Heterogeneous recruitment abilities to RNA polymerases generate superlinear and sublinear scaling of gene expression level with cell volume

Jie Lin
Center for Quantitative Biology and Peking-Tsinghua Joint Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China
(Dated: January 26, 2021)

Many experiments have shown that the expression levels of most genes are typically proportional to cell volumes in growing cells. However, a subset of genes often exist and they exhibit nonlinear scaling of expression levels with cell volume, and their corresponding mRNA and protein concentrations are therefore not constant as the cell volume increases. While these genes with changing concentrations often have important biological functions such as cell cycle regulation, the biophysical mechanism underlying the nonlinear scaling between the gene expression level and the cell volume is still unclear. In this manuscript, we show that the nonlinear scaling is in fact a direct consequence of heterogeneous recruitment abilities of promoters to RNA polymerases. Those genes with weaker (stronger) recruitment abilities compared with the average ability spontaneously exhibit superlinear (sublinear) scaling with cell volume. Our model makes several predictions that are consistent with experimental observations, such as a correlation between the expression levels and the degrees of nonlinear scaling with cell volume among genes.

Homeostasis of gene expression level in a dynamically growing cell volume is widely observed across the domains of biology, namely, the concentrations of mRNAs and proteins of most genes are approximately constant [1–10]. Since the cell volume usually grows exponentially [9, 11, 12], this implies exponential growth of mRNA and protein copy numbers as well and a linear scaling between the gene expression level and cell volume exists. Theoretically, the linear scaling between gene expression level and exponentially growing cell volume has been shown as a direct consequence of the limiting nature of gene expression machinery: RNA polymerases (RNAPs) and ribosomes [13], in agreement with experimental observations [10, 14–16]. However, along with the constant concentrations of mRNAs and proteins of most genes, there are a subset of genes exhibiting nonconstant concentrations as the cell volume increases. These genes are often crucial for cell cycle and size regulation as they allow cells to sense their sizes based on the concentrations of proteins that have different scaling behaviors with cell volume, such as Whi5 and Cln3 in *Saccharomyces cerevisiae* [17–19]. Other examples include DNA-binding proteins such as histone proteins, whose expression levels scale with the total DNA amount instead of cell volume [20]. A fundamental question then arises: if a linear scaling between gene expression level and cell volume are by default for most genes, how can cells achieve nonlinear scaling for a subset of genes with cell volume in the meantime? In this paper, we show that the superlinear and sublinear scaling of gene expression level is in fact a direct consequence of the heterogeneous recruitment abilities of promoters to RNA polymerases. Given a unimodal distribution of recruitment abilities, those genes with the recruitment abilities of their promoters below (above) the average spontaneously exhibit superlinear (sublinear) scaling with cell volume, while the majority of genes with recruitment abilities near the average exhibit approximately linear scaling.

In the following, we first introduce a coarse-grained whole-cell level model of gene expression in which the promoters of all genes have the same recruitment abilities to RNAPs, which is analytically solvable. Then we consider a scenario in which all genes except a small subset of genes have the same recruitment abilities and show that the expression levels of those special genes can exhibit superlinear or sublinear scaling with cell volume depending on their relative magnitudes of recruitment abilities compared with the majority of genes. Then we extend the simplified model to allow a continuous distribution of recruitment abilities and show that our simplified model can quantitatively capture this more realistic scenario. Genes with recruitment abilities below the average value naturally exhibit superlinear (sublinear) scaling with cell volume and active regulation is not necessary for nonlinear scaling.

Results

Model of gene expression at the whole-cell level. We consider a coarse-grained model of gene expression and the mRNA production rate (how fast mRNAs are produced) of one particular gene labeled by index i is proportional to its gene copy number and the probability for its promoters to be bound by RNAPs

$$k_{n,i} = \Gamma_{n,i} g_i \frac{c_{n,free}}{c_{n,free} + K_{n,i}}. \tag{1}$$

Here g_i is the gene copy number of gene i, $c_{n,free}$ is the concentration of free RNAPs in the nucleus (see an schematic of the model in Figure 1). We assume the probability for one promoter to be bound by a RNAP follows

