

# Origins of extrinsic variability in eukaryotic gene expression

## Supplementary Information

Dmitri Volfson<sup>1,2,\*</sup>, Jennifer Y. Marciniak<sup>1,\*</sup>, William J. Blake<sup>3</sup>, Natalie Ostroff<sup>1</sup>, Lev S. Tsimring<sup>2</sup> & Jeff Hasty<sup>1</sup>

<sup>1</sup> Department of Bioengineering, University of California San Diego, La Jolla, CA 92093

<sup>2</sup> Institute for Nonlinear Science, University of California San Diego, La Jolla, CA 92093

<sup>3</sup> Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, MA 02215 and

\*These authors contributed equally to this work

### I. VARIABILITY RESULTS FOR GENES ON DIFFERENT CHROMOSOMES

The dynamics of GFP expression from our localized identical promoters could be tightly correlated due to their proximal location[S1]. We investigated this possibility by constructing two new strains for comparison. In addition to our strain with a single copy on chromosome II (utilized above), we constructed a strain with a single copy on chromosome IV, and a strain with two copies on chromosomes II and IV, respectively. While we observe that the mean values depend on the location at small induction levels, the strain with two copies is on average expressing as much as both copies separately (Fig. 1a). Interestingly, we find a similar trend for the standard deviations, and this leads to a coefficient of variation independent of copy number (Fig. 1b). These results imply that the origin of the variability is not temporal correlations arising from the expression of adjacent genes.

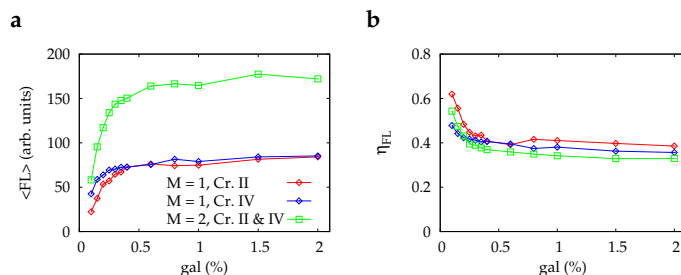


FIG. 1: Constructs on different chromosomes. **a**, Induction curves for two strains with copy number  $M = 1$  located on different chromosomes along with a strain with copy number  $M = 2$  of which one copy of *his3 $\Delta$ 200* is integrated at the *GAL1-10* locus and one copy *trp1 $\Delta$ 63* at the *TRP1* locus. **b**, The collapse of the CV as a function of galactose implies the extrinsic variability does not arise from locality.

### II. INTRINSIC AND EXTRINSIC NOISE LEVELS FROM MULTIPLE COPY NUMBER MEASUREMENTS

The scalings given in the main text are only observed if the genetic fluctuations were of purely intrinsic or extrinsic nature. In general, when both intrinsic and extrinsic fluctuations are present, the relative contributions of intrinsic and extrinsic sources of variability can be calculated from the variances of observed fluorescence fluctuations for different copy number experiments. For example, the variances of the intrinsic ( $V_i$ ) and extrinsic ( $V_e$ ) noise contributions can be found from the variances of the fluorescence fluctuations for a single ( $V_1$ ) and double ( $V_2$ ) insertion strains according to simple formulas  $V_i = 2V_1 - V_2/2$ ;  $V_e = V_2/2 - V_1$ . These formulas are obtained under assumption that both copies of GFP gene have identical strength, however they can be easily generalized to the case of non-equal expression. Furthermore, similar formulas can be derived for an arbitrary pair of strains with different copy numbers of constitutive genes. In the latter case, we find  $V_i = (n^2V_m - m^2V_n)/(mn^2 - nm^2)$ ;  $V_e = (mV_n - nV_m)/(mn^2 - nm^2)$ .

Using these formulas one can calculate the levels of intrinsic and extrinsic noise in the GFP expression and estimated the measurement errors.

### III. VARIABILITY OF GATED SUBPOPULATIONS

We calculated coefficient of variation for subpopulations of cells within narrow windows of cell sizes separated in Fig. 2a by thin vertical lines. As seen in Fig. 2b, the coefficient of variation for subpopulations with small distribution of cell sizes (as measured by the forward scatter in flow cytometer) is up to 50% smaller than that for the whole population. This observation implied that cell growth and division should play an important role in the observed variability of gene expression.

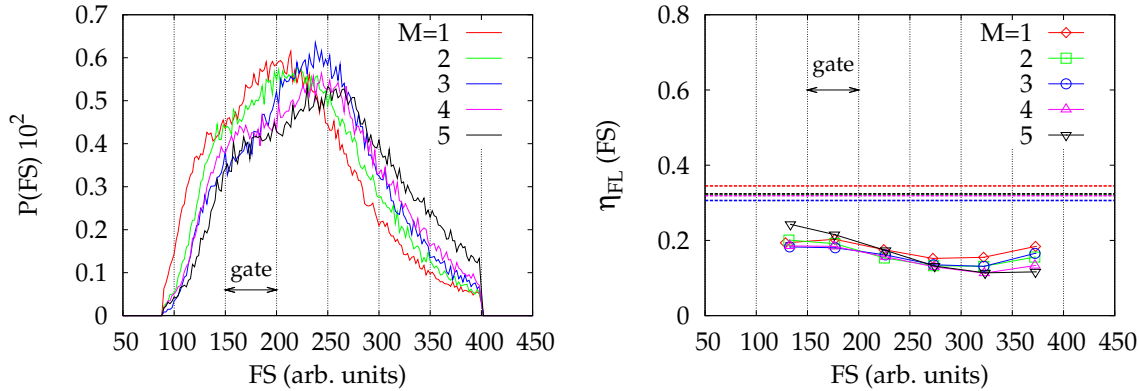


FIG. 2: Cell size distributions (a) and coefficients of variation for *gfp* expression measured within narrow gates separated by thin vertical lines, for several values of the copy number  $M$ .

