volume becomes larger, mRNA may get saturated as well. The limiting factor for translation is now the mRNA copy number (Fig. 4c). The threshold PTD ratio for the transition from Phase 2 to Phase 3 is (Methods)

$$y_2 = \frac{k_0}{\tau n_i \phi_i}.$$  

Here $r_i$ is the upper bound of the number of ribosomes that can work on one mRNA. In this phase, the translation rate is proportional to the mRNA number and the protein number grows linearly as $p_i = k_s n_i \phi_i r_i$, with a linear growth rate proportional to the gene number. Therefore, within the assumption that the cell volume is determined by the total protein number, the cell volume grows linearly as well with the linear growth rate proportional to the total gene number,

$$\mu_1 = k_s k_0 \tau n_i r_i \sum_i g_i,$$

and therefore proportional to the genome copy number. As in Phase 2, the mRNA concentration decreases as the cell volume grows, however, the protein concentration is still constant with the average protein concentration equal to the gene allocation fraction ($(c_i) = \phi_i$, Methods). In Phase 3, even though the cell volume grows linearly, the population still grows exponentially with a population growth rate. However, there is no general relation between the ribosomal fraction in the proteome and the population growth rate, in contrast to the growth law in Phase 1 and 2. We summarize the predicted phase diagram of cellular growth in Fig. 4a.

To gain some sense regarding the parameters associated with our proposed phase diagram, we estimate the PTD ratio of \textit{E. coli}. Considering the typical cell volume of \textit{E. coli} as \textsl{1} $\mu$m$^3$, the protein density as $3 \times 10^6$ proteins/$\mu$m$^3$ and the total number of protein-coding genes in \textit{E. coli} as 4000\textsuperscript{52}, we estimate the protein-to-DNA ratio for \textit{E. coli} as $\gamma \approx 1000$. Estimates of the two threshold values of PTD ratios (see Methods) suggest that $y_1 \sim 1500$ and $y_2 \sim 20,000$.

These estimates suggest that wild-type \textit{E. coli} cells are found in Phase 1, but close to Phase 2. We remark that the actual threshold values of PTD ratio for the transitions between different growth phases may be affected by other factors, e.g., the heterogeneous size of genes, but we propose that the general scenario of the transition from Phase 1 to Phase 3 as the protein-to-DNA ratio increases should be generally applicable. As the PTD ratio increases, we predict a transition from exponential growth to linear growth for protein number and cell volume (Supplementary Fig. 5). We propose future experiments to study the potential transition from exponential to linear growth of cell volume, for example using filamentous \textit{E. coli} where cell division and gene replication are inhibited. Similar experiments can also be done for larger cells, e.g., mammalian cells, in which the transition from exponential growth to linear growth of cell volume may be easier to achieve. Preliminary results from experiments measuring the growth of cell mass of mammalian cells\textsuperscript{56} and yeast cells\textsuperscript{56} indeed show a crossover from exponential growth to linear growth when the cell mass is above a threshold value, consistent with our prediction.

It has been shown in bacteria that there are excess RNAPs nonspecifically bound to DNA\textsuperscript{40,41}. In the Methods, we consider a modified model taking into account the partitioning of RNAPs to free RNAPs, elongating RNAPs, promoter-bound RNAPs and nonspecifically bound RNAPs. The transcription rate is determined by the concentration of free RNAPs through Michaelis-Menten kinetics\textsuperscript{40,65}. We find that our conclusions remain intact with an approximately constant fraction of actively transcribing RNAPs in the total RNAPs for Phase 1 (Supplementary Fig. 6). The effect of nonspecifically bound RNAPs is therefore to renormalize the transcription constant $k_s$ in Phase 1 (Eq. (1)) by a constant factor. The transition from Phase 1 to Phase 2 is qualitatively unaffected (Supplementary Fig. 7) and the threshold PTD ratio $y_1$ (Eq. (9)) from Phase 1 to Phase 2 is changed by a constant factor (Methods). We note that alternative mechanisms of gene saturation can occur upon introducing the different classes of RNAPs, through the saturation of free RNAPs and the Michaelis-Menten kinetics (Methods).