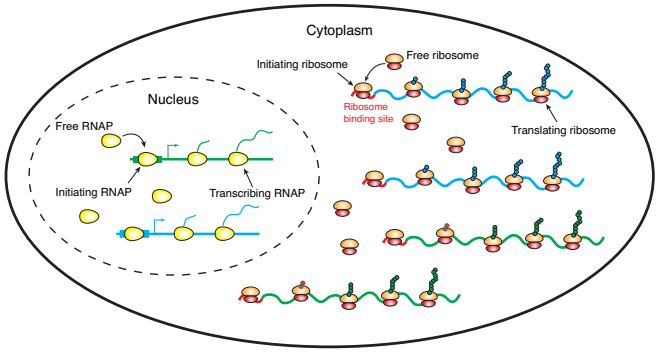

FIG. 1. A coarse-grained whole-cell model of gene expression. Inside the nucleus, genes compete for free RNAPs to bind to their promoters and for simplicity we only show two genes in this schematic. The binding probability of free RNAPs on the promoters depend on the free RNAP concentration and the recruitment abilities of the promoters. In the cytoplasm, mRNAs compete for free ribosomes to bind to their ribosome binding sites and the binding probability depends on the free ribosome concentrations and the recruitment abilities of the mRNAs to ribosomes.

the Michaelis-Menten equation $P_{n,b} = c_{n,free}/(c_{n,free} + K_n)$ where $K_{n,i}$ is the Michaelis-Menten constant which is inversely proportional to the binding rate of RNAPs on the promoters (see detailed derivations in Supplementary Information SI A). In the following, we use $1/K_{n,i}$ as a metric of the recruitment abilities of genes to RNAPs: a larger $K_{n,i}$ represents a smaller recruitment ability. $\Gamma_{n,i}$ is the initiation rate of transcription of gene i: the rate that a promoter-bound RNAP start transcribing the gene and producing mRNA. Note that we mainly discuss eukaryotic cells in this work in which the transcription and translation processes are spatially separate and nonspecific binding of RNAPs to DNA are mostly irrelevant [21, 22]. As we also mainly discuss the transcription of mRNA, RNAPs here refer to RNAP II. Real transcription processes in eukaryotic cells are complex and may involve enhancers and TATA-binding proteins [23, 24], which are all coarse-grained into the effective Michaelis-Menten constant within our model.

Because the typical time scale for a RNAP to finish transcribing a gene is below one minute [25], much shorter than the typical doubling times, we assume that the dynamics of RNAPs along each gene is in the steady state so that the outgoing flux of RNAPs that finish transcribing a gene and the initiation rate of transcription are balanced. Therefore the maximum number of elongating RNAPs along gene i becomes $\Lambda_{n,i} = \Gamma_{n,i} L_i / v_n$ which can happen when the promoter is constantly occupied by RNAP. Here L_i is the length of the gene i and v_n is the elongation speed of RNAP which we approximate as a constant.

A simplified model in which all genes share the same recruitment ability. In the following we consider a simplified scenario assuming the promoters of each gene has the same recruitment ability to RNA polymerases so that $K_{n,i} = K_n$ for all i, which is analytically solvable (later we will relax this assumption). Using the conservation of the total number of RNAPs including free RNAPs, initiating RNAPs and transcribing RNAPs, we obtained the self-consistent equation to determine the fraction of free RNAPs in all RNAPs:

$$\sum_i g_i (1 + \Lambda_{n,i}) \frac{c_n F_n}{c_n F_n + K_n} = n - n F_n. \tag{2}$$

Here n is the total number of RNAPs, F_n is fraction of free RNAPs. c_n is the concentration of total RNAPs in the nucleus and we assume that the nuclear volume is proportional to the total cell volume, consistent with experimental observations [26]. Let us now consider a thought experiment in which all RNAPs suddenly become free, one would expect that these free RNAPs will have a strong tendency to rebind to the promoters otherwise a large fraction of RNAPs will be idle, which is clearly inefficient in normal cellular physiological states. Therefore, we assume that $K_n \ll c_n$ and it is straightforward to find that in this case, the F_n solving Eq. (2) must be much smaller than 1 if $n < n_c$ where $n_c = \sum_i g_i (1 + \Lambda_{n,i})$, which is the maximum of number of RNAPs the entire genome can hold (Figure 2A). In this scenario, almost all RNAPs are either bound to the promoter or transcribing ($F_n \ll 1$). Taking $F_n = 0$ in Eq. (2), we express the binding probability as n/n_c and using Eq. (1), we find that the mRNA production rate now becomes proportional to the number of RNAPs,

