Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/4/4/e1701775/DC1

Supplementary Materials for

Heritable stress response dynamics revealed by single-cell genealogy

Meenakshi Chatterjee and Murat Acar

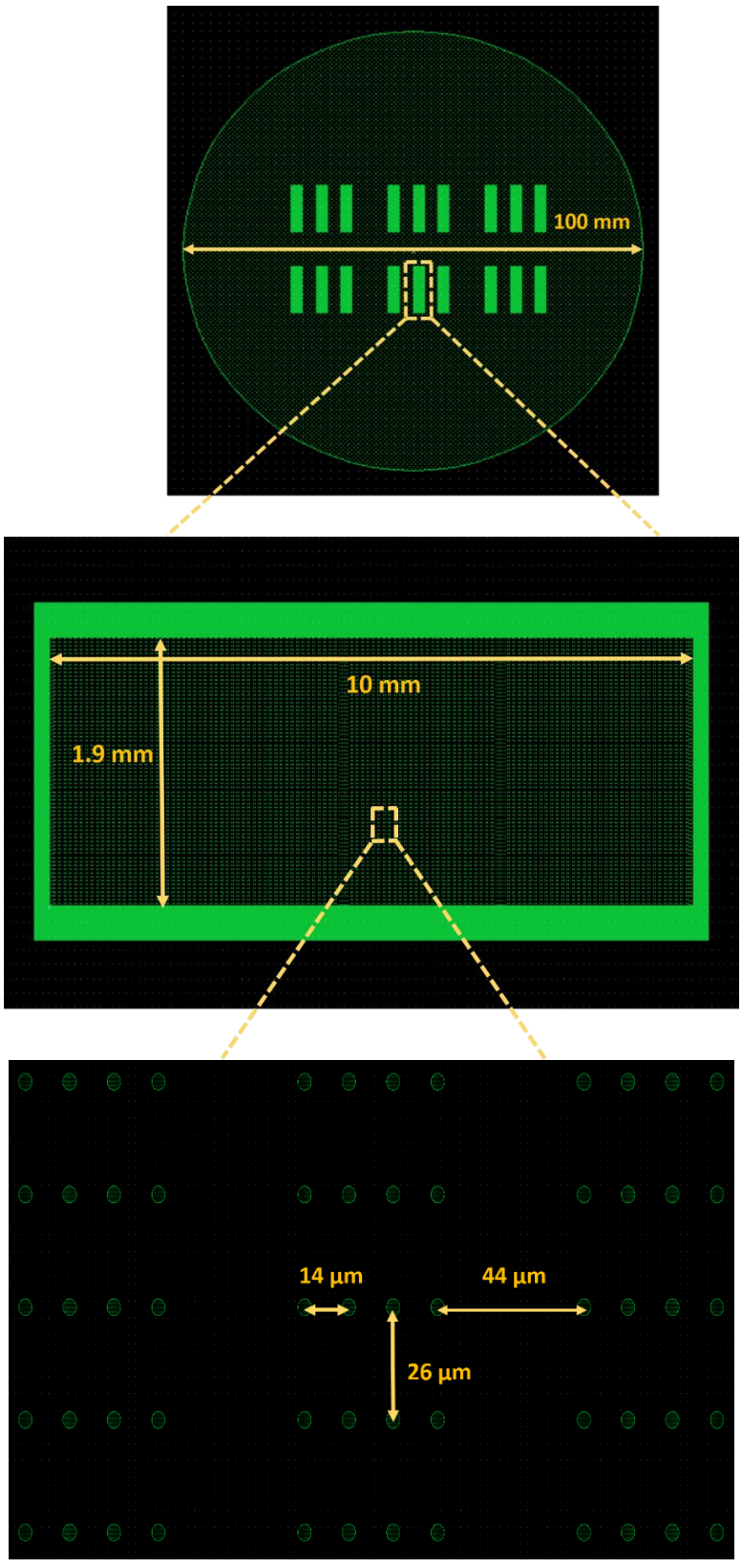
Published 18 April 2018, *Sci. Adv.* **4**, e1701775 (2018)
DOI: 10.1126/sciadv.1701775

This PDF file includes:

- fig. S14. Integral of Msn2 nuclear localization in different stress conditions.
- fig. S15. Dissimilarity analysis of different localization features of Msn2 is not affected by spatial proximity between cell pairs.
- fig. S16. Time dependent analysis of inheritance of Msn2 amplitude.
- fig. S17. Similarity of Msn2 localization spikes as a function of time for M-D cell pairs in different glucose concentrations (related to Fig. 7B).
- fig. S18. Correlation analysis between Msn2 localization features and cellular growth rate.
- fig. S19. Mean squared error for Lasso solution as a function of different values of regularization or shrinkage parameter.
- fig. S20. Lasso analysis is not sensitive to the choice of threshold used for Msn2 nuclear localization quantification.
- fig. S21. System and measurement noise for different stress environments.
- fig. S22. Block diagram of system identification and prediction steps.
- fig. S23. Predicting CFP expression levels using cross-validation.
- fig. S24. Sample polynomial fits for single-cell CFP trajectories.
- table S1. Population doubling times of cells in different glucose concentrations obtained from fig. S2 (A to C).
- table S2. Population doubling times calculated from OD_{600} measurements of cells grown in batch, using a shaker-incubator.
- table S3. Values of parameters obtained from fitting data in fig. S6 to a sigmoidal function.
- table S4. The *P* values comparing duration of Msn2 nuclear localization between 2 and 0.25% glucose, as well as 2 and 0.1% glucose across all threshold levels.
- table S5. The *P* values comparing amplitude (A), frequency (B), and duration (C) of Msn2 nuclear localization between 2 and 0.25% glucose, as well as 2 and 0.1% glucose across different cell generations.
- table S6. The *P* values obtained from Mann-Whitney *U* test.
- table S7. The *P* values obtained from Mann-Whitney *U* test.
- table S8. The *P* values obtained from Mann-Whitney *U* test.
- table S9. Parameter values extracted from the linear state-space model's application to the data obtained from 0.25 and 0.1% glucose experiments.
- References (*42–53*)

I. SUPPLEMENTARY FIGURES

A



B

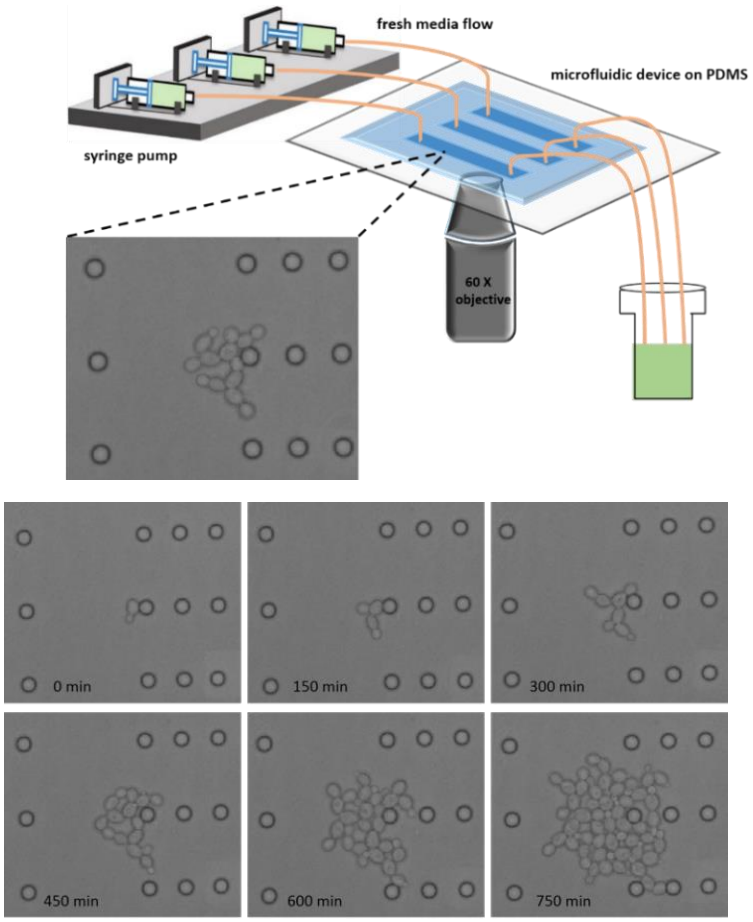


fig. S1. Design of a microfluidic platform and a typical experimental setup. A. Design of the microfluidic platform used in lineage dependent experiments to trap families of single cells. All designs have been made using the Layout Editor software. Top: Six identical microfluidic chips, each with three independent flow channels, have been drawn on a circular area representing the diameter of the silicon wafer. Middle: A single microfluidic chip. Bottom: A local area of a flow channel displayed in a zoomed-in fashion. **B**. Top: A schematic representation of a typical experimental set up. Bottom: Bright field images of cells growing into a micro-colony beginning from one ancestor cell. Only images at an interval of 150 minutes are shown here. A time-lapse movie actually acquired bright field images every 2.5 minutes.

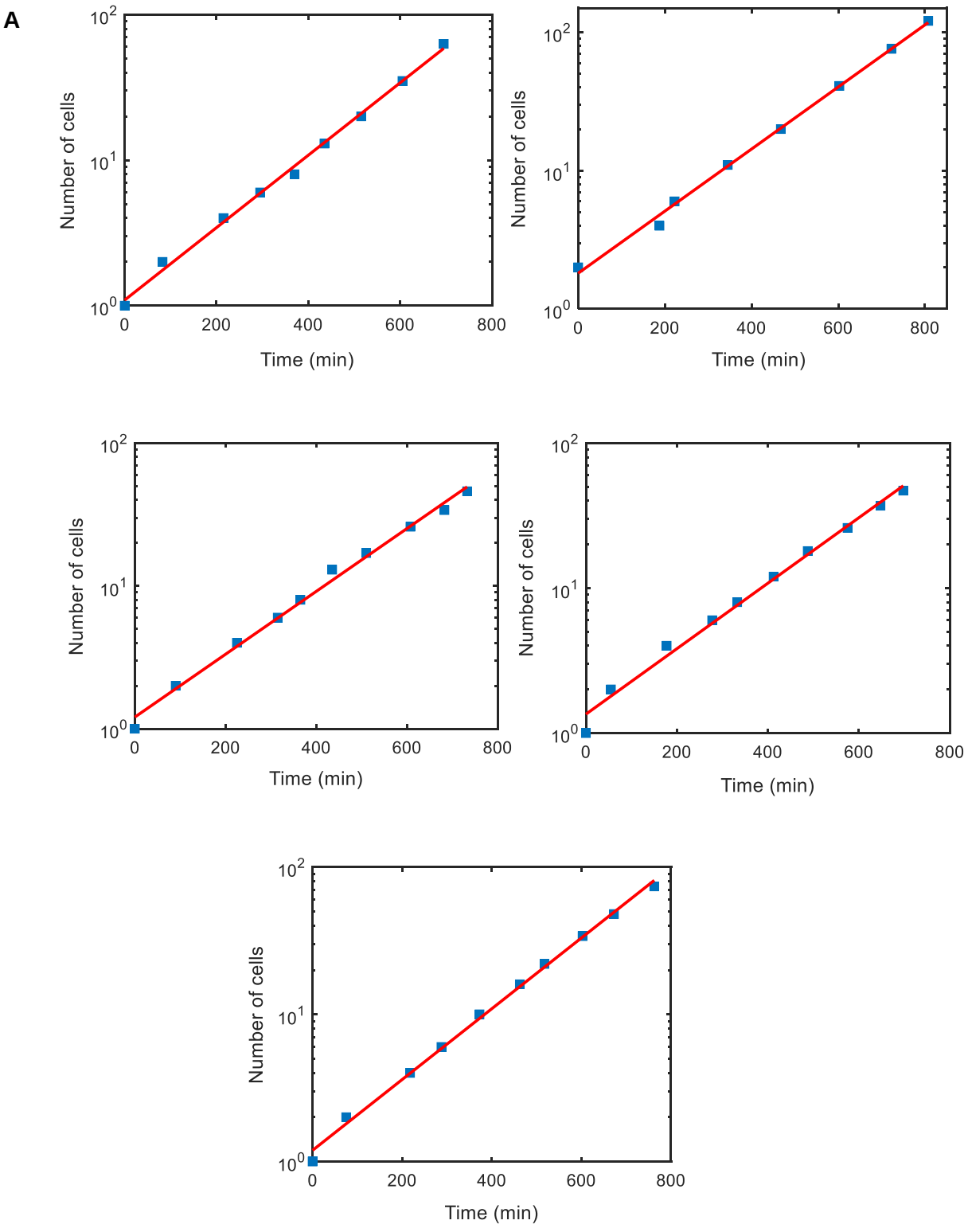


fig. S2. Growth of cell populations for cells grown in 2, 0.25, and 0.1% glucose. A. Growth of cell populations in 5 independent experiments in 2% glucose. Blue squares in a given panel denote $\log_{10}$(number of cells) measured at a given time point. A linear fit (denoted by red solid line) to the data points shows exponential cellular growth over time. The doubling time obtained after fitting five independent experiments separately and averaging them is shown in the first column of table S1. The r-square values of goodness of fit for the five fits are 0.997, 0.9962, 0.9932, 0.9897 and 0.996.

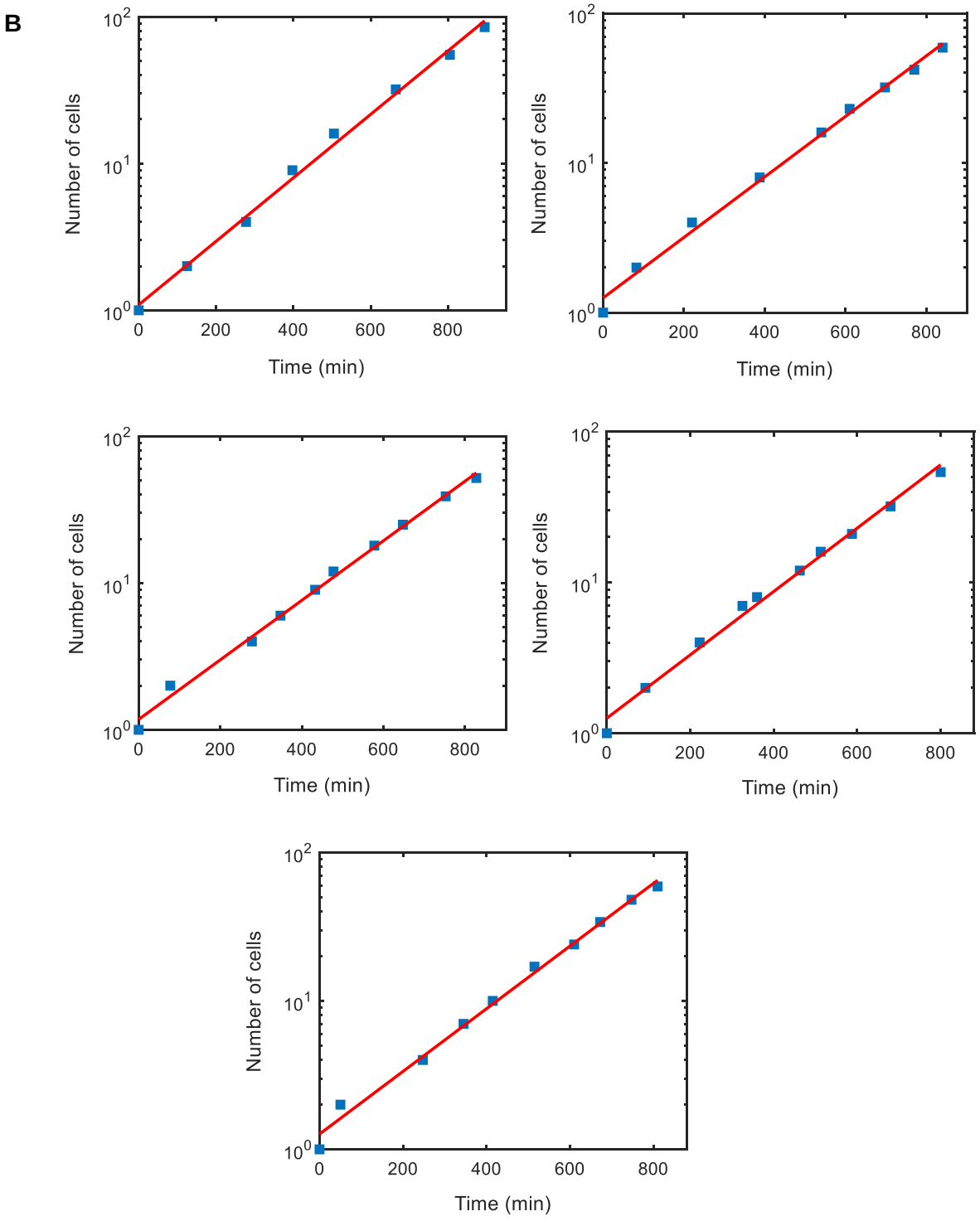


fig. S2B. Growth of cell populations in 5 independent experiments in 0.25% glucose. Blue squares in a given panel denote $\log_{10}$(number of cells) measured at a given time point. A linear fit (denoted by red solid line) to the data points shows exponential cellular growth over time. The doubling time obtained after fitting five independent experiments separately and averaging them is shown in the second column of table S1. The r-square values of goodness of fit for the five experiments are 0.9953, 0.9942, 0.9948, 0.9919 and 0.9925.

