*nature*

# Regulated cell-to-cell variation in a cell-fate decision system

Alejandro Colman-Lerner[1]*, Andrew Gordon[1]*, Eduard Serra[1], Tina Chin[1], Orna Resnekov[1], Drew Endy[2], C. Gustavo Pesce[1] & Roger Brent[1]

**Here we studied the quantitative behaviour and cell-to-cell variability of a prototypical eukaryotic cell-fate decision system, the mating pheromone response pathway in yeast. We dissected and measured sources of variation in system output, analysing thousands of individual, genetically identical cells. Only a small proportion of total cell-to-cell variation is caused by random fluctuations in gene transcription and translation during the response ('expression noise'). Instead, variation is dominated by differences in the capacity of individual cells to transmit signals through the pathway ('pathway capacity') and to express proteins from genes ('expression capacity'). Cells with high expression capacity express proteins at a higher rate and increase in volume more rapidly. Our results identify two mechanisms that regulate cell-to-cell variation in pathway capacity. First, the MAP kinase Fus3 suppresses variation at high pheromone levels, while the MAP kinase Kss1 enhances variation at low pheromone levels. Second, pathway capacity and expression capacity are negatively correlated, suggesting a compensatory mechanism that allows cells to respond more precisely to pheromone in the presence of a large variation in expression capacity.**

Biological systems are composed of physical constituents that constrain their performance. However, some aspects of system performance, including cell-to-cell variation, are often regulated by active mechanisms[1–8]. The study of variation in the behaviour of genetically identical cells goes back as far as Delbrück[9], who measured differences in the numbers of phage T1 produced by individual, singly infected *E. coli*.

Recently, a number of studies have used fluorescent protein reporters to study cell-to-cell variation in gene expression[10–16]. For example, variation in gene expression among genetically identical bacteria has been studied by measuring the correlation in expression of two different fluorescent protein reporter genes under control of the same promoters[11]. Cell-to-cell variation resulted from both stochastic fluctuations in the expression of each reporter protein (termed 'intrinsic noise') and differences in the levels of cellular components needed for expression of both reporters (termed 'extrinsic noise'), and the results suggested that some components of extrinsic noise affect gene expression in general[11]. The component of extrinsic noise that affects overall gene expression has recently been quantified in *Escherichia coli*[12,13]. This type of gene expression analysis has also been performed in yeast[14], and revealed that intrinsic noise contributes little to cell-to-cell variation in gene expression. Cell-to-cell variation in the expression of two non-identical promoters was correlated, consistent again with the idea that some extrinsic noise is a result of global differences in gene expression. Others have shown that changes in the amount of transcription and translation affect the amount of overall cell-to-cell variation in the expression of a single reporter[15,16].

Here we studied cell-to-cell variation, not in gene expression, but in the quantitative output of a cell-fate decision system: the pheromone response pathway in the yeast *Saccharomyces cerevisiae*. In haploid cells of the **a** mating type, α-factor (a pheromone secreted by cells of the α mating type) triggers a fate decision to switch from normal, vegetative growth to the initiation of mating events, including induction of gene transcription, cell cycle arrest and changes in morphology. The pathway is a prototypical eukaryotic signal transduction system that includes a G-protein-coupled receptor and a MAP kinase cascade[17] (Fig. 1a).

To study cell-to-cell variation in the workings of this decision system, we used pheromone-induced expression of fluorescent protein reporter genes as a readout. We realized that cell-to-cell differences in the levels of fluorescent proteins would convolve differences in the operation of the signal transduction pathway with cell-to-cell differences in gene expression from the reporters. To distinguish between and quantify these two contributions, we generated a series of yeast strains containing genes for yellow and cyan fluorescent protein (YFP and CFP). We compared the results from experiments in which YFP and CFP were controlled by identical α-factor-responsive promoters with results from experiments in which YFP was driven by an α-factor-responsive promoter and CFP by an α-factor-independent promoter (Fig. 1b, c).

We constructed an analytical framework to guide the design and interpretation of these experiments. We considered the α-factor response pathway and the means used to measure its activity (reporter gene expression) as a single system composed of two connected subsystems: 'pathway' and 'expression' (Fig. 1a). In each subsystem, we distinguished two sources of variation: stochastic fluctuations and cell-to-cell differences in 'capacity'. Capacity depends on the number, localization and activity of proteins that transmit the signal (pathway capacity) or express genes into proteins (expression capacity), and is determined by the state of the cells at the start of the experiment. We limited the term 'noise' to refer to the variation due to stochastic fluctuations in subsystem function that occur during the experiment (for example, spontaneous differences in the occurrence and timing of discrete probabilistic chemical reactions). By distinguishing these two sources, we modify the
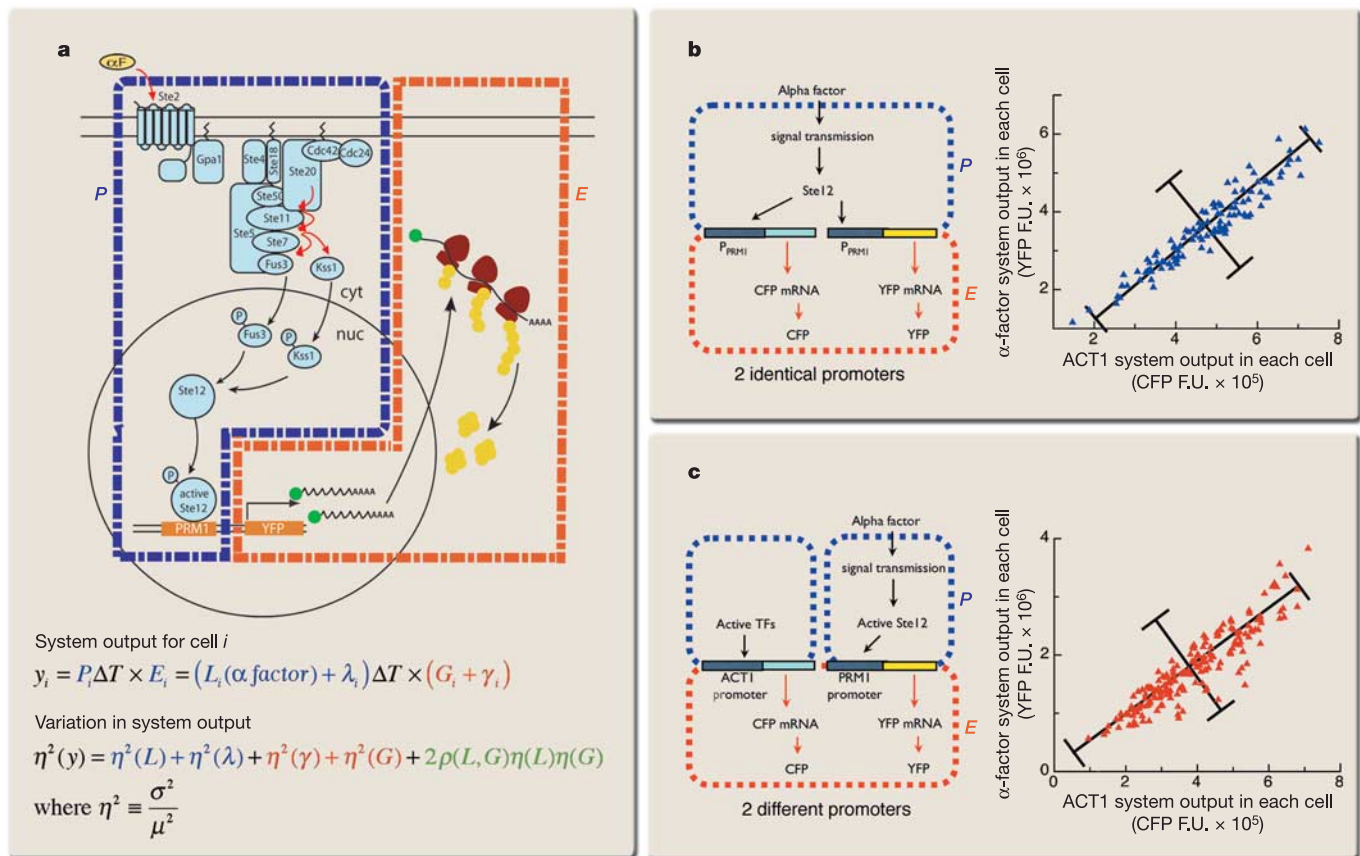
terminology used in ref. 11, which used the term noise to refer to both sources of variation.

