Contents

S1 Supplemental Data

Figure S1

Figure S1. Construct and technical validation, related to Figure 1. **A, left panel** Nexcelom cell counting software with maximum declustering settings versus **A, right panel** the same cells counted with limited declustering. Edge cases tend to be called as a single cell with aggressive declustering. **B-F** Steady state distributions of SSA1^P-H-GFP, SSA1^P-Y-GFP, single and double copy strains at 22°C and 37°C, and SSA1^P-GFP at 22°C as measured by flow cytometry. Black dashed lines are autofluorescence in the GFP negative background strain, BY4743. Distributions are not deconvoluted from autofluorescence. **G** Plotted are biological replicates of single copy SSA1^P-GFP strains during heat shock induction. Once induction was observed, cycloheximide was added (time zero) and plotted versus a no-cycloheximide control. Fit of blocked samples assuming the same degradation rate for folded molecules as unfolded yields a maturation rate of $.0803 \pm .0016 \text{min}^{-1}$.
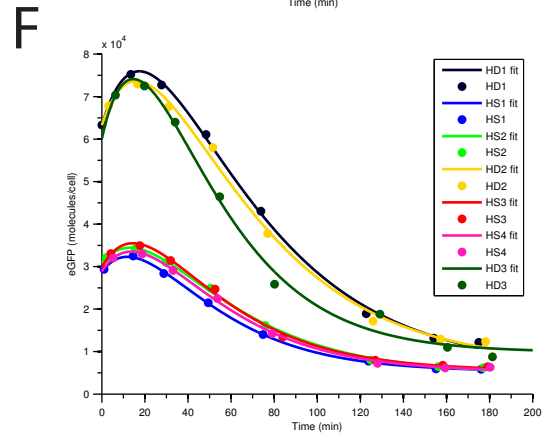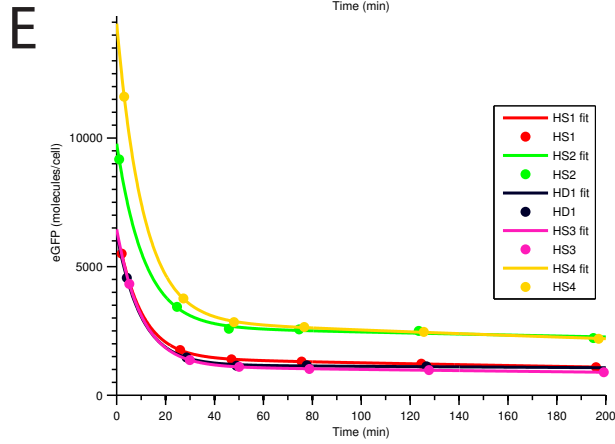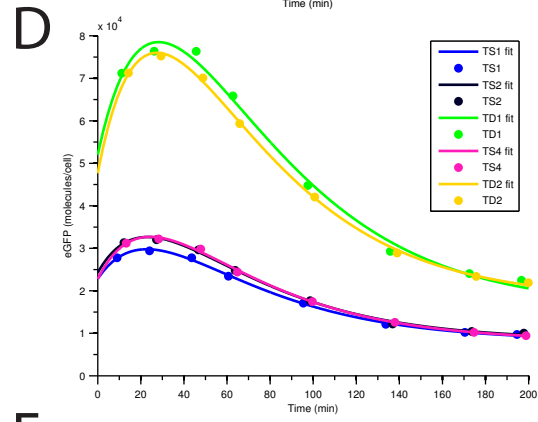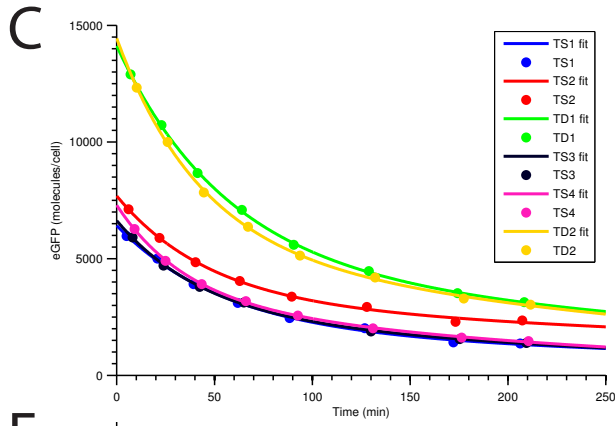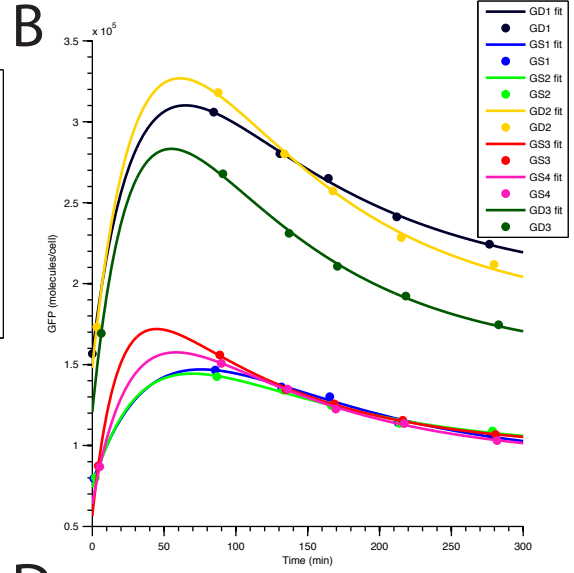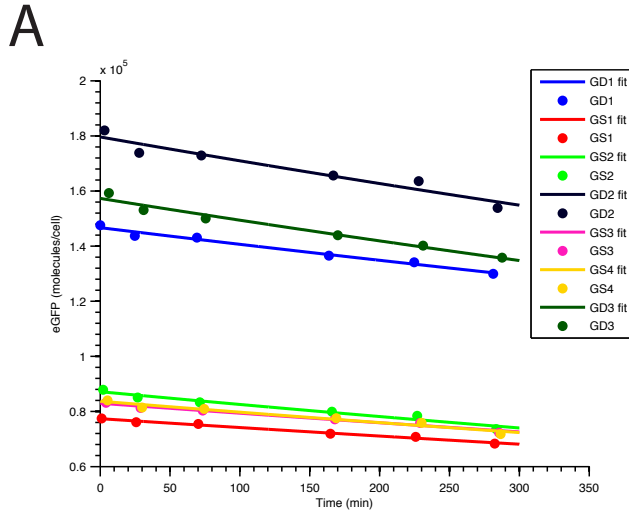
Figure S2. Example RNA and protein decay curves for SSA1^P-H-GFP and SSA1^P-Y-GFP, related to Figure 2. Each trajectory represents a biological replicate with: G=eGFP, S=single-copy strain, D=double-copy strain, and numbers indicate specific replicates. Trends were fit to each line separately and resulting parameter estimates averaged to generate the error bars in Figure 3, main text. The shown trajectories are data conditioned on the largest forward scatter bin; this process was repeated for all bins. **A and B** cyclohexamide block and heat-shock relaxation decay curves for SSA1^P-GFP. **C and D** cyclohexamide block and heat-shock relaxation decay curves for SSA1^P-Y-GFP. **E and F** cyclohexamide block and heat-shock relaxation decay curves for SSA1^P-H-GFP.
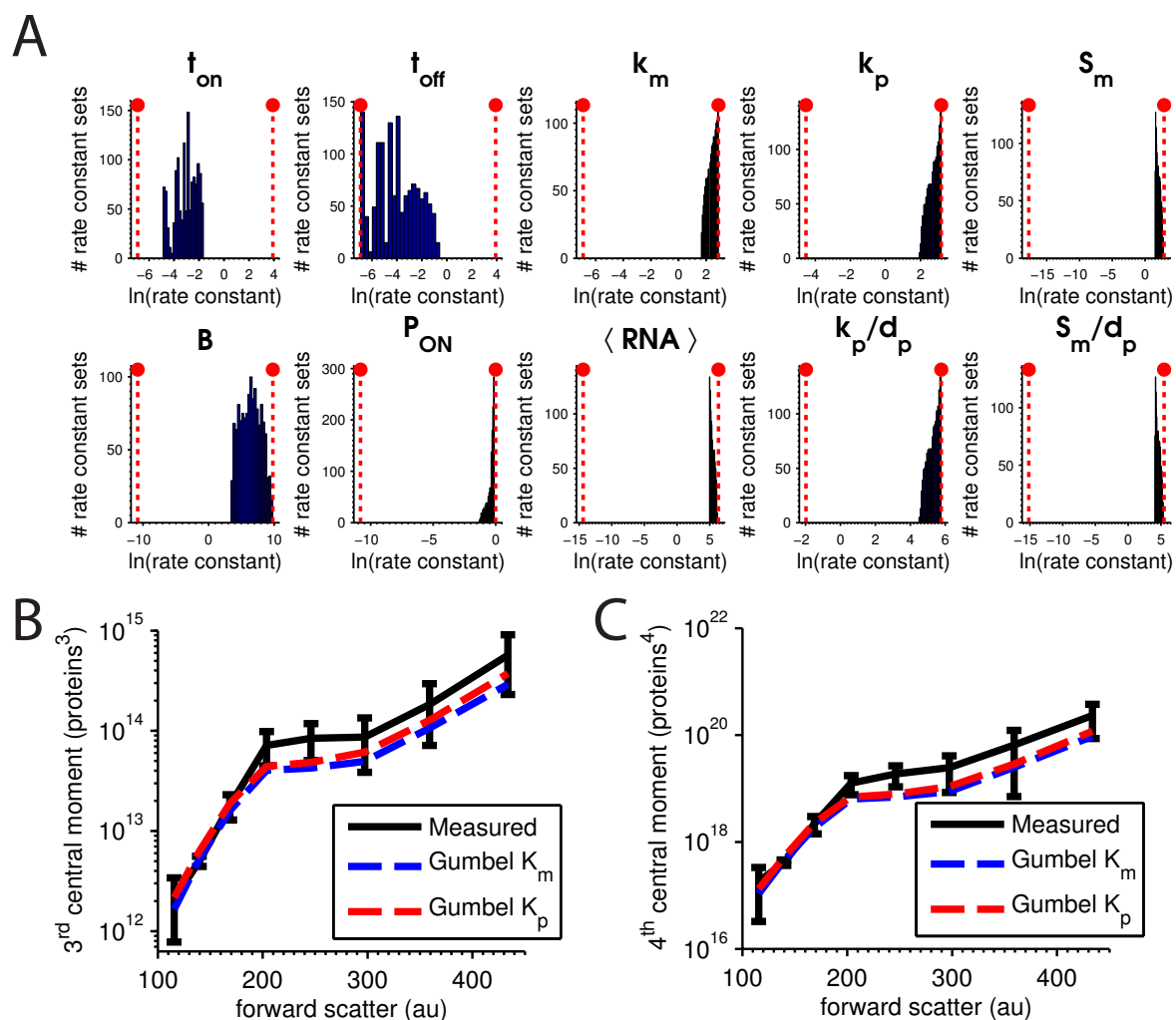
Figure S3