C

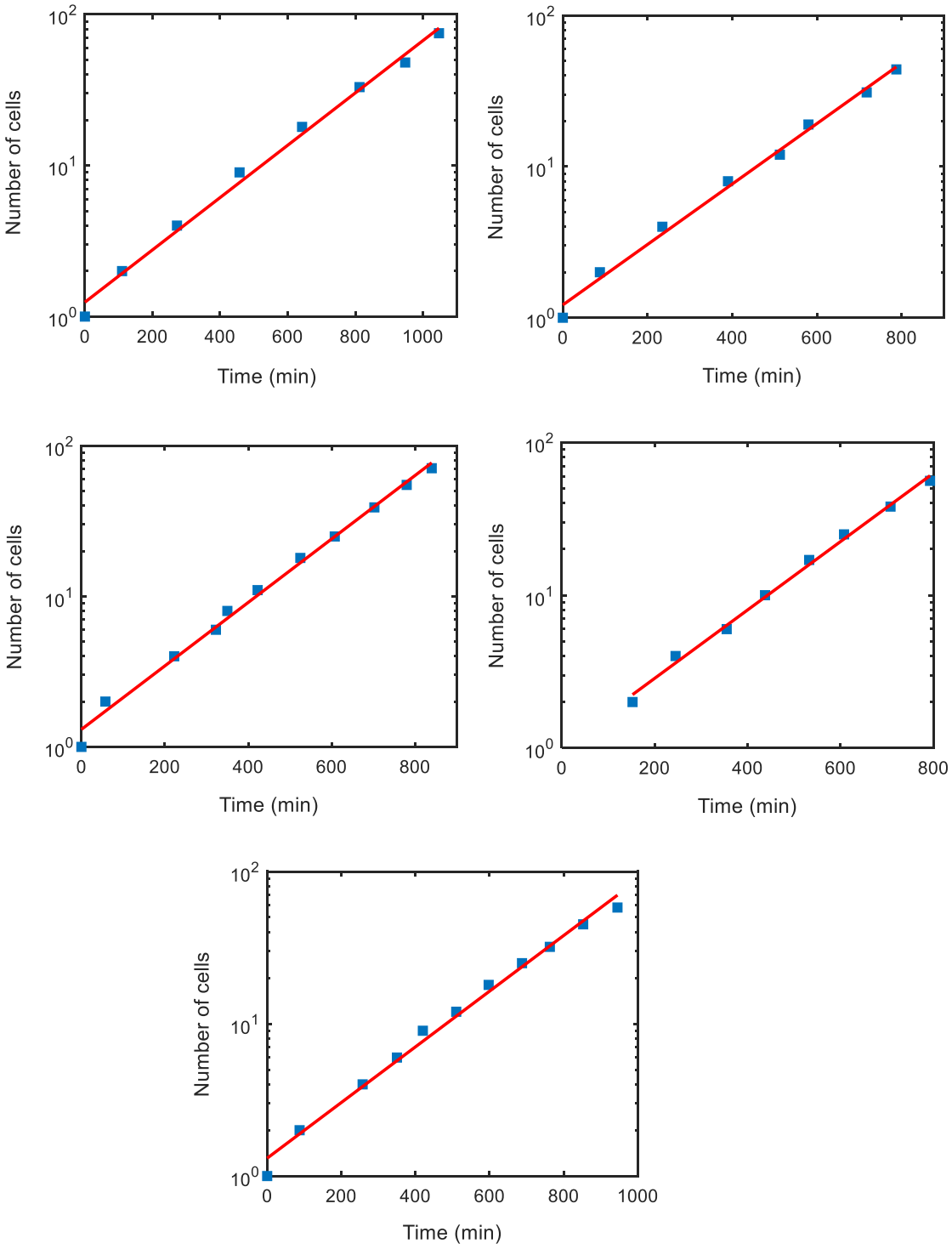


fig. S2C. Growth of cell populations in 5 independent experiments in 0.1% glucose. Blue squares in a given panel denote $\log_{10}$(number of cells) measured at a given time point. A linear fit (denoted by red solid line) to the data points shows exponential cellular growth over time. The doubling time obtained after fitting five independent experiments separately and averaging them is shown in the third column of table S1. The r-square values of goodness of fit for the five experiments are 0.9932, 0.9932, 0.9931, 0.9953 and 0.9904.

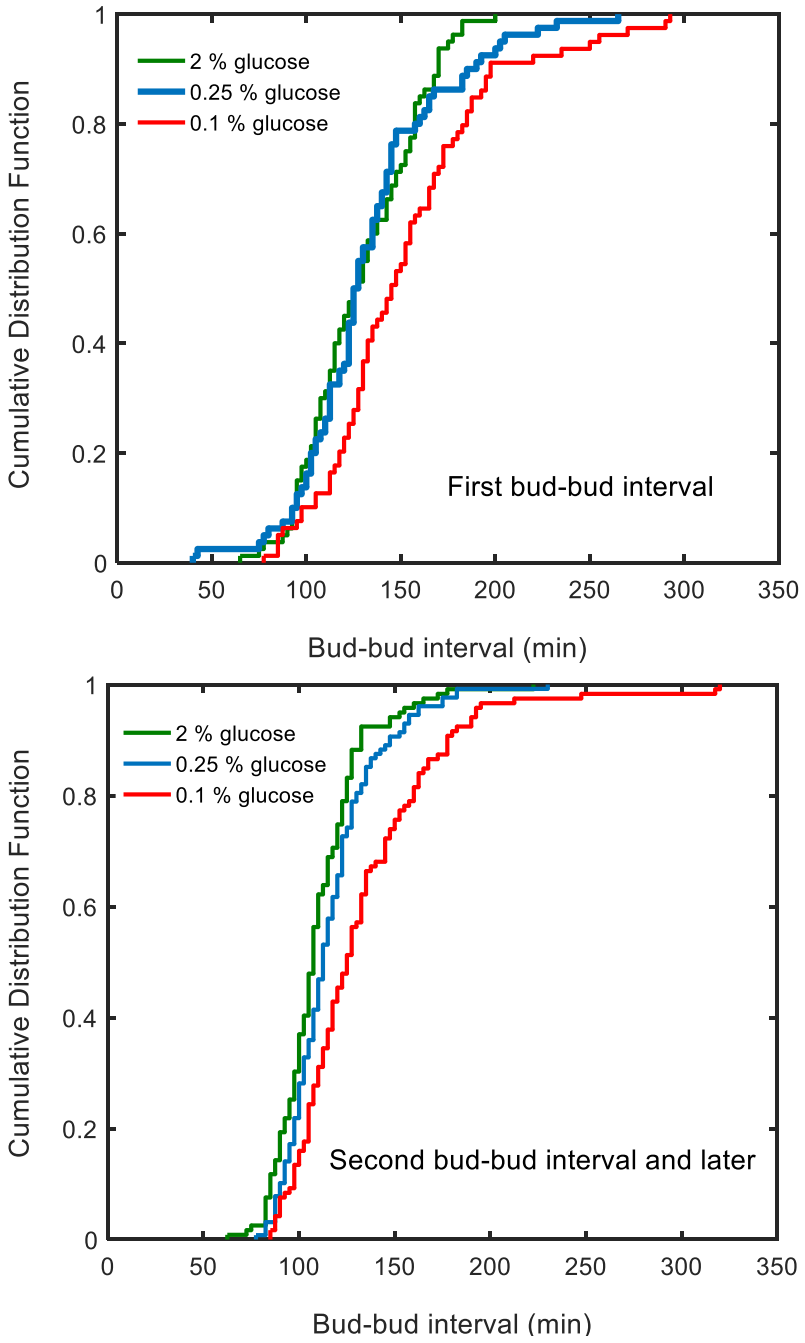
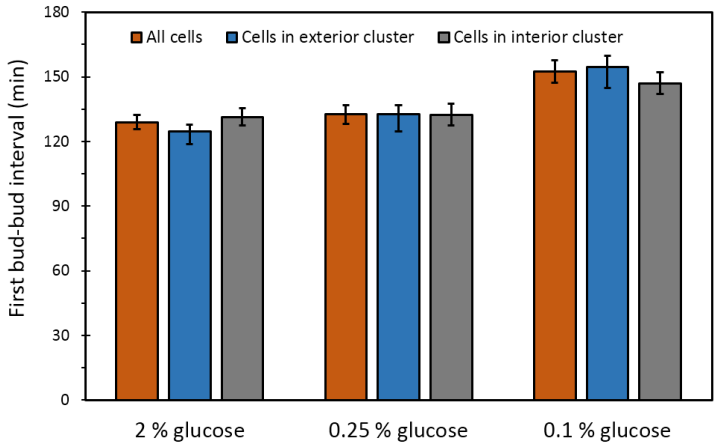


fig. S3. Cumulative distribution function of budding intervals measured from individual cells grown in 2, 0.25, and 0.1% glucose. Panel on top shows the distribution for only the time interval between occurrence of first and second bud in single cells. No significant difference ($p = 0.51$) in the first bud-bud interval were observed between 2% glucose and 0.25% glucose. Significant differences in the first bud-bud interval were observed between 2% glucose and 0.1% glucose ($p = 1.69 \times 10^{-4}$), as well as between 0.25% glucose and 0.1% glucose ($p = 0.0034$). Panel on bottom shows the distribution for subsequent budding intervals in single cells. Significant difference ($p = 0.029$) in the subsequent bud-bud intervals was observed between 2% glucose and 0.25% glucose. Significant differences were also observed in the subsequent bud-bud intervals between 2% glucose and 0.1% glucose ($p = 6.24 \times 10^{-8}$), as well as between 0.25% glucose and 0.1% glucose ($p = 7.16 \times 10^{-5}$).

A



B

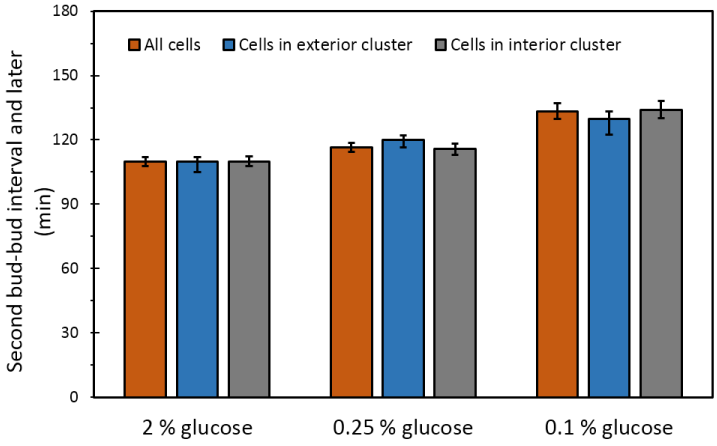


fig. S4. Budding interval durations of spatially separate cells grown in 2, 0.25, and 0.1% glucose. A. The mean and s.e.m. (N=37-45) of the first bud-bud interval duration for: all cells, cells in the exterior of the cluster, and cells in the interior of the cluster. No significant difference in first bud-bud times were observed between cells in the three categories in any glucose concentration, showing that any potential non-uniformity of the light experienced by the cells does not lead to growth differences. The p-values for 2% glucose: 0.342, 0.525, and 0.632 between cells in interior and exterior of a cluster, between all cells and cells in exterior cluster, and between all cells and cells in interior cluster, respectively. The p-values for 0.25% glucose: 0.982, 0.988, and 0.99 between cells in interior and exterior of a cluster, between all cells and cells in exterior cluster, and between all cells and cells in interior cluster, respectively. The p-values for 0.1% glucose: 0.501, 0.846, and 0.461, between cells in interior and exterior of a cluster, between all cells and cells in exterior cluster, and between all cells and cells in interior cluster, respectively. **B**. The mean and s.e.m. (N=40-81) of the subsequent bud-bud interval durations as averaged over cells in the three spatial categories for each glucose concentration. No significant difference in subsequent bud-bud times were observed between cells in the three spatial categories in any glucose concentration. The p-values for 2% glucose: 1, 1, and 1, between cells in interior and exterior of a cluster, between all cells and cells in exterior cluster, and between all cells and cells in interior cluster, respectively. The p-values for 0.25% glucose: 0.306, 0.384, and 0.8, between cells in interior and exterior of a cluster, between all cells and cells in exterior cluster, and between all cells and cells in interior cluster, respectively. The p-values for 0.1% glucose: 0.601, 0.649, and 0.9036, between cells in interior and exterior of a cluster, between all cells and cells in exterior cluster, and between all cells and cells in interior cluster, respectively. For each of the five independent experiments in each glucose condition, cells in the interior and exterior of a cluster have been defined in the following way: i. for each movie, when

all cells were born and formed a cluster, the centroid of each cell was identified; ii. approximating the perimeter of the cluster of all cells as a circle, the centroid of that circle was obtained by averaging centroids of all cells in the cluster. To determine the radius of this circle, distance between the centroid of the circle and centroid of each cell was calculated. Radius of the cluster circle was then determined as the maximum of these distances; iii. finally, within the circle corresponding to the entire cell cluster, an exterior and interior circle and associated areas were constructed (they were concentric). The area of the interior circle was defined as half of the area of the full cluster (the radius of the interior circle was $(1/\sqrt{2})$ of the radius of the full circle). All cells that fell within the interior circle/area were called “cells in the interior cluster”, while all the remaining cells were called “cells in the exterior cluster”.

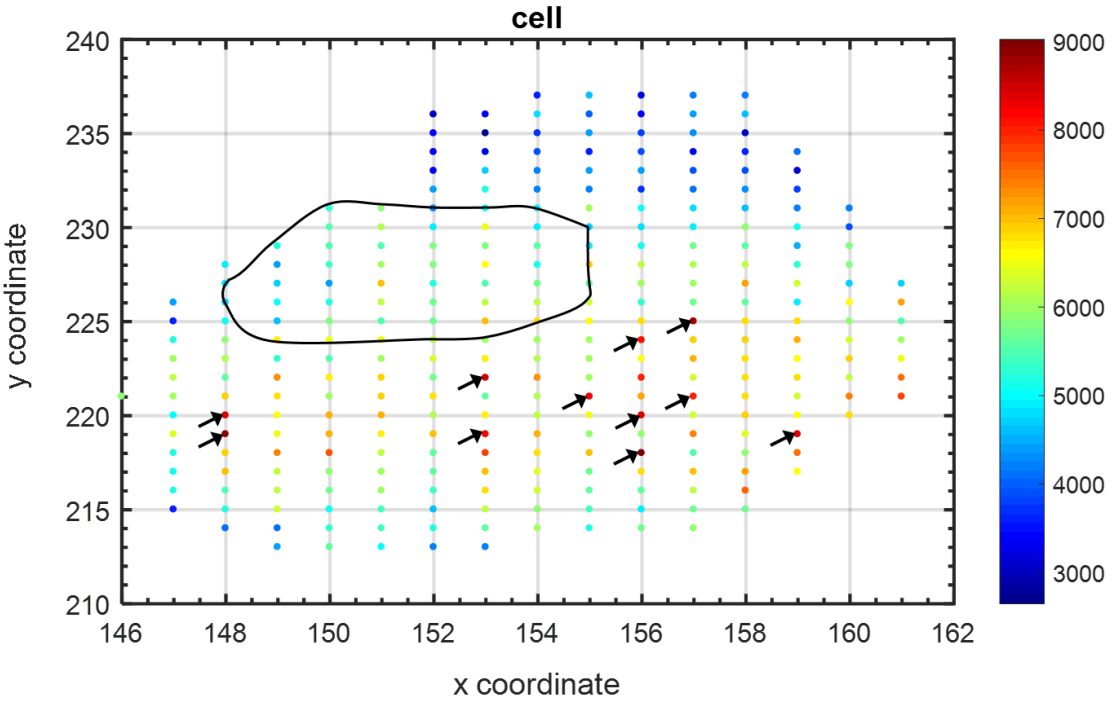


fig. S5. Example of erroneous approximation of Msn2 nuclear localization intensity. Approximation of nuclear intensity by averaging the intensity of brightest pixels in the whole cell may give erroneous estimates, especially when the localization is weak. Figure shows the scatter plot of pixels in a cell. Color bar shows the intensity of pixels (a.u.). Position of the nucleus of the cell, as determined by a using nuclear marker, is denoted by a black polygon. Brightest 11 pixels in the cell are shown with arrows.