If we could stimulate a cell numerous times by going back in time and repeating the experiment (thereby guaranteeing the same initial state), the average system output for all the repeated trials would be a measure of the expectation value of the output. According to our framework, this value would depend on pathway capacity and expression capacity. Any differences in system output on individual trials would arise from random fluctuations during each trial in the number of molecules and the workings of the machineries (1) transmitting the signal (transmission noise) and (2) transcribing the reporter messenger RNA and translating it into protein (expression noise). If we performed the above thought experiment

on a different cell, a different average value for system output might be obtained. This cell-to-cell difference in average system output would be the manifestation of a cell-to-cell difference in the capacities of the two subsystems—caused, for example, by a pre-existing difference in the number of molecules that transmit the signal or express proteins from genes.

## Analytical framework

The first subsystem, the α-factor response pathway, includes all steps that lead to activation of transcription, which depends on the activity of DNA-bound transcription factors. The input to the pheromone pathway is α-factor and the output depends on the total amount of active transcription factor (Ste12) bound upstream of the α-factor-



System output for cell $i$

$$y_i = P_i \Delta T \times E_i = \left(L_i(\alpha\,factor) + \lambda_i\right)\Delta T \times \left(G_i + \gamma_i\right)$$

Variation in system output

$$\eta^2(y) = \eta^2(L) + \eta^2(\lambda) + \eta^2(\gamma) + \eta^2(G) + 2\rho(L,G)\eta(L)\eta(G)$$

where $\eta^2 \equiv \dfrac{\sigma^2}{\mu^2}$

**Figure 1 | Quantifying sources of cell-to-cell variation. a**, The mating pheromone response system and analytical framework for decomposition. Diagram shows proteins in the yeast cell membrane, cytoplasm (cyt), and nucleus (nuc). Events in the blue box are classified as the pathway subsystem. The binding of α-factor to the receptor Ste2 causes dissociation of the heterotrimeric G protein α-subunit Gpa1 from the Ste4–Ste18 dimer (βγ-subunits). Ste4 recruits the scaffold protein Ste5 to the membrane, and Ste5 binds the MEKK Ste11, the MEK Ste7 and the MAPK Fus3. The PAK kinase Ste20 initiates the MAPK cascade by activating Ste11, which activates Ste7, which in turn activates the MAP kinases Fus3 and Kss1. Phosphorylated Fus3 and Kss1 leave Ste5 and translocate to the nucleus, where they activate the transcription factor Ste12. At a given concentration of α-factor, the amount of activated Ste12 on the promoter is the 'pathway subsystem output' $P$ (defined in the text). Events in the red box are classified as the expression subsystem, quantified by $E$ (defined in the text). $E$ includes transcription initiation, mRNA elongation and processing, nuclear export and cytoplasmic protein translation. The total system output—the amount of fluorescent reporter protein $y$ produced in any cell $i$—depends on $P$, $E$, α-factor concentrations and the duration of stimulation $\Delta T$. To measure cell-to-cell variation in the population we used the normalized variance $\eta^2$, which is decomposed into separate additive terms that represent different sources of cell-to-cell variation as described in the text. **b**, Type I experiment, measuring gene expression noise ($\gamma$). In strains containing two identical

α-factor-responsive promoters driving the YFP and CFP reporter genes, the same pathway (blue box) and expression machinery (red box) controls the production of reporter proteins. We stimulated TCY3096 cells with a high concentration (20 nM) of α-factor and collected YFP and CFP images after 3 h. Each cell is represented by a single symbol showing its YFP and CFP signals (in F.U. or fluorescent units). The uncorrelated variation between YFP and CFP can be seen as the width of the minor axis, which is orthogonal to the 45° diagonal major axis (lines in black); it is caused only by stochastic variation in gene expression ($\gamma$). We used the orthogonal scatter as a measure of $\eta^2(\gamma)$, here 0.002 ± 0.0001. See Table 1 and Supplementary Fig. S4 for the gene expression noise shown by other promoters and other α-factor concentrations. **c**, Type II experiment, measuring variation in pathway subsystem output ($P$) and expression capacity ($E$). In strains containing different promoters driving the YFP and CFP reporter genes, different subsystems (blue boxes) regulate the activity of the DNA-bound transcription factors, but the subsystem enabling expression of the reporter genes (red box) is the same. We stimulated TCY3154 cells as in the type I experiment above. Variation in expression capacity affected only the correlated variation (the dispersion of points along the major axis, or the 45° diagonal). The uncorrelated variation (the dispersion of points along the minor axis) is due to the gene expression noise measured from type I experiments and to cell-to-cell variations in the pathway subsystems for each promoter.

responsive reporter gene, for the time elapsed since addition of input, $\Delta T$. For any individual cell $i$, the pathway output is $P_i \Delta T$. The term $P_i$ is the average output per unit time, which is the sum of $L_i$, the expectation value of this average (which we call pathway power) and $\lambda_i$, the stochastic fluctuations in $P_i$. Pathway power $L$ is a function of the input and pathway capacity. Note that the same analysis can apply to any pathway that leads to activation of transcription factors bound upstream of genes (including constitutive or 'housekeeping' genes) that one might not normally think of as responsive to signals.

The second subsystem, reporter gene expression, includes all events from transcription initiation through to the accumulation of protein. For $\alpha$-factor reporter genes, the input to this subsystem is the output from the $\alpha$-factor pathway subsystem, and the output of the expression subsystem is the amount of mature fluorescent reporter protein. For any individual cell $i$, expression output per unit of promoter activity ($E_i$) is the sum of $G_i$, the expectation value of $E_i$, and $\gamma_i$, the stochastic fluctuations of $E_i$. $G_i$ measures the ability

of cell $i$ to express the reporter protein from the gene; it is independent of the input level (see below) and is determined primarily by expression capacity, a global cellular property that equally affects expression of other genes. From here on, we will refer to $G$ as expression capacity. In contrast, the value of $\gamma$ refers only to the stochastic fluctuations in the levels of reporter protein.

We assumed that induction of reporters (and the other genes induced by $\alpha$-factor[18]) did not significantly decrease cell-wide gene expression capacity. Two facts supported this assumption. First, at a given stimulus (concentration and treatment duration of $\alpha$-factor), cells with two or three copies of the $\alpha$-factor-inducible reporter produced a corresponding two- or threefold increase in fluorescent protein levels (data not shown). Second, after $\alpha$-factor treatment, expression of fluorescent proteins controlled by constitutive promoters was unchanged (Supplementary Fig. S1 and data not shown). These findings suggest that the overall working of the expression subsystem is independent of the pheromone pathway subsystem.