Figure S3. Modeling results for SSA1$^{\mathbf{P}}$-GFP at 37°C, related to Figure 3. **A** ACES ensemble solutions for each parameter, fitted to GFP distribution at 37°C using only mean and variance but constraining fits to measured degradation rates, $d_p = .0078 min^{-1}$, $d_m = .033 min^{-1}$. Hyperparameters are P_{on} (probability that a promoter is active, $t_{on}/(t_{on} + t_{off})$), S_m (RNA synthesis rate, $P_{on}k_m$), B (burst size, number of RNAs produced per ON duration, k_m/t_{off}), $\langle RNA \rangle$ (average count of RNAs, S_m/d_m), k_p/d_p (ratio of average protein count to average RNA count), and S/d_p (number of RNAs produced per cell cycle if protein decay driven by dilution). **B and C** Comparison of extrinsic K_m and extrinsic K_p models versus measured skewness (B) and kurtosis (C).
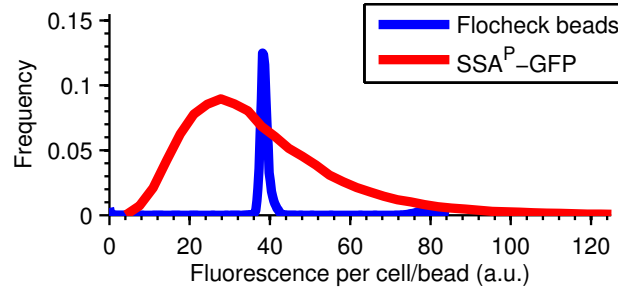
Figure S4

Figure S4. SSA1$^{\mathbf{P}}$-GFP expression at 37°C versus Flow Check microspheres, related to Figure 1. Distribution of fluorescence signal measured from Flow Check microsphere calibration beads plotted against SSA1$^{\mathrm{P}}$-GFP expression at 37°C.

Table S1. Intrinsic and extrinsic decomposition of SSA1$^\mathbf{P}$-Y-GFP and SSA1$^\mathbf{P}$-H-GFP variance, related to Figure 1

	I_v (prot^2)	E_v (prot^2)	fraction intrinsic	Mean (prot/cell)
SSA1$^\mathbf{P}$-Y-GFP 37°C	1.21E+06	9.81E+06	0.1100	4921.5
SSA1$^\mathbf{P}$-Y-GFP 22°C	2.98E+06	4.97E+06	0.375	2796.1
SSA1$^\mathbf{P}$-H-GFP 37°C	1.41E+06	9.04E+06	0.135	4170.4
SSA1$^\mathbf{P}$-H-GFP 22°C	5.55E+06	6.78E+06	0.45	3114.2
SSA1$^\mathbf{P}$-GFP 37°C	1.62E+08	1.93E+09	0.0775	81738
SSA1$^\mathbf{P}$-GFP 22°C	1.52E+07	8.47E+07	0.1525	9745.6

Data S1. Degradation data for SSA1^P-GFP, SSA1^P-Y-GFP and SSA1^P-H-GFP, related to Experimental Procedures. Contains single exponential and double exponential fits for SSA1^P-GFP, SSA1^P-Y-GFP and SSA1^P-H-GFP RNA and protein degradation experiments.

Data S2. AIC ranking for arbitrary distributions fit to cytometry data, related to Experimental Procedures. Spreadsheets contain AIC rankings of best statistical distribution fit to the experimental distributions produced from the first plate of 96 random promoters in the strong, weak, and no promoter library from Mogno et al, "TATA is a modular component of synthetic promoters." Genome Research, 2010.

Data S3. Model values across forward scatter, related to Figure 2. Contains measurements (mean, intrinsic variance, extrinsic variance, d_m, and d_p) for synchronized slices across forward scatter, along with the associated model-fit parameters. Sheet1 contains a detailed key.