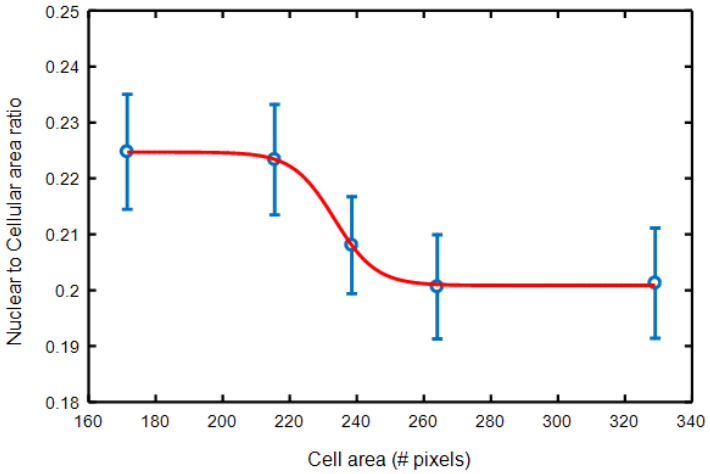


fig. S6. Ratio of nuclear to cellular area as a function of the cell area. Small cells have a relatively larger nuclear-to-cellular area ratio while large cells have a smaller ratio. Data was binned and fitted to the following sigmoidal function: $y = a_1 + \frac{a_2 - a_1}{1+10^{(a_3 - x)*a_4}}$, where, x is the area of a cell and y is the ratio of nuclear area to cell area. The values of parameters a_1, a_2, a_3, a_4 are given in table S3. Details of analysis are given in the Supplementary Materials and Methods.

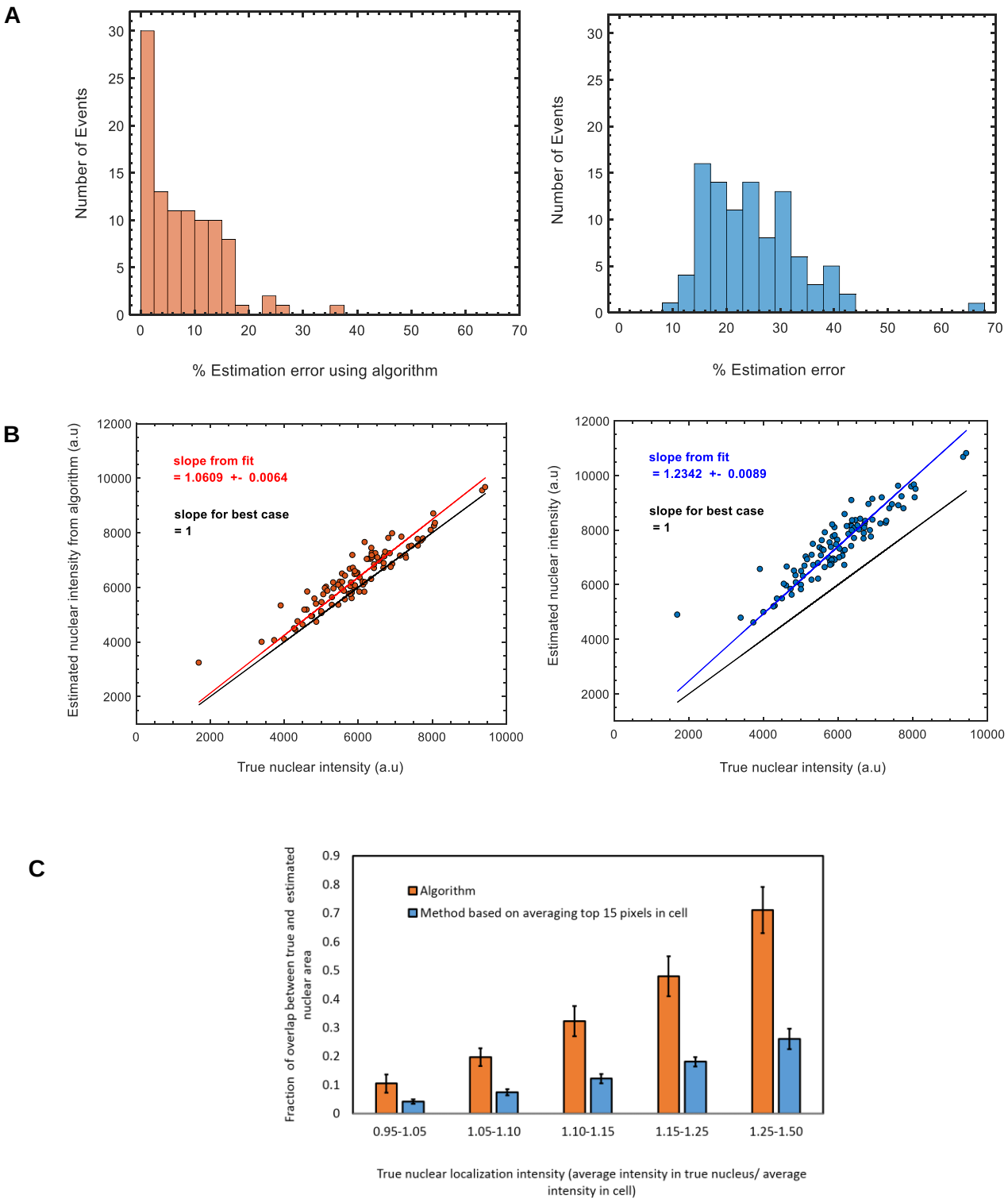


fig. S7. Performance of algorithm in estimation of nuclear localization of Msn2 from whole cell, in the absence of a nuclear marker. A. Histogram of error in estimation of Msn2 localization in individual cells, using our algorithm (left) and using a method in literature based on averaging intensities of bright pixels in a cell (right). Details of how error in estimation of Msn2 localization is calculated for each cell is mentioned in the Supplementary Materials and Methods. The average

estimation error is only 8.46±1.07% using our algorithm (compared to 26.4±1.88% when taking average of top 15 YFP pixel intensities in single cells). **B**. Left: Plot of nuclear intensity in a cell estimated using our algorithm vs true nuclear intensity of that cell. Number of cells used for analysis is 100. Slope of the linear fit deviates from that of the best case only by approximately 6%. Right: Plot of nuclear intensity estimated from averaging top 15 pixels in a cell vs true nuclear intensity of that cell. Slope of the linear fit in this case deviates from the best case by approximately 23%. **C**. Fraction of overlap between true nuclear area and estimated nuclear area (either using our algorithm or the method in literature based on average brightest 15 YFP pixels in whole cell) as a function of nuclear localization intensity. Nuclear localization of single cells has been binned, and the range of localization intensity chosen for each bin (bin width) is shown on the x-axis. As the strength of nuclear localization signal increases, higher is the algorithm's power to capture true nuclear pixels, thus giving rise to a higher amount of fractional overlap. The algorithm performed better than the method in literature, regardless of the strength of the localization signal.

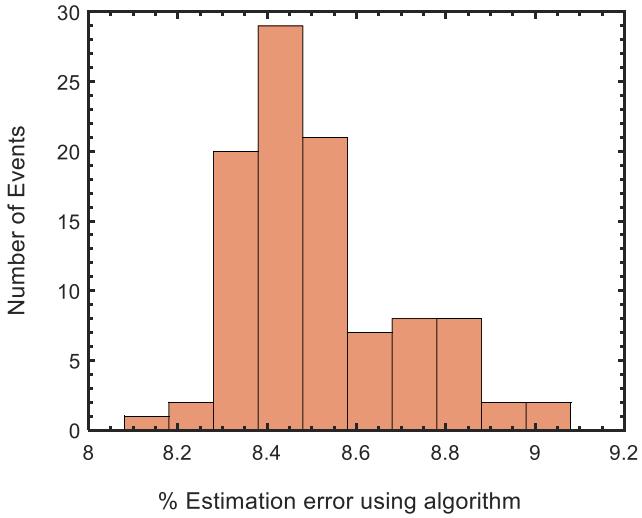


fig. S8. Estimation of nuclear localization of Msn2 using the algorithm is not sensitive to the variability in ratio of nuclear to cellular area at the single-cell level. For each cellular area in fig. S6, the corresponding ratio of nuclear to cellular area was given a value randomly sampled from a uniform distribution with an interval ranging from ± 1 standard deviation of mean ratio. Nuclear localization was then estimated using the algorithm and performance of the algorithm was determined by calculating the average estimation error, as before. The method was then repeated 100 times, and average estimation error was obtained each time. Histogram of average estimation error for each of the 100 runs is shown. The average estimation error from a total of 100 runs is 8.52±0.18% and shows a negligible increase in error compared to the case when only mean values of the ratio of nuclear to cellular area (which was 8.46±1.07% as shown in fig. S7A) was used.

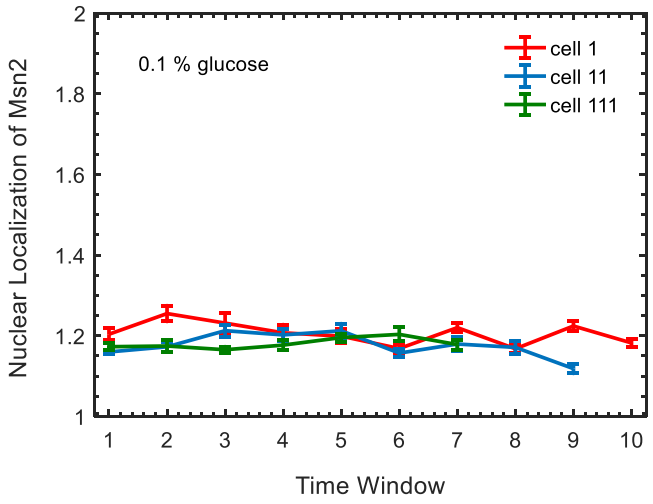


fig. S9. No photobleaching or significant drop in Msn2 signal was observed over the course of an experiment. Using the data plotted in Fig. 1E, the Msn2 signal of a cell (measured from its birth to the end of the experiment in 0.1% glucose condition) was divided into windows of 100 minutes each. Mean of each cell's Msn2 signals in a window was calculated. Error bars denote standard deviation of the mean. The p-value between the first averaged data point and the last one for all three cells were greater than 0.05 (0.6, 0.61 and 0.08), and hence do not reflect a significant change in average intensity of Msn2.

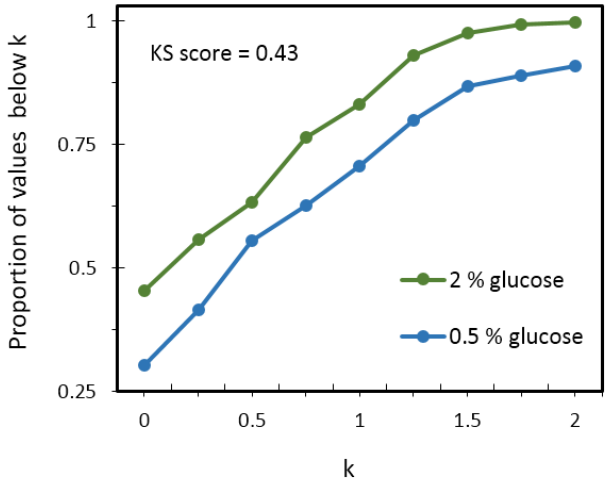


fig. S10. Proportion of cells having localization values below a given threshold k as a function of different thresholds k. A threshold (or a value of k in this figure) represents the number of standard deviation above the mean of localization values from 2% glucose dataset. For example, k = 0 implies that the threshold is taken as mean of the localization values from 2% glucose; k = 1 implies that the threshold is taken as (mean +1 standard deviation) of localization values from 2% glucose dataset. A two sample Kolmogorov-Smirnov (K-S) test, (which is a nonparametric test for equality of two probability distributions) yields a score of 0.43 (p value< 0.001). The test confirms that the above two distributions obtained from the two glucose concentrations are different. Detectable and similar difference in localization values between the two glucose concentrations is observed to appear for a threshold between 0 and 2 standard deviation of the mean of localization of values of all cells in 2% glucose condition.

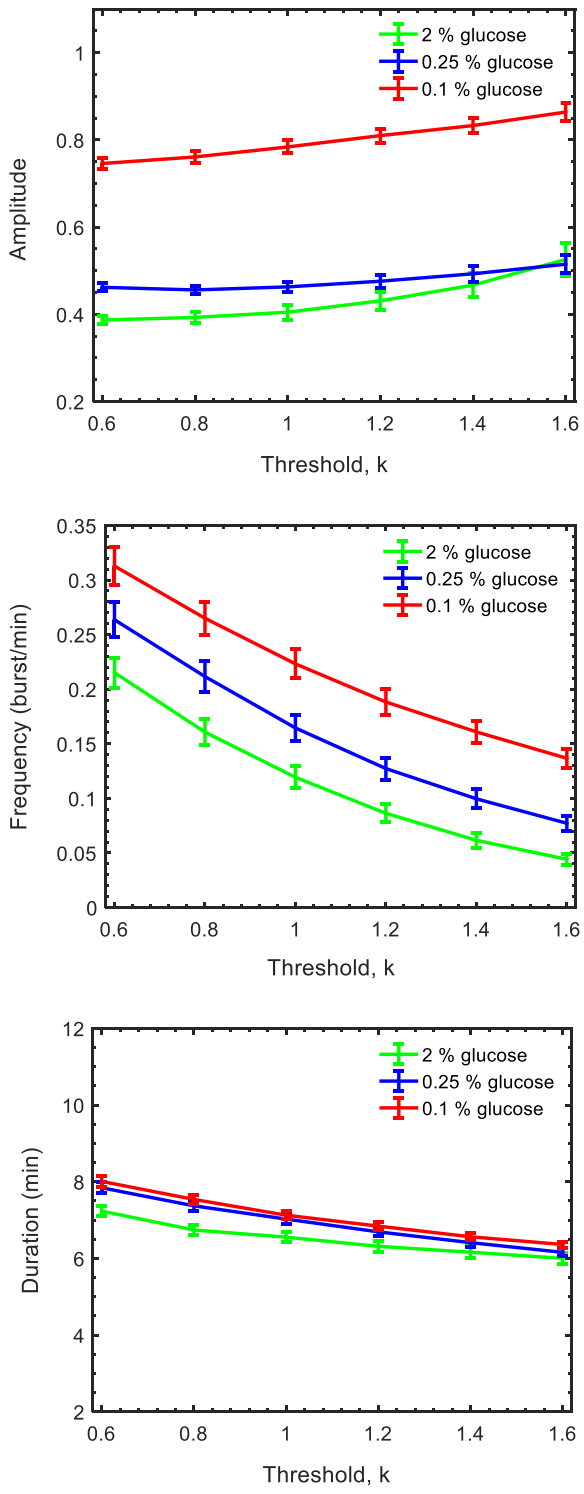
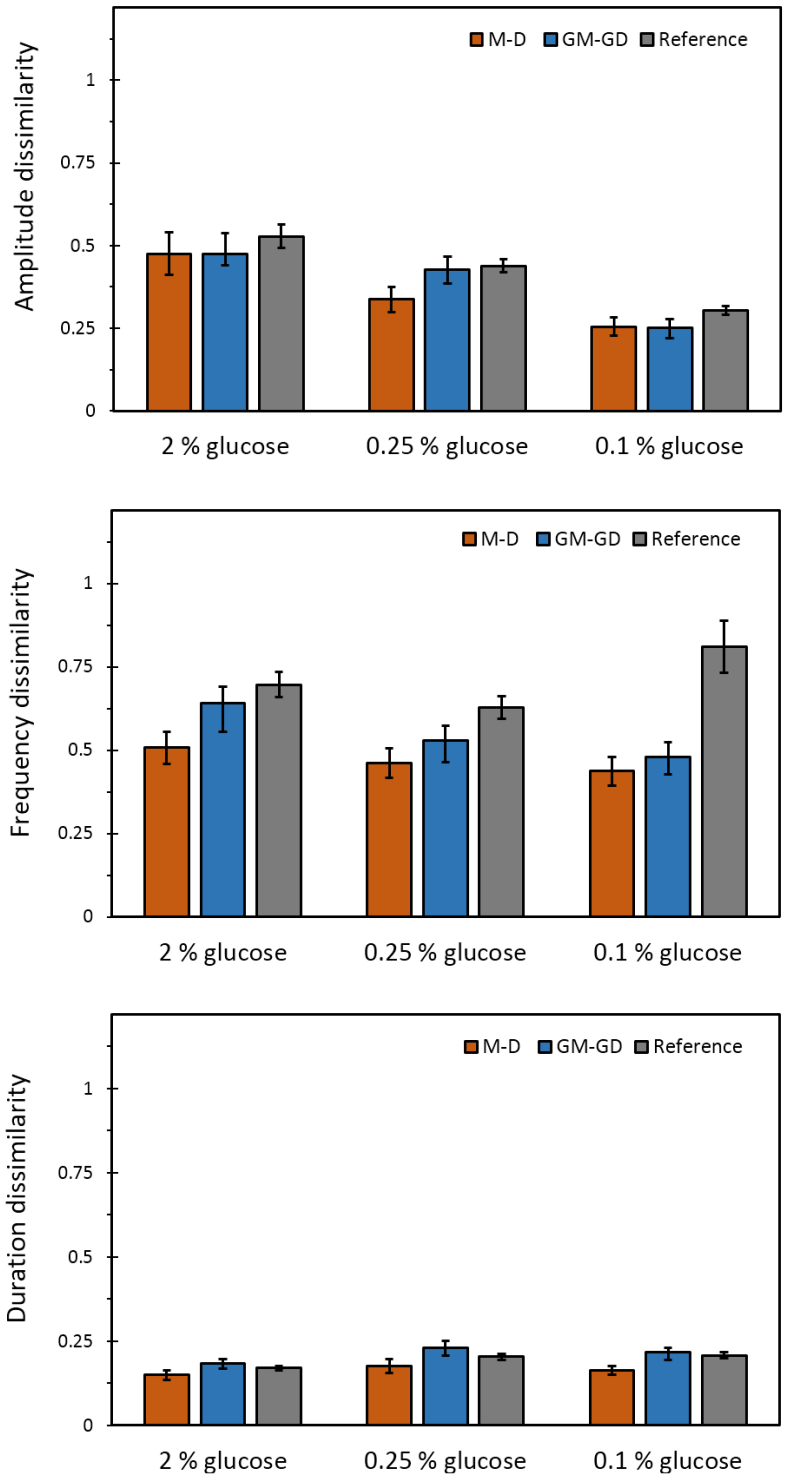