$$k_{n,i} = \Gamma_{n,i} \frac{g_i}{\sum_j g_j (1 + \Lambda_{n,j})} n. \tag{3}$$

This linear scaling between the mRNA production rate and the number of RNAPs is the root of the linear scaling between mRNA level and cell volume as we explain in the following. A similar model as Eq. (1) also applies to the translational process (Figure 1, see an alternative formulation of translation model in Ref. [27]). For simplicity, we assume the recruitment abilities of mRNAs to ribosomes and their translation initiation rates as the same for different genes, therefore the ratios between the protein production rates (how fast proteins are produced) of different genes are simply equal to the ratios between their mRNA copy numbers. This is consistent with the fact that protein levels are often primarily determined by mRNA levels [28]. From Eq. (3), we find that besides those common factors, mRNA copy numbers are set by their corresponding gene copy numbers g_i, their lifetimes $\tau_{m,i}$, and the initiation rates $\Gamma_{n,i}$. Therefore, when the gene copy numbers are fixed, the mass fractions of proteins are approximately constant (SI B). We also note that the above conclusion is valid even we relax the assumption that the recruitment abilities of mRNAs

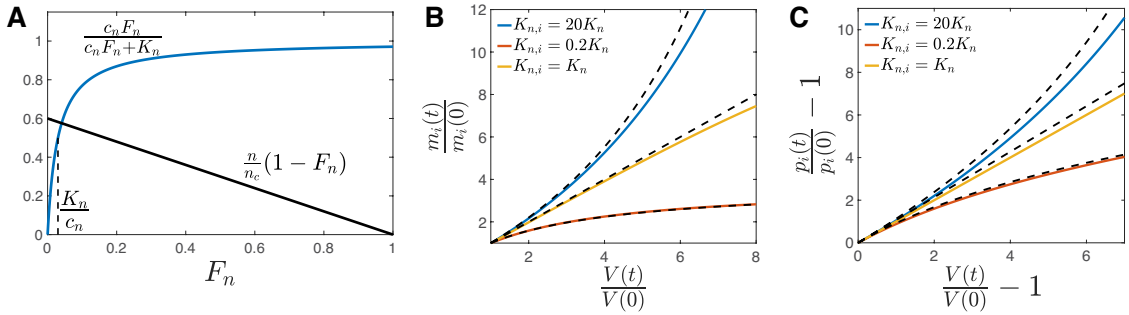

FIG. 2. (A) Based on the conservation of total number of RNAPs, a self-consistent equation, Eq. (2), is derived for the case of homogeneous recruitment abilities to RNAPs. Assuming $K_n/c_n \ll 1$, the solution F_n must be much smaller than 1 if $n < n_c$. (B) We consider two special genes with recruitment abilities, $K_{n,i} = 20K_n$, $K_{n,i} = 0.2K_n$. The majority of genes have $K_{n,i} = K_n$. The mRNA levels of these two genes show superlinear and sublinear scalings with cell volume, in agreement with the theoretical predictions (dashed lines). The mRNA number and cell volume are normalized by their initial values. (C) Same analysis as (B) for non-degradable proteins.

to ribosomes and their translation initiation rates are the same (SI B).

Given constant mass fractions of proteins, the total number of RNAPs is proportional to the total protein mass, which is known to be highly correlated with the total dry mass and cell volume [29, 30]. Within our gene expression model, we therefore assume a constant ratio between the total protein mass (M) and cell volume (V). Therefore, the mRNA production rate is proportional to the cell volume as well, and this scenario of gene expression has been called Phase 1 of gene expression [13], which is the typical state of cells in normal conditions [6, 14, 31] and the main focus of this work. When the number of RNAPs exceeds the threshold value n_c, the linear scaling between mRNA numbers and cell volume breakdowns for all genes and the growth mode of cell volume also deviates from exponential growth (Phase 2). If the ribosome number exceeds some threshold value, the cell volume eventually grows linearly [31] (Phase 3, see Ref. [13] and detailed derivations in the SI B). We also discuss the effects of nonspecific binding of RNAPs on the transition between Phase 1 and Phase 2 in SI C in which the condition of Phase 1 is much more stringent.