S2 Supplemental Experimental Procedures

S2.1 Calibration of flow cytometry in absolute protein units

We quantified single-cell GFP counts and concentrations by performing flow cytometry, single-cell counting by imaging cytometry, and fluorescence correlation spectroscopy on individual reference samples. Yeast strain SSA1^P-GFP was used for all reference calibration experiments because expression was relatively stable during the preparatory steps. An autofluorescent control strain, BY4743 was grown in parallel. For both strains, a single colony was inoculated into $5mL$ synthetic complete (SC) media and grown overnight in a rotating drum (60 RPM, 37°C). In the morning, optical density was measured and both SSA1^P-GFP and BY4741 were diluted to an optical density (OD600) of .025 in $50mL$ of SC in $250mL$ flasks and placed on a floor shaker (300RPM, 37°C). After 6 hours, optical densities were taken and generally found to be in close agreement at an O.D. of .4, corresponding to a 90 minute division time at 37°C.

The two cultures were spun down in $50mL$ falcon tubes at 3000RPM for 3 minutes, then resuspended in 1 to 2 mL of water, depending on the application. At this point, each sample was split three ways, with $20\mu L$ going to cell counting, $100\mu L$ going to flow cytometry, and the remainder being split into technical replicates for lysis and fluorescence correlation spectroscopy.

S2.1.1 Cell lysis

Cells were aliquoted into technical replicates. For the calibration experiments used for all cytometry measurements reported, $250\mu L$ aliquots (V_{load}) were separated into microcentrifuge tubes and spun down at $3000g$. After discarding the supernatant, the cells were resuspended in lysis buffer. The most reproducible lysis method in our hands was a chemical lysis kit, CelLytic Y (Sigma-Aldrich, C4482). We used the kit protocol with the following proportions and additions. A mastermix was made immediately prior to spinning the cells down to minimize the time between flow cytometry and lysis. A mastermix consists of the following final concentrations 5mM DTT, $>50U/mL$ Benzonase Nuclease (EMD Millipore, 70746-4), 1X cOmplete Mini, EDTA-free (Roche Applied Science, 11-836-170-001), and CelLytic Y to the appropriate volume. Because of the various arbitrary units, and our deviation from exact specifications, the following is an example mastermix: $54\mu L$ DTT (.1M), $2.5\mu L$ Benzonase, $154\mu L$ cOmplete Mini, and $870\mu L$ CelLytic Y for a final volume of $1080\mu L$. We resuspended the pelleted cells in $120\mu L$ (V_{lys}) of lysis mastermix, briefly vortexed each tube, and placed the microcentrifuge tubes on a shaker (30°C, 180RPM)

for 20 minutes. The cells were then spun down at 20,000g for 20 minutes, and $110\mu L$ of supernatant was pipetted into new microcentrifuge tubes for protein quantification by fluorescence correlation spectroscopy (FCS). The cell pellet and remaining supernatant were resuspended in $100\mu L$, and the volume measured to determine the pellet size and original volume.

S2.1.2   Fluorescence correlation spectroscopy (FCS)

Supernatants from cell lysis were kept on ice and measured by FCS the same day. Samples were loaded on Nunc Lab-Tek Chambered Coverglass (Thermo Scientific, 155411) and imaged on a Zeiss Axiovert 200M with ConfoCor 2 and LSM 510 modules (Carl Zeiss-Evotec, Jena, Germany) and a 40X water-immersion objective. Excitation was with an argon laser line at 488nm and emission was monitored by bandpass filter in the range 505nm to 530nm. All autocorrelation analyses were performed with the accompanying commercial LSM 510 / ConfoCor 2 software (Carl Zeiss-Evotec, Jena, Germany). We assumed the simplest autocorrelation model: diffusive motion for a single species with accounting for triplet states.

S2.1.3   Flow cytometry for calibration experiment

$100\mu L$ of SSA1$^\text{P}$-GFP and BY4741 were each taken and split into 4 samples of $25\mu L$ and diluted in $175\mu L$ of SC media. Each $200\mu L$ sample was loaded onto a Costar round bottom 96-well plate in alternating fashion. We then loaded the GFP calibration protocol which, once set, was used for all following experiments. See Methods: Flow cytometry for general equipment and running details. For the calibration experiments, the wells were run for 120 seconds continuously and in parallel to cell lysis. This allowed us to observe how much expression changes during the the processing. If there were even modest changes in expression, we used the average fluorescence measured in the well closest in time to the point of suspending the cells in lysis buffer.

S2.1.4   Cell counting

At the usual density of OD .4, cells were diluted into water at ratios 1:10 and 1:40 and loaded into disposable $20\mu L$ Nexcelom Counting Chambers (Nexcelom Bioscience, CHT4-SD100-002). We measured 4 technical replicates at each dilution on an imaging cellometer (Cellometer X2, Nexcelom Bioscience) to compute cell concentration (C_{nex}). Using the Nexcelom counting software (version 2.1.4.2) we implemented the following custom parameters: cell diameter (2-9 microns), roundness (0.1), contrast enhancement

(0.4), decluster edge factor (0.1), decluster th factor (0.5), and background adjustment (1.0). These settings tended to 'decluster,' i.e. separate joined cells only when there is a significant gap between cells. Thus, higher concentrations tended to underestimate the actual cell count. When significant disagreement arose, we accepted the lower dilution replicates' estimate of cell count. We favored minimal declustering because we sought to maximize agreement between what the flow cytometer counts as a cell, and what the cellometer counts as a cell. With high declustering settings, the cellometer often calls not-quite-divided cells as separate, whereas our flow cytometer counts as one cell any unit that passes across the laser. See Fig. S1A for a comparison of declustering settings.

S2.1.5   Absolute count calibration analysis

The previous experiments generate the following quantities:

1. A - distribution of autofluorescence from BY4743 (flow cytometry)

2. AG - distribution of GFP expression with autofluorescence (flow cytometry)

3. C_{nex} - concentration of cells per mL (imaging cellometer)

4. V_{load} - volume of cells loaded

5. V_{lys} - volume of lysis solution

6. V_{illum} - microscope beam illumination volume (fluorescence correlation spectroscopy)

7. N_{fcs} - number of fluorescent molecules per illumination volume (fluorescence correlation spectroscopy)

The assumption underlying all fluorescence measurements is that a linear factor k_{ct} relates protein count to fluorescence.

$$k_{ct} = \frac{\langle \# \text{ proteins/cell} \rangle}{\langle \text{fluorescence/cell} \rangle} \tag{S1}$$

That is, measurements of mean protein count per cell, and the mean fluorescence per cell from the same population is sufficient to compute k_{ct}. With the conversion factor in hand, the ith cell with fluorescence f_i is expected to have $k_{ct} f_i$ proteins.