fig. S11. Quantification of Msn2 localization dynamics is robust to the choice of threshold used for localization quantification. Amplitude (top), frequency (middle), duration (bottom) of Msn2 burst as a function of different thresholds k. Each value of threshold k represents the number of standard deviation above the mean of localization values from 2% glucose experiments. Five replicates of 2% glucose experiments were used for determination of threshold. All p-values comparing duration between 2% glucose and 0.25% glucose, as well as 2% and 0.1% glucose across different values of thresholds is given in table S4.

A


M-D
GM-GD
Reference
Amplitude dissimilarity
1
0.75
0.5
0.25
0
2 % glucose
0.25 % glucose
0.1 % glucose
M-D
GM-GD
Reference
Frequency dissimilarity
1
0.75
0.5
0.25
0
2 % glucose
0.25 % glucose
0.1 % glucose
M-D
GM-GD
Reference
Duration dissimilarity
1
0.75
0.5
0.25
0
2 % glucose
0.25 % glucose
0.1 % glucose

B

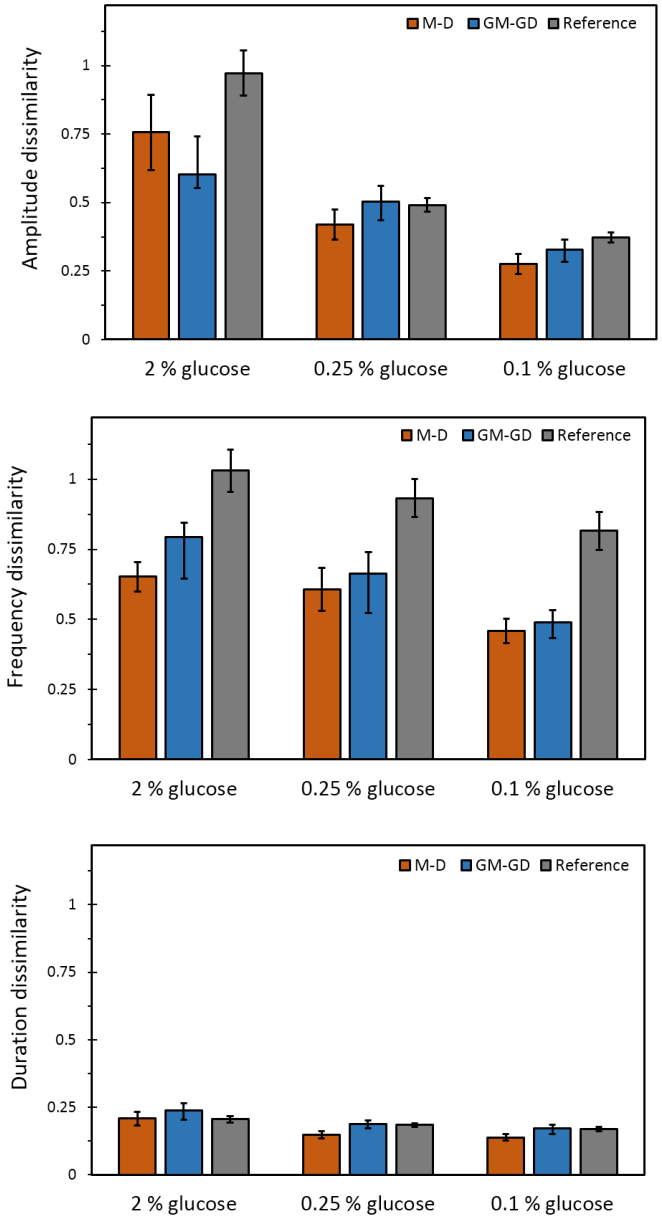
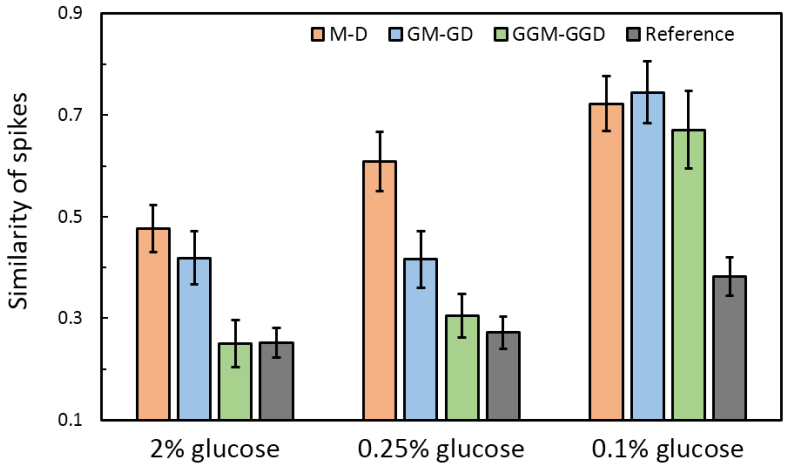


fig. S12. Heritability analysis of different localization features of Msn2 and analysis of inheritance of dynamical patterns of Msn2 localization is not sensitive to the choice of threshold used for Msn2 localization quantification. (**A** and **B**) Heritability analysis of different localization features of Msn2 is not sensitive to the choice of threshold used for Msn2 localization quantification. Single cell trajectories of Msn2 localization were thresholded either at 0.8 standard deviation (A) or at 1.2 standard deviation (B) of the mean of localization traces from 2% glucose, and analysis of heritability of Msn2 amplitude, frequency, and duration were performed for each case. Error bars represent standard error of mean. Table S6B-C gives a list of all p-values obtained by comparing amplitude, frequency and duration of Msn2 localization between M-D pairs, GM-GD pairs and Reference pairs for each glucose concentration, and across different threshold levels.

C



D

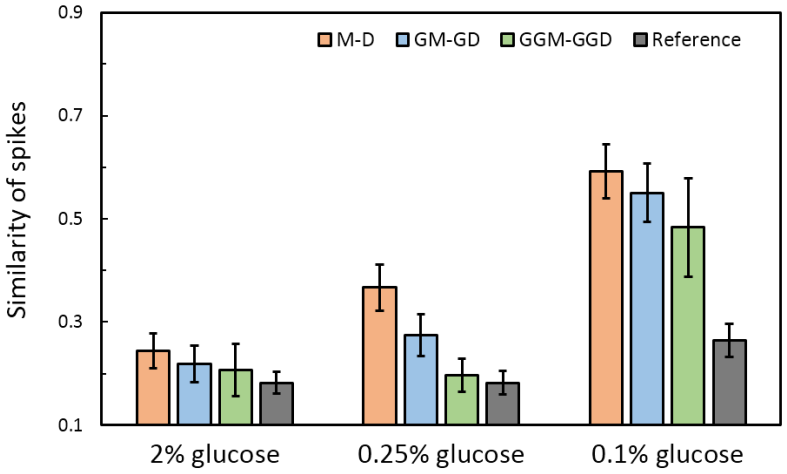


fig. S12. (**C** and **D**) Analysis of inheritance of dynamical patterns of Msn2 localization is not sensitive to the choice of threshold used for Msn2 localization quantification. Single cell trajectories of Msn2 localization were thresholded either at 0.8 standard deviation (C) or at 1.2 standard deviation (D) of the mean of localization traces from 2% glucose and similarity of Msn2 spikes was computed in each case. Error bars were obtained from bootstrapping. Table S8B-C gives a list of all p-values obtained by comparing similarity of Msn2 spikes between M-D, GM-GD, GGM-GGD and Reference pairs for each glucose concentration and across different threshold levels.

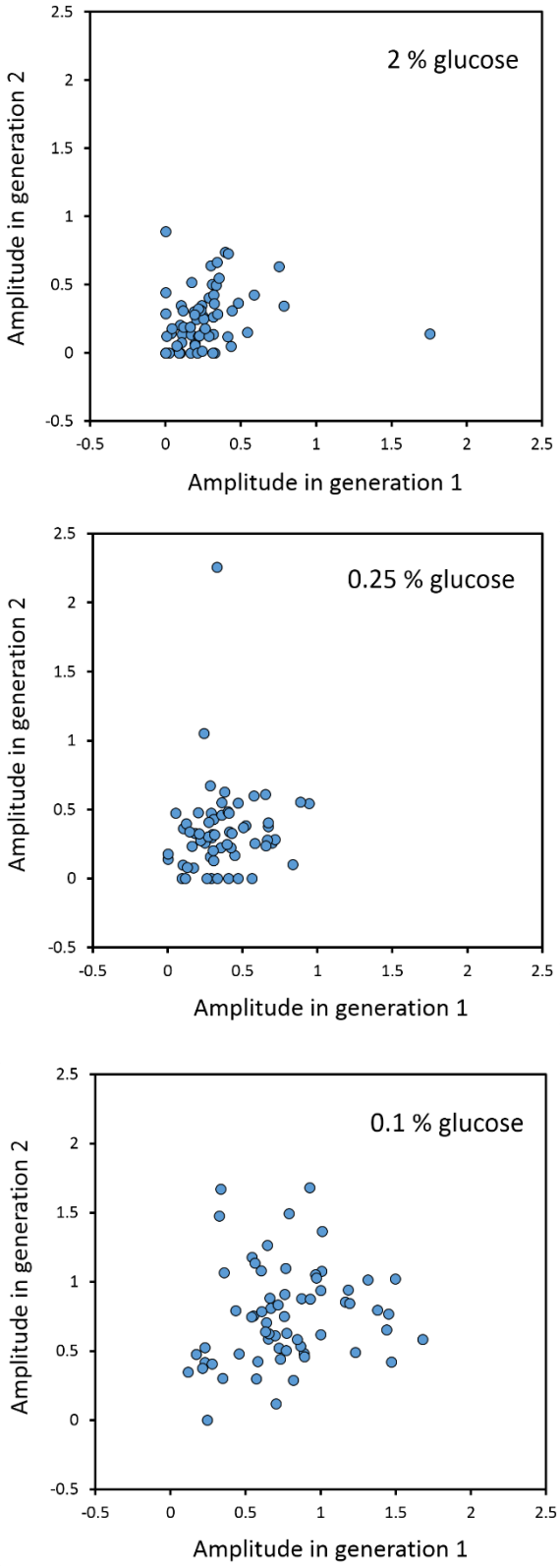


fig. S13. Lack of correlation in localization amplitude between the first and second generation of the same cell. No significant correlation was observed for any concentration (p=0.14, p=0.43, p=0.22 for 2%, 0.25%, and 0.1% glucose respectively). One cell generation was defined as the time interval between the start of two consecutive S-phases for the same cell.

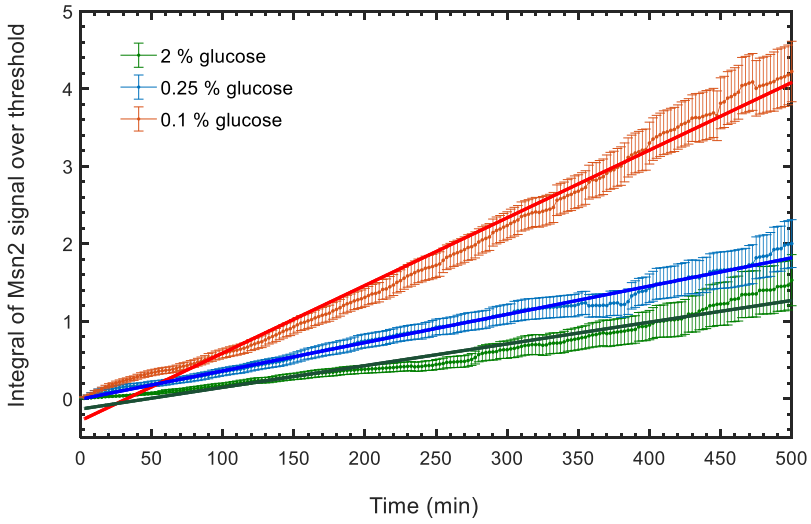


fig. S14. Integral of Msn2 nuclear localization in different stress conditions. For each glucose concentration and at each time point, the integral of each cell's Msn2 localization value (above the threshold) was calculated since the cell's birth. Then, for each time point, the integrands were averaged over all cells present at each time point (normalized by the number of cells present at each time point). Error bars at a given time represent standard error of mean at that time. As an example, integral of Msn2 localization averaged over all cells at 500 minutes, are 1.53$\pm$0.33, 2$\pm$0.31, and 4.22$\pm$0.39 for 2%, 0.25%, and 0.1% glucose respectively. Dark solid line for each glucose concentration represents a linear fit to the averaged values for that concentration. The slopes of the fitted lines are 2.8×10^{-3}, 3.65×10^{-3} and 8.73×10^{-3} for 2%, 0.25% and 0.1% glucose respectively. The r-square values of goodness of fit are 0.956, 0.991 and 0.988 for 2%, 0.25% and 0.1% glucose respectively.