\textbf{Discussion}

In this work, we propose a coarse-grained model of stochastic gene expression incorporating cell volume growth and cell division. In the first part, we consider the biological scenario that RNAPs are limiting for transcription and ribosomes are limiting for translation. In other words, neither DNA nor mRNA is saturated. We find that the limiting nature of ribosomes in the translation process leads to the exponential growth of protein numbers. The limiting nature of RNA polymerase and its exponential growth lead to the exponential growth of protein numbers. Homeostasis of protein concentrations originates from the fact that ribosomes make all proteins. Homeostasis of mRNA concentration comes from the resulting bounded concentration of RNAPs. Our model is consistent with the constancy of mRNA and protein concentration as the genome copy number varies since the transcription rate depends on the relative fraction of genes in the genome rather than its absolute number\textsuperscript{22}.

During DNA replication, we find that the gene allocation fraction $\phi_i$ for one specific gene doubles after the gene is replicated but decreases afterwards since other genes are replicated as well and compete for RNAPs. This prediction can be tested by monitoring the mRNA concentration and the copy number of one gene throughout the cell cycle. Furthermore, we extend our model to more general cases in which DNA and mRNA can be saturated by an excess of RNAP and ribosome. We find three possible phases of cellular growth as the protein-to-DNA ratio $\gamma$ increases. A transition from exponential growth to linear growth of protein number and cell volume is predicted. In the future, it will be interesting to study the interplay between the global interactions which are the focus of this work and local...
interactions between genes. Our model provides an alternative model to constant rate models to study genetic networks, which would be advantageous when cell cycle effects are important.

Methods
Derivation of protein and mRNA concentrations. We define the fraction of mRNA i in the total mRNA pool as \( f_i = m_i / \sum m_i \), and the concentration of protein and mRNA of gene \( i \) as \( c_i = m_i / V \). We denote the RNAP and ribosome concentrations as \( c_r \) and \( c_p \), respectively. According to Eqs. (1)–(3), the deterministic equations of the above variables then become

\[
\begin{align*}
\frac{df_i}{dt} &= \frac{k_i n}{\sum m_j} (f_i - f_j) \\
\frac{dc_r}{dt} &= k_i c_r (f_i - c_i) = \mu (f_i - c_i). \\
\frac{dc_p}{dt} &= \frac{1}{\tau} (k_p c_p c_r - (1 + \mu \tau) c_p^2).
\end{align*}
\]

Using the condition that mRNA degradation time is much smaller than the doubling time \( (\mu \tau < 1) \), we find the fixed points for the dynamics of \( f_i \) and \( c_r \). These are, \( f_i = c_i = \phi_i \) and \( c_p^0 = k_i \phi_i c_r \). Replacing \( f_i \) by \( \phi_i \) and \( c_r \) by \( \phi_w \), we obtain the approximate version of the above equations, Eqs. (7a) and (7b).

Simulations of independent growth model. In the growth model corresponding to Fig. 2d, we assume the protein number and cell volume grow exponentially and independently,

\[
\frac{dp}{dt} = (1 + \xi_p(t)) \rho
\]

\[
\frac{dV}{dt} = (1 + \xi_v(t)) V.
\]

Here, \( \xi_p(t), \xi_v(t) \) are white noise terms, with the auto-correlation function, \( \langle \xi_p(0) \xi_p(t) \rangle = A_p \delta(t) \). In Fig. 2d of the main text, we choose \( A_p = A_v = 1 \).

Simulations of growing cell model. We simulated Eqs. (1)–(3), fixing \( r_0, n_0, \lambda_0, \phi \), \( f_0, I_0, \tau \) as well as the growth rate \( \mu \). Other parameters are inferred given the above values, e.g., \( \phi_w = n_0 \phi f_0 r_0 \), \( \mu = \mu (f_0 f_j) \), \( k_i = k_i f_0 r_0 (b_0 n_i) \). We fix the time step \( \delta t \) so that the probability for one event to happen during a time step is smaller than 0.1. We track one of the daughter cells after cell division.