From our protocol and values listed above, k_{ct} can be computed. The denominator of Eq. S2, mean population fluorescence is measured by subtracting mean autofluorescence from total fluorescence:

$\langle F \rangle = \langle AG \rangle - \langle A \rangle$. The numerator can be further decomposed into $\langle N \rangle = N_{fcs}/(V_{illum}C_{lyse})$, where V_{illum} is the beam illumination/detection volume, and C_{lyse} is the final concentration of cells lysed. By performing FCS on rhodamine 6G (Sigma-Aldrich, R4127) of known concentration V_{illum} was determined to be $.25fL \pm .02$. Lysed cells per volume was computed by combining the initial cell concentration ascertained by imaging cellometry (C_{nex}) with loading volume V_{load} and resuspension volume V_{lys}, so that $C_{lyse} = C_{nex}V_{load}/V_{lys}$. Together we get:

$$k_{ct} = \frac{\langle N \rangle}{\langle F \rangle} = \frac{1}{\langle AG \rangle - \langle A \rangle} \frac{N_{fcs}V_{lys}}{V_{illum}V_{load}C_{nex}} \tag{S2}$$

S2.1.6   Quantitative Western blot validation of FCS

We validated our FCS measurements by quantitative Western blot analysis. Two GFP-positive technical replicates were prepared as above with the samples split immediately after cell counting. Four samples were also taken from the autofluorescent control. These samples were used to generate a standard curve in the blot. Specifically, we spiked in known quantities of purified GFP (Clontech Laboratories, 632502) into the autofluorescent cell samples immediately before lysis. Post-lysis, we spun down at high velocity as above and extracted the supernatant. The GFP-positive samples were split, with one fraction measured by FCS, and the other fraction used for the Western.

With the standard curve samples, the remaining GFP-positive samples, and matched resuspensions of the insoluble debris, we performed two Western blots. Samples were boiled with 4X SDS loading buffer (0.2M Tris/HCl, pH 7, 8% SDS, 40% glycerol, 20% β-mercaptoethanol) and then analyzed by SDS-PAGE and immunoblotting. Immunoblots were developed with SuperSignal West Pico Chemiluminescent Substrate (Thermo-Scientific, Rockford, IL) and exposed to autoradiography film. To detect GFP, we used rabbit polyclonal antibody (ab6556, Abcam, Cambridge, MA) and goat anti-rabbit HRP-conjugated secondary antibody (A0545, Sigma-Aldrich, St. Louis, MO).

Bands were quantified using pixel densitometry estimates obtained from ImageJ following recommended practices [1], with the exception that the lack of dynamic range in pixel density led us to integrate total band intensity. We found that arbitrary choices of background level cutoff, image exposure, and band bounding box size heavily influenced the shape of the curve and values of the density, but had little impact on the final interpolated value.

S2.2 Treatment of autofluorescence

Solutions for isolating mean expression from background autofluorescence include spectral compensation [2], chemical treatment [3], and/or usage of a reporter-negative control. With live cells, we ruled out chemical treatment, and two-wavelength (510nm, 600nm) emission spectral decomposition was deemed inappropriate since applying it resulted in a significant number of cells with negative expression. This latter observation is likely because the tail of the GFP emission spectrum extends past 600nm [4]. We therefore used autofluorescent controls in all experiments and subtracted out the autofluorescent component to mean expression.

Autofluorescent contribution to variance was more troublesome than accounting for mean effects. Close examination of autofluorescence reveals that the exact treatment, and associated assumptions, significantly affects estimation of intrinsic and extrinsic noise for lower expressing constructs SSA1$^\mathrm{P}$-H-GFP and SSA1$^\mathrm{P}$-Y-GFP. We find that autofluorescence likely correlates with expression with $\rho = .2 \pm .1$. We used $\rho = .2$ for all analyses (manuscript in preparation).

S2.3 Flow cytometry

Flow cytometry was performed with a Cytomics FC500 MPL (Beckman Coulter, Brea, CA). The MPL extension on our cytometer enabled 96-well plate reads. Excitation was with a uniphase argon ion 488nm 20mW laser, and emission was ultimately recorded via a 510nm bandpass filter. Flow rate was set to medium, limiting measurement to at most 3000 cells/second. Before every experiment, two calibration measurements were obtained by flowing Flow-Check Flourospheres (Beckman Coulter, 6605359) which excite with the same 488nm laser and emit at a wavelength capturable by the same bandpass filter. These checks allowed us to assess instrument drift, which could affect our FCS-based calibration. As we developed the method we found that only the yearly cytometer maintenance significantly affected calibration readings, after which we re-calibrated absolute counts by FCS.

Flow cytometry data was exported via FlowJo (Tree Star, Ashland OR) to tab delimited text files. All analysis was done via MATLAB (v2011a, The MathWorks, Inc., Natick, MA) using custom code. Cytometry data post-processing was limited to outlier removal. To accomplish this, we fit a lognormal distribution to the data. With the best fit parameters, we calculated PDF values of the lognormal distribution corresponding to the cutoff $P(F > f) = 4 \times 10^{-6}$ on either extreme. This setting comes from

the following thought experiment. In the largest single experiment we did, we measured approximately 4 million cells. This cutoff admits that at most one cell in that experiment should be beyond the cutoff. Any additional cells must be outliers. We found that most outliers likely arose from either cells sticking together (as indicated by unusually high forward scatter), or contamination from the previous well ($<.1\%$).

S2.3.1 Steady state measurements

The following protocol was used for steady state expression measurements of all strains. In late morning, colonies were picked and inoculated into $5\,mL$ Synthetic Complete media (US Biological, #D9515) in $15\,mL$ falcon tubes and placed in a rotating drum (60 RPM) at either 22C or $37°$C. 10-12 hours later, optical densities were measured for all cultures and diluted to OD600=.025 in $5\,mL$ SC media, and replaced on same rotator. The following morning, optical densities were assessed again, and diluted into $30\,mL$ SC media at OD600=.025 in $250\,mL$ baffled flasks and placed on a floor shaker (300RPM) at $22°$C or $37°$C. For $37°$C cultures, optical densities were taken 6 hours later and were usually ~OD600=.4, corresponding to a division time of 90 minutes at $37°$C. At $22°$C, we observed OD600=.4 reached after 11-12 hours, corresponding to a ~170 minute division rate. We generally saw close agreement among optical densities between replicates. Outliers with an optical density differing by more than a factor of 2 were discarded. Steady state was confirmed by taking two flow cytometry measurements spaced 20-40 minutes apart. Close agreement in expression between these technical replicates was taken as an indication of steady state.

S2.3.2 Cycloheximide block experiments

To cultures initially at $22°$C or $37°$C steady state we added $1\mu L$ of 50mg/mL cycloheximide (Sigma-Aldrich, C7698) per $1\,mL$ culture [5]. We marked this as time zero. Culture flasks were maintained on their shakers at their initial temperatures for the duration. Flow cytometry time points were taken for several hours until fluorescence leveled out.

S2.3.3 Induction experiments

We initiated all induction experiments after cultures attained $22°$C steady state. SC media was pre-warmed to $40°$C. To heat shock, we added $5\,mL$ culture to $25\,mL$ $40°$C media in new pre-warmed $250\,mL$ baffled flasks, which we estimated should immediately bring the resulting culture to ~$37°$C. Inoculation into the

warmer media was taken as time zero. Flasks were then placed on a floor shaker also pre-warmed to 37°C (300RPM), and measurements were taken immediately, and every 15-25 minutes after. To minimize exposure to room temperature, flasks were removed two at a time, $100\mu L$ of culture was taken for each sample, and then the flasks were quickly replaced. Time values for each observation were taken as the flow-cytometry reported time-zero for each well read, thus accounting for the few minutes it takes to read 6-12 wells.

S2.3.4 *In vivo* maturation of GFPmut3

Reporter maturation influences investigators ability to interpret stochastic expression. We sought to minimize the impact of maturation by selecting a rapidly folding fluroescent protein. Although GFPmut3 maturation has been measured *in vitro* [6], we wondered how well it folds *in vivo*. To measure maturation in live cells we performed the following experiment.