A more realistic model in which genes can have different recruitment abilities. We now consider a more realistic scenario in which the recruitment abilities to RNAPs of different genes can be different. We start from a simple scenario in which all genes have the same recruitment ability $1/K_n$ except one special gene i has a recruitment ability $1/K_{n,i}$. Since the contribution of the particular gene to the global allocation of RNAPs is negligible, the probability for the promoters of most genes to be bound is still $P_{n,b} = c_nF_n/(c_nF_n + K_n) \approx n/n_c$ from which we find that the free RNAP concentration and the mRNA production rate for the particular gene i becomes

$$k_{n,i} = \Gamma_{n,i} g_i \frac{K_n n}{K_{n,i} n_c - (K_{n,i} - K_n) n}. \tag{4}$$

As n gets closer n_c (with $F_n \ll 1$ still satisfied), we find that the mRNA production rate of gene i exhibits a superlinear (sublinear) dependence on the RNAP number and cell volume, if $K_{n,i} > K_n$ ($K_{n,i} < K_n$). For mRNAs with short lifetimes, the mRNA copy numbers are simply equal to their production rates multiplied by their lifetimes. If the corresponding proteins also have short lifetimes, their copy numbers will be proportional to the mRNA production rates as well. More importantly, we find that the nonlinear scaling in mRNA copy numbers also propagates to non-degradable proteins, suggesting that the nonlinear scalings in protein copy number are insensitive to their lifetimes (see detailed derivations in SI D).

In the following we discuss a more realistic scenario that is a continuous distribution of $K_{n,i}$ to reflect the heterogeneity of the promoter structures. In this case, we propose that the nonlinear scaling, Eq. (4), is still approximately valid for any gene if K_n is replaced by $\langle K_{n,i} \rangle$, the average value of $K_{n,i}$ among all genes after some appropriate weightness (later we show that the appropriate average is weighted by the protein mass fractions). Therefore, those genes with $K_{n,i}$ considerably larger (smaller) than $\langle K_{n,i} \rangle$ exhibit superlinear (sublinear) scaling, while genes with $K_{n,i}$ near the average exhibit approximately linear scaling. To quantify the nonlinear degree of volume-expression scaling of each gene, we introduce two parameters β and α respectively for mRNAs and non-degradable proteins (see detailed derivations and also discussions about degradable proteins in SI D) so that their scalings with cell volume become

$$\widetilde{m}_i(t) = \widetilde{V}(t) \frac{1 + \beta_i}{1 + \beta_i \widetilde{V}(t)}, \tag{5}$$

$$\Delta \widetilde{p}_i(t) = C_i \ln(1 + \alpha_i \Delta \widetilde{V}(t)). \tag{6}$$

Here $\widetilde{m}_i(t) = m_i(t)/m_i(0)$, $\widetilde{V}(t) = V(t)/V(0)$, $\Delta\widetilde{V}(t) =$