Gene dosage effects. In reality, the gene allocation fraction \( \phi \) changes during the cell cycle due to the finite duration of DNA replication. In this section we introduce the modified version of the gene expression model incorporating DNA replication. Although our model is general, we focus on DNA replication in bacteria for concreteness, specifically E. coli where this process is very well characterized. We expect our conclusions to be generally valid. Furthermore, we refine our model for cell division, assuming that the initiator protein triggers the initiation of DNA replication rather than cell division, with the threshold \( I \) proportional to the number of origins of replication \( 2^{28} \) (the number of which doubles at each initiation). We assume that the cell division takes place a fixed time \( C + D \) after initiation of the DNA replication, where \( C, D \) are respectively the time for DNA replication and the time between the completion of DNA replication and cell division. The number of origins reduce by half at each cell division. Other details are the same as in the main text. Each gene doubles its copy number during the C period, and we choose this gene replication time to be randomly and uniformly distributed across all genes. When a gene \( i \) replicates,

\[
\phi_i \rightarrow 2\phi_i
\]

where the second equation accounts for the normalization of the gene allocation fraction. We choose the experimentally reported \( C \) and \( D \) and cell doubling time from ref.\(^{29}\). In Fig. 3a, we simulate the model by tracking one daughter cells. In Fig. 3b, we track all the cells in an exponentially growing population, which starts from 100 cells to 5000 cells.

Simulations of gene activation. Generalized the constitutive expressed genes considered in the main text to include a single regulated gene by considering a random telegraph process of the effective gene copy number \( g \),

\[
k_p^* = \frac{g_0}{k_p^* (c_p^2)}
\]

Here the gene deactivation rate \( k_p^\ast \) is constant, and the activation rate is set by the concentration of transcription factor through positive regulation, \( k_p^\ast = k_p^\beta \). Here, \( k_p = \mu (f_1 - \phi_1) \). When \( \phi_1 \) is active, the corresponding gene allocation fraction \( \phi_r = \phi_1 (1 - \phi_1) \), and when it becomes deactivated \( \phi_r = 0 \). Note that here we only consider one regulated gene \( i \), but the changing gene allocation of gene \( i \) also affects other genes’ allocation fraction. We simulate the model in Phase 1, and the deactivation of gene \( i \) increases other genes’ allocation fraction as \( \phi_i = \phi_i (1 - \phi_i) \). Since ribosomes are typically more translate because mRNAs are saturated. Since ribosomes are typically more expensive to make than other proteins\(^{30,31}\), we assume the biological scenario, Eq. (24) will be satisfied. From Eq. (21) and using \( n_\phi \sum_p \rho_i = \phi_i \), we obtain the threshold PTD ratio for the transition from Phase 1 to Phase 2,

\[
y_1 = \frac{\phi \rho}{\phi \rho_i}
\]

In Phase 2, the average mRNA concentration becomes

\[
(\phi^0) = \frac{k_p^\beta p \rho i}{V / \phi} = \frac{k_p^\beta p \rho i \sum g_i}{V / \phi} = \frac{k_p^\beta p \rho i}{\phi / \rho_i}
\]

which is inversely proportional to the protein-to-DNA ratio.

From Eq. (23) and using \( r / \sum p_i = \phi_i \), we obtain the transition PTD ratio from Phase 2 to Phase 3 as,

\[
y_2 = \frac{k_p^\beta p \rho i}{\phi / \rho_i}
\]

In Phase 3, the mRNA concentration is the same as Phase 2. Because the cell volume is the sum of all proteins, the protein concentration is the same as Phase 2 and Phase 1
Estimation of the threshold protein-to-DNA ratios for E. coli. We approximate the upper bound of RNAP number working on a single gene as roughly equal to the number of RNAPs on a typical gene (~10^-3 base pairs) when half of the gene is occupied. The linear size of RNAP is about 5 nm, and the length of one base pair is about 0.3 nm, leading to the estimate of τ ~ 30. A similar calculation for the upper bound of ribosome on a single mRNA leads to τ ~ 10 since ribosome's linear size is about 3 times larger than RNA36.