We then described system output, the amount of reporter protein $y$ in cell $i$, as the product of $P_i$, the average pathway subsystem output per unit time, $\Delta T$, the time since addition of $\alpha$-factor, and $E_i$, the expression output per unit of pathway subsystem output:

$$y_i = P_i \Delta T \times E_i \qquad (1)$$

where $P_i = L_i + \lambda_i$ and $E_i = G_i + \gamma_i$. For $\alpha$-factor response, $P$ might vary with $E$, but $E$ would not vary with $P$. For example, a higher $E$ might increase (or decrease) $P$ if a higher $E$ leads to an increased (or decreased) ratio of positive regulators to negative regulators of the pathway. This potential dependency generates the correlation term in equation (2) below.

We defined variation ($\eta^2$) in system output among cells as the 'normalized variance', the variance ($\sigma^2$) divided by the mean squared ($\mu^2$) (that is, $\eta^2 = \sigma^2/\mu^2$). As derived in Supplementary Materials, the variation in $y$ for a population of cells is described by the sum of the individual sources of variation, plus a correlation term:

$$\eta^2(y) = \overbrace{\eta^2(L) + \eta^2(\lambda)}^{\eta^2(P)} + \overbrace{\eta^2(\gamma) + \eta^2(G)}^{\eta^2(E)} + 2\rho(L,G)\eta(L)\eta(G) \qquad (2)$$

where $\eta^2(L)$ is the variation in pathway power, $\eta^2(\lambda)$ is transmission noise, $\eta^2(G)$ is variation in the expression capacity, $\rho(L,G)$ is the correlation coefficient between $L$ and $G$, and $\eta^2(\gamma)$ is expression noise. This last term is equivalent to 'intrinsic noise', and total variation $\eta^2(y)$ minus $\eta^2(\gamma)$ is equivalent to 'extrinsic noise', as used in ref. 11. The term $\eta^2(P)$ is the variation in average pathway subsystem output per unit time, which, because pathway subsystem output is given by $P\Delta T$ and $\Delta T$ is the same for every cell, is equivalent to variation in pathway subsystem output. The term $2\rho(L,G)\eta(L)\eta(G)$ accounts for a possible correlation between expression capacity and pathway power, as discussed above. This term increases or decreases the total variation depending on the sign of the correlation coefficient. We assumed here that the pathway and expression subsystems do not share molecular components, and therefore that stochastic fluctuations in the expression of the reporter protein ($\gamma$) are not correlated with the pathway subsystem ($L$ or $\lambda$); similarly, $\lambda$ is not correlated to the expression subsystem ($G$ or $\gamma$). Thus, $\lambda$ and $\gamma$ do not appear in any correlation terms.

## From analytical framework to experimental design

To measure the contributions of the different sources of cell-to-cell variation in system output, we constructed haploid yeast strains that contained different combinations of two promoters driving the expression of YFP and CFP. We used strains with YFP and CFP reporters driven by the $\alpha$-factor-responsive $P_{PRM1}$ promoter[19] or by the $\alpha$-factor-independent $P_{ACT1}$ promoter (Fig. 1 and Supplementary Table S1). We treated cells attached to the glass bottom of microtitre wells with $\alpha$-factor and captured images at intervals using an inverted epifluorescence microscope and a CCD camera. Custom
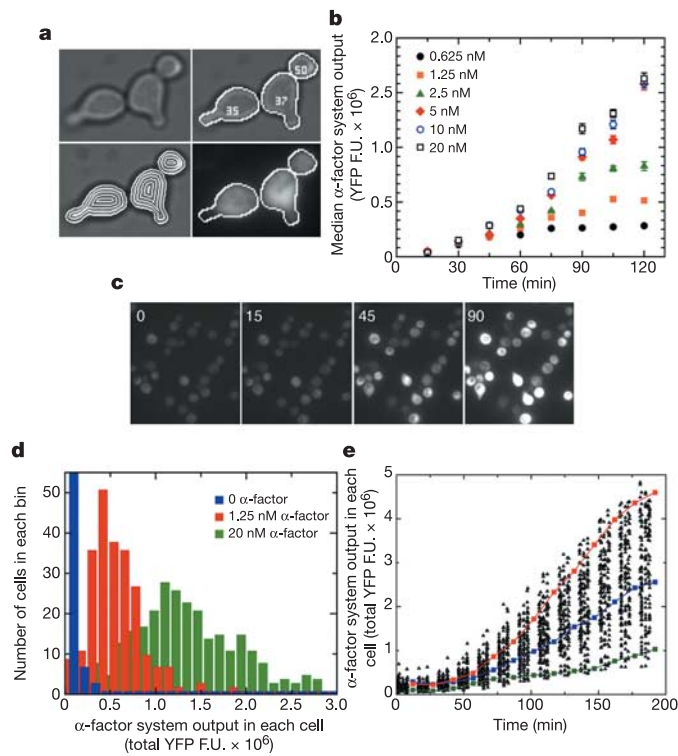
**Figure 2 | Quantification of system output in single cells during $\alpha$-factor response.** We treated ACLY387 ($P_{PRM1}$-YFP) cells with the indicated concentrations of $\alpha$-factor at time zero, acquired YFP images every 15 min for 3 h and analysed them using Cell-ID. **a**, A sampling of cells exposed to 100 nM $\alpha$-factor for 75 min. For a bright-field image (top left), we used Cell-ID to locate, determine the perimeter of, and number (top right) each cell, then draw consecutively tighter annuli (bottom left, white lines show every other annulus) to calculate cell volume. The boundary contour of the cells was then transferred to the corresponding fluorescence image (bottom right) and the fluorescence intensity of the enclosed pixels was summed. **b**, Time-dependent dose-response. Data correspond to the median system output $\pm$ s.e.m.; $n = 400$–$600$ cells. **c**, YFP fluorescence images of cells at the indicated times (in min), showing cell-to-cell variation in system output. **d**, Distribution of system output in populations of yeast exposed to $\alpha$-factor for 2 h. **e**, System output of individual cells treated with 20 nM $\alpha$-factor. We tracked three fields of cells treated identically. Images of each field were captured every 15 min, generating three 'columns' of cells at each interval. Each triangle represents the output of a single cell at a single time point. The trajectories of a representative strong (red), medium (blue) and weak (green) responding cell are shown, connected by lines and shifted rightward by 8 min to aid visualization. System output is measured in fluorescent units (F.U.), with one F.U. corresponding to approximately 2.5 photons hitting the CCD chip.

image-analysis software (Cell-ID 1.0) was used to extract measurements of individual cells from the images and to correct the measured fluorescence intensity to account for photobleaching and for the fact that cells of different size have a different fraction of their volume in focus (Fig. 2a and A.G., A.C.-L., T.C., K. Benjamin & R.B., submitted manuscript). We measured YFP and CFP fluorescence in large numbers of genetically identical cells treated with uniform amounts of α-factor, and computed the cell-to-cell variation in fluorescence. Despite differences in size, these cells each had one YFP and one CFP gene in their genome, and so to measure reporter gene activity we calculated total fluorescence per cell rather than fluorescence per unit volume. (In Fig. 3c and Supplementary Fig. S2, we examined the relationship between total reporter protein per cell and cell volume.)