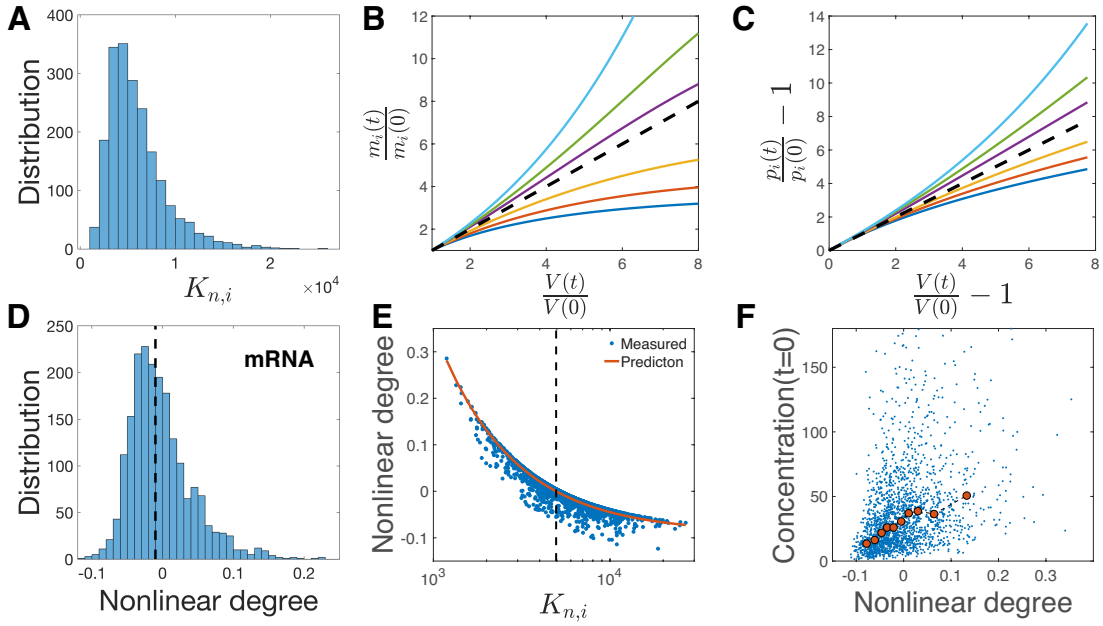

FIG. 3. (A) Distribution of the Michaelis-Menten constants used in the simulations, which we take as a lognormal distribution with the coefficient of variation (standard deviation/mean) equal to 0.5. (B) Examples of mRNA numbers *vs.* cell volume including superlinear and sublinear scaling. The dashed line is the $y = x$ line. (C) Examples of protein number increment *vs.* cell volume increment. The dashed line is the $y = x$ line. Here we show the results for non-degradable proteins and our results are equally valid for degradable proteins (Figure S4). (D) Distribution of the measured nonlinear degrees β of mRNA numbers from numerical simulations. The dashed line marks the location of the median value of the nonlinear degrees. (E) We compare the theoretically predicted nonlinear degrees of mRNA numbers and the measured one from numerical simulations. The dashed line marks the location of the average $\langle K_{n,i} \rangle$ weighted by protein mass fractions at which the nonlinear degree is predicted to be zero. (F) We plot the mRNA concentrations at time zero *v.s.* their corresponding nonlinear degrees. The red points are the median values after binning.

$\widetilde{V}(t) - 1$, and $\Delta\widetilde{p}_i(t) = p_i(t)/p_i(0) - 1$. All the variables are measured relatively to their values at time $t = 0$. C_i is a constant depending on the particular gene. Positive (negative) β_i and α_i represent sublinear (superlinear) scaling behaviors of mRNA copy numbers and protein copy numbers. They are related to the recruitment abilities of their corresponding genes to RNAPs (see the full expressions in SI D) and when $K_{n,i}$ is close to the mean Michaelis-Menten constant $\langle K_{n,i} \rangle$, we find that $\beta_i \propto \langle K_{n,i} \rangle - K_{n,i}$, $\alpha_i \propto \langle K_{n,i} \rangle - K_{n,i}$. Note that we return to linear scaling when $\beta_i \to 0$ and $\alpha_i \to 0$.

Numerical tests and analyses of experimental data. To test our theoretical predictions, we simulated a fictitious cell with 2000 genes including RNAP and ribosome which we coarse-grained as single proteins. Simulation details and parameters are summarized in SI E and Table S1. To avoid the effects of time-dependent gene copy numbers, which is certainly important but also complicates our analyses, we considered the scenario in which the cell volume grows without DNA replication, e.g., cells arrested in G1 phase which is a common experimental protocol to study the effects of cell volume on gene expression [19, 31].

We first simulated the simplified model with homogeneous recruitment abilities to RNAPs and confirmed the exponential growth of cell volume, and the linear scaling between mRNA copy numbers, protein copy numbers, and cell volume (Figure S2). We then simulated the case when all genes share the same recruitment ability K_n but two genes have different abilities which are respectively smaller and larger than K_n. Our theoretical predictions for the expression levels of mRNAs, non-degradable proteins (Figure 2B, C) and degradable proteins (Figure S3) assuming short mRNA lifetimes, match the simulation results reasonably well. Deviations are expected since the actual mRNA lifetimes are finite.

Finally, we simulated the more realistic scenario in which $K_{n,i}$ is continuously distributed, which we modeled as a lognormal distribution (Figure 3A) and our conclusions are independent of the specific choice of the form of distribution. Examples of volume dependence of several mRNAs and proteins are shown in Figure 3B, C. Nonlinear degrees of mRNAs and proteins are measured based on Eqs. (5, 6) and the resulting distribution of mRNAs is shown in Figure 3D, and those of proteins are shown in Figure S4. To test our theoretical predictions on the relations between the nonlinear degrees β, α and $K_{n,i}$ (Eqs. S34, S36), we also need the expression of the appropri-