We take φn = 0.2 according to the ref.30, and estimate the gene allocation fraction of RNAP to be φn ~ 0.02 since the number of RNAPs in E. coli is roughly 10% of the number of ribosomes62. We estimate the life time of mRNA as 5 min62. We estimate the transcription rate of one RNAP by considering two potential limiting steps in transcription and taking the slower one. First, assuming the initiation of transcription is diffusion limited, we could estimate the time scale for one RNAP to bind the transcription site as Δt ~ 1 μm^2/(0.2 μm^2/μs) ~ 5 s using the measured diffusion constant of RNAP41,67. Second, we could also estimate the elongation time as the typical length of gene divided by the elongation rate of RNAP, Δt ~ 1000 nt/50 nt/s ~ 20 s63. Taking the slower time scale from the above two calculations, we estimate kT as 0.05 s^-1. Finally, we compute γT and γS using the above estimated parameters, and obtain γT ~ 1500, γS ~ 20,000.

Effect of nonspecifically bound RNAPs. Previous studies on bacteria have shown that there are excess RNAPs bound nonspecifically to the genome and modeled their kinetics40,41. In this section, we consider a modified model to take into account nonspecifically bound RNAPs. For our purpose, we consider a simplified model with four classes of RNAPs, namely, (1) elongating RNAPs, nE (ii) nonspecifically bound to a promoter, nψ (iii) RNAPs nonspecifically bound to DNA, nns (iv) free RNAPs, nfree. We assume a Michaelis-Menten relation for the number of promoter-bound RNAPs and nonspecifically bound RNAPs40,65.

\[ n_p = G \frac{c_{free}}{c_{free} + K_s} \]

(28)

\[ n_{ns} = G_{ns} \frac{c_{free}}{c_{free} + K_{ns}} \]

(29)

Here c_{free} is the concentration of free RNAPs and K_s, K_{ns} are the Michaelis constants. G is the total number of genes and G_{ns} is the number of nonspecific binding sites per gene. Note that c_{free} = c_{free}^\psi+c_{free}^\ns with the concentration of total RNAPs and F_free the fraction of free RNAPs in the total RNAP pool. For simplicity, we assume one promoter for each gene.

The number of elongating RNAPs is then given by:

\[ n_E = n_\psi K_{at} + n_{ns} K_{ps} \]

(30)

where K_{at} is the transition rate from promoter-bound RNAPs to elongating RNAPs, τ_{at} is the time for transcribing a gene, and Δ = K_{at} τ_{at} is the ratio between the number of elongating RNAPs and promoter-bound RNAPs40. We consider parameter regimes motivated by typical biological scenarios, in particular (1) the number of sites for nonspecific binding is much larger than the number of promoters, (2) nonspecific binding of RNAPs to DNA is much weaker than the specific binding to the promoters, (3) the number of promoter-bound RNAPs is small compared with elongating RNAPs.

Using n = n_p + n_{ns} + n_{free} + n_\psi we can find a self-consistent equation for F_free

\[ c_{free} = (1 + \Lambda) \left( 1 - \frac{c_{free}^\psi}{c_{free}^\ns} \right) + \frac{2}{\Lambda} F_{free} \]

(31)

Here γ is the protein-to-DNA (PTD) ratio, i.e., the ratio between the total number of proteins (equivalent to cell volume V within our model) and the total gene number, γ = V/G. We can use Eq. (31) to compute the fraction of free RNAPs given a PTD ratio and use Eq. (28) - (30) to compute the fraction of elongating RNAPs, F_e and nonspecifically bound RNAPs, F_free.

Since the left side of Eq. (31) monotonically decreases from c_{free} to 0 as F_{free} increases from 0 to 1 and the right side of Eq. (31) monotonically increases as F_{free} increases, we find that as the protein-to-DNA ratio increases, F_{free} increases. Previous studies have shown that F_{free} ~ 0.1 for bacteria40,41 and support the assumption that F_{free} is smaller or comparable to K_{s}40,65 (note that nonspecific binding is characterized by K_{ns} larger than K_{s}). In the following, we first assume a small F_{free} and c_{free} ~ K_s, and show this leads to a behavior qualitatively similar to Phase 1 of the main text, albeit with a renormalization of the transcription constant. K_s. We also later discuss the situations when these conditions break down, and show how they lead to a behavior consistent with Phase 2. The transcription rate for one specific gene with an effective gene copy number g, can be written as

\[ TR_i = \frac{n_E^i}{\tau_i} = \frac{1}{\phi} n_E^i \]