### IV. REDUCED MODEL FOR GALACTOSE SIGNALING PATHWAY

The galactose signaling pathway in yeast in its simplified form is shown in Fig. 3 (see, e.g. Acar et al., 2005, Aauri et al, 2004). The galactose is transported into the cell by Gal2 and binds and activates Gal3. Gal3, in turn, binds and sequesters inhibitor protein Gal80, which otherwise binds to the constitutive activator Gal4 and renders it unable to bind to the promoter of Gal1 and other Gal genes.

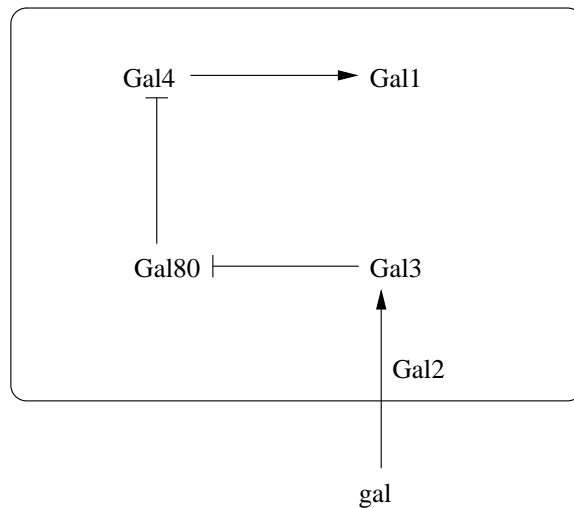
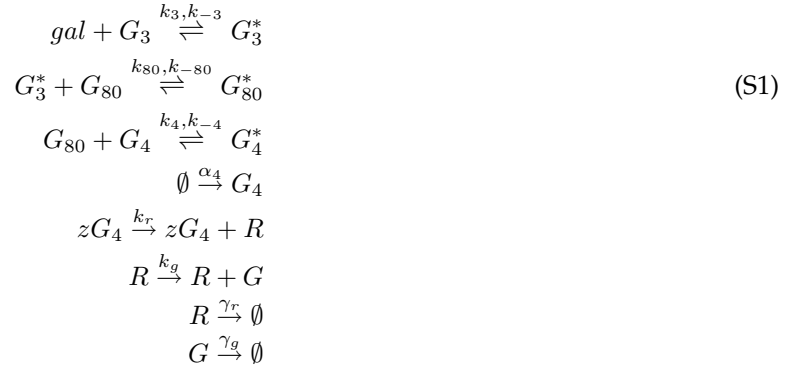


FIG. 3: Primary reactions in the galactose signaling pathway

Thus, the reactions can be formulated as follows



Here  $gal$  denotes the galactose,  $G_3, G_4, G_{80}$  denote the corresponding Gal proteins,  $G_3^*$  is the activated Gal3 proteing bound with galactose,  $G_{80}^*$  is the Gal80 protein bound with the activated Gal3, and  $G_4^*$  is the Gal4-Gal80 complex. We take into account the reaction of transcription of the GFP mRNA which we denote as  $R$  and translation of the mRNA into the GFP protein  $G$ . We also take into account the degradation of Gal4, mRNA and GFP, but neglect the production and degradation of other proteins thereby assuming that their total concentrations remain constant. We also ignore the dimerization reactions of Gal4 and Gal80. Multiplier  $z$  models the cooperativity in Gal4 production, according to experiments,  $z = 8$  as four dimers of Gal4 are needed to to bound to the Gal1 promoter in order to initiate transcription.

The mass action equations corresponding to reactions (S1) read as follows

$$\dot{g}_3^* = k_3[gal]g_3 - k_{-3}g_3^* \tag{S2}$$

$$\dot{g}_{80}^* = k_{80}g_3^*g_{80} - k_{-80}g_{80}^* \tag{S3}$$

$$\dot{g}_4^* = k_4g_4g_{80} - k_{-4}g_4^* \tag{S4}$$

$$\dot{g}_4 = \alpha_4 - k_4g_4g_{80} + k_{-4}g_4^* \tag{S5}$$

$$\dot{r} = k_r NH(g_4) - \gamma_r r \tag{S6}$$

$$\dot{g} = k_g r - \gamma_g g \tag{S7}$$

plus conservation conditions  $g_3 + g_3^* = t_3, g_{80} + g_{80}^* = t_{80}$  (we took into account that there are  $N$  identical copies of  $P_{Gal1} - GFP$  gene pairs).