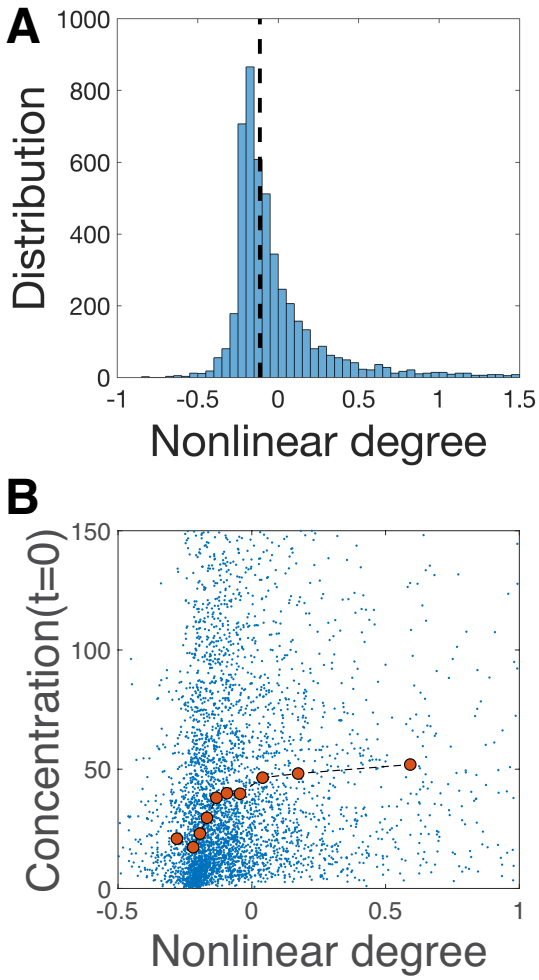

FIG. 4. (A) Distribution of the nonlinear degrees of cell volume scaling of mRNAs for *S. cerevisiae* among genes. The data is from Ref. [19]. The dashed line marks the location of the median value of the nonlinear degrees. (B) A positive correlation between the nonlinear degrees of mRNAs and their initial concentrations. The red data points are median values after binning.

ate average $\langle K_{n,i} \rangle$. We found that the appropriate average Michaelis-Menten constant is the one weighted by the initial protein mass fractions, $\langle K_{n,i} \rangle = \sum_i \phi_i K_{n,i}$, where ϕ_i is the mass fraction of proteins i in the entire proteome at time $t = 0$. The above expression of $\langle K_{n,i} \rangle$ leads to a good agreement between numerical simulations and theoretical predictions (Figure 3E). We set the Michaelis-Menten constants of RNAP and ribosome as the average value to ensure that their concentrations are approximately constant as the cell volume increases.

We note that the recruitment abilities not only determines the nonlinear degree of volume scaling, but also affects the mRNA production rate since a higher recruitment ability also enhance the RNAP binding probability to promoter. Our model therefore predicts a positive correlation between the initial concentrations of mRNAs and their nonlinear degrees, which is confirmed by the numerical simulations (Figure 3F and also true for proteins, see Figure S4). Our results suggest that those genes with sublinear scaling and smaller $K_{n,i}$ contribute more in the weighted average of $K_{n,i}$, therefore $\langle K_{n,i} \rangle < \overline{K_{n,i}}$ where $\overline{K_{n,i}}$ is the average over all genes with equal weightness. Therefore, genes with $K_{n,i} \approx \overline{K_{n,i}}$ are expected to exhibit superlinear scaling (Figure 3E). We note that this does not imply most genes show superlinear scaling and in fact we find that the median value of nonlinear degrees is still close to zero which suggests the majority of genes show approximately linear scaling (Figure 3D).

We also analyzed the experimental data set from Ref. [19] where the volume dependence of mRNA levels are measured for *Saccharomyces cerevisiae* across the genome. We calculated the nonlinear degrees of mRNA scaling with cell volume using Eq. (5) and obtained the resulting distribution (Figure 4A). Interestingly, we also found a positive correlation between the nonlinear degrees β and the initial mRNA concentrations (Figure 4B). Similar as our simulations, the median value of nonlinear degrees is also close to zero in the experiments, suggesting that the majority of genes show approximately linear scaling. In conclusions, analyses of experimental data appear to be consistent with our theoretical predictions although alternative mechanisms on the correlation between the concentrations and the nonlinear degrees may exist [19].

Discussion

For abnormally large cells, the number of RNA polymerases or ribosomes may exceed some threshold values so that the bottlenecks of gene expression become the templates of gene expression: gene copy numbers and mRNA copy numbers [14, 31]. However, in typical cellular physiological states, cells are far below the thresholds and the limiting bottlenecks of gene expression are the copy numbers of RNAPs and ribosomes [10, 15, 16]. We prove that in this case, if the promoters of all genes share the same recruitment ability, the gene expression levels of all genes, should exhibit liner scaling with cell volume both at the transcription and translation level. We then extend this simple scenario to consider the more realistic case in which the recruitment abilities among genes are continuously distributed. We analytically derive the dependence of the mRNA production rate on the RNAP copy number, which is in turn proportional to the cell volume. We show that for genes with recruitment abilities below (above) the average exhibit superlinear (sublinear) scaling with cell volume, which is a natural consequence of the heterogeneous distribution of recruitment abilities. In other words, active regulation is not necessary to achieve the nonlinear scaling for a subset of genes.