(32)

Here g/G = φ is the gene allocation fraction of gene i, and F_e is the fraction of elongating RNAPs in the total RNAP pool. In the limit of a small F_{free} and c_{free} ~ K_s, F_e becomes a constant.

\[ F_e \approx \frac{n_{ns}^i + n_{ps}^i}{1 + \Lambda + \phi G K_s/K_{ns}} \]

(33)

Therefore, the transcription rate corresponds to Phase 1 of our model, TR = kT φ ψ n with k_T = k_T F_e.

In the limit that all RNAPs are actively transcribing and no nonspecifically bound RNAPs, Λ >> 1 and gns = 0, we go back to the situation that all RNAPs are working with k_T = 1/τ_T = k_T. Therefore we conclude that the introduction of nonspecifically bound RNAPs does not affect our model qualitatively, and its effect is to renormalize the transcription constant k_T in Eq. (1) by a constant factor, k_T = k_T F_e.

We simulate a single lineage of growing cells using the full model (with partitioning of RNAPs and gene replication). We set the parameters as: A = 50, g_{ns} = 1000, k_{ns} = 1 min^{-1}, G = 2000 (before gene duplication), K_s = 0.8 (c) where (c) is the total protein concentration within the cell which we set to 1 throughout the paper. Our conclusions are independent of the specific values of parameters. The gene allocation fractions are the same as the main text and the average RNAP concentration during the cell cycle (c_T) = φn ~ 0.02. The fractions of elongating RNAPs and nonspecifically bound RNAPs are approximately constant with a small cell cycle modulation (the coefficient of variation is of the order of 0.01), consistent with the above results since F_{free} is small (Supplementary Fig. 6a). We also find a linear scaling between mRNA number and cell volume, consistent with Phase 1 of our model (Supplementary Fig. 6b).

We next consider the transition from Phase 1 (RNAP limiting) to Phase 2 (gene limiting) by also testing the saturation of genes is due to the steric hindrance of elongating RNAPs with a minimum distance between two successive RNAPs, we can find the threshold PTD ratio from Phase 1 to Phase 2. Since the fraction of elongating RNAPs is constant, we can compute the threshold PTD ratio using n_E/G = n (where n is the upper bound of the number of RNAPs that can work on one gene simultaneously), and obtain

\[ \gamma_T = \frac{1 + \Lambda + \phi G K_s/K_{ns}}{\Lambda} \]

(34)

We find that the introduction of nonspecifically bound RNAPs does not affect our main results qualitatively and it changes the threshold PTD ratio from Phase 1 to Phase 2, Eq. (9), by a constant factor, which is the inverse of the fraction of elongating RNAPs.

In Supplementary Fig. 7, we show the transition from Phase 1 to Phase 2 in the proposed experiment in which cell division is blocked, in the presence of nonspecific binding. At each point in time we solve the self-consistent equation of the partitioning of RNAPs, Eq. (31). The plot shows both the deterministic dynamics of average mRNA number (black lines) as well as the results of the stochastic simulations (red/blue solid lines). We compare the dynamics for two cells with different fixed genome sizes. Initially, the two cells are in Phase 1 and therefore have identical mRNA number proportional to cell volume. When they exceed their respective threshold PTD ratio, the mRNA number begins to saturate and becomes limited by the gene number, therefore the cell with a twice larger genome size has twice more mRNAs (Phase 2).

So far we assumed that F_{free} is small and c_{free} is smaller than the Michaelis constant. In the following we relax these assumptions. We find that the introduction of the different RNAP classes introduces alternative mechanisms for gene saturation, that lead to behavior consistent with Phase 2.