In the following we assume that reactions (S2),(S3) are in a local equilibrium, and neglect the time derivatives in the l.h.s. of these equations. Then we get local balance conditions,

$$g_3^* = K_3[gal]g_3 = \frac{t_3}{(K_3[gal])^{-1} + 1}, \quad g_{80} = [K_{80}g_3^*]^{-1}g_{80}^* = \frac{t_{80}}{1 + K_{80}g_3^*} \tag{S8}$$

We also assume that the concentrations of Gal3-gal and Gal80 are small ( $g_3^* \ll t_3, g_{80} \ll t_{80}$ ) and neglect the unbinding reaction of the Gal4-Gal80 complex. Then we obtain,

$$g_3^* \approx K_3[gal]t_3, \quad g_{80} \approx \frac{t_{80}}{K_{80}g_3^*} \approx [K_3K_{80}[gal]t_3]^{-1}t_{80} \tag{S9}$$

Denoting  $\gamma_4 = k_4(K_3K_{80}[gal]t_3)^{-1}t_{80}$ , we obtain model equations in the text.

## V. CELLULAR POPULATION MODEL OF BUDDING YEAST (MODEL A)

Here we formulate and analyze a deterministic model which couples a mesoscopic description of cell growth and division with a microscopic description of gene expression. Our starting point is the classical model of growth and unequal division introduced by Hartwell and Unger [S2]. We consider an asynchronous population of cells which grow exponentially, bud, and eventually divide in asymmetric fashion, i.e. the size of a mother after division is larger than that of a daughter. The process of growth and divisions

repeats indefinitely producing a structured population of cells of multiple generations where each cell is characterized by a position within the cell cycle and genealogical age.

Parallel to the volume growth each cell is constantly producing GFP. This production obeys its own dynamics between divisions, which is independent of the growth process. The coupling between the growth and production comes about through the partitioning of the GFP content of the mother before division into daughter cells by specifying the initial conditions for the subsequent dynamics of GFP production.

Following the original model [S2] we make several simplifying assumptions. First, we only distinguish among two generations of cells, daughters and mothers, thus neglecting the difference between mothers of second generation and higher. Second, we assume that independent of the generation cells are growing with the same rate. Next, we make an assumption that both daughters and mothers divide when reach the same size and in the same proportion. Regarding the process of GFP production we make a strong assumption that it is produced with a constant rate between the divisions and distributed among daughters in the same proportion as volume. The former approximation is relevant for the case when mRNA dynamics is fast and degradation of GFP is negligible (see the main text), while the latter may be due to fast diffusion of GFP molecules.

Our simplified model of the formation of structured population of budding yeast cells ignores numerous complications of both cell growth and GFP production sides [S3]. The rationale behind considering such an over simplified system is that our model is analytically tractable and thus allows straightforward interpretation of the results which provide adequate predictions of our experimental observations.

Let us formulate the model explicitly. We describe the dynamics of the size growth and GFP production between the divisions with a set of differential equations

$$\dot{x} = \alpha x, \quad (\text{S10})$$

$$\dot{y} = \gamma, \quad (\text{S11})$$

where  $x$  denotes the size which also characterizes a position of a cell within cycle,  $y$  stands for number of GFP molecules in a cell,  $\alpha$  is the rate of an (exponential) growth of the cell size and  $\gamma$  is the GFP production rate. After the cell reaches the size  $x_0$  it divides asymmetrically and so does GFP,

$$x_0 \rightarrow \kappa x_0 + \xi x_0, \quad y(x_0) \rightarrow \kappa y(x_0) + \xi y(x_0), \quad \text{with } \xi + \kappa = 1 \quad (\text{S12})$$

For definiteness we assume  $\kappa < 1/2$ , so that daughter size and GFP content after division are  $\kappa x_0$  and  $\kappa y(x_0)$  respectively while the rest of cell volume and GFP remains with mother.

We are interested in stationary distributions which are generated by the evolution of a cell population according to rules (S10)-(S12) in the large-time asymptotic limit. Therefore, we consider a joint probability function  $p(x, y, t)$  such that  $p(x, y, t) dx dy$  is a probability of finding an arbitrary cell with size within  $[x, x + dx]$  and GFP within  $[y, y + dy]$  at time  $t$ . According to (S10)-(S12) such probability obeys the following Liouville equation

$$\begin{aligned} \partial_t p(x, y, t) + \alpha \partial_x (x p(x, y, t)) + \gamma \partial_y p(x, y, t) = \\ = \alpha x_0 [\kappa^{-1} \delta_{x, \kappa x_0} p(x_0, \kappa^{-1} y, t) + \xi^{-1} \delta_{x, \xi x_0} p(x_0, \xi^{-1} y, t)] - \\ - \alpha x_0 \int_{y'} dy' p(x_0, y', t) p(x, y, t) \end{aligned} \quad (\text{S13})$$

This equation describes the time evolution of the probability  $p(x, y, t)$  generated by the ensemble of identical dynamical systems (S10),(S11) with different initial conditions. Here the terms on the left side of the equation comprise the full derivative of the distribution function describing its evolution along the trajectories (flux lines);  $\delta_{a,b} = \delta(a - b)$  is the Dirac delta-function. The the first two terms on the right-hand side stand for the influx of cells due to divisions which occur at two i fixed sizes ( $x = \kappa x_0$  for daughters and  $x = \xi x_0$  for mothers); the last term ensures conservation of the probability which could be checked by the integration of both sides of the equation over the stripe  $\{x, y\} \in [0, x_0] \times [0, \infty)$ .