Massive transcriptional upregulation of *SSA1* during heat shock (22°C to 37°C) provides an excellent opportunity to study maturation. We followed gene expression of SSA1$^{\text{P}}$-GFP over time during heat shock. During the initial rapid rise in expression, and before transcriptionally-mediated repression sets in, we added cycloheximide to block translation. We subsequently tracked expression of cycloheximide blocked samples, with the hypothesis being that any continued appearance of protein will be due to maturation of our reporter. Fig. S1G demonstrates that blocked cultures completely level out approximately 30 minutes after cycloheximide addition. The following model was used to fit these traces.

$$\frac{dU}{dt} = -(k_f + d_p)U \tag{S3}$$

$$\frac{dP}{dt} = k_f U - d_p P \tag{S4}$$

Here, U represents unfolded protein and P represents visible protein. We assumed both protein forms degrade at rate d_p, and that folding is a first order process. Fitting this model yields the traces shown in Fig. S1G, and yields an *in vivo* maturation rate of $.0803 \pm .0016\text{min}^{-1}$. This corresponds to a half-maturation time of ~8.63 minutes, consistent with the *in vitro* measurement of ~7.5 minutes.

S2.4 Modeling for parameter estimation

For the purposes of parameter estimation, we fit experimental data to the first moments of a simple stochastic model

$$\frac{dP(N_M, N_P)}{dt} = k_m P(N_M - 1, N_P) \tag{S5}$$

$$+ k_p N_M P(N_M, N_P - 1) \tag{S6}$$

$$+ d_m(N_M + 1)P(N_M + 1, N_P) \tag{S7}$$

$$+ d_P(N_P + 1)P(N_M, N_P + 1) \tag{S8}$$

$$- (k_m + k_p N_M + d_M N_M + d_p N_P)P(N_M, N_P) \tag{S9}$$

where N_M and N_P represent the numbers of mRNA and protein, respectively. Here, k_m represents constant transcription, k_p represents the translation rate, and d_m and d_p capture mRNA and protein unimolecular decay rates. The first moments of this stochastic model represent average cell behavior.

$$\frac{d\langle N_M \rangle}{dt} = k_m - d_m \langle N_M \rangle \tag{S10}$$

$$\frac{d\langle N_P \rangle}{dt} = k_p \langle N_M \rangle - d_p \langle N_P \rangle \tag{S11}$$

Units of all rate constants are min^{-1}.

S2.4.1 Estimating protein degradation rates

The principle enabling protein degradation rate measurement is that blocking translation decouples RNA from protein dynamics. Setting k_p to zero and solving Eq. S11 yields a single exponential:

$$\langle N_P \rangle_t = \langle N_P \rangle_0 e^{-d_p t} \tag{S12}$$

We also considered a two-step degradation process. Let N_{P1} represent the number of destabilized reporter molecules degrading at rate d_1, N_{P2} represent the more stable form which degrades at rate d_2,

and allow their conversion at rate c_{12}. If both forms are visible, then degradation proceeds as:

$$\langle N_{P1}\rangle_t + \langle N_{P2}\rangle_t = \frac{e^{-d_2 t}\left(c_{12}\langle N_{P1}\rangle_0 e^{t(-d_1-f)+d_2 t} - d_1\langle N_{P2}\rangle_0 + d_2\langle N_{P2}\rangle_0 - c_{12}\langle N_{P2}\rangle_0 - c_{12}\langle N_{P1}\rangle_0\right)}{-d_1 + d_2 - c_{12}}$$

$$\text{(S13)}$$

$$+ \langle N_{P1}\rangle_0 e^{t(-d_1-c_{12})} \qquad\qquad\qquad\qquad\qquad\qquad \text{(S14)}$$

Using the cycloheximide block protocol above, we obtained decay curves for all three constructs at 37°C (example trajectories in Fig. S2A-F). Consistent with other reports, we observed inefficient active degradation of untagged GFP. With a single-exponential model, we saw acceptable fits with a degradation rate of $4.76 \times 10^{-4} \pm 4.73 \times 10^{-5}\text{min}^{-1}$. This corresponds to a ~1400 minute GFP half-life. However, the average growth rate estimated during this experiment yielded $7.4 \times 10^{-4} \pm 3.08 \times 10^{-4}\text{min}^{-1}$, which overlaps the estimated degradation rate. We conclude that the dominant mode of decay for GFP is dilution. At 37°C we typically observed a ~90 minute doubling time, and a ~170 minute doubling time at 22°C.

SSA1$^{\text{P}}$-H-GFP and SSA1$^{\text{P}}$-Y-GFP strains demonstrated remarkably accelerated decay, however, our reporters did not degrade as a single exponential process. This was consistent with two modes of decay. Although biphasic decay is necessary to fit the cycloheximide traces for SSA1$^{\text{P}}$-H-GFP and SSA1$^{\text{P}}$-Y-GFP, the actual second decay mode should, as above, be determined by the growth rate. We also fit both constructs' decay curves' first few data points to a simple exponential decay model under the hypothesis that initial decay will be less influenced by the global, reactive effects of cyclohexamide. These experiments yielded two estimates for the decay constants for both SSA1$^{\text{P}}$-Y-GFP and SSA1$^{\text{P}}$-H-GFP tagged GFPs. Both are reported in Data S1.

The initial d_p for SSA1$^{\text{P}}$-Y-GFP was $.01085 \pm .00085\text{min}^{-1}$ and the double exponential decay fast-mode parameter was $.022 \pm .0038\text{min}^{-1}$. For SSA1$^{\text{P}}$-H-GFP, the initial d_p was $.0394 \pm .00437\text{min}^{-1}$ and the fast-mode decay constant from the double exponential model was $.095 \pm .0053\text{min}^{-1}$.

S2.4.2 Estimating RNA degradation rates

To measure RNA decay rates we fit to the full temporal solution of Eq. S11.

$$\langle N_P \rangle_t = \frac{d_p k_p k_m - d_m k_p k_m}{d_p^2 d_m - d_p d_m^2} \tag{S15}$$

$$+ \langle N_P \rangle_0 \frac{d_p e^{d_m t} - d_m e^{d_m t}}{d_p e^{d_p t + d_m t} - d_m e^{d_p t + d_m t}} \tag{S16}$$

$$+ \langle N_M \rangle_0 \frac{k_p e^{d_p t} - k_p e^{d_m t}}{d_p e^{d_p t + d_m t} - d_m e^{d_p t + d_m t}} \tag{S17}$$

$$+ \frac{d_m k_p k_m e^{d_m t} - d_p k_p k_m e^{d_p t}}{d_p^2 d_m e^{d_p t + d_m t} - d_p d_m^2 e^{d_p t + d_m t}} \tag{S18}$$

In the case that complete shutoff is achieved, k_m is set to zero and Eq. S18's first and fourth summands denoted become zero. If shutoff is incomplete, better estimates of RNA decay would be obtained by allowing for partial shutoff (non-zero k_m).

Upon temperature shift from 22°C to 37°C, SSA1 exhibits a large pulse of expression [7]. Both the initial rise in expression and subsequent shutoff are likely mediated by increases and decreases in active heat shock transcription factor (HSF) [7–9], suggesting that the *SSA1* promoter should confer a similar response on our reporter gene. Young *et al.* observed an initial rise in RNA expression within minutes of temperature shift [9], followed by RNA decay starting ~20 minutes after induction.