We first relax the assumption that F_{free} ~ 1 (assuming that c_{free} << K_s still holds). Because the transcription rate for a specific gene is

\[ TR_i = \frac{n_E^i}{\tau_i} = \frac{1}{\phi} n_E^i \]

(35)

when F_{free} is comparable to 1, the transcription rate will be saturated as well and proportional to the gene copy number g. From Eq. (31), we find that F_{free} = γ (1/K_s + 1/K_{ns}) γ + γ, therefore the threshold PTD ratio for F_{free} to be comparable to 1 is

\[ \gamma_T = \frac{1 + \Lambda + \phi G K_s/K_{ns}}{\Lambda} \]

(36)

Second, we can relax the assumption that c_{free} << K_s (assuming F_{free} << 1 still holds). When c_{free} = K_s, the transcription rate can be saturated as well. Using F_{free} = γ (1/K_s + 1/K_{ns}) γ + γ, we can obtain the corresponding threshold PTD ratio as

\[ \gamma_T = \frac{1 + \Lambda + \phi G K_s/K_{ns}}{\Lambda} \]

(37)

We remark that the particular mechanism which drives the cell to Phase 2 (gene limiting) is the one with the smallest threshold. Comparison between Eqs. (34),
(36) and (37) shows that when $n_\text{on}K < K$ and $n_\text{off} < A$, genes get saturated due to stochastic hindrance. Reference26 shows that for wild type E. coli in fast growth conditions the mRNA levels in the cell do not change when the DNA amount is lowered. Within our model this is consistent with Phase 1, and inconsistent with Phase 2, thus suggesting that $F_\text{free} \ll 1$ and $F_\text{free} \ll K$, as discussed earlier.

Data availability
The data that support the findings of this study are available from the authors on request.

Received: 23 January 2018 Accepted: 17 September 2018
Published online: 29 October 2018