In what follows we shall construct the stationary solution of Eq. (S13). According to the renewal equations (S12), the dynamics of GFP do not affect the dynamics of cell growth, and the distribution of the cell sizes may be obtained irrespectively of the GFP distribution. In fact, performing integration of the equation (S13) over all possible values of  $y$ , we find for  $p_v(x) = \int_0^\infty dy' p(x, y')$

$$\frac{d}{dx} [x p_v(x)] = [\delta_{x, \kappa x_0} + \delta_{x, \xi x_0}] x_0 p_{v0} - x_0 p_{v0} p_v(x) \quad (\text{S14})$$

where  $p_{v0} = p_v(x_0)$ . This ordinary differential equation may be readily solved. Omitting details we present the final result for the stationary distribution of sizes  $p_v(x)$

$$p_v(x) = \begin{cases} 0, & x_0 < x < \kappa x_0, \\ \kappa x_0 x^{-2}, & \kappa x_0 \leq x < \xi x_0, \\ x_0 x^{-2}, & \xi x_0 \leq x \leq x_0, \end{cases} \quad (\text{S15})$$

and  $p_{v0} = x_0^{-1}$ . This solution can be easily verified by direct substitution in Eq.(S14). It is presented in Fig. 4(a) along with the results of numerical simulation of the system (S10)-(S12). Obviously  $p_v = 0$  for  $x > x_0$ , since we do not allow cells to grow beyond the maximum size  $x_0$ . Similarly, we notice that in the stationary state there is a cut-off size  $\kappa x_0$  below which  $p_v = 0$ , this is due to the absence of influx of the offspring for  $x < \kappa x_0$ . An important feature of the size distribution is the presence of two distinct peaks located at  $x = \kappa x_0$  and  $x = \xi x_0$ , they correspond to influx of cells right after divisions, and formally generated by point-source terms in the Eq.(S14) (first two terms on the right side).

Let us now find the stationary joint probability  $p(x, y)$  which should satisfy the following partial differential equation

$$\begin{aligned} \alpha \partial_x (xp(x, y)) + \gamma \partial_y p(x, y) = \\ = -\alpha p(x, y) + \alpha x_0 [\kappa^{-1} \delta_{x, \kappa x_0} p(x_0, \kappa^{-1} y) + \xi^{-1} \delta_{x, \xi x_0} p(x_0, \xi^{-1} y)] \end{aligned} \quad (\text{S16})$$

which is the stationary version of Eq. (S13) simplified with the help of Eq. (S15).

The solution of Eq. (S16) can be found by the method of characteristics [S4]. The characteristic curves are phase trajectories of the following system

$$\frac{dx}{ds} = \alpha x, \quad (\text{S17})$$

$$\frac{dy}{ds} = \gamma, \quad (\text{S18})$$

$$\frac{dp}{ds} = -\alpha p(x, y) + \alpha x_0 [\kappa^{-1} \delta_{x, \kappa x_0} p(x_0, \kappa^{-1} y) + \xi^{-1} \delta_{x, \xi x_0} p(x_0, \xi^{-1} y)], \quad (\text{S19})$$

so that solution of the original PDE(S16) is found implicitly as a solution of the ODE (S19) along the family of trajectories of the subsystem (S17,S18). The latter are readily found for arbitrary initial conditions  $\hat{x}, \hat{y}$ ,

$$y - \hat{y} = a [\ln x - \ln \hat{x}], \text{ where } a = \gamma \alpha^{-1} \quad (\text{S20})$$

However, we have to find the family of phase trajectories which correspond to the stationary distribution of cell. In this case the ‘‘outflux’’ of cell through the line  $x = x_0$  in  $(x, y)$ -plane has to be exactly compensated by the influxes through the lines  $x = \kappa x_0$  and  $x = \xi x_0$ . It is easy to see that the family of phase trajectories corresponding to this flux conservation form a ax-like shape shown in Figure 4(b) by shading. This family can be parametrized by the parameter  $\mu \in [0, 1]$ , so the cell trajectories are given by the following initial conditions

$$\hat{x} = \kappa x_0, \quad \hat{y} = y_{03} + \mu(y_{04} - y_{03}), \quad \text{daughters} \quad (\text{S21})$$

$$\hat{x} = \xi x_0, \quad \hat{y} = y_{01} + \mu(y_{02} - y_{01}), \quad \text{mothers} \quad (\text{S22})$$

where  $y_{01}, y_{02}$  and  $y_{03}, y_{04}$  are lower and upper bounds of the amount of GFP for mothers and daughters in the beginning of cell cycle. Now to specify the trajectories completely we have to find  $y_{01}, y_{02}, y_{03}, y_{04}$  from the closure conditions. For example, substituting  $\hat{x} = \xi x_0, \hat{y} = y_{03}$  and  $x = x_0, y = y_{03}/\xi$  into Eq.(S20), we obtain  $y_{01}$ , etc. The result reads as follows:

$$y_{01} = a \ln (\xi^{-1}) \xi \kappa^{-1}, \quad (\text{S23})$$

$$y_{02} = a \ln (\kappa^{-1}), \quad (\text{S24})$$

$$y_{03} = a \ln (\xi^{-1}), \quad (\text{S25})$$

$$y_{04} = a \ln (\kappa^{-1}) \kappa \xi^{-1}. \quad (\text{S26})$$

Now in order to find the joint probability distribution we need to solve Eq. (S19) along the characteristics specified above. There is however a simpler way which explores the following remarkable property of the two-dimensional distribution: when  $y$  is between the “limiting” characteristics  $p(x, y) = \text{const}(x)$  and  $p(x, y) = 0$  elsewhere. To demonstrate this we introduce four functions  $f(y) = p(\xi x_0, y), y \in [y_{01}, y_{02}]$ ,  $q(y) = p(\kappa x_0, y), y \in [y_{03}, y_{04}]$ ,  $r(y) = p(x_0, y), y \in [y_{11}, y_{12}]$ , and  $h(y) = p(x_0, y), y \in [y_{13}, y_{14}]$  which represent the dependence of  $p(x, y)$  along the vertical segments connecting limiting characteristics (see Fig. 4(b)) in the beginning and the end of the cell cycle. Now we make use of the conservation of the probability flux from the segment  $x_0 \times [y_{11}, y_{14}]$  into the segment  $\kappa x_0 \times [y_{03}, y_{04}]$  which reads

$$\int_{y_{11}}^{y_{13}} dy' r(y') + \int_{y_{13}}^{y_{14}} dy' h(y') = \int_{y_{11}}^{y_{14}} dy' h(\kappa y') = \int_{y_{11}}^{y_{14}} dy' r(\xi y') \quad (\text{S27})$$

where the two last integrals are obtained by the change of variables from  $\int_{y_{03}}^{y_{04}} dy' \kappa q(y')$  and  $\int_{y_{01}}^{y_{02}} dy' \xi f(y')$  with the help of relations  $y_{14} = y_{04}/\kappa = y_{02}/\xi$ ,  $y_{11} = y_{03}/\kappa = y_{02}/\xi$  and Eq. (S15). The system of equations (S27) comprises the map of  $p(x_0, y), y \in [y_{11}, y_{14}]$  on itself and can be thought of as a Perron-Frobenius equation. In our case there is a simple solution which satisfies (S27)

$$p(x_0, y) = \begin{cases} \text{const}, & y_{11} \leq y \leq y_{14}, \\ 0, & \text{otherwise} \end{cases} \quad (\text{S28})$$