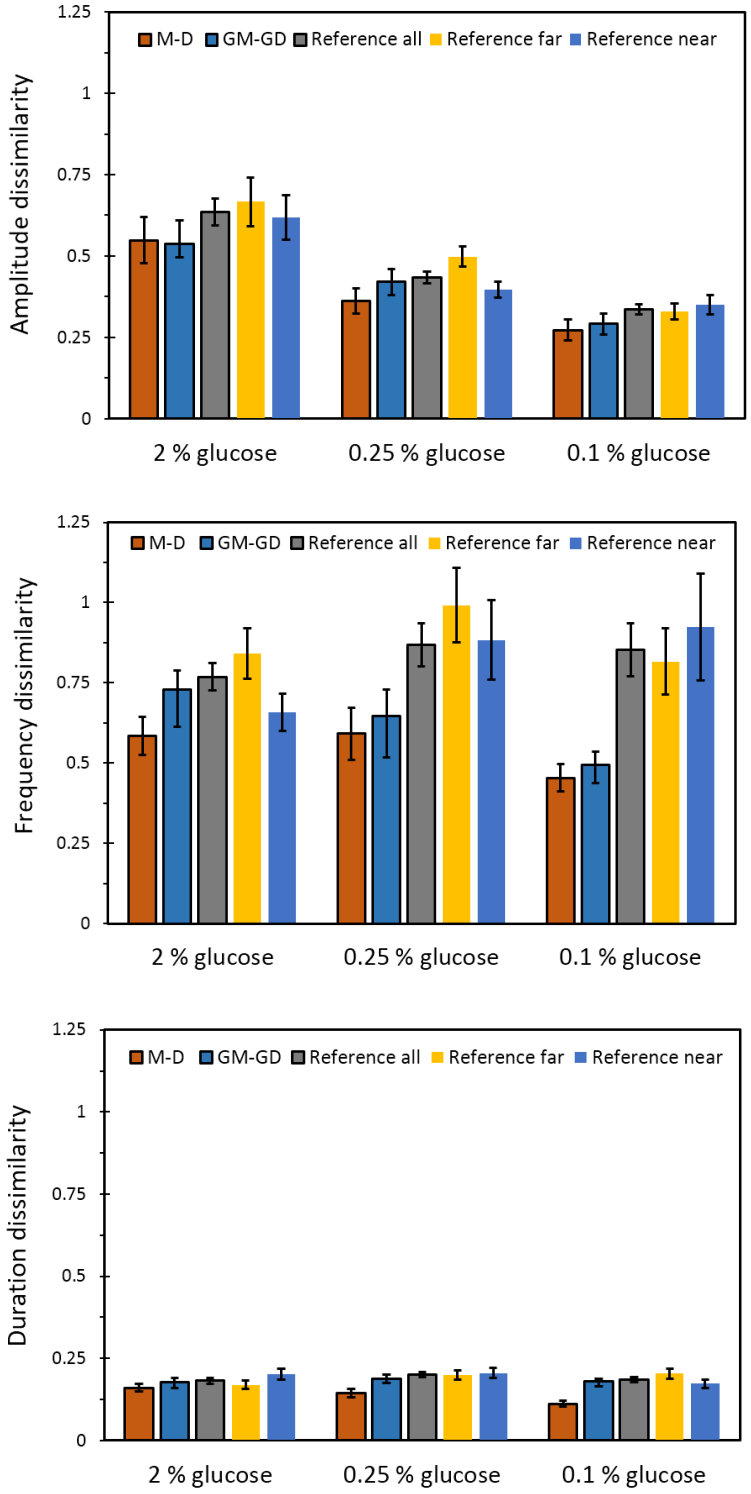


fig. S15. Dissimilarity analysis of different localization features of Msn2 is not affected by spatial proximity between cell pairs. Single cell trajectories of Msn2 localization were thresholded at 1 standard deviation of the mean of localization traces from 2% glucose, and analysis of heritability of Msn2 amplitude (top), frequency (middle), and duration (bottom) were performed for each case. Values for M-D, GM-GD and *Reference all* for top, middle and bottom panels have been reproduced from M-D, GM-GD and *Reference* pairs in Fig. 6B, 6C, 6D, respectively. Additionally,

two new sets of reference cell pairs were constructed from the *Reference all* category: i) reference cell pairs which were spatially far from each other (named as *Reference far*) and ii) reference cell pairs which were spatially close to each other (named as *Reference near*). For each experiment in a glucose condition, the distance between centroids of the two cells in each reference pair was first calculated and the centroid-centroid distances for all reference pairs were then sorted. Reference cell pairs having centroid-centroid distances within the lowest 35% were considered near to each other and defined as *Reference near*. Reference cell pairs having centroid-centroid distances within the highest 35% were considered spatially far away from each other and defined as *Reference far*. The remaining 30% of the total reference cell pairs, whose centroid-centroid distances belonged to the middle range, were not used and hence ensured a larger discrimination between the spatially far and near categories of reference cell pairs. Error bars represent standard error of mean (N = 120-150 for *Reference far* and *Reference near*). Table S7 provides a list of all p-values obtained by comparing amplitude, frequency and duration of Msn2 localization between *Reference all, Reference far* and *Reference near* cell pairs for each glucose condition. Indicating that the analysis of Msn2 localization features is not influenced by spatial distance between cell pairs, no significant differences were observed between the three *reference* categories for any glucose conditions, except for 0.25% glucose for the feature of localization amplitude between *Reference far* and *Reference near* categories (p-value: 0.01), as well as between *Reference all* and *Reference far* categories (p-value: 0.03). Although these two p-values are less than the 0.05 cut-off, we note that we did not use any correction for multiple-hypothesis testing (for the sake of extra stringency) and applying such a correction (e.g., the Benjamini-Hochberg procedure) will cause these p-values to become insignificant.

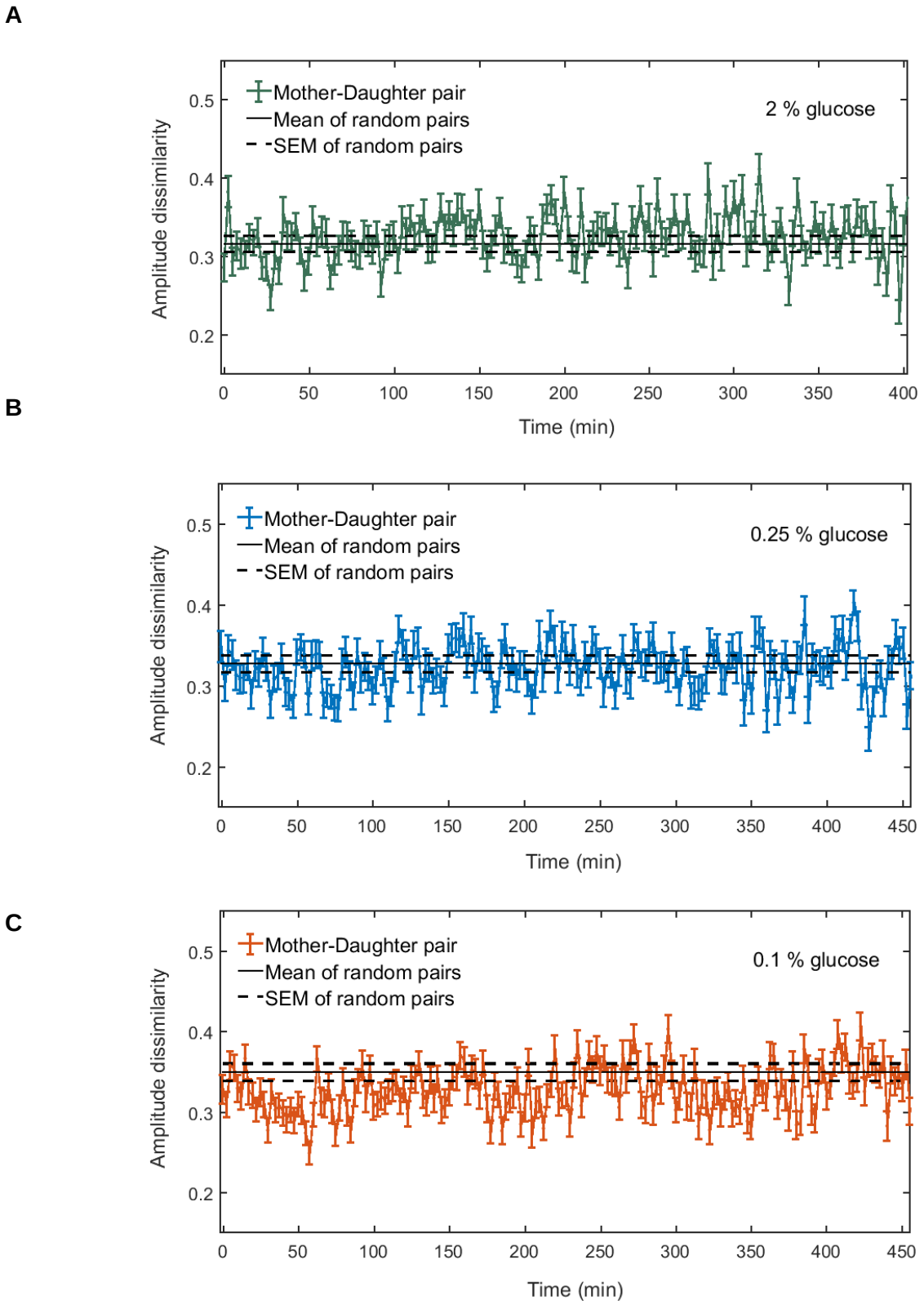


fig. S16. Time dependent analysis of inheritance of Msn2 amplitude. Mean of pairwise Msn2 amplitude difference between mother and daughter cells as a function of time following cell division in 2% glucose (**A**), 0.25% glucose (**B**) and 0.1% glucose (**C**). Solid black line

represents the mean of pairwise distances of reference pair of cells chosen at random without regard to lineage relationships. Dashed black lines above and below the solid black line represent the standard error of mean.

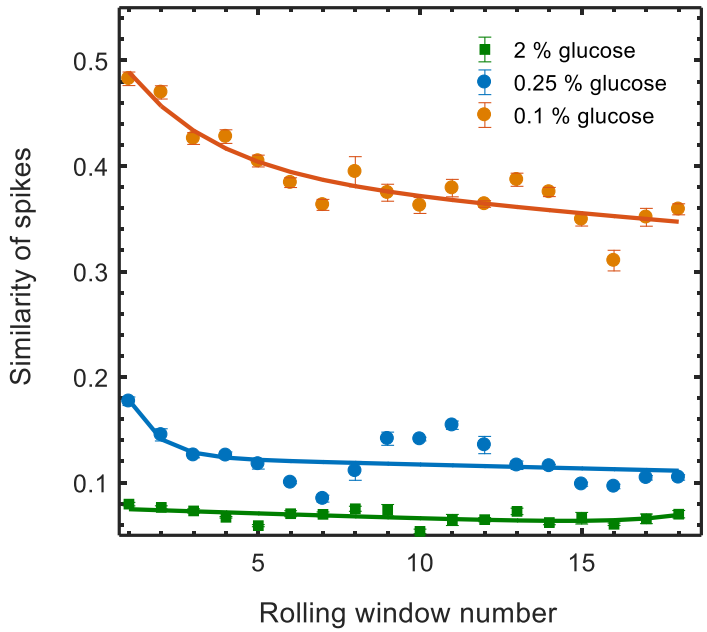
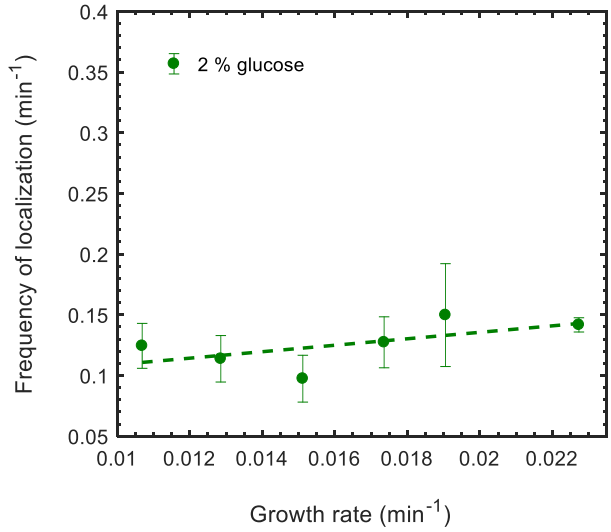
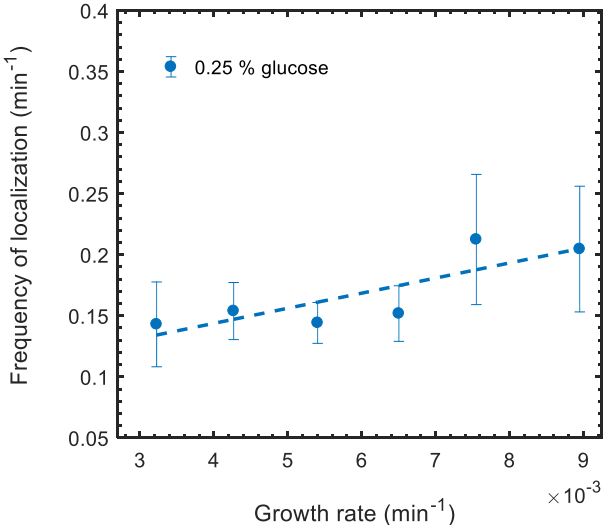
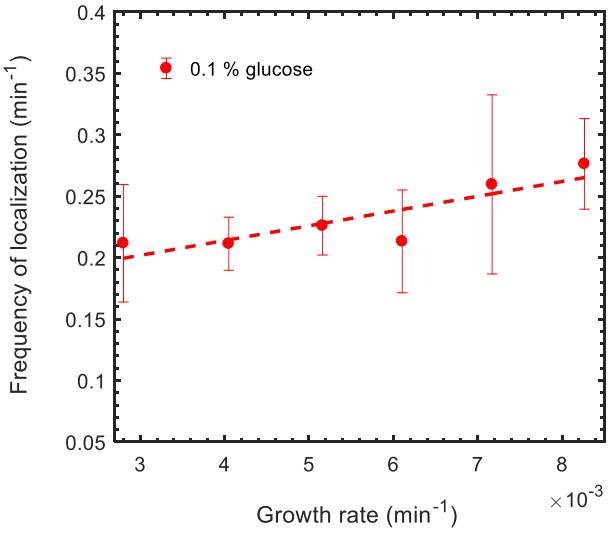


fig. S17. Similarity of Msn2 localization spikes as a function of time for M-D cell pairs in different glucose concentrations (related to Fig. 7B). As before, similarity in spike pattern for any M-D pair was calculated starting from the time the daughter cell was born (as indicated by the blue dashed line in Fig. 7A). Then using a time window of length 150 minutes, and sliding or rolling the window in every time step (of 2.5 minutes), the similarity in spikes was computed between M-D cell pairs for each rolling time window. Similarity in spikes in 5 consecutive rolling windows were then averaged for smoothing purposes. Error bars represent standard error of mean calculated from the 5 values. As an example, to obtain the similarity of spikes at rolling window number 1 in the above figure: similarity in spikes between M-D pairs were first calculated for a particular glucose concentration in 5 time windows spanning [0-150 minutes], [2.5-152.5 minutes], [5-155 minutes], [7.5-157.5 minutes], [10-160 minutes] and then averaged over these 5 time windows. Similarly, to obtain the similarity of spikes at rolling window number 2 in the above figure: similarity in spikes between M-D pairs were first calculated for the particular glucose concentration in 5 time windows spanning [12.5-162.5 minutes], [15-165 minutes], [17.5-167.5 minutes], [20-170 minutes], [22.5-172.5 minutes] and then averaged over these 5 time windows. The similarity of spikes at other rolling windows were also calculated in a similar fashion. The similarity in spikes is higher for higher stress environment and decreases with time both in the 0.1% and 0.25% glucose environments, with the decreasing trend being the most obvious for 0.1% glucose. Similarity in spikes in 2% glucose is much lower to begin with and remains flat over time. Circles represents data and solid lines represent an exponential fit to the data values.

A


2 % glucose
Frequency of localization (min-1)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0.01
0.012
0.014
0.016
0.018
0.02
0.022
Growth rate (min-1)


0.25 % glucose
Frequency of localization (min-1)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
3
4
5
6
7
8
9
×10-3
Growth rate (min-1)


0.1 % glucose
Frequency of localization (min-1)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
3
4
5
6
7
8
×10-3
Growth rate (min-1)

B

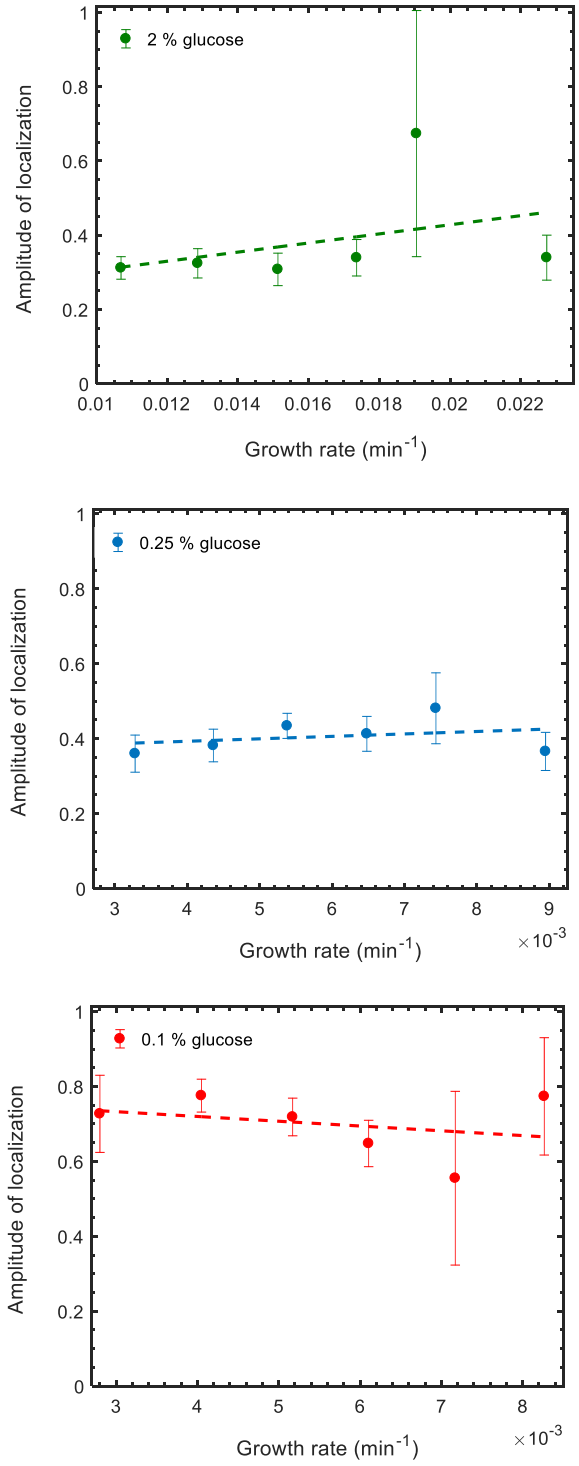


fig. S18. Correlation analysis between Msn2 localization features and cellular growth rate. Correlations in Msn2 localization frequency (**A**) and amplitude (**B**) with growth rate for different glucose concentrations. Growth rate of single cells were first calculated and data was binned into windows of fixed width. Error bars denote standard error of means for each such bin. Correlations in localization frequency emerge as stress intensity is increased with the correlation coefficients equal to 0.829 (p = 0.04) and 0.861 (p = 0.02) in 0.25% and 0.1% glucose, respectively. No significant correlation for localization frequency was observed for 2% glucose (p = 0.19). No significant

correlation for localization amplitude was observed for any glucose concentration (p = 0.46, p = 0.57, p = 0.55 for 2% glucose, 0.25% glucose and 0.1% glucose respectively).