References
1. Paulsson, J. Models of stochastic gene expression. Phys. Life Rev. 2, 157–175 (2005).
2. Kærn, M., Elston, T. C., Blake, W. J., & Collins, J. J. Stochasticity in gene expression: from theories to phenotypes. Nat. Rev. Genet. 6, 451–464 (2005).
3. Pedraza, J. M. & van Oudenaarden, A. Noise propagation in gene networks. Science 307, 1965–1969 (2005).
4. Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell 135, 216–226 (2008).
5. Thattai, M. & van Oudenaarden, A. Intrinsic noise in gene regulatory networks. Proc. Natl Acad. Sci. 98, 8614–8619 (2001).
6. Sanchez, A. & Golding, I. Genetic determinants and cellular constraints in noisy gene expression. Science 342, 1188–1193 (2013).
7. Wang, P. et al. Robust growth of escherichia coli. Curr. Biol. 20, 1099–1103 (2010).
8. Campos, M. et al. A constant size extension drives bacterial cell size homeostasis. Cell 159, 1433–1446 (2014).
9. Taher-Araghi, S. et al. Cell-size control and homeostasis in bacteria. Curr. Biol. 25, 385–391 (2015).
10. Cermak, N. et al. High-throughput measurement of single-cell growth rates using serial microfluidic mass sensor arrays. Nat. Biotechnol. 34, 1052–1059 (2016).
11. Eun, Y.-J. et al. Archaeal cells share common size control with bacteria despite noisier growth and division. Nat. Microbiol. 3, 148 (2018).
12. Elliott, S. & McLaughlin, C. Rate of macromolecular synthesis through the cell cycle of the yeast saccharomyces cerevisiae. Proc. Natl Acad. Sci. 75, 4384–4388 (1978).
13. Cooper, S. Distinguishing between linear and exponential cell growth during the division cycle: single-cell studies, cell-culture studies, and the object of cell-cycle research. Theor. Biol. Med. Model. 3, 10 (2006).
14. Di Talia, S., Skottheim, J. M., Bean, J. M., Siggia, E. D. & Cross, F. R. The effects of molecular noise and size control on variability in the budding yeast cell cycle. Nature 448, 947–951 (2007).
15. Godin, M. et al. Using buoyant mass to measure the growth of single cells. Nat. Methods 7, 387–390 (2010).
16. Sinclair, W. K. & Ross, D. W. Modes of growth in mammalian cells. Biophys. J. 95, 1056–1070 (1999).
17. Crissman, H. A. & Steinkamp, J. A. Rapid, simultaneous measurement of DNA, protein, and cell volume in single cells from large mammalian cell populations. J. Cell Biol. 99, 766 (1983).
18. Kempe, H., Schwabe, A., Crémazy, F., Verschure, P. J. & Bruggeman, F. J. The volumes and transcript counts of single cells reveal concentration homeostasis and capture biological noise. Mol. Biol. Cell 26, 797–804 (2015).
19. Padovan-Merhar, O. et al. Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. Mol. Cell 58, 339–352 (2015).
20. Ietswaart, R., Rosa, S., Wu, Z., Dean, C. & Howard, M. Cell-size-dependent transcription of lfc and its antisense long non-coding rna coolax explains cell-to-cell expression variation. Cell 140, 622–635.e9 (2017).
21. Brenner, N. et al. Single-cell protein dynamics reproduce universal fluctuations in cell populations. Eur. Phys. J. E 38, 1–9 (2015).
22. Zheng, X.-y & O’Shea, E. K. Cyanobacteria maintain constant protein concentration despite genome copy-number variation. Cell Rep. 19, 497–504 (2017).
23. Friedman, N., Cai, L. & Xie, X. S. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. Phys. Rev. Lett. 97, 168302 (2006).
24. Shahrezaei, V. & Swain, P. S. Analytical distributions for stochastic gene expression. Proc. Natl Acad. Sci. 105, 17256–17261 (2008).
25. Cai, L., Friedman, N. & Xie, X. S. Stochastic protein expression in individual cells at the single molecule level. Nature 440, 358–362 (2006).
57. Walden, M., Fange, D., Lundius, E. G., Baltekin, O. & Elf, J. The synchronization of replication and division cycles in individual E. coli cells. Cell 166, 729–739 (2016).
58. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. Science 297, 1183–1186 (2002).
59. Bar-Even, A. et al. Noise in protein expression scales with natural protein abundance. Nat. Genet. 38, 636–643 (2006).
60. Dar, R. D., Razooky, B. S., Weinberger, L. S., Cox, C. D. & Simpson, M. L. The low noise limit in gene expression. PLoS ONE 10, 1–21 (2015).
61. Thomas, P., Terradot, G., Danos, V. & Weisse, A. Stochasticity of cellular growth: sources, propagation and consequences. Preprint at https://www.biorxiv.org/content/early/2018/02/18/267658 (2018).
62. Milo, R. & Phillips, R. Cell biology by the numbers (Garland Science, New York, 2015).
63. Liu, X., Oh, S. & Kirschner, M. Personal Communications. (2018).
64. Neurohr, G & Amon, A. Personal Communications. (2018).
65. Patrick, M., Dennis, P. P., Ehrenberg, M. & Bremer, H. Free rna polymerase in escherichia coli. Biochimie 119, 80–91 (2015).
66. Ho, P.-Y. & Amir, A. Simultaneous regulation of cell size and chromosome replication in bacteria. Front. Microbiol. 6, 662 (2015).
67. Bratton, B. P., Mooney, R. A. & Weiss, J. C. Spatial distribution and diffusive motion of rna polymerase in live escherichia coli. J. Bacteriol. 193, 5138–5146 (2011).
68. Churchward, G., Bremer, H. & Young, R. Transcription in bacteria at different dna concentrations. J. Bacteriol. 150, 572–581 (1982).

Acknowledgements
We thank Naama Barkai, Ido Golding, Andreas Hilfinger, Po-Yi Ho, Meriem El Karoui, Andrew Murray, Johan Paulsson, Leonardo A. Sepúlveda, and Sven van Teeffelen for useful discussions related to this work. AA thanks the A.P. Sloan foundation, the Milton Fund, the Volkswagen Foundation and Harvard Dean’s Competitive Fund for Promising Scholarship for their support. JL was supported by the George F. Carrier fellowship and the National Science Foundation through the Harvard Materials Research Science and Engineering Center (DMR-1420570).

Author contributions
All authors conceived the work, carried out the work, and jointly wrote the manuscript.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-06714-z.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018