We further show that the nonlinear scaling between the mRNA production rates and cell volume propagates to the mRNA copy numbers and proteins copy numbers. All of our theoretical predictions were confirmed by numerical simulations.

Because recruitment abilities are also positively correlated with gene expression levels, our model implies that highly (weakly) expressed genes are more likely to exhibit sublinear (superlinear) scaling with cell volumes. Our predictions is consistent with recent experimental data in *S. cerevisiae* although alternative mechanisms of the correlation between the nonlinear degrees and expression levels may exist [19]. Our results can have far reaching implications as the nonlinear scaling of gene expression level often has important biological functions, such as cell cycle regulation and cell size sensing. Our results suggest that sensing the concentration differences among a group of proteins as a measure of cell volume is probably the easiest option cells can take to achieve cell size regulation.

[1] Harry A Crissman and John A Steinkamp, "Rapid, simultaneous measurement of dna, protein, and cell volume in single cells from large mammalian cell populations," The Journal of Cell Biology **59**, 766 (1973).

[2] SG Elliott and CS McLaughlin, "Rate of macromolecular synthesis through the cell cycle of the yeast saccharomyces cerevisiae," Proceedings of the National Academy of Sciences **75**, 4384–4388 (1978).

[3] Hanna Salman, Naama Brenner, Chih-kuan Tung, Noa Elyahu, Elad Stolovicki, Lindsay Moore, Albert Libchaber, and Erez Braun, "Universal protein fluctuations in populations of microorganisms," Phys. Rev. Lett. **108**, 238105 (2012).

[4] Samuel Marguerat and Jrg Bhler, "Coordinating genome expression with cell size," Trends in Genetics **28**, 560 – 565 (2012).

[5] Hermannus Kempe, Anne Schwabe, Frédéric Crémazy, Pernette J Verschure, and Frank J Bruggeman, "The volumes and transcript counts of single cells reveal concentration homeostasis and capture biological noise," Molecular Biology of the Cell **26**, 797–804 (2015).

[6] Olivia Padovan-Merhar, Gautham P Nair, Andrew G Biaesch, Andreas Mayer, Steven Scarfone, Shawn W Foley, Angela R Wu, L Stirling Churchman, Abhyudai Singh, and Arjun Raj, "Single mammalian cells compensate for differences in cellular volume and dna copy number through independent global transcriptional mechanisms," Molecular Cell **58**, 339–352 (2015).

[7] Robert Ietswaart, Stefanie Rosa, Zhe Wu, Caroline Dean, and Martin Howard, "Cell-size-dependent transcription of flc and its antisense long non-coding rna coolair explain cell-to-cell expression variation," Cell Systems **4**, 622 – 635.e9 (2017).

[8] Xiao-yu Zheng and Erin K O'Shea, "Cyanobacteria maintain constant protein concentration despite genome copy-number variation," Cell Reports **19**, 497–504 (2017).

[9] Benjamin D Knapp, Pascal Odermatt, Enrique R Rojas, Wenpeng Cheng, Xiangwei He, Kerwyn Casey Huang, and Fred Chang, "Decoupling of rates of protein synthesis from cell expansion leads to supergrowth," Cell Systems **9**, 434–445 (2019).

[10] Xi-Ming Sun, Anthony Bowman, Miles Priestman, Francois Bertaux, Amalia Martinez-Segura, Wenhao Tang, Chad Whilding, Dirk Dormann, Vahid Shahrezaei, and Samuel Marguerat, "Size-dependent increase in rna polymerase ii initiation rates mediates gene expression scaling with cell size," Current Biology **30**, 1217 – 1230.e7 (2020).

[11] Manuel Campos, Ivan V Surovtsev, Setsu Kato, Ahmad Paintdakhi, Bruno Beltran, Sarah E Ebmeier, and Christine Jacobs-Wagner, "A constant size extension drives bacterial cell size homeostasis," Cell **159**, 1433–1446 (2014).

[12] Nathan Cermak, Selim Olcum, Francisco Feijo Delgado, Steven C Wasserman, Kristofor R Payer, Mark A Murakami, Scott M Knudsen, Robert J Kimmerling, Mark M Stevens, Yuki Kikuchi, Arzu Sandikci, Masaaki Ogawa, Vincent Agache, Francois Baleras, David M Weinstock, and Scott R Manalis, "High-throughput measurement of single-cell growth rates using serial microfluidic mass sensor arrays," Nature Biotechnology **34**, 1052–1059 (2016).