Taking advantage of this, and using the protocol above we measured induction and subsequent falloff of expression in response to heat shock. Starting from the first decrease in slope during induction, we modeled decay as above. Fits obtained were excellent (Fig. S2B,D,F) and resulted in RNA degradation rates in close agreement with each other, despite very different protein degradation rates. Bulk RNA degradation rates are reported in Data S1.

S2.5 RNA and protein degradation rates across forward scatter

These analyses were performed as described in the main text with the following additional details. To make estimated degradation rates within a forward scatter interval comparable, the same forward scatter bins were used for all experiments. This necessitated ensuring a sufficient number of observations (cells) were present in all forward scatter bins for all experiments. The 8 bins shown in the main text figure (Fig. 2A) correspond to 8 bins that in all experiments contain at least 500 cells.

For each bin within each experiment, the models above were fit as above and recorded separately for each biological replicate. Fitting was with MATLAB's FIT() function. Averaging and error was

determined subsequent to each replicates' fitting to generate the error bars in Figure 3, main text. Example trajectories and fits for one forward scatter bin for each degradation experiment are shown in Fig. S2. The values obtained for each replicate are represented in Data S3. RNA data were pooled for better accuracy given their similarity in average rates.

S2.5.1 Estimation of intrinsic and extrinsic noise by 1- and 2-copy reporter strains

The one- versus two-copy fluorophore method for decomposing total expression variance into intrinsic and extrinsic components depends on the following relationship [10]. For $\text{Var}(F_1 + F_2)$ where F_1 and F_2 are identical copies of the same fluorescent reporter gene in every cell, $\text{Var}(F_1 + F_2) = 2\text{Var}(F) + 2\text{Cov}(F_1, F_2)$. Using the 1-copy variance with the two-copy variance then, intrinsic (I_v) and extrinsic variability (E_v) are, respectively:

$$I_v = 4\text{Var}(F) - \text{Var}(F_1 + F_2)/2 = \text{Var}(F) - \text{Cov}(F_1, F_2) \tag{S19}$$

$$E_v = \text{Var}(F) - I_v = \text{Cov}(F_1, F_2) \tag{S20}$$

Stewart-Ornstein *et al.* observed this approach was significantly more error-prone than the two-color method for decomposing noise [11]. Our initial findings agreed: if we computed I_v and E_v from biological replicates, then held out one observation, estimates of both noises changed significantly. This was particularly troublesome with the SSA1$^\text{P}$-GFP constructs.

We reasoned that variability in the estimate could arise from two different sources of error. On one hand, error could be technical shot noise from the instrument or from slightly different biological conditions experienced by the replicates. In that situation, the only solution would be to make more measurements and average. Alternatively, we reasoned that since mean expression is related to expression variance both genome-wide and within promoter context [12–14], perhaps differences between biological replicates corresponds to a conserved perturbation in mean and variance simultaneously.

Plotting mean versus variance for the SSA1$^\text{P}$-H-GFP constructs revealed essentially shot noise; but here, the holdout strategy has the smallest effect on estimates of intrinsic and extrinsic noise. Alternatively, the same plot drawn from SSA1$^\text{P}$-GFP construct replicates demonstrates a remarkably linear relationship between mean and variance for one-copy and two-copy constructs. Therefore, high error introduced by

the hold-out strategy for this experiment is essentially explained by local excursions away from the true steady state mean, which has an equally concerted, and much larger impact on variance.

This local mean-variance relationship suggests a way to reduce error in the intrinsic-extrinsic decomposition. First, we could simply draw a line connecting the two-copy genes (Eq. S21), and a similar line for one-copy mean vs. variance (Eq. S22). Assaying the resulting line fits at $S_M = X$ and $D_M = 2X$ should alleviate the effect of slightly different means causing large differences in estimated variances.

$$D_v = aD_M + b \tag{S21}$$

$$S_v = cS_M + d \tag{S22}$$

However, more statistical power can be obtained by combining the regression into a single dataset. This is achieved by remapping the double-copy observation of variance, D_v by $D_v^* = D_v/k$, and mean double-copy measurements D_m by $D_m^* = D_m/2$. This remapping captures the essential features we already know relate single copy to double copy mean and variance observations. Namely, bulk mean expression of the double copy variants is on average exactly two times greater than single copy expression, after autofluorescence correction. For variance, an unknown coefficient k relates variance of the double copy observations to single copy observations.

In fact, k represents the balance of intrinsic and extrinsic noise; double copy variance can be at most four times single copy variance, and at least two times single copy variance.

We hypothesized that there exists a value k such that the remapping of double copy variances causes M_v^* and D_v^* to trace the same line. That is, $D_v^* = cD_m^* + d$. Functionally this corresponds to the assumption that the local mean (and variance) excursions can affect the *magnitude* of both types of variance but not their relative contribution to total variance. Code that minimized a linear regression r^2 by iteratively testing values of $2 \leq k \leq 4$ generated excellent estimates of intrinsic and extrinsic decompositions that were robust to 2-fold hold out–see SEM for intrinsic (Iv) and extrinsic (Ev) variance in Data S3 for untagged GFP.

S2.6 Exploratory modeling

S2.6.1 Variance normalization reduces the information in stochastic measurements

Absolute protein expression level adds concrete information to our understanding of stochastic gene expression. If we consider P_i to be a random variable representing protein stochasticity, mean expression scales with constant k_{ct}, since $\mathrm{E}[k_{ct}F_i] = k_{ct}\mathrm{E}[F_i]$, while variance scales with k_{ct}^2, since $\mathrm{Var}(k_{ct}F_i) = k_{ct}^2\mathrm{Var}(F_i)$. To ameliorate this discrepancy, investigators commonly normalize variance using the coefficient of variation, $CV = \sqrt{k_{ct}^2 Var(F_i)}/(k_{ct}E[F_i])$ or CV^2, as this eliminates the pesky protein count conversion factor. A side effect of this practice is to reduce the predictive power of stochastic models, since CV measurements are not independent of mean expression effects.

To illustrate this fact, one need only to study how relative fluorescence impedes the use of an elegant, mechanistic solution to a simple stochastic model of protein distributions. The Friedman *et al.* analytical solution to protein stochasticity enables investigators to link $CV^2 = d_p/k_m$ and Fano noise ($F = \sigma^2/\mu = k_p/d_m$) to mechanistically relevant parameter ratios [15]. However, use of relative fluorescence in these calculations yields:

$$\frac{k_{ct}^2 \mathrm{Var}(P_i)}{(k_{ct}\mathrm{E}[P_i])^2} = \frac{d_p}{k_m} \tag{S23}$$

$$\frac{k_{ct}^2 \mathrm{Var}(P_i)}{k_{ct}\mathrm{E}[P_i]} = \frac{k_{ct}k_p}{d_m} \tag{S24}$$

The result is that only the d_p/k_m ratio can be abstracted without a measure of absolute protein count; the k_p/d_m parameter is inaccessible. This motivated us to measure protein stochasticity in units of protein count.

S2.6.2 Models incorporating extrinsic noise

Intrinsic and extrinsic contributions to total variance are often written as a formula [13, 16]:

$$\sigma_{tot}^2 = \sigma_{intr}^2 + \sigma_{extr}^2 \tag{S25}$$

However, this representation obscures the relationship between the two quantities. The full version of the Law of Total Variance is:

$$\text{Var}(X) = \underbrace{\text{E}[\text{Var}(X|e)]}_{\text{Intrinsic variance}} + \underbrace{\text{Var}(\text{E}[X|e])}_{\text{Extrinsic variance}} \tag{S26}$$

where e corresponds to a set of environmental variables that jointly vary with a random variable X.