We performed two types of experiments. In type I experiments, we used cells with identical promoters driving the expression of YFP and CFP (Fig. 1b). Although both constructs were controlled by the same pathway and shared the same expression subsystem, they were two separate genes. Therefore, differences in the levels of CFP and YFP in each cell should be due to expression noise $\eta^2(\gamma)$ (ref. 11) (and also to fluctuations in the amount of active transcription factor at each promoter; see Supplementary Materials section 5.3) and not to transmission noise $\eta^2(\lambda)$. In type II experiments, we used cells with different promoters driving the expression of CFP and YFP (Fig. 1c). These experiments differed from Type I experiments in that the two promoters were now controlled by different and independent pathways. Therefore, differences in the levels of CFP and YFP in each cell should be due not only to expression noise but also to transmission noise and differences in promoter-specific pathway subsystem power. In both types of experiments, the CFP and YFP genes shared the same expression subsystem. Therefore, cell-to-cell variation in expression capacity $G$ only caused cell-to-cell differences in CFP and YFP levels, not differences within a given cell (Fig. 1c).

Mathematical analysis of measurements from the two types of experiments allowed us to quantify variation in expression capacity $\eta^2(G)$, expression noise $\eta^2(\gamma)$, variation in pathway subsystem output $\eta^2(P)$ (a measure that combined $\eta^2(L)$ and $\eta^2(\lambda)$, which we have not separated experimentally), and the covariance term $2\rho(L,G)\eta(G)\eta(L)$ (see Supplementary Materials), and it also suggested approaches for examining mechanisms that might regulate the different sources of variation.

### Large variation in expression capacity

In cells exposed to high concentrations of α-factor (20 nM), we detected induced fluorescence within the first 30 min (Fig. 2b). Much of this delay was caused by the slow maturation of the YFP and CFP fluorophores ($T_{1/2} = 39$ and 49 min for YFP and CFP, respectively) (A.G., A.C.-L., T.C., K. Benjamin & R.B., submitted manuscript). YFP and CFP mature at slightly different rates, but this did not affect the results presented below (see Supplementary Materials, section 3). The measurements were sensitive to low pathway activation—we readily detected output in single cells stimulated with 0.1 nM α-factor, a concentration 30 times lower than that needed for half-maximal output (3 nM, which is a good match to the published $K_d$ of α-factor for its receptor[20]). Cells differed greatly in system output; the top 5% of cells showed approximately fourfold higher output than the bottom 5% of cells. The distribution of system output in the population was roughly bell-shaped at all concentrations tested (Fig. 2d and not shown), indicating that system output as measured by our reporter shows a graded (as opposed to an all-or-nothing) response to α-factor (see also Supplementary Fig. S3).

Variation in α-factor system output ($\eta^2$) was relatively constant over time (Fig. 3a). This suggests that most of the variation was due to cell-to-cell differences already present at the time of addition of α-factor. If the observed variation were caused by accumulation of signal transmission noise or gene expression noise, the standard deviation should grow as the square root of the mean (as expected for Poisson processes) rather than linearly, as our data demonstrated (Fig. 3b). Consistent with this interpretation, only a small proportion ($1.2 \pm 0.1\%$) of the observed variation was caused by gene expression noise. This was shown by the narrow dispersion of the data points along the minor axis in Fig. 1b, which shows YFP and CFP levels derived from genes controlled by identical $P_{PRM1}$ promoters. We also observed low gene expression noise for several other, α-factor-independent reporters (Supplementary Table S3 and Supplementary Fig. S4). Notably, variation in α-factor pathway subsystem output ($\eta^2(L) + \eta^2(\lambda)$) was also small (17% of the total) in cells stimulated with high concentrations of α-factor (20 nM). This was derived from the dispersion of the points along the minor axis of Fig. 1c, where YFP and CFP levels from the $P_{PRM1}$ and $P_{ACT1}$ promoters are shown, respectively. The bulk of the total variation ($\eta^2 = 0.17 \pm 0.02$, 3 h after addition of 20 nM α-factor) was caused by cell-to-cell differences in expression capacity ($\eta^2(G) = 0.14 \pm 0.02$), as shown by the wider dispersion of data points along the major compared to the minor axis in Fig. 1c ($\rho_{YFP,CFP} = 0.88 \pm 0.05$). Details of these calculations are provided in the Supplementary Information.

We obtained consistent results using two other constitutive promoters ($P_{STE5}$ and $P_{BMH2}$) to drive CFP expression (data not
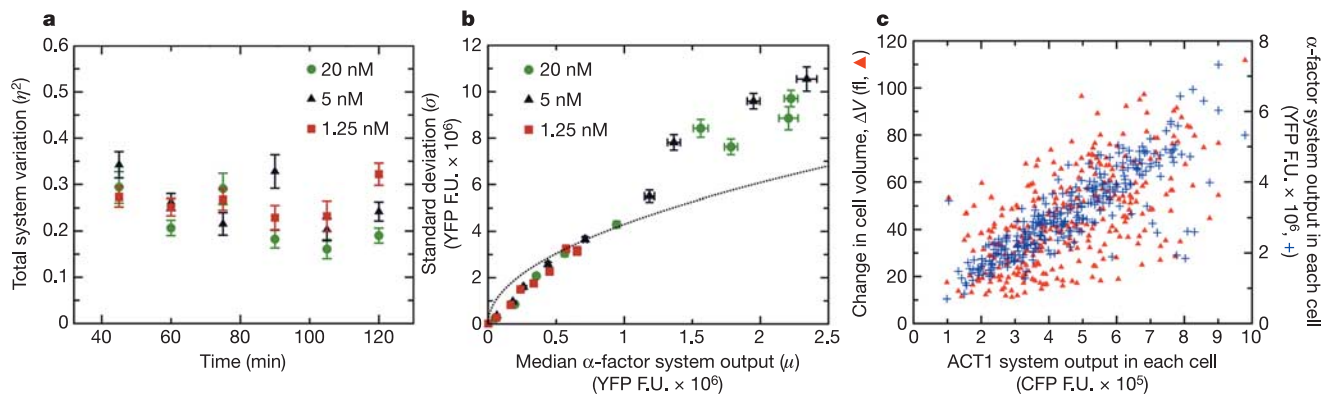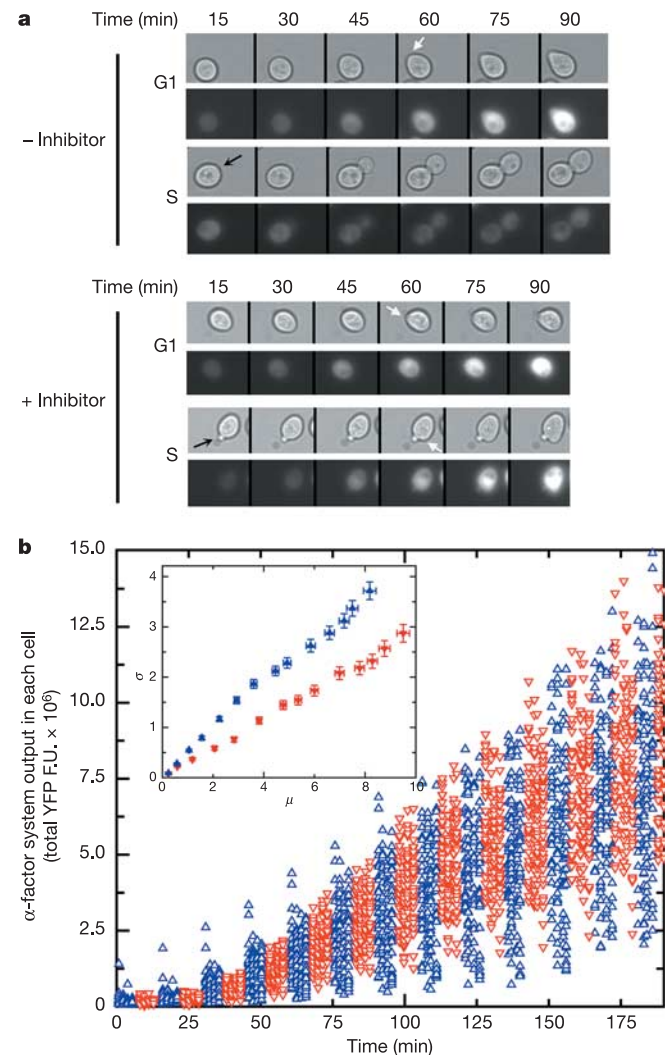