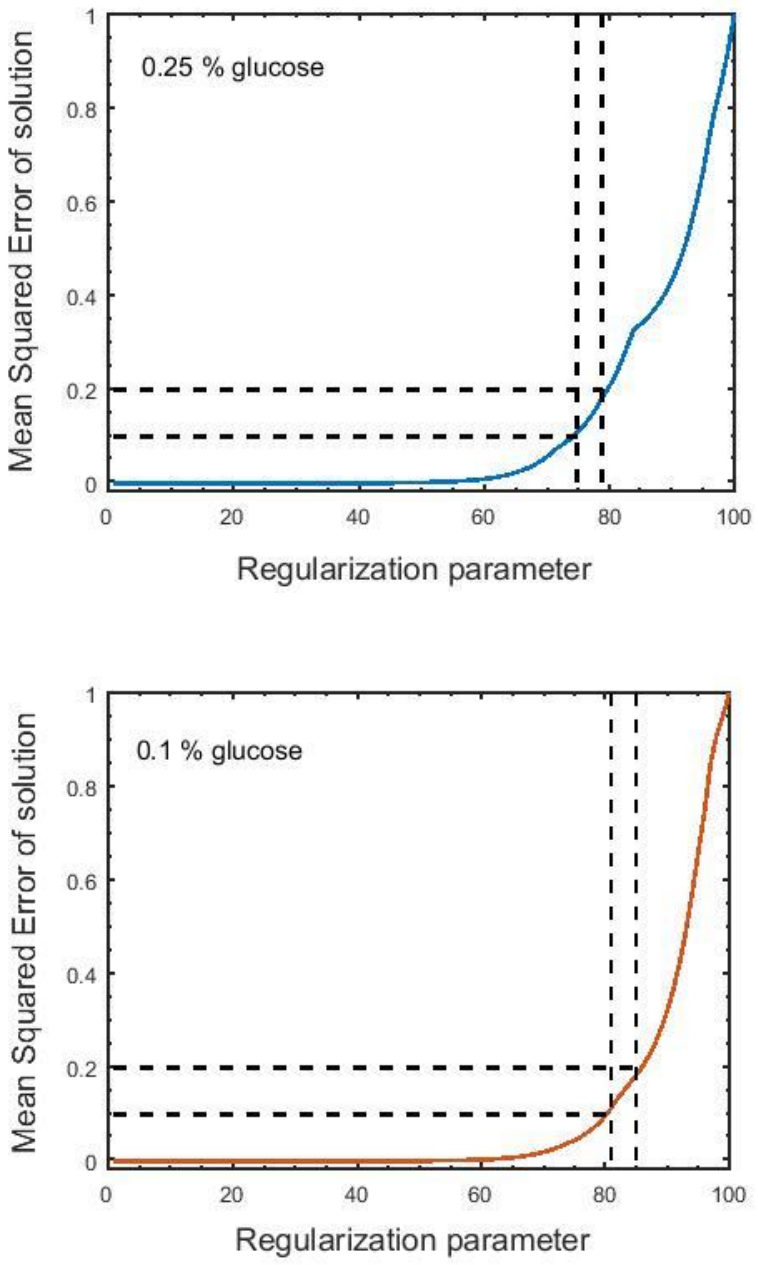
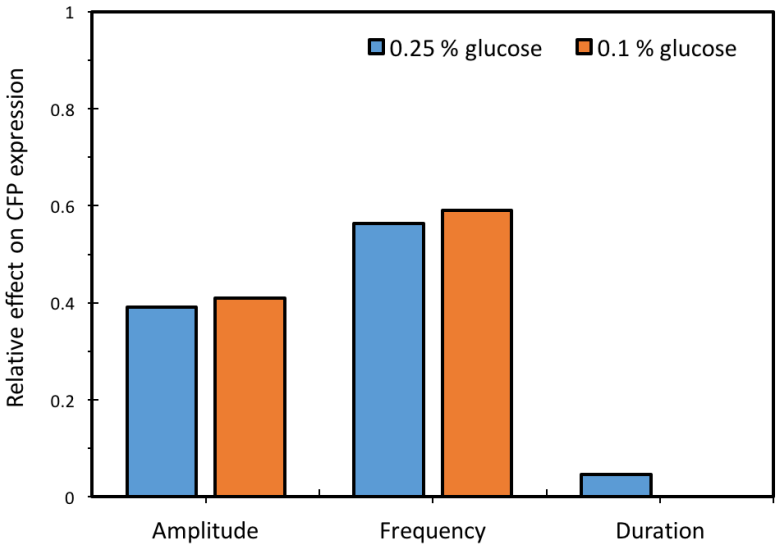


fig. S19. Mean squared error for Lasso solution as a function of different values of regularization or shrinkage parameter. Top panel: 0.25% glucose. Bottom panel: 0.1% glucose.

A



B

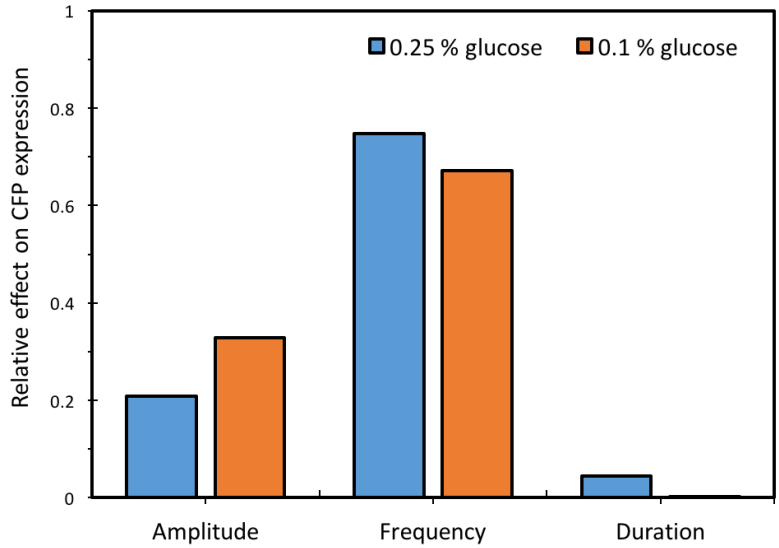


fig. S20. Lasso analysis is not sensitive to the choice of threshold used for Msn2 nuclear localization quantification. Single cell trajectories of Msn2 localization were thresholded either at 0.8 standard deviation (**A**) or at 1.2 standard deviation (**B**) of the mean of localization traces from 2% glucose, and Lasso was implemented for each case to ascertain which features of Msn2 influence CFP expression. As in the case for the original threshold using 1 standard deviation (shown in Fig. 8B), duration does not affect CFP, while amplitude and frequency influences CFP expression.

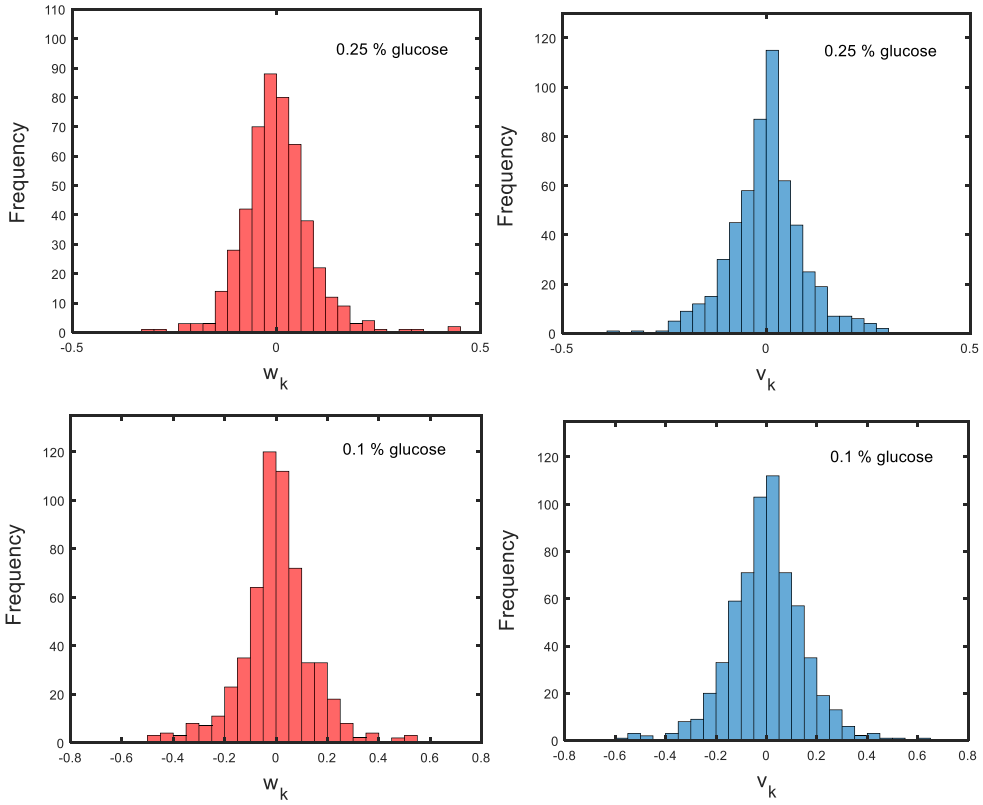


fig. S21. System and measurement noise for different stress environments. System noise w_k and measurement noise v_k in 0.25% (top) and 0.1% (bottom) glucose follows a Gaussian distribution with zero mean.

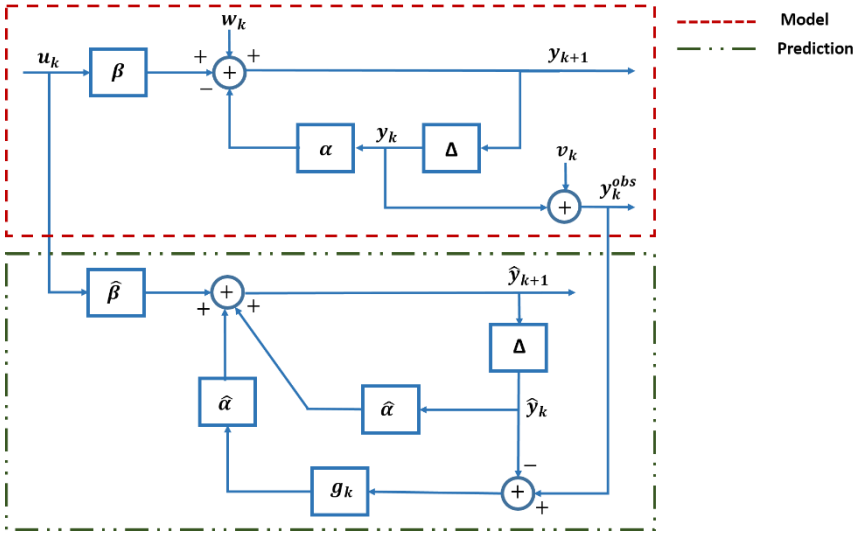


fig. S22. Block diagram of system identification and prediction steps. Red dashed area denotes the block diagram for the linear state-space model, obtained by combining the system model (eq. [5]) and the observation model (eq. [6]). Description and mathematical equations for the state-space model is given in the section called System identification: discrete linear dynamical system. Green dashed area denotes the block diagram for the linear prediction using Kalman filter (eq. [7]). Description and mathematical equations of prediction is given in the section called Model prediction: Kalman filter.

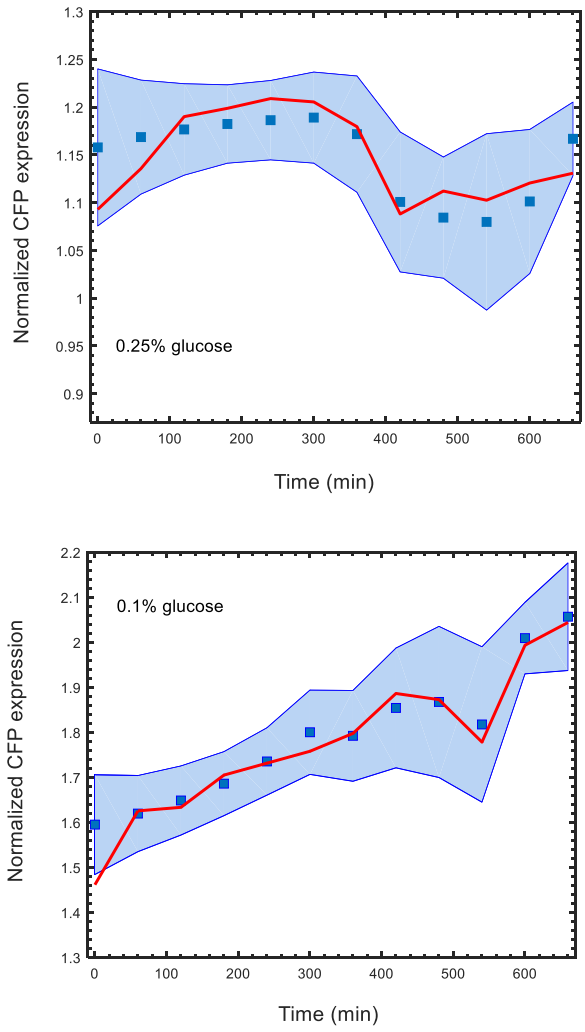


fig. S23. Predicting CFP expression levels using cross-validation. Prediction results for CFP expression for 0.25% (top) and 0.1% (bottom) glucose experiments using cross validation. CFP levels were measured in single cells at every 60 minutes. CFP expression was then normalized by the average CFP expression obtained from all cells grown in the 2% glucose condition. The normalized single-cell CFP time traces were next divided into training set and test set for the purpose of cross-validation. The training set was constructed by randomly selecting Msn2-YFP nuclear localization and corresponding CFP time series of 75% of the total number of cells. The test set comprised the Msn2-YFP nuclear localization and corresponding CFP time series of the remaining 25% of all cells. Cross validation was performed by learning the parameters of the linear state-space model from the randomly-chosen training set data and CFP expression was then predicted on the corresponding test set data (an example is shown above for each glucose concentration). Blue square at a given time point represents the average CFP expression across all cells in that test-set measured at that time point. Blue shaded area denotes the standard error of mean, extended between 60-min intervals as guide to the eye. The red points denote the CFP expression predicted by the model (at every 60 minutes) and averaged across the cells of the test set (red lines connecting red points during the 60-min intervals are guide to the eye). Prediction using cross validation is discussed in the section called “Robustness of identification and prediction: cross-validation”.