We can use this relationship to analytically separate intrinsic and extrinsic noise for simple models. We start with the simplest model of gene expression with no ON and OFF states. Mean and variance for this model are:

$$\text{E}[N_P] = \frac{k_m k_p}{d_m d_p} \tag{S27}$$

$$\text{Var}(N_P) = \frac{k_m k_p (d_m + d_p + k_p)}{d_m d_p (d_m + d_p)} \tag{S28}$$

Since the parameters are constants, mean and variance are also constant for a given set of parameters. This represents biochemical intrinsic noise only.

If we instead postulate that k_m is itself a random variable with distribution K_m, then Eqs. S27 and S28 become distributions themselves. Using the definitions of noise decomposition from Eq. S26 we arrive at the equations shown in the main text, and reproduced here:

$$\text{E}[N_P] = \frac{k_p}{d_p d_m} \text{E}[K_m] \tag{S29}$$

$$I_v(N_P) = \frac{k_p (d_p + d_m + k_p)}{d_p d_m (d_p + d_m)} \text{E}[K_m] \tag{S30}$$

$$E_v(N_P) = \frac{k_p^2}{d_p^2 d_m^2} \text{Var}(K_m) \tag{S31}$$

where N_p is the random variable representing the number of proteins in the basic model.

If we instead let k_p be a random variable K_p, Eqs. S27 and S28, and again invoke the law of total variance, we obtain:

$$\mathrm{E}[N_p] = \mathrm{E}[\mathrm{E}[N_p|K_p]] = \frac{k_m}{d_m d_p}\mathrm{E}[K_p] \tag{S32}$$

$$\mathrm{Var}_{intr} = \mathrm{E}[\mathrm{Var}[N_p|K_p]] = \frac{k_m}{d_m d_p(d_m + d_p)}(\mathrm{Var}[K_p] + \mathrm{E}[K_p]^2 + (d_m + d_p)\mathrm{E}[K_p]) \tag{S33}$$

$$\mathrm{Var}_{extr} = \mathrm{Var}[\mathrm{E}[N_p|K_p]] = \frac{k_m^2}{d_m^2 d_p^2}\mathrm{Var}[K_p] \tag{S34}$$

S2.6.3 K_m as a Generalized Extreme Value distribution

Although a number of distributions come close to capturing stochastic expression shape, the only distribution that captures tail behavior well was the Generalized Extreme Value Distribution. Fits generally gave scale parameter values of very close to zero. Setting the scale value to zero results in a Gumbel-type extreme value distribution with parameters μ_{GEV} and σ_{GEV}. With the measurements obtained, we can solve for both parameters along with k_p:

$$\mu_{GEV} = -\frac{d_m d_p \mathrm{E}[N_p]\left((d_m + d_p)\mathrm{E}[N_p](-I_v + \mathrm{E}[N_P])\phi + \gamma\sqrt{E_v}\sqrt{(d_m + d_p)^2(I_v - \mathrm{E}[N_P])^2\phi}\right)}{(d_m + d_p)^2(I_v - \mathrm{E}[N_P])^2\phi} \tag{S35}$$

$$\sigma_{GEV} = \frac{d_m d_p\sqrt{E_v}\mathrm{E}[N_P]}{\sqrt{(d_m + d_p)^2(I_v - \mathrm{E}[N_P])^2\phi}} \tag{S36}$$

$$k_p = \frac{(d_m + d_p)(I_v - \mathrm{E}[N_P])}{\mathrm{E}[N_P]} \tag{S37}$$

where $I_v(N_P)$ and $E_v(N_P)$ were shortened to I_v and E_v respectively, $\phi = \frac{\pi}{6}$, and γ is the Euler-Mascheroni constant (~.5772). Although the Gumbel distribution is supported on $(-\infty, \infty)$, parameter estimations for μ_{GEV} and σ_{GEV} showed MATLAB-computed density values of zero at all x-values of $K_m = 0$ or less.

S2.6.4 K_m as an inverse Gaussian distribution

Among common two-parameter distributions, the inverse Gaussian distributions generally appeared in either the second or third positions among rankings of distributions by Akaike Information Criterion. Moments of the inverse Gaussian were much easier to work with. Unlike the Gumbel distribution above, the inverse Gaussian is correctly supported on $[0, \infty)$. An intrinsic-extrinsic model with K_m being inverse Gaussian distributed $K_M = IG(\mu_{IG}, \sigma_{IG})$ yields the following relationships:

$$\mu_{IG} = \frac{d_p d_m \mathrm{E}[N_P]^2}{(d_p + d_m)(I_v - \mathrm{E}[N_P])} \tag{S38}$$

$$\sigma_{IG} = \frac{d_p d_m \mathrm{E}[N_P]^4}{(d_p + d_m)E_v(I_v - \mathrm{E}[N_P])} \tag{S39}$$

$$k_p = \frac{(d_p + d_m)(I_v - \mathrm{E}[N_P])}{\mathrm{E}[N_P]} \tag{S40}$$

We found this model reliably underestimates distribution skewness which is likely why it consistently underperforms the Gumbel distribution (above) in terms of AIC.

S2.7 Simulations to validate hybrid model against experimental data

S2.7.1 Extrinsic K_m model

The following steps enabled determination of model-predicted higher moments. For each parameter set (forward scatter bin), we randomly drew 100,000 values from a Gumbel distribution with parameters μ_k and σ_k. Each number can be thought of as k_m for a single cell. We then employed analytical moment solutions for the model above for each hierarchical parameter set $(k_{m,i} \in K_m, k_p, d_m, d_p)$ to obtain 100,000 "single-cell" moment sets for the first four central moments [17]. These distributions represent $\mathrm{E}[N_p|e]$, $\mathrm{Var}(N_p|e)$, $\mathrm{Skew}(N_p|e)$ and $\mathrm{Kurt}(N_p|e)$, where e corresponds to the specific condition, the set of rates $k_{m,i} \in K_m$, k_p, d_m, and d_p.

These moment distributions, conditioned on K_m, were further analyzed using the law of total variance to verify variance estimates matched with the input. Higher order moments were analyzed as follows. Like variance (Eq. S26), skewness (third central moment) and kurtosis (fourth central moment) can also be decomposed into intrinsic and extrinsic components:

$$\mathrm{Skew}(N_p) = \mathrm{E}[\mathrm{Skew}(N_p|e)] + \mathrm{Skew}(\mathrm{E}[N_p|e]) + 3\mathrm{Cov}(\mathrm{E}[N_p|e], \mathrm{Var}(N_p|e)) \tag{S41}$$

$$\mathrm{Kurt}(N_p) = \mathrm{E}[\mathrm{Kurt}(N_p|e)] + \mathrm{Kurt}(\mathrm{E}[N_p|e]) + 4\mathrm{Cov}(\mathrm{E}[N_p|e], \mathrm{Skew}(N_p|e))$$

$$+ 6\mathrm{Cov}(\mathrm{E}[N_p|e], \mathrm{E}[N_p|e]\mathrm{Var}(N_p|e)) - 6\mathrm{E}[N_p]\mathrm{Cov}(\mathrm{E}[N_p|e], \mathrm{Var}(N_p|e)) \tag{S42}$$

These terms represent moments of the simulated moment distributions $\mathrm{E}[N_p|e]$, $\mathrm{Var}(N_p|e)$, $\mathrm{Skew}(N_p|e)$

and Kurt$(N_p|e)$, and were thus calculated directly to generate the moments across forward scatter plots (Fig. S3B and C). Note again, skewness and kurtosis refer to 3rd and 4th central moments.