**Figure 3 | Cell-to-cell variation is dominated by initial differences between cells.** We treated ACLY387 cells ($P_{PRM1}$-YFP) with α-factor as in Fig. 2 and measured cell-to-cell variation. **a**, Total variation $\eta^2(y)$ over time after addition of α-factor. **b**, Standard deviation $\sigma$ was plotted against median system output $\mu \pm$ s.e.m. for populations of cells treated with the indicated concentrations of α-factor for different times. The dotted line corresponds to a square root function of $\mu$ (forced to pass through the point $x = 1 \times 10^6$, $y = 4 \times 10^6$), the relationship expected if most of the variation in the population originated by stochastic processes over the course of the experiment. **c**, We treated TCY3154 cells ($P_{PRM1}$-YFP, $P_{ACT1}$-CFP) with 20 nM α-factor for 3 h. Data correspond to single-cell values of ACT1 system output (total CFP at 3 h minus total CFP at time zero) versus the change in volume, $\Delta V$ (left $y$ axis, red, $\rho_{CFP,\Delta V} = 0.77$), and versus α-factor system output (right $y$ axis, blue, $\rho_{CFP,YFP} = 0.89$).

shown). Moreover, we observed a strong correlation between YFP and CFP in strains with reporters controlled by two different α-factor-independent promoters (Supplementary Fig. S5 and data not shown). The above results suggest that in yeast, cell-to-cell variation in gene expression is dominated by differences in expression capacity among cells, and that expression capacity is a global cellular feature controlling expression of many or most genes.

As cells with high expression capacity produce protein more rapidly, we reasoned that cells with high expression capacity might increase faster in volume (larger $\Delta V$). To test this, we examined the correlation between cell volume, the rate of change in cell volume and the expression of fluorescent reporters. There was a low correlation between $P_{ACT1}$-derived CFP and initial volume (Supplementary Fig. S2). However, we observed a significant correlation between

$P_{ACT1}$-derived CFP and $\Delta V$ ($\rho = 0.77$, Fig. 3c), suggesting that expression capacity might help to determine the rate of increase in cell volume. However, the fact that the correlation of $P_{ACT1}$-CFP with $P_{PRM1}$-YFP ($\rho = 0.89$) was significantly better than the correlation of $P_{ACT1}$-CFP with $\Delta V$ indicates that other factors, uncorrelated with expression capacity, also influenced $\Delta V$.

### Effect of cell cycle on variation

One conspicuous difference among exponentially growing yeast is their cell-cycle position. To determine the effects of cell-cycle position on cell-to-cell variation in system output, we measured variation in system output in yeast arrested from cycling. We replaced the cyclin-dependent kinase Cdc28 with an engineered variant (Cdc28-as2)[21] that is sensitive to the chemical inhibitor 1-NM-PP1 (4-amino-1-(tert-butyl)-3-(1′-naphtylmethyl)pyrazolo[3,4-D]pyrimidine). In the absence of inhibitor, Cdc28-as2 cells behave the same as wild-type cells (not shown), but addition of $10\,\mu M$ inhibitor arrests Cdc28-as2 cells at the G2/M transition[21]. We stimulated cells containing Cdc28-as2 and the $P_{PRM1}$-YFP and $P_{ACT1}$-CFP reporters (TCY3154 cells) with 20 nM α-factor with or without inhibitor, and followed them over time. Visual examination revealed that without inhibitor, cells with small buds or those about to initiate bud formation showed a delay in pathway induction (Fig. 4a), consistent with reports that cells at the G1/S transition cannot respond to pheromone because a cyclin-dependent kinase complex inhibits the MAP kinase cascade[22–24]. With inhibitor, cells began to induce the pathway at almost the same time (Fig. 4a), reducing total α-factor system output variation by 45% (from $\eta^2(y) = 0.19 \pm 0.03$ to $\eta^2(y) = 0.11 \pm 0.01$, Fig. 4b). Addition of inhibitor did not alter cell-to-cell variation in $P_{ACT1}$-CFP signal (not shown), indicating that inhibition of Cdc28 reduced variation specific to the α-factor pathway subsystem and not variation in the gene expression subsystem. We obtained similar results by synchronizing the cell cycle in late mitosis using the Cdc15-2ts mutant (not shown), suggesting that all the measured cell-cycle-dependent pathway variation was due to Cdc28 activity.

In addition to eliminating a source of variation, inhibiting Cdc28 allowed us to study system output at low concentrations of α-factor. Only concentrations of α-factor above 2.5 nM caused uniform cell-cycle arrest. At lower concentrations, cells that continued to divide diluted the reporter protein into daughter cells, complicating the measurement of cell-to-cell variation in system output. However, in the presence of Cdc28-as2 inhibitor, all cells arrest.

### Large pathway power variation at low α-factor concentrations

The above experiments determined that at high concentrations of α-factor (>20 nM), less than 25% of cell-to-cell variation in system output was due to differences in pathway subsystem output, and more than 75% was due to differences in expression capacity (Table 1). To determine whether the relative contributions of the sources of variation in system output depended on the concentration of α-factor, we stimulated TCY3154 cells with lower α-factor concentrations in the presence of Cdc28 inhibitor. At low concentrations, we observed a reduced correlation between $P_{PRM1}$-YFP and $P_{ACT1}$-CFP ($\rho_{YFP,CFP} = 0.94 \pm 0.01$ at 20 nM and $\rho_{YFP,CFP} = 0.72 \pm 0.02$ at 1.25 nM, Fig. 5a). A control experiment demonstrated that α-factor concentration did not affect correlation between two α-factor-independent promoters (Supplementary Fig. S5). At low concentrations of α-factor, a substantial amount of the total variation in system output was caused by differences in pathway subsystem output $P$ (Table 1 and Supplementary Fig. 6a). At 1.25 nM α-factor, $\eta^2(P)$ accounted for 59% of the total output, but at 20 nM, $\eta^2(P)$ accounted for only 22% of the total (Table 1). This suggests that high levels of input might conceal pre-existing cell-to-cell differences in pathway capacity.