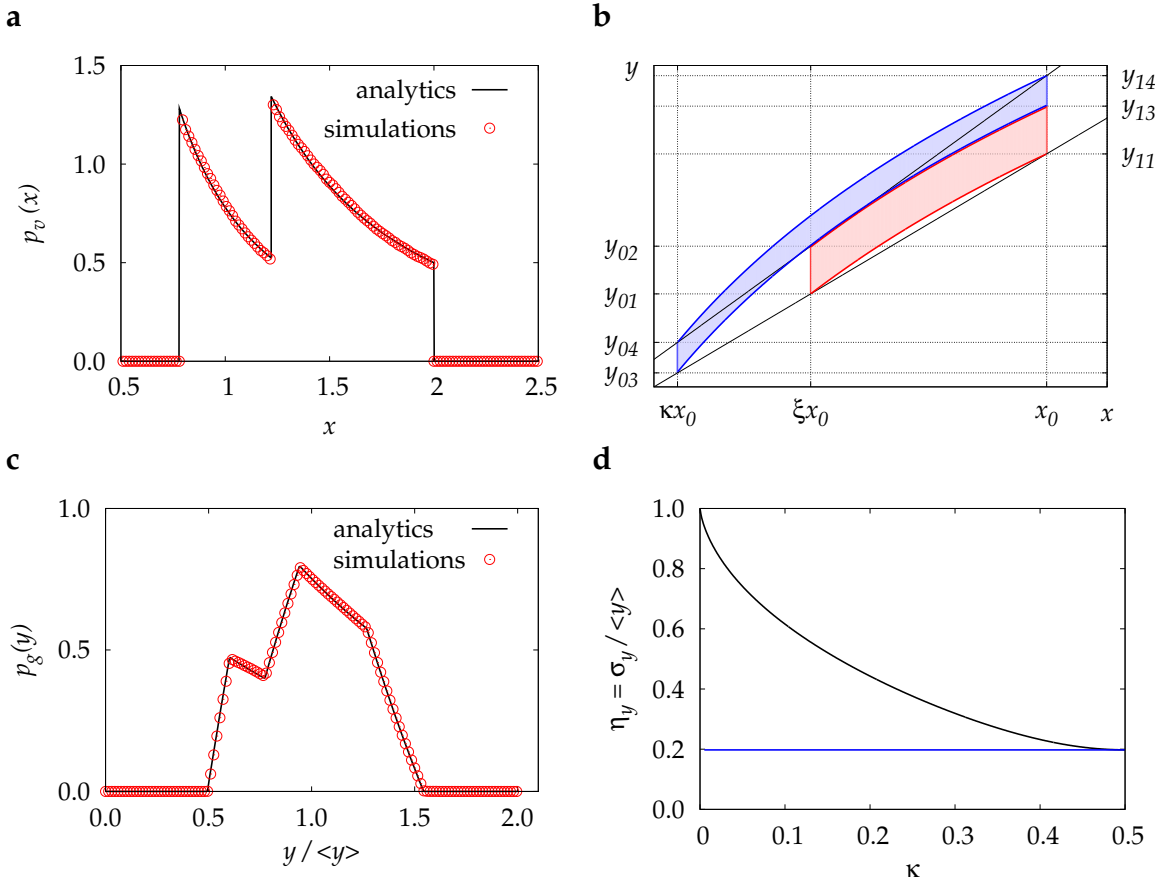


FIG. 4: (a) distribution of sizes from Model A (S13)(line) along with the results of numerical simulations of the equations (S10-S12)(symbols) (b) Characteristics of the stationary solution of the Eq. (S13), (c) distribution of GFP from Model A given by the Eq. (S30) (line) along with the results of numerical simulations of the equations (S10-S12)(symbols) (d) Model A predicts  $\eta_y$  - coefficient of variation of GFP to be independent of the galactose and specified by the parameter  $\kappa$  - daughter/mother size ratio.

which suggest that size dependence may be factorized. Using Eq. (S15) we obtain the functional form of the two-dimensional distribution

$$p_{vg}(x, y) = \Delta^{-1} p_v(x) p_g(y), \quad \Delta = y_{04} - y_{03} = \frac{a}{\xi} \ln \left( \frac{\xi^\xi}{\kappa^\kappa} \right) \quad (\text{S29})$$

Finally, integrating (S29) over  $x$  within the limiting characteristics we find the  $y$ -dependence,

$$p_g(y) = (\xi \Delta)^{-1} \begin{cases} 0 & 0 \leq y < y_{03} \\ (e^{-\frac{y}{a}} - \xi) & y_{03} \leq y < y_{04} \\ e^{-\frac{y}{a}} \left( \xi \kappa^{-\frac{\kappa}{\xi}} - 1 \right) & y_{04} \leq y < y_{01} \\ \left( \kappa^{-\frac{\kappa}{\xi}} \xi - 1 - \kappa \xi^{-\frac{\xi}{\kappa}} \right) e^{-\frac{y}{a}} + \kappa^{-1} & y_{01} \leq y < y_{02} \\ e^{-\frac{y}{a}} \left( \kappa \xi^{-\frac{\xi}{\kappa}} - \xi \kappa^{-\frac{\kappa}{\xi}} \right) & y_{02} \leq y < y_{11} \\ \left( -\kappa + e^{-\frac{y}{a}} \kappa^{-\frac{\kappa}{\xi}} \right) \xi & y_{11} \leq y < y_{14} \\ 0 & y_{04} \leq y \end{cases} \quad (\text{S30})$$

Distribution (S30) is presented in Fig.(4)(c) along with the results of the simulations which produce the identical result.

Now using Eq.(S30) we can find the mean and the standard deviation of GFP distribution,

$$\langle y \rangle = a \quad (\text{S31})$$

$$\sigma_y = az(\kappa) \quad (\text{S32})$$

$$\eta_y = \left[ 1 - \frac{2}{3\kappa\xi} \left( \kappa^2 \ln(\kappa)^2 + \kappa\xi \ln(\kappa) \ln(\xi) + \xi^2 \ln(\xi)^2 \right) \right]^{1/2} \quad (\text{S33})$$

In Fig. (4)(d) we present the coefficient of variation for GFP as a function of the partition ratio  $\kappa$ .  $\eta_y$  monotonously decreases as  $\kappa$  approaches the limit of symmetric division  $\kappa = 1/2$ . The asymptotic value of  $\eta_y$  for equal division is  $\lim_{\kappa \rightarrow 1/2} z(\kappa) = (1 - 2(\ln 2)^2)^{1/2} \approx 0.198$ .

For the comparison with the experiment we used an effective ratio of the volumes of daughter after division to mother before the division,  $\kappa = 0.34$ , representing the weighted average of  $\kappa$  among several generations [S3] (see Figure 1g in Main Text).

## VI. MODEL B

Model B was solved numerically using a hybrid technique[S10] whereby the stochastic dynamics of GAL4p protein have been modeled with the Direct Gillespie algorithm[S11], the intermediate reactions of dimerization and binding of GAL4p dimers were eliminated employing quasi-stationary assumptions for the sake of simplicity and the dynamics of *mRNA* and *GFP* were simulated using the rate equations presented in Box 1. For the population dynamics we used previously reported measurements regarding the structure of the population of *Saccharomyces cerevisiae*[S3]. Namely, for the first generation, we used  $\kappa = 0.4$  and for subsequent generations  $\kappa = 0.3$ , the mean size of the cells at divisions was chosen to be 1.089, 1.179, 1.268, for genealogical ages from 2 to 4 respectively and 1.357 for all older generations; and the size at division for a particular cell was drawn from a narrow gaussian distribution near these mean values with coefficient of variation 0.15. The parameters of the simulations in units of average growth rate are,  $k_r = 80$ ,  $\gamma_m = 4$ ,  $k_g = 4$ ,  $\gamma_g = 0$ ,  $k_4 = 8$ ,  $\gamma_4 = 0.5 \text{ gal}^{-1}$ ,  $K = 0.05$ . The protocol of the simulations closely resembled the experimental procedure. We started from a small collection of cells (typically  $10^3$ ), with Gaussian distributions of sizes and quantities of GAL4p, mRNA, and GFP near the expected mean values. This small population was "grown" until it reached 100,000 cells. The state of the population was recorded within a narrow time window right after that. We checked that this measurement was consistent with the time-ensemble average over a longer period of time indicating that the evolution of the population reached the stationary state.