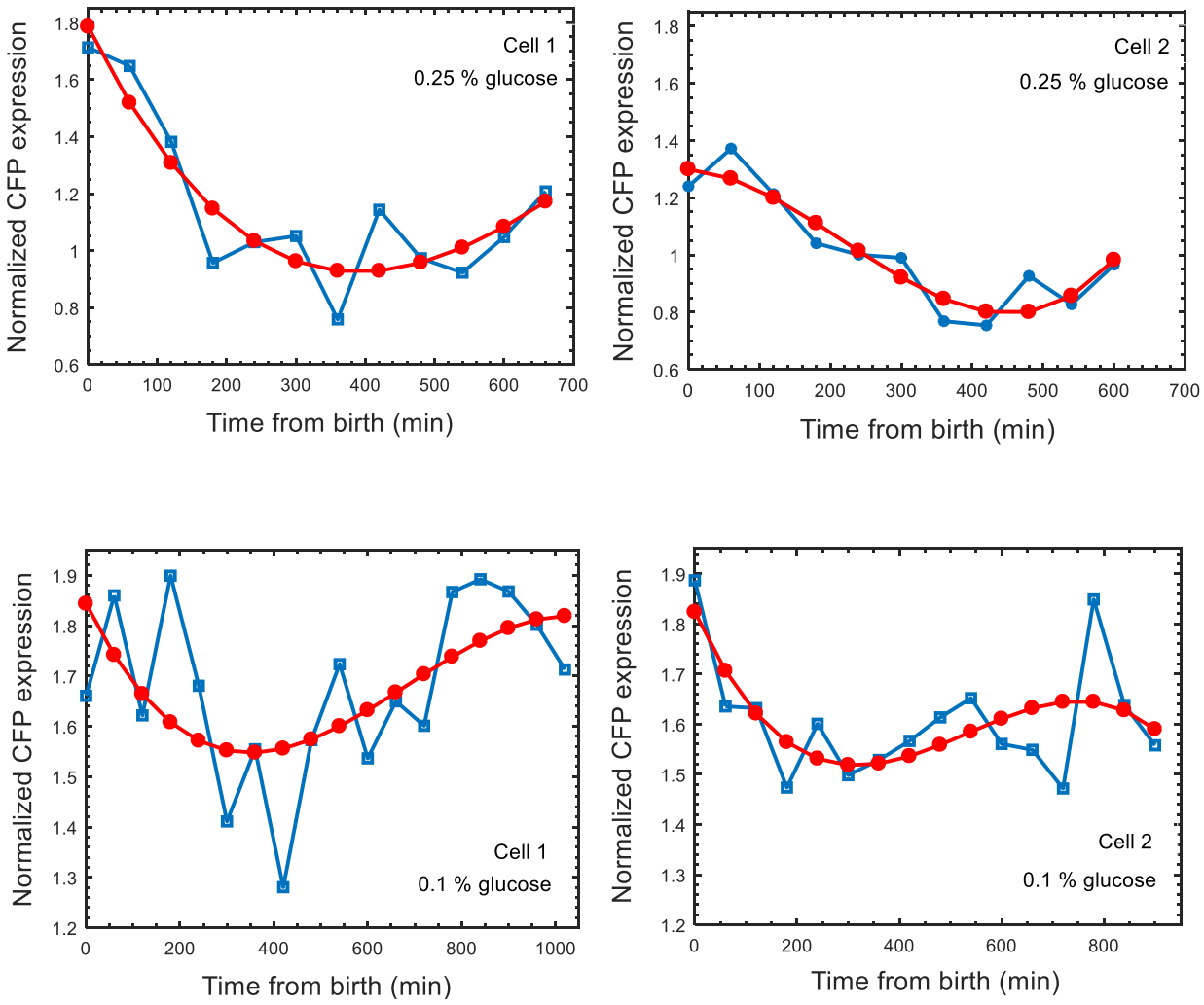


fig. S24. Sample polynomial fits for single-cell CFP trajectories. Blue squares denote CFP expression trajectories, measured at every 60 minutes and normalized by the average CFP expression obtained from all cells in 2% glucose, in two sample cells grown in the 0.25% (top) or 0.1% glucose (bottom) conditions. Blue squares are connected by solid blue lines as a guide to eye. A third degree polynomial function $c_o + c_1 t + c_2 t^2 + c_3 t^3$, where t denotes time, was fit to the CFP trajectories and the resultant fit is shown in red. The goal of the fitting process was to obtain a specific functional form that could be used as a smooth representation of the data. A third degree polynomial was the lowest-order polynomial with reasonably good fitting ability on the data.

II. SUPPLEMENTARY TABLES

table S1. Population doubling times of cells in different glucose concentrations obtained from fig. S2 (A to C). Experiments involved acquiring both bright field and fluorescence images (five independent experiments for each glucose concentration) and population doubling times were calculated for each such experiment. Population doubling times obtained from fits in fig. S2A-C in each experiment has been averaged over 5 independent experiments for each glucose concentration. Errors represent standard error of mean (N=5) for 5 independent experiments for each glucose concentration.

2% glucose	0.25% glucose	0.1% glucose
135.25 $\pm$ 6.85 (min)	158.48 $\pm$ 1.05 (min)	169.48 $\pm$ 10.06 (min)

table S2. Population doubling times calculated from OD_{600} measurements of cells grown in batch, using a shaker-incubator. Two replicate experiments, each comprising multiple OD_{600} measurements, were performed for each glucose concentration. Errors represent standard error of mean (N=2) for two replicates in each glucose concentration. The r-square values for goodness of fit determination for the two replicate experiments were: 0.9971 and 0.9992 for 2% glucose, 0.9960 and 0.9937 for 0.25% glucose, and 0.9975 and 0.9940 for 0.1% glucose.

2% glucose	0.25% glucose	0.1% glucose
88.47 $\pm$ 4.31 (min)	118.76 $\pm$ 1.26 (min)	123.35 $\pm$ 8.84 (min)

table S3. Values of parameters obtained from fitting data in fig. S6 to a sigmoidal function.

a_1	0.2247
a_2	0.2009
a_3	233.2659
a_4	0.0701

table S4. The *P* values comparing duration of Msn2 nuclear localization between 2 and 0.25% glucose, as well as 2 and 0.1% glucose across all threshold levels. Data in fig. S11 was used to compute the p-values. Unless otherwise specified, hypothesis testing has been performed using Welch's t-test for all figures, except for Fig. 6, Fig. 7, fig. S11, fig. S12 and fig. S15 where Mann-Whitney U test was performed since the default assumption of normality of dataset did not hold in these cases.

Threshold, k	p-value comparing 2% and 0.25% glucose	p-value comparing 2% and 0.1% glucose
0.6	0.0321	0.000231
0.8	0.000643	0.00000035
1	0.0022	0.00000232
1.2	0.0020	0.00000217
1.4	0.0043	0.00000565
1.6	0.0197	0.0000137

table S5. The *P* values comparing amplitude (A), frequency (B), and duration (C) of Msn2 nuclear localization between 2 and 0.25% glucose, as well as 2 and 0.1% glucose across different cell generations. Data in Fig. 4 was used to compute the p-values.

A

Generation number	p-value comparing 2% and 0.25% glucose	p-value comparing 2% and 0.1% glucose
0	0.0338	7.22×10^{-9}
1	0.0088	2.35×10^{-16}
2	0.2	2.75×10^{-11}
3	0.469	1.63×10^{-6}

B

Generation number	p-value comparing 2% and 0.25% glucose	p-value comparing 2% and 0.1% glucose
0	0.0012	3.1×10^{-5}
1	0.0246	1.47×10^{-4}
2	0.0571	6.78×10^{-7}
3	0.3742	2.70×10^{-5}

C

Generation number	p-value comparing 2% and 0.25% glucose	p-value comparing 2% and 0.1% glucose
0	0.0076	1.7×10^{-4}
1	0.0047	2.55×10^{-5}
2	0.4741	0.0161
3	0.2881	0.0037

table S6. The *P* values obtained from Mann-Whitney *U* test. The p-values were obtained using the data in Fig. 6 (**A**), in fig. S12A (**B**), and in fig. S12B (**C**).

A

Features of Msn2 localization	Glucose concentration	p-value between M-D and GM-GD	p-value between M-D and Reference	p-value between GM-GD and Reference
Amplitude	2%	0.1527	1.4×10^{-9}	1.95×10^{-4}
	0.25%	0.0733	3.8×10^{-11}	0.0429
	0.1%	0.4917	2.14×10^{-10}	3.7×10^{-5}
Frequency	2%	0.333	2.77×10^{-9}	2.42×10^{-6}
	0.25%	0.468	6.83×10^{-23}	4.24×10^{-14}
	0.1%	0.645	4.87×10^{-28}	4.18×10^{-16}
Duration	2%	0.7054	1.7×10^{-5}	0.0369
	0.25%	0.0081	3.94×10^{-10}	0.0186
	0.1%	8.12×10^{-4}	6.01×10^{-16}	4.65×10^{-4}

B

Features of Msn2 localization	Glucose concentration	p-value between M-D and GM-GD	p-value between M-D and Reference	p-value between GM-GD and Reference
Amplitude	2%	0.0511	2.32×10^{-11}	0.4189
	0.25%	0.0152	9.12×10^{-12}	3.26×10^{-4}
	0.1%	0.9596	3.07×10^{-9}	6.7×10^{-6}
Frequency	2%	0.2482	1.48×10^{-7}	8.5×10^{-5}
	0.25%	0.3087	9.32×10^{-13}	0.0045
	0.1%	0.4025	6.44×10^{-27}	2.84×10^{-16}
Duration	2%	0.0953	4.53×10^{-5}	0.8431
	0.25%	0.006	6.87×10^{-9}	0.1326
	0.1%	0.1244	6.34×10^{-8}	0.0031

C

Features of Msn2 localization	Glucose concentration	p-value between M-D and GM-GD	p-value between M-D and Reference	p-value between GM-GD and Reference
Amplitude	2%	0.303	2.09×10^{-19}	1.62×10^{-20}
	0.25%	0.0772	1.02×10^{-10}	0.0015
	0.1%	0.179	6.75×10^{-17}	1.17×10^{-6}
Frequency	2%	0.5298	8.81×10^{-18}	9.47×10^{-16}
	0.25%	0.4797	1.63×10^{-24}	8.94×10^{-20}
	0.1%	0.6519	1.62×10^{-24}	4.45×10^{-16}
Duration	2%	0.7719	2.05×10^{-4}	0.0272
	0.25%	0.0281	1.04×10^{-8}	0.2532
	0.1%	0.358	7.99×10^{-9}	0.3766

table S7. The *P* values obtained from Mann-Whitney *U* test. The p-values were obtained using the data in fig. S15.

Features of Msn2 localization	Glucose concentration	p-value between Reference far and Reference near	p-value between Reference all and Reference far	p-value between Reference all and Reference near
Amplitude	2%	0.7844	0.9068	0.8686
	0.25%	0.0124	0.0276	0.4251
	0.1%	0.8770	0.8176	0.9584
Frequency	2%	0.1207	0.4883	0.1929
	0.25%	0.2845	0.3110	0.7416
	0.1%	0.1832	0.4666	0.3715
Duration	2%	0.4468	0.8231	0.5131
	0.25%	0.8899	0.8751	0.7184
	0.1%	0.1199	0.3302	0.4003

table S8. The *P* values obtained from Mann-Whitney *U* test. The p-values were obtained using the data in Fig. 7 (**A**), in fig. S12C (**B**), and in fig. S12D (**C**).

A

p-value for similarity in Msn2 spikes between cell pairs	2% glucose	0.25 % glucose	0.1% glucose
p-value between M-D and GM-GD	0.452	0.49	0.925
p-value between M-D and GGM-GGD	0.157	0.234	0.46
p-value between MD and Reference	1.66×10^{-5}	1.32×10^{-8}	5.54×10^{-15}
p-value between GM-GD and GGM-GGD	0.389	0.629	0.405
p-value between GM-GD and Reference	1.7×10^{-5}	3.31×10^{-8}	2.18×10^{-13}
p-value between GGM-GGD and Reference	6.89×10^{-4}	1.38×10^{-5}	1.95×10^{-8}

B

p-value between cell pairs	2% glucose	0.25 % glucose	0.1% glucose
p-value between M-D and GM-GD	0.562	0.067	0.753
p-value between M-D and GGM-GGD	0.120	0.11	0.878
p-value between MD and Reference	7.32×10^{-11}	1.88×10^{-13}	3.03×10^{-15}
p-value between GM-GD and GGM-GGD	0.341	0.828	0.611
p-value between GM-GD and Reference	5.74×10^{-10}	1.88×10^{-13}	8.02×10^{-13}
p-value between GGM-GGD and Reference	3.6×10^{-7}	2.99×10^{-9}	1.64×10^{-8}

C

p-value for similarity in Msn2 spikes between cell pairs	2% glucose	0.25 % glucose	0.1% glucose
p-value between M-D and GM-GD	0.687	0.343	0.533
p-value between M-D and GGM-GGD	0.812	0.248	0.215
p-value between MD and Reference	2.85×10^{-10}	1.41×10^{-12}	1.19×10^{-15}
p-value between GM-GD and GGM-GGD	0.939	0.655	0.487
p-value between GM-GD and Reference	5.85×10^{-9}	6.25×10^{-12}	3.52×10^{-14}
p-value between GGM-GGD and Reference	3.29×10^{-5}	2.31×10^{-8}	1.58×10^{-9}

table S9. Parameter values extracted from the linear state-space model's application to the data obtained from 0.25 and 0.1% glucose experiments. The negative values of the parameter $\hat{\beta}_2$ can be explained by that the frequency and amplitude are not truly independent features affecting the downstream gene expression, based on the application of the model. A cell experiencing a high nuclear-localization amplitude might be expected to take more time to de-localize that high amplitude, leading to a high number of above-threshold localization events per unit time.

Parameter values	0.25% glucose	0.1% glucose
$\hat{\alpha}$	0.994	1.005
$\hat{\beta}_1$	0.025	0.030
$\hat{\beta}_2$	-0.059	-0.019
$\bar{y}_0$	1.109	1.495
σ_0^2	0.053	0.122
σ_w^2	0.008	0.020
σ_v^2	0.008	0.022

III. SUPPLEMENTARY MATERIALS and METHODS

Microfluidics and time lapse microscopy

A customized microfluidic chip was designed using Layout Editor software and built using standard techniques of plasma etching followed by replica molding with polydimethylsiloxane (PDMS) (*42,43*) (fig. S1). Each microfluidic chip consisted of three flow channels. Each flow channel had a length of 10 mm, width of 1.9 mm and a height of approximately 4 μm. By designing the height of a channel to have the same height as an average haploid cell, we ensured cells were trapped inside the channel and grew in a mono layer. Further, the entire channel was populated with circular support structures, of radius 2 μm each, and arranged in multiple blocks, with a block to block spacing of 40 μm. The circular structures were spaced at a center-to-center distance of 14 μm from each other within each block.

A PDMS based microfluidic chip was made using replica molding by mixing PDMS curing agent and PDMS elastomer in a ratio of 1:10 for 5 minutes. The mixture was poured on a silicon wafer, and desiccated in vacuum for 4 hours until no air bubble, resulting from the mixing process, remained. It was then cured at 150^0C for 20 min so that the PDMS mixture solidifies. The cured PDMS was then allowed to cool down. After cooling, the area of the PDMS containing the microfluidic design (here the area containing three flow channels) was peeled off the wafer using a razor blade. Inlet and outlet holes were punched on each flow channel. The punched PDMS chip was then bonded to a custom made and pre-cleaned cover glass using air plasma activation. Finally, tygon tubing was inserted into the inlet and outlet holes of each channel.

The experimental protocol followed to acquire a time lapse movie, using PDMS based microfluidic chip is detailed in a step-by-step fashion below. The strain MC01 was used for all experiments in this study.

i. Yeast cells were grown overnight at 30^o C in a shaking incubator, for 18 hours in low fluorescence synthetic drop out medium with amino acids and desired concentration of glucose (either 2%, 0.25%, or 0.1% glucose), to a final OD_{600nm} of 0.1.

ii. Exponentially growing cells were then diluted 200 fold to make the cell density suitable for being loaded into the microfluidic chip.

iii. Before loading the cells, a solution of concanavalin A (1.5mg/ml), $CaCl_2$ (5mM), $MnCl_2$ (5mM) with pH between 6-7 was prepared. Each channel in the microfluidic chip was then primed by flowing 25 μl of this solution from inlet of the chip at 5ul/min, using a 5 ml syringe. The syringe was connected to the chip through the inlet tubing and driven by a syringe pump. Priming with the solution of conA, a lectin that binds to the saccharides on the surface of a cell, further facilitated long term retention of cells which have already been trapped by the height of the channel.

iv. The conA treated chip was then incubated at room temperature for 20-30 minutes, after which it was mounted on a wide field inverted Nikon Microscope coupled with EMCCD camera. To prevent shifting of the chip in the stage throughout the experiment, it was securely taped to the stage mount.

v. Diluted and exponentially growing cells from step (ii), were then loaded into the channel at 10 ul/min for 10-12 minutes. The stage of the microscope was then moved around to locate a position in the channel where only a single cell was trapped in a field of view. Care was taken to avoid overloading and hence clogging of the inlet.

vi. Once a stage position having a single cell was located, media flow was begun and maintained at a constant flow rate of 1.5 µl/min over the entire timespan of the experiment.

vii. A 60X objective (N.A 1.4, oil Ph3, Plan-Apochromat) acquired 16-bit images of a time-lapse movie. Each movie started with a single ancestor cell and continued until the starting cell grew into a cluster of cells (confined within a monolayer) over a timespan of 15-18 hours.

viii. One microfluidic chip comprising three channels was used per experiment, and one field of view (or stage location) was acquired for each channel in that experiment by an Andor DU-897 camera. Bright field images were acquired every 2.5 min for a z-stack of 3 z-planes which consisted of the focal plane as well as the planes 1.4 µm above and below it. To measure localization, images in the YFP channel were acquired every 2.5 min, also for a z-stack comprising the focal plane $\pm$ 1.4 µm. To measure Msn2 dependent gene expression, CFP images were acquired every 60 minutes at only the focal plane (i.e. using a single z plane). The acquisition routine was performed by NIS elements software coupled to the microscope. The microscope moves between different stage positions or fields of view and maintains the focal plane, using Nikon Perfect Focus. The exposure times for YFP, CFP and mRFP (mRFP used only for fig. S5-S8 and fig. S10 for purpose of validating the accuracy of nuclear estimation algorithm) were 300 ms, 175 ms, 300 ms respectively. These exposure times were chosen such that a modest signal to noise ratio was achieved, while avoiding light induced photo toxicity in cells to ensure healthy cell growth and division. The sola lamp model used is SE-5LCR-SA. Filter sets used were optimized for the detection of YFP (Chroma 49003-ET-EYFP, ET500/20x (EX), ET535/30m (EM)), CFP (Chroma 49001-ET-ECFP, ET436/20x (EX), ET480/40m (EM)) and mRFP (Chroma 49008-ET-mCherry, ET560/40x (EX), ET630/75m (EM)).