[13] Jie Lin and Ariel Amir, "Homeostasis of protein and mrna concentrations in growing cells," Nature Communications **9**, 4496 (2018).

[14] Jacob Zhurinsky, Klaus Leonhard, Stephen Watt, Samuel Marguerat, Jürg Bähler, and Paul Nurse, "A coordinated global control over cellular transcription," Current Biology **20** (2010).

[15] M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa, "Interdependence of Cell Growth and Gene Expression: Origins and Consequences," Science **330**, 1099 (2010).

[16] Eyal Metzl-Raz, Moshe Kafri, Gilad Yaakov, Ilya Soifer, Yonat Gurvich, and Naama Barkai, "Principles of cellular resource allocation revealed by condition-dependent proteome profiling," Elife **6**, e28034 (2017).

[17] Kurt M Schmoller, JJ Turner, M Kõivomägi, and Jan M Skotheim, "Dilution of the cell cycle inhibitor whi5 controls budding-yeast cell size," Nature **526**, 268–272 (2015).

[18] Athanasios Litsios, Daphne H. E. W. Huberts, Hanna M. Terpstra, Paolo Guerra, and Matthias Heinemann, "Differential scaling between g1 protein production and cell size dynamics promotes commitment to the cell division cycle in budding yeast," Nature Cell Biology **21** (2019).

[19] Yuping Chen, Gang Zhao, Jakub Zahumensky, Sangeet Honey, and Bruce Futcher, "Differential scaling of gene expression with cell size may explain size control in budding yeast," Molecular Cell **78**, 359 – 370.e6 (2020).

[20] Kora-Lee Claude, Daniela Bureik, Petia Adarska, Abhyudai Singh, and Kurt M Schmoller, "Transcription coordinates histone amounts and genome content," bioRxiv (2020).

[21] MARIE T Killeen and JACK F Greenblatt, "The general transcription factor rap30 binds to rna polymerase ii and prevents it from binding nonspecifically to dna." Molecular and Cellular Biology.

[22] Robert G. Roeder, "The role of general initiation factors in transcription by rna polymerase ii," Trends in Biochemical Sciences **21**, 327 – 335 (1996).

[23] Chin-Tong Ong and Victor G Corces, "Enhancer function: new insights into the regulation of tissue-specific gene expression," Nature Reviews Genetics **12**, 283–293 (2011).

[24] Andrew D Basehoar, Sara J Zanton, and B Franklin Pugh, "Identification and distinct regulation of yeast tata box-containing genes," Cell **116**, 699–709 (2004).

[25] Ron Milo and Rob Phillips, *Cell biology by the numbers* (Garland Science, 2015).

[26] Frank R Neumann and Paul Nurse, "Nuclear size control in fission yeast," The Journal of Cell Biology **179**, 593–600 (2007).

[27] Sophia Hsin-Jung Li, Zhiyuan Li, Junyoung O Park, Christopher G King, Joshua D Rabinowitz, Ned S Wingreen, and Zemer Gitai, "Escherichia coli translation strategies differ across carbon, nitrogen and phosphorus limitation conditions," Nature microbiology **3**, 939 (2018).

[28] Yansheng Liu, Andreas Beyer, and Ruedi Aebersold, "On the dependency of cellular protein levels on mrna abundance," Cell **165**, 535–550 (2016).

[29] Andrea K. Bryan, Vivian C. Hecht, Wenjiang Shen, Kristofor Payer, William H. Grover, and Scott R. Manalis, "Measuring single cell mass, volume, and density with dual suspended microchannel resonators," Lab Chip **14**, 569–576 (2014).

[30] Markus Basan, Manlu Zhu, Xiongfeng Dai, Mya Warren, Daniel Sévin, Yi-Ping Wang, and Terence Hwa, "Inflating bacterial cells by increased protein synthesis," Molecular Systems Biology **11** (2015), 10.15252/msb.20156178.

[31] Gabriel E Neurohr, Rachel L Terry, Jette Lengefeld, Megan Bonney, Gregory P Brittingham, Fabien Moretto, Teemu P Miettinen, Laura Pontano Vaites, Luis M Soares, Joao A Paulo, *et al.*, "Excessive cell growth causes cytoplasm dilution and contributes to senescence," Cell **176**, 1083–1097 (2019).