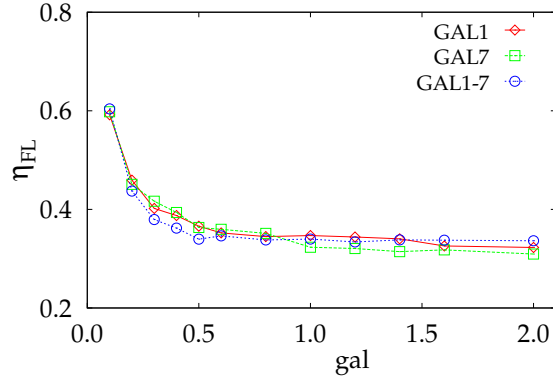


FIG. 5: Coefficient of variation as a function of galactose for three promoters regulated by GAL4p. Two strains contain GAL1p and GAL7p fusions, respectively, while a third strain contains both a GAL1p and GAL7p fusion. The collapse is predicted by the model.

## VII. GAL4P REGULATION OF OTHER PROMOTERS

If the extrinsic variability observed from the *GAL1* promoter at low galactose is due to the GAL4p regulator, an interesting model prediction is that other GAL4p-regulated promoters should also exhibit extrinsic noise at low galactose. We tested this prediction with the construction of two additional strains. In the first strain, we fused GFP to GAL7p, whose promoter is activated by GAL4p. In a second strain, we fused to both GAL7p and to GAL1p. As predicted by the model, we observed scaling which indicates that extrinsic noise dominates the system at low galactose levels (Fig. 5). While it is still certainly possible that other upstream factors are responsible for the observed downstream variations, these results provide additional support for our hypothesis that variability in the GAL4p activator is a dominant factor.

## VIII. PLASMID CONSTRUCTION

The first plasmid was constructed by combining the integrative feature of pRS405 (Stratagene), and the *GAL1-yEGFP* portion of pESC1-yG. Each plasmid was digested using the restriction enzyme PvuII and the pertinent pieces were ligated together to make pRS51-yG. Plasmids pRS31-yG, pRS41-yG, and pRS61-yG differ from pRS51-yG in their yeast selective markers (see Table I). All plasmids contain an ampicillin resistance marker and the ColeI replication origin. All restriction enzymes and T4 DNA ligase were purchased from New England Biolabs. Plasmids were introduced into *Escherichia coli* XL10-Gold (Stratagene) by using a standard heat shock transformation protocol [S6]. All bacterial cells were grown in LB (Fisher) with 100  $\mu\text{g}/\text{mL}$  of ampicillin (Sigma).

## IX. FUSION STRAINS

Yeast-enhanced Green Fluorescence Protein (yEGFP) was fused to several yeast genes that were found to be highly expressed during normal growth conditions [S9] and three genes of the galactose utilization pathway (see Table III). Fusion strains were created by amplifying pKT128 [S5] using the Phusion<sup>TM</sup> High-Fidelity DNA Polymerase (Finnzymes) PCR system following the manufacturer's instructions. Each primer consists of 40 base pairs that overlap with the yeast gene of interest in order to encourage homologous recombination. The portion of the plasmid that is amplified contains yEGFP and the *SpHIS5* selective marker. The PCR product was introduced into YPH500 using heat shock.

Plasmid Name	Description
pRS31-yG	integrating vector with <i>GAL1</i> promoter driving <i>yEGFP</i> ; his3 marker
pRS41-yG	integrating vector with <i>GAL1</i> promoter driving <i>yEGFP</i> ; trp1 marker
pRS51-yG	integrating vector with <i>GAL1</i> promoter driving <i>yEGFP</i> ; leu2 marker
pRS61-yG	integrating vector with <i>GAL1</i> promoter driving <i>yEGFP</i> ; ura3 marker

TABLE I: Summary of integration plasmids.