If expression capacity and pathway subsystem output were independent, then the larger variation in pathway subsystem output



**Figure 4 | The cell-cycle kinase Cdc28 causes a large part of the variation in α-factor system output.** We treated TCY3154 cells ($P_{prm1}$-YFP, $P_{ACT1}$-CFP, *cdc28-as2*) with 20 nM α-factor in the presence or absence of 10 μM 1-NM-PP1 Cdc28-as2 inhibitor at time zero, collected YFP and CFP images, and quantified the fluorescent signals over time as in Fig. 2. **a**, Bright-field and YFP images of cells treated with 20 nM α-factor in the presence or absence of inhibitor, showing that Cdc28 inhibition allows cells in the S phase of the cell cycle to produce YFP in response to α-factor (black arrows mark the budding site). White arrows mark the first time point with a visible mating projection (shmoo tip) **b**, Alpha-factor system output of individual cells treated with 20 nM α-factor in the presence (red) or absence (blue) of 10 μM inhibitor. For visual clarity, the red symbols have been shifted 8 min to the right. Inset shows the standard deviation ($\sigma$) of the population versus the median ($\mu$) system output ± standard error on both axes (YFP F.U. × 10$^6$).

**Table 1** | α-factor concentration, Fus3 and Kss1 regulate cell-to-cell variation in pathway subsystem output

| Strain | Promoters | α-factor (nM) | Total variation $\eta^2(y)$ ($\times 10^{-3}$) | Gene expression noise $\eta^2(\gamma_{YFP})$ ($\times 10^{-3}$) | Variation in pathway subsystem output $\eta^2(P_{YFP})$ ($\times 10^{-3}$) | Variation in G (+covariance) ($\times 10^{-3}$) | $\rho_{CFP,YFP}$* |
|---|---|---|---|---|---|---|---|
| TCY3154 (WT) | $P_{PRM1}$-YFP versus $P_{ACT1}$-CFP | 1.25 | 132 ± 12 | 5.62 ± 0.014 (4.26 ± 0.10) | 78.1 ± 8.0 (59.2 ± 6.1) | 48 ± 14 (36 ± 11) | 0.725 ± 0.021 |
| GPY3262 (Δfus3) | $P_{PRM1}$-YFP versus $P_{ACT1}$-CFP | 1.25 | 125 ± 16 | 4.21 ± 0.14 (3.39 ± 0.11) | 83.0 ± 11.0 (66.4 ± 9.2) | 38 ± 19 (30 ± 15) | 0.565 ± 0.049 |
| GPY3263 (Δkss1) | $P_{PRM1}$-YFP versus $P_{ACT1}$-CFP | 1.25 | 152 ± 17 | 2.76 ± 0.14 (1.82 ± 0.09) | 54.6 ± 7.7 (35.8 ± 5.0) | 95 ± 18 (62 ± 12) | 0.774 ± 0.029 |
| TCY3154 (WT) | $P_{PRM1}$-YFP versus $P_{ACT1}$-CFP | 20 | 115 ± 12 | 2.01 ± 0.14 (1.74 ± 0.12) | 24.8 ± 5.3 (21.6 ± 4.7) | 88 ± 13 (77 ± 11) | 0.867 ± 0.021 |
| GPY3262 (Δfus3) | $P_{PRM1}$-YFP versus $P_{ACT1}$-CFP | 20 | 128 ± 12 | 2.83 ± 0.14 (2.20 ± 0.11) | 60.3 ± 7.1 (47.0 ± 5.5) | 65 ± 14 (51 ± 11) | 0.712 ± 0.031 |
| GPY3263 (Δkss1) | $P_{PRM1}$-YFP versus $P_{ACT1}$-CFP | 20 | 142 ± 24 | 2.40 ± 0.14 (1.693 ± 0.097) | 36.9 ± 9.1 (26.0 ± 6.4) | 103 ± 25 (72 ± 18) | 0.827 ± 0.032 |

WT, wild type.
*Correlation coefficient between YFP and CFP.
Distribution of total cell-to-cell variation $\eta^2(y)$ among different sources. We treated strains TCY3154, GPY3262 and GPY3263 with the indicated α-factor concentration and 10 μM 1-NM-PP1, and collected YFP and CFP images at 15-min intervals. The amount of variation due to the different sources of variation was calculated as explained in the main text and Supplementary Materials. Data correspond to the measurement 3 h after α-factor addition. The percentage of total variation $\eta^2(y)$ is given in parentheses. Error measurements are s.e.m. In addition to the shown errors, we associate a 10–15% systematic uncertainty with the reported numbers for the variation in pathway subsystem output and gene expression noise, due to the omission of higher-order terms in equation (2) (see Supplementary Materials).
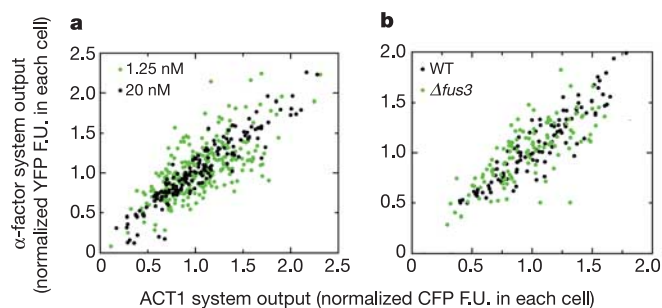
$\eta^2(P)$ observed at low concentrations compared to high concentrations would sum with constant variation in expression capacity $\eta^2(G)$ and expression noise $\eta^2(\gamma)$ to result in a larger total variation $\eta^2(y)$ at low concentrations. Instead, $\eta^2(y)$ remained relatively constant with concentration (Table 1 and Supplementary Fig. S6). The fact that total variation does not increase significantly at low concentrations compared to high concentrations indicates that there is a negative correlation between G and L in equation (2), implying that cells with low expression capacity have a higher than expected pathway subsystem output and vice versa, and that cell-to-cell variation in pathways is regulated.

**Fus3 and Kss1 regulate pathway power variation**
Alpha-factor signal can be transmitted independently by two MAP kinases, Fus3 and Kss1 (ref. 17) (Fig. 1a). Fus3 and Kss1 both can phosphorylate the transcription factor Ste12, but they also phosphorylate distinct substrates[25–27]. In addition, unphosphorylated Kss1 can bind and inhibit Ste12 (ref. 28). We reasoned that cell-to-cell differences in the relative levels of activated Fus3 and Kss1 might lead to different levels of active Ste12 and therefore to cell-to-cell



**Figure 5** | **Genetic control of cell-to-cell variation in pathway subsystem output $\eta^2(P)$. a**, Alpha-factor regulates variation in pathway subsystem output. We treated TCY3154 cells with 20 nM (black) or 1.25 nM (green) α-factor and 10 μM 1NM-PP1. **b**, Fus3 reduces variation in pathway subsystem output. We treated TCY3154 (wild type, black) or GPY3262 (Δfus3, green) cells with 20 nM α-factor and 10 μM 1-NM-PP1. Data correspond to the output of the α-factor system versus the ACT1 system 3 h after stimulation. YFP and CFP F.U. were normalized to the median of each population to allow the overlaying of data with different means. The increased variation in pathway subsystem output at low concentrations of α-factor in wild-type cells (**a**) and at high concentrations in Δfus3 cells (**b**) is manifested as a wider spread along the minor axis.

differences in pathway subsystem output. To test this idea, we derived strains from TCY3154 cells that lacked either Fus3 or Kss1, and stimulated the cells with high (20 nM) or low (1.25 nM) concentrations of α-factor. At high α-factor concentrations, the average responses of Δfus3 cells and Δkss1 cells were nearly the same as that of wild-type cells (within 20%). At low concentrations, the average response of Δfus3 cells was nearly the same as that of wild-type cells (within 5%). However, Δkss1 cells showed a stronger response than wild-type cells (~2-fold higher, data not shown), consistent with previous reports[29].