S2.7.2 Extrinsic K_p model

The same intrinsic and extrinsic moments of skew and kurtosis also hold for the extrinsic K_p model. We first solved the experimental data for the Gumbel hyperparameters for K_p assuming k_m was fixed using Eqs. S32-S34. Then using the same analytical moment solutions as above we simulated the hierarchical parameter set $(k_{p,i} \in K_p, k_m, d_m, d_p)$ to obtain 100,000 "single-cell" moment sets for the first four central moments. Both the extrinsic K_p and extrinsic K_m model yielded similar moment predictions despite having different forms of the intrinsic variance equation (Fig. S3B and C).

S2.8 Analysis of RNA count distributions from Gandhi *et al.*

Following the same steps in the main text, the extrinsic K_m model for RNA posits a distribution of transcription rates K_m with a constant degradation rate. The various parameters can be solved using intrinsic and extrinsic moments and a measurement of d_m.

$$\mathrm{E}[N_m] = \frac{E[K_m]}{d_m} \tag{S43}$$

$$\mathrm{Var}_{intr} = \frac{1}{d_m^2}\mathrm{Var}[K_m] \tag{S44}$$

$$\mathrm{Var}_{extr} = \frac{1}{d_m}\mathrm{Var}[K_m] \tag{S45}$$

Therefore, the mean and variance of the extrinsically distributed transcription rate K_m is:

$$\mathrm{E}[K_m] = \mathrm{E}[N_m]d_m \tag{S46}$$

$$\mathrm{Var}[K_m] = d_m^2(\mathrm{Var}[N_m] - E[N_m]) \tag{S47}$$

Finally, substituting the GEV parameters σ_{km} and μ_{km} and solving, we get:

$$\sigma_{km} = \sqrt{\frac{6\text{Var}[K_m]}{\pi^2}} \tag{S48}$$

$$\mu_{km} = \text{E}[K_m] - \sigma_{km}\gamma \tag{S49}$$

where γ is the Euler-Mascheroni constant. For each RNA distribution depicted in Fig. 7 (main text) from Gandhi *et al.* [18], the mean and variance was calculated ($\text{Var}[N_m]$ and $\text{E}[N_m]$). RNA degradation rates were obtained from Wang *et al.* [19]. Resulting fits were produced by drawing 10,000 k_m values from $GEV(\sigma_{km}, \mu_{km})$ and dividing by the corresponding d_m.

References

[1] Gassmann, M., Grenacher, B., Rohde, B. & Vogel, J. Quantifying Western blots: Pitfalls of densitometry. *Electrophoresis* **30**, 1845–1855 (2009).

[2] Lichten, C. A., White, R., Clark, I. B. N. & Swain, P. S. Unmixing of fluorescence spectra to resolve quantitative time-series measurements of gene expression in plate readers. *BMC biotechnology* **14**, 11 (2014). URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3917901\&tool= pmcentrez\&rendertype=abstract.

[3] Mosiman, V. L., Patterson, B. K., Canterero, L. & Goolsby, C. L. Reducing cellular autofluorescence in flow cytometry: An in situ method. *Communications in Clinical Cytometry* **30**, 151–156 (1997).

[4] Cormack, B. P., Valdivia, R. H. & Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33–8 (1996). URL http://www.ncbi.nlm.nih.gov/pubmed/8707053.

[5] Mateus, C. & Avery, S. V. Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. *Yeast (Chichester, England)* **16**, 1313–23 (2000). URL http://www.ncbi.nlm.nih.gov/pubmed/11015728.

[6] Iizuka, R., Yamagishi-Shirasaki, M. & Funatsu, T. Kinetic study of de novo chromophore maturation of fluorescent proteins. *Analytical biochemistry* **414**, 173–8 (2011). URL http://www.ncbi.nlm.nih. gov/pubmed/21459075.

[7] Slater, M. R., Craig, E. a. & Slatert, M. R. Transcriptional regulation of an hsp70 heat shock gene in the yeast Saccharomyces cerevisiae. *Molecular and cellular biology* **7**, 1906–16 (1987). URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=365295\&tool=pmcentrez\&rendertype=abstract`.

[8] Park, H. O. & Craig, E. a. Positive and negative regulation of basal expression of a yeast HSP70 gene. *Molecular and cellular biology* **9**, 2025–33 (1989). URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=362995\&tool=pmcentrez\&rendertype=abstract`.

[9] Young, M. R. M. & Craig, E. A. E. Saccharomyces cerevisiae HSP70 heat shock elements are functionally distinct. *Molecular and cellular biology* **13** (1993). URL `http://mcb.asm.org/content/13/9/5637.short`.

[10] Chabot, J. R., Pedraza, J. M., Luitel, P. & van Oudenaarden, A. Stochastic gene expression out-of-steady-state in the cyanobacterial circadian clock. *Nature* **450**, 1249–52 (2007). URL `http://www.ncbi.nlm.nih.gov/pubmed/18097413`.

[11] Stewart-Ornstein, J., Weissman, J. S. & El-Samad, H. Cellular noise regulons underlie fluctuations in Saccharomyces cerevisiae. *Molecular cell* **45**, 483–93 (2012). URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3327736\&tool=pmcentrez\&rendertype=abstract`.

[12] Bar-Even, A. *et al.* Noise in protein expression scales with natural protein abundance. *Nature genetics* **38**, 636–43 (2006). URL `http://dx.doi.org/10.1038/ng1807`.

[13] Newman, J. R. S. *et al.* Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**, 840–6 (2006). URL `http://www.ncbi.nlm.nih.gov/pubmed/16699522http://dx.doi.org/10.1038/nature04785`.

[14] Taniguchi, Y. *et al.* Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)* **329**, 533–8 (2010). URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2922915\&tool=pmcentrez\&rendertype=abstract`.

[15] Friedman, N., Cai, L. & Xie, X. Linking Stochastic Dynamics to Population Distribution: An Analytical Framework of Gene Expression. *Physical Review Letters* **97**, 1–4 (2006). URL `http://link.aps.org/doi/10.1103/PhysRevLett.97.168302`.

[16] Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science (New York, N.Y.)* **297**, 1183–6 (2002). URL `http://www.sciencemag.org/cgi/content/abstract/297/5584/1183http://www.ncbi.nlm.nih.gov/pubmed/12183631`.

[17] Sherman, M. S. & Cohen, B. a. A computational framework for analyzing stochasticity in gene expression. *PLoS computational biology* **10**, e1003596 (2014). URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4014403\&tool=pmcentrez\&rendertype=abstract`.

[18] Gandhi, S. J., Zenklusen, D., Lionnet, T. & Singer, R. H. Transcription of functionally related constitutive genes is not coordinated. *Nature structural & molecular biology* **18**, 27–34 (2011). URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3058351\&tool=pmcentrez\&rendertype=abstracthttp://dx.doi.org/10.1038/nsmb.1934`.

[19] Wang, Y. *et al.* Precision and functional specificity in mRNA decay. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 5860–5865 (2002).