Strain	Description
YPH500	WT strain [S7]
JYM101	pRS31-yG integrated at <i>GAL1-10</i> locus
JYM113	pRS31-yG and pRS41-yG integrated at <i>GAL1-10</i> locus
JYM154	pRS31-yG, pRS51-yG, and pRS61-yG integrated at <i>GAL1-10</i> locus
JYM198	pRS31-yG, pRS41-yG, pRS51-yG, and pRS61-yG integrated at <i>GAL1-10</i> locus
JYM199	pRS31-yG, pRS41-yG, pRS51-yG, and pRS61-yG integrated at <i>GAL1-10</i> locus
JYM203	pRS41-yG integrated at TRP1 locus
JYM213	pRS31-yG integrated at <i>GAL1-10</i> locus; pRS41-yG integrated at TRP1 locus
FUS001	<i>yEGFP</i> fused to <i>FBA1</i>
FUS002	<i>yEGFP</i> fused to <i>YEF3</i>
FUS003	<i>yEGFP</i> fused to <i>RPP2B</i>
FUS004	<i>yEGFP</i> fused to <i>SOD1</i>
FUS005	<i>yEGFP</i> fused to <i>RPS26B</i>
FUS006	<i>yEGFP</i> fused to <i>HSP82</i>
FUS007	<i>yEGFP</i> fused to <i>RPP2A</i>
FUS008	<i>yEGFP</i> fused to <i>TSA1</i>
FUS009	<i>yEGFP</i> fused to <i>GAL1</i>
FUS010	<i>yEGFP</i> fused to <i>GAL7</i>
FUS011	<i>yEGFP</i> fused to <i>GAL10</i>

TABLE II: Summary of yeast strains.

## X. GENE EXPRESSION EXPERIMENTS

Yeast strains are plated from frozen stock on agar plates with synthetic drop-out (SD) medium supplemented for selection of transformants containing 2% glucose and grown for 40-48 hours. One colony is chosen and grown in liquid SD medium with appropriate supplements and 2% raffinose. Raffinose is used as the primary carbon source to avoid the inhibitory effects of glucose on the galactose-utilization pathway. To prepare cells for the flow cytometer, they are pelleted, resuspended in filtered PBS (Sigma), and vortexed to reduce clumps.

Gene	Function
FBA1	Fructose 1,6-bisphosphate aldolase: catalyses the reversible formation of fructose-1,6 bis-phosphate from dihydroxyacetone phosphate and glyceraldehyde-3-phosphate.
YEF3	Translational elongation factor 3: stimulates the binding of aminoacyl-tRNA to ribosomes.
RPP2B	Ribosomal protein P2 beta: regulates the accumulation of P1 in the cytoplasm
SOD1	Cu, Zn superoxide dismutase: catalyzes the dismutation of superoxide anion
RPS26B	Ribosomal protein S26B: protein component of the small (40S) ribosomal subunit
HSP82	Heat shock protein 90: a cytoplasmic chaperone that is required for pheromone signaling and negative regulation of Hsf1p
RPP2A	Ribosomal protein P2 alpha: regulates the accumulation of P1 in the cytoplasm
TSA1	Thioredoxin-peroxidase: reduces H <sub>2</sub> O <sub>2</sub> and alkyl hydroperoxides
GAL1	Galactokinase: phosphorylates alpha-D-galactose to alpha-D-galactose-1-phosphate
GAL7	Galactose-1-phosphate uridyl transferase: synthesizes glucose-1-phosphate and UDP-galactose from UDP-D-glucose and alpha-D-galactose-1-phosphate
GAL10	UDP-glucose-4-epimerase: catalyzes the interconversion of UDP-galactose and UDP-D-glucose in galactose metabolism; also catalyzes the conversion of alpha-D-glucose or alpha-D-galactose to their beta-anomers

TABLE III: Summary of yeast genes fused to *yEGFP* [S8].

- 
- [S1] Cohen, B. A. , Mitra, R. D., Hughes, J. D., Church, G. M. A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nat. Genetics*, **26**, 183-186 (2000).
- [S2] Hartwell, L. H. , Unger, M. W. Unequal division in *Saccharomyces cerevisiae* and its implications for the control of cell division. *J. Cell Biol.* **75**, 422-35 (1977).
- [S3] Woldringh, C. L. , Huls P.G, Vischer, N. O. E., Volume growth of daughter and parent cells during the cell cycle of *Saccharomyces cerevisiae* a/ $\alpha$  as determined by image cytometry. *J. Bacter.* **75**(10), 3174-3181 (1993).
- [S4] Metz, J.A.J., Diekmann, O. The dynamics of physiologically structured populations. *Lect. Notes Biomath.* **68**, Springer-Verlag, Berlin, 1986
- [S5] Sheff, M. A., Thorn, K. S. Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae* *Yeast* **21**(8), 661-67 (2004)
- [S6] Sambrook, J., Fritsch, E.F., & Maniatis, T. The dynamics of physiologically structured populations. *Molecular Cloning: A Laboratory Manual* , Cold Spring Harbor Lab. Press, Plainview, NY, 1989
- [S7] Sikorski, R. S. and P. Hieter, 1989 A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19-27 (1989)
- [S8] The information presented in this table may be found in *Saccharomyces Genome Database*, <http://www.yeastgenome.org>
- [S9] Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E.K., Weissman J.S. Global analysis of protein expression in yeast *Nature* **425**, 737-741 (2003)
- [S10] Adalsteinsson, D., McMillen, D., Elston, T. Biochemical Network Stochastic Simulator (BioNetS): Software for stochastic modeling of biochemical networks *BMC Bioinformatics* **5**, 24 (2004).
- [S11] Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions *J. Phys. Chem.* **81**, 2340-2361 (1977).