Relative to wild-type, Δfus3 cells had higher pathway variation at high α-factor concentrations and the same pathway variation at low concentrations. In contrast, relative to wild-type cells, Δkss1 cells showed the same pathway variation at high concentrations and lower pathway variation at low concentrations (Fig. 5b and Table 1). We performed western blots with an antibody that recognized the active forms of Fus3 and Kss1 (which are phosphorylated on two sites) to assess the relative activities of these protein kinases at high and low concentrations of α-factor. At high concentrations, we observed a larger phospho-Fus3 to phospho-Kss1 ratio than at low concentrations (data not shown), indicating that at high concentrations signal transmission depends more strongly on Fus3. Taken together, these results suggest that signal transmission is less variable when more dependent on Fus3 (at high α-factor concentrations or when Kss1 is absent) than when more dependent on Kss1 (at low concentrations or when Fus3 is absent).

**Discussion**
The approach described above is generally applicable to dissecting the sources of cell-to-cell variation for cellular processes for which the output can be measured with transcriptional reporters. Here we quantified the contributions of four sources of variation: cell-to-cell differences in pathway power or ability to transmit a signal (L), cell-to-cell differences in expression capacity or the ability to express proteins from genes (G), and noise in the operations of the pathway (λ) and the gene expression (γ) subsystems. We constructed an analytical framework to distinguish pathway activation from reporter gene expression, and used it to guide the design and interpretation of experiments. The framework relied on a subdivision of the process in general terms, rather than representation of the molecular mechanisms that underlie it. The use of such formalism has allowed us to model a non-steady-state process and to devise ways to measure pre-existing differences in the ability of cells to induce the pathway and express genes. In contrast to, for example, the rigorous treatment of mRNA and protein dynamics as coupled processes[30], we collapsed all

steps from reporter gene transcription to fluorescence output into a single composite process. This formalism can be modified as new experimental techniques allow the measurement of different processes, and as deeper molecular understanding leads us to further subdivide the causes of cell-to-cell variation.

We have shown that about half of the cell-to-cell variation is due to pre-existing differences in the cell-cycle position of individual cells at the time of pathway induction. This would be expected on the basis of previous work showing that, when yeast cells are about to commit to a new round of cell division, a complex between Cdc28 and Cln2 inhibits activation of the MAP kinase cascade in response to mating pheromone[22]. Our results extend this earlier work by suggesting that there are no relevant slow processes (such as transcription) mediating the Cdc28/Cln2-dependent inhibition of pathway activation and that the substrate(s) of Cdc28/Cln2 is relatively short-lived, as we observed the same reduction in variation whether we pre-incubated the cells with the Cdc28 inhibitor or added it simultaneously with α-factor (data not shown).

Another large component of the variation in system output is due to cell-to-cell differences in the capacity of cells to express proteins from genes ($G$), whereas little is due to noise in gene expression ($\eta^2(\gamma)$). Genetically identical cells had different $G$ values, perhaps due to differences in the numbers of ribosomes or RNA polymerase II complexes, or the cellular energy level. Our findings, combined with recent studies in *E. coli*[12,13], suggest that variations in $G$ are found in both eukaryotes and prokaryotes.

All processes that depend on levels of gene expression should be sensitive to cell-to-cell differences in expression capacity, as we showed for the process of cell volume changes. As *S. cerevisiae* cells initiate daughter formation when they reach a critical volume[31], cells with high expression capacity should also reproduce more rapidly and account for a greater proportion of newly formed daughter cells in the population. However, we find the distribution of expression capacity in an exponentially growing population is stable over time (data not shown), suggesting that expression capacity is not strongly heritable, as reported for *E.coli*[12]. Although expression capacity might change with cell age, our experiments show that age alone cannot account for high variation in expression capacity. In our exponentially growing populations most of the cells are young; approximately 60% are newborn daughters and half of the remainder (20% of the total) are cells that have given birth only once.

Notably, we found that variation in output of the pathway subsystem changed with α-factor input: at high concentrations, output variation was low, whereas at low concentrations, variation was high. We expected the noise component of this variation $\eta^2(\lambda)$ to behave like gene expression noise $\eta^2(\gamma)$, decreasing with increasing mean system output (Supplementary Fig. S4). However, the output variation of the pathway subsystem did not decrease with increasing mean system output over time, suggesting that pathway subsystem output variation is dominated by cell-to-cell differences in pathway power ($L$) rather than noise $\eta^2(\lambda)$.

Our results indicate that the amount of variation in the pathway subsystem output is regulated by the MAP kinases Fus3 and Kss1. The Fus3-dependent reduction in variation might be due to Fus3 autoregulatory negative feedback. For example, activated Fus3 induces the protein phosphatase Msg5 that dephosphorylates and inactivates Fus3 (ref. 32) but not Kss1 (ref. 33). Such feedback mechanisms would tend to equalize the levels of active Fus3 between cells. The Kss1-dependent increase in cell-to-cell variation might result from inputs to Kss1 from filamentation[34] and cell wall integrity[35] pathways. Thus, increasing the relative activity of Kss1 (compared to Fus3) may make the α-factor pathway more sensitive to variation in these other inputs. It may be advantageous for cells to regulate variation in cell-fate decisions. At high levels of α-factor, the decision to respond is clear; cells should respond as best as they can in order to mate; and pathway subsystem output depends predominantly on Fus3. At low levels, however, the decision to

respond might have to be dependent on factors in addition to the level of α-factor. Consequently, the pathway might rely more heavily on Kss1, which can integrate the α-factor pathway with other cellular information-processing pathways.

We imagine that cell-to-cell variation in general cellular capacities (such as expression capacity) creates circumstances that can distort the transmission of signals and provides selective pressure for the evolution of specialized, compensatory mechanisms that enable cells to generate less biased, more uniform responses. Such a compensatory regulatory mechanism might be responsible for the negative correlation we observed between expression capacity and pathway power. This correlation decreases the effect of differences in expression capacity on the α-factor system output. We predict that some of the 'feedback' and inhibitory genes that modulate pathway subsystem output[17] might function in this compensatory mechanism.

We undertook this work as a step towards predicting the quantitative output of a cell-fate decision system in response to defined perturbations. Although many of the molecular components that comprise this system are known, the mechanisms that control its quantitative behaviour are not. Our experiments have defined two such mechanisms, and have begun to identify genes that affect their function. We hope that the combination of physiological experimentation enabled by new measurement tools and existing molecular and genetic methods will allow us to gain greater insight into the mechanisms that regulate quantitative variation. Understanding mechanisms that regulate global capability to express proteins from genes might have applicability to protein expression and the engineering of biological systems. Understanding the mechanisms that constrain variation in cell-fate decision systems might also enable new therapeutic interventions, for example to narrow the distribution of cellular responses to a pro-apoptotic anti-cancer drug.

## METHODS

Nucleic acid and yeast manipulations were performed as previously described[36,37]. Derivation of yeast strains (Supplementary Table S1) from YAS245-5C (ref. 38) and the protein methods used are detailed in the Supplementary Materials. Quantification of system output from single cells and measurements of cell-to-cell variation were performed using time-lapse fluorescent microscopy followed by data analysis using Cell-ID (A.G., A.C.-L., T.C., K. Benjamin & R.B., submitted manuscript).

1.  Sternberg, P. W. & Horvitz, H. R. Pattern formation during vulval development in *C. elegans*. *Cell* **44**, 761–772 (1986).
2.  Priess, J. R. & Thomson, J. N. Cellular interactions in early *C. elegans* embryos. *Cell* **48**, 241–250 (1987).
3.  Kimble, J. & Hirsh, D. The postembryonic cell lineages of the hermaphrodite and male gonads in *Caenorhabditis elegans*. *Dev. Biol.* **70**, 396–417 (1979).
4.  Kimble, J. Alterations in cell lineage following laser ablation of cells in the somatic gonad of *Caenorhabditis elegans*. *Dev. Biol.* **87**, 286–300 (1981).
5.  Karp, X. & Greenwald, I. Post-transcriptional regulation of the E/Daughterless ortholog HLH-2, negative feedback, and birth order bias during the AC/VU decision in *C. elegans*. *Genes Dev.* **17**, 3100–3111 (2003).
6.  Doe, C. Q. & Goodman, C. S. Early events in insect neurogenesis. II. The role of cell interactions and cell lineage in the determination of neuronal precursor cells. *Dev. Biol.* **111**, 206–219 (1985).
7.  Jan, Y. N. & Jan, L. Y. Maggot's hair and bug's eye: role of cell interactions and intrinsic factors in cell fate specification. *Neuron* **14**, 1–5 (1995).
8.  Hoang, T. The origin of hematopoietic cell type diversity. *Oncogene* **23**, 7188–7198 (2004).
9.  Delbrück, M. The burst size distribution in the growth of bacterial viruses (bacteriophages). *J. Bacteriol.* **50**, 131–135 (1945).
10. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
11. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
12. Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S. & Elowitz, M. B. Gene regulation at the single-cell level. *Science* **307**, 1962–1965 (2005).
13. Pedraza, J. M. & van Oudenaarden, A. Noise propagation in gene networks. *Science* **307**, 1965–1969 (2005).

14. Raser, J. M. & O'Shea, E. K. Control of stochasticity in eukaryotic gene expression. *Science* **304**, 1811–1814 (2004).

15. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature Genet.* **31**, 69–73 (2002).

16. Blake, W. J., Kærn, M., Cantor, C. R. & Collins, J. J. Noise in eukaryotic gene expression. *Nature* **422**, 633–637 (2003).

17. Dohlman, H. G. & Thorner, J. W. Regulation of G protein-initiated signal transduction in yeast: paradigms and principles. *Annu. Rev. Biochem.* **70**, 703–754 (2001).

18. Roberts, C. J. *et al.* Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. *Science* **287**, 873–880 (2000).

19. Heiman, M. G. & Walter, P. Prm1p, a pheromone-regulated multispanning membrane protein, facilitates plasma membrane fusion during yeast mating. *J. Cell Biol.* **151**, 719–730 (2000).

20. Jenness, D. D., Burkholder, A. C. & Hartwell, L. H. Binding of alpha-factor pheromone to *Saccharomyces cerevisiae* a cells: dissociation constant and number of binding sites. *Mol. Cell. Biol.* **6**, 318–320 (1986).

21. Bishop, A. C. *et al.* A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* **407**, 395–401 (2000).

22. Oehlen, L. J. & Cross, F. R. G1 cyclins CLN1 and CLN2 repress the mating factor response pathway at Start in the yeast cell cycle. *Genes Dev.* **8**, 1058–1070 (1994).

23. Oehlen, L. J. & Cross, F. R. Potential regulation of Ste20 function by the Cln1-Cdc28 and Cln2-Cdc28 cyclin-dependent protein kinases. *J. Biol. Chem.* **273**, 25089–25097 (1998).

24. Wassmann, K. & Ammerer, G. Overexpression of the G1-cyclin gene *CLN2* represses the mating pathway in *Saccharomyces cerevisiae* at the level of the MEKK Ste11. *J. Biol. Chem.* **272**, 13180–13188 (1997).

25. Breitkreutz, A., Boucher, L. & Tyers, M. MAPK specificity in the yeast pheromone response independent of transcriptional activation. *Curr. Biol.* **11**, 1266–1271 (2001).

26. Chou, S., Huang, L. & Liu, H. Fus3-regulated Tec1 degradation through SCFCdc4 determines MAPK signalling specificity during mating in yeast. *Cell* **119**, 981–990 (2004).

27. Bao, M. Z., Schwartz, M. A., Cantin, G. T., Yates, J. R. & Madhani, H. D. III Pheromone-dependent destruction of the Tec1 transcription factor is required for MAP kinase signalling specificity in yeast. *Cell* **119**, 991–1000 (2004).

28. Bardwell, L., Cook, J. G., Zhu-Shimoni, J. X., Voora, D. & Thorner, J. Differential regulation of transcription: repression by unactivated mitogen-activated protein kinase Kss1 requires the Dig1 and Dig2 proteins. *Proc. Natl Acad. Sci. USA* **95**, 15400–15405 (1998).

29. Sabbagh, W. Jr, Flatauer, L. J., Bardwell, A. J. & Bardwell, L. Specificity of MAP kinase signalling in yeast differentiation involves transient versus sustained MAPK activation. *Mol. Cell* **8**, 683–691 (2001).

30. Paulsson, J. Summing up the noise in gene networks. *Nature* **427**, 415–418 (2004).

31. Johnston, G. C., Pringle, J. R. & Hartwell, L. H. Coordination of growth with cell division in the yeast *Saccharomyces cerevisiae*. *Exp. Cell Res.* **105**, 79–98 (1977).

32. Doi, K. *et al.* MSG5, a novel protein phosphatase promotes adaptation to pheromone response in *S. cerevisiae*. *EMBO J.* **13**, 61–70 (1994).

33. Andersson, J., Simpson, D. M., Qi, M., Wang, Y. & Elion, E. A. Differential input by Ste5 scaffold and Msg5 phosphatase route a MAPK cascade to multiple outcomes. *EMBO J.* **23**, 2564–2576 (2004).

34. Madhani, H. D. & Fink, G. R. The riddle of MAP kinase signalling specificity. *Trends Genet.* **14**, 151–155 (1998).

35. Lee, B. N. & Elion, E. A. The MAPKKK Ste11 regulates vegetative growth through a kinase cascade of shared signalling components. *Proc. Natl Acad. Sci. USA* **96**, 12679–12684 (1999).

36. Ausubel, F. M. *et al.* (eds) *Current Protocols in Molecular Biology* (Wiley, New York, 2004).

37. Guthrie, C. & Fink, G. R. (eds) *Methods in Enzymology. Guide to Yeast Genetics and Molecular Biology* (Academic, San Diego, 1991).

38. Colman-Lerner, A., Chin, T. E. & Brent, R. Yeast Cbk1 and Mob2 activate daughter-specific genetic programs to induce asymmetric cell fates. *Cell* **107**, 739–750 (2001).

**Author Contributions** A.C.-L. and A.G. conceived the framework, developed the experimental methods, performed most of the experiments and analysed the results. R.B. provided input regarding problem choice, experimentation and interpretation. A.C.-L., A.G. and R.B. wrote the paper and stand as guarantors of its findings. T.C. made most of the plasmid and yeast strains. E.S. made some plasmids and yeast strains, and some of the measurements in Fig. 2. C.G.P. made the observation that Fus3 regulates pathway variation and collaborated in interpreting its biological implications. O.R. and A.C.-L. made the observation that activated Fus3/Kss1 ratios are α-factor-dependent. D.E. suggested analysis of the stochastic fluctuations in the system and helped with describing the framework.

**Author Information** Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.C.-L. (colman-lerner@molsci.org), A.G. (gordon@molsci.org) and R.B. (brent@molsci.org).