Cellular growth rates and doubling times

Population growth rate (γ) was measured by counting the number of cells at several time points (t) over the whole time span of an experiment. The data was then fit to an exponential function: $N_0 e^{\gamma t}$, where N_0 is the initial number of cells at time $t = t_0$. Population doubling time (t_d) was then calculated from growth rate as: $t_d = \frac{\ln 2}{\gamma}$. Semi-logarithm plots of exponential growth of cell populations in 5 independent experiments for each glucose concentration and their corresponding linear fits to the logarithmic y-axis have been shown in fig. S2A-C. The population doubling time of cells obtained from the fits have been averaged over 5 independent experiments for each glucose concentration and has been shown in table S1.

Population doubling times were also calculated from OD_{600} measurements of cells grown in batch, using a shaker-incubator. Two replicate experiments were performed for each glucose concentration. Each replicate experiment involved multiple measurements of OD_{600} in the logarithmic growth phase of the yeast cells. An exponential function was then fit to the OD_{600} measurements for each experiment to obtain the growth rate (γ) and doubling time was calculated from the growth rate. Doubling times obtained from independent exponential fits from two replicates for each glucose concentration were then averaged and has been shown in table S2.

Measurement of the division times of a single cell require determining the exact time of occurrence of cytokinesis, i.e. recording the time instances at which a daughter cell bud separated from its mother compartment to form two independent cell regions. Because a daughter cell could still be touching the mother even after the latter had undergone cell division, there is generally no clear visual cue to identify exact occurrence of cytokinesis from the bright field images. It is, however, observable by fluorescent labeling of septins which localize at the bud neck. To avoid phototoxicity in cells by imaging an additional florescent marker over long time periods, we instead measured the time interval between the beginning of two consecutive S-phases of a cell and quantified its doubling time in this manner. The first appearance of the bud served as a morphological cue to detect the start of the S-phase. The durations between two consecutive S-phases of each cell were recorded for each cell from its birth to the end of an experiment or movie. The resulting average doubling times across all generations of all cells in each of the three glucose environments were: 119.28$\pm$5.87min for 2% glucose, 122.84$\pm$2.21min for 0.25% glucose, 140.41$\pm$7.06min for 0.1% glucose, where errors are SEM from five independent experiments (movies). The reason for why these doubling-times are shorter compared to the doubling-times obtained from fig. S2 is the following: in fig. S2 we analyze how the number of single cells increase during each experiment and fit it to an exponential function. As we could not use a cytokinesis marker, the number of cells plotted on the y-axis of fig. S2 had to be determined conservatively (i.e. even when there was a daughter cell which looked like has already completed the cytokinesis, we waited a few more frames to really call it an independent cell and increase the cell-count in the population by +1). As a result, the growth rates obtained from fig. S2 slightly underestimate the speed of doubling, while the above single-cell-based average doubling times are more accurate. However, based on the goodness of fits, fig. S2 still provides useful information by showing that the doubling speeds of the cell populations in each movie are similar between the early parts of the movie and its later parts. As a final remark, compared to the doubling-times obtained from batch-culture experiments, the longer doubling times measured by watching the cells in microscopy experiments were not unexpected due to the bright field and fluorescence light exposures on cells during the movies.

Analysis of fluorescent images in single cells

Quantification of CFP expression in single cells

CFP signal in a single cell was quantified similar to that described previously *(9)*. CFP reporter expression of each cell was first obtained by quantifying the average pixel intensity of the whole cell, as CFP is distributed uniformly inside the whole cell. Average CFP intensity of each cell

was then subtracted from the average background intensity calculated from an arbitrary area in the CFP image without cells.

Quantification of Msn2 localization in single cells from YFP images

Several metrics have been employed in literature to measure the nuclear localization of Msn2 in a cell. The most direct method of them involves tagging a nuclear protein of the cell to a fluorescent marker and then quantifying the average pixel intensities of the fluorescent channel to which Msn2 is tagged in the cell. In absence of nuclear marker, localization have been defined from averaging brightest 10 or 15 pixel intensity of the cell (*8,28*). Despite the simple nature of these metrics, they can potentially give rise to erroneous estimates of nuclear intensity, as some bright pixels in the cell used in quantification may not actually belong to the nucleus. This could particularly pose a problem when the strength of the stress stimulus is weak, and Msn2 is predominantly cytoplasmic. Figure S5 gives an example of such a scenario. We developed an algorithm to estimate localization of Msn2, wherein we integrate the information of the spatial location of the brighter pixels of the cell to compute a new localization metric, in contrast to solely focusing on their intensity values. This increases the chance that the chosen set of bright pixels belong to the nucleus of the cell. We validated that our metric approximates the true nuclear intensity significantly better than previously employed metrics. Our algorithm is described as follows:

i. First, we experimentally characterized how nuclear area of a cell varies with its corresponding cellular area. It is known that cells grown in glucose-limited media have smaller sizes compared to the size of cells grown in the optimal glucose concentration *(44,45*). Studies (*46*) have shown that the nuclear area is proportional to cellular area with smaller cells displaying a relatively larger nuclear-to-cellular-area ratio compared to the ratio displayed in larger cells. Therefore, to prevent a bias in estimation of Msn2 nuclear localization in cells of different sizes, we first performed a separate experiment to calculate the ratio of nuclear to cellular area from a sample of 100 cells.

ii. Cells had mRFP tagged to a nuclear protein Nhp6a, in addition to Msn2 being tagged to YFP. Bright field, YFP, and mRFP images were acquired for a z-stack of 3 z-planes comprising the focal plane and the planes 1.4 µm above and below it. Quantification of cellular and nuclear area was performed similar to that described previously *(46*). Cellular area was measured based on the maximum of the total number of cellular pixels from the three z-planes. Nuclear area was measured by taking the maximum number of nuclear pixels from the three z-planes taken on the mRFP data. Data was binned and fitted to the following sigmoidal function (fig. S6)

$$y = a_1 + \frac{a_2 - a_1}{1 + 10^{(a_3 - x) * a_4}}$$

where, x is the area of a cell and y is the ratio of nuclear area to cell area. The values of parameters a_1, a_2, a_3, a_4 are given in table S3. Thus, for every new cellular area

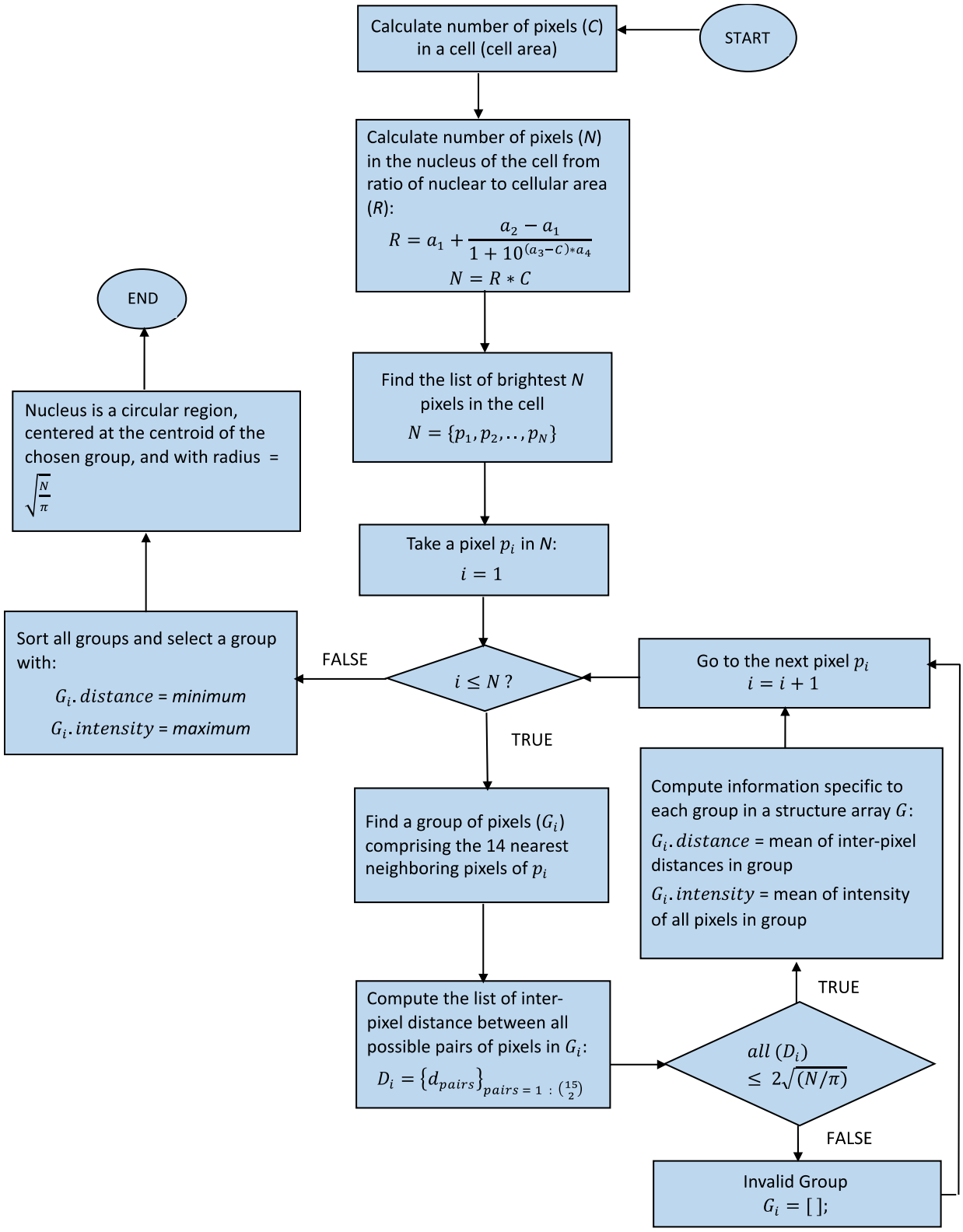
measured, the corresponding nuclear area is calculated from the above sigmoidal function.

iii. Next, for a cell with nuclear area of *N* pixels, for each of the *N* pixels, a group of 15 pixels was chosen (composed of the pixel plus its 14 nearest neighbors). For each group of 15 pixels, two parameters were calculated: a) the average YFP pixel intensity of the group, and b) the mean of inter-pixel distances, as a measure of connectedness of that group. Groups with mean inter-pixel distances higher than the estimated nuclear diameter were not considered as valid groups and rejected by the algorithm. All valid groups of pixels were then sorted to select the group having the minimum mean inter-pixel distance and the maximum average YFP pixel intensity. We also tried two other distance metrics: a) minimum variance of inter-pixel distances, and b) minimum diameter, with diameter defined as the maximum distance between a pair of pixels in a group. All three methods gave similar results.

iv. Nucleus was then estimated as a circular area with its center located at the centroid of the group of 15 pixels, which had minimum mean inter-pixel distance and the maximum average YFP pixel intensity. The diameter of that circular area corresponded to the estimated nuclear diameter equal to $2\sqrt{\frac{N}{\pi}}$.

v. Finally, Msn2 localization in a single cell was defined as: average YFP pixel intensity in the estimated nucleus of a cell divided by average YFP pixel intensity in the whole cell.

A detailed flowchart of the entire algorithm used for estimating nuclear localization of Msn2 from the whole cell Msn2-YFP snapshots is given below:



To gauge the accuracy of our estimation, we collected a sample of 100 cells grown in a low-stress environment (0.5% glucose), and measured Msn2-YFP and NhP6a-mRFP levels. This

sample involved cells from all phases of cell cycle as well as all possible cell sizes. We calculated the true localization of Msn2 for each cell using the direct metric: average YFP pixel intensity in true nucleus (say loc^i_{true}, for cell i) detectable by NhP6a-mRFP. We compared the true localization value of each cell with the localization estimated from the same cell using the above algorithm (say loc^i_{est}, for cell i). Estimation error (say err^i, for cell i) of the algorithm in each cell was quantified by

$$error^i = \frac{|\ loc^i_{true} - loc^i_{est}\ |}{loc^i_{true}}$$

Histogram of estimation error of YFP nuclear localization in single cells using our algorithm and a conventional method (based on quantifying average of the top 15 YFP pixels in a cell) is shown in fig. S7A. The average estimation error was quantified by averaging the estimation error over all cells

$$error = \frac{1}{100}\sum_{i=1}^{100} error^i$$

The average estimation error was found to be only 8.46 $\pm$ 1.07% (compared to 26.4 $\pm$ 1.88% when taking the average of the top 15 YFP pixel intensities in cells). A comparison between a cell's estimated nuclear Msn2 intensity and its corresponding true intensity is given in fig. S7B. The slope of the linear fit is 1.0609 using our algorithm, and is very close to the slope of the best-possible linear fit. Further, a quality-control for estimation of nuclear area using the algorithm vs that obtained from using the nuclear marker was also performed. To do so, Msn2 nuclear localization in single cells were first binned into different localization intensities. For each such bin, the fractional overlap between true nuclear area (obtained from using the nuclear marker), and estimated nuclear area (obtained from either our algorithm or methods in literature based on averaging only brightest 15 pixels in whole cell) were calculated and shown in fig. S7C. The fractional overlap or the ability of the algorithm to capture true nuclear pixels, increases with the increase in strength of Msn2 nuclear localization signal in cells. The algorithm also showed better performance compared to methods in literature for all levels/strengths of localization signals. Finally, we also ensured that the estimation of nuclear intensity is not sensitive to the variability in the ratio of nuclear to cellular area, as observed in fig. S8. For cellular area of each of the 100 cells in fig S6, the corresponding ratio of nuclear to cellular area was randomly given a number sampled from a uniform distribution with an interval ranging from $\pm$ 1 standard deviation of mean ratio. An estimation error for each of the 100 cells were then calculated using the above equation for $error$. This process was repeated 100 times, and 100 estimation errors were obtained in each run. The estimation error averaged over 100 such runs was found to be 8.52% $\pm$ 0.18% (fig. S8).

Threshold determination for Msn2 nuclear localization quantification in single cells

We validated that the threshold used for burst quantification of Msn2 in single cells, is a reasonable choice to determine occurrence of nuclear localization events. To validate this, we performed two independent experiments where cells in a microfluidic chip were either induced with a stress of 0.5% glucose, or with 2% glucose. Bright Field, Msn2-YFP, and NhP6a-mRFP images were acquired in each channel for a z-stack of 3 z-planes, comprising the focal plane and 1.4 µm above and below it. Msn2 localization value of each cell in each experiment was then calculated by: average YFP pixel intensity in nucleus divided by average YFP pixel intensity in whole cell. Figure S10 shows a plot of proportion of localization values of 100 cells which are below a given threshold, as a function of the different thresholds obtained from data in 2% glucose. Detectable and similar difference in localization values between the two concentrations appeared for a threshold between 0 and 2 standard deviation of the mean of localization values of all cells in 2% glucose condition. Thus one standard deviation above the mean of the Msn2 localization values obtained for cells from 2% glucose, is a reasonable choice of threshold to determine Msn2 nuclear localization.

In all lineage time-lapse experiments, Msn2 localization trajectory for every cell was thresholded at 1 standard deviation of the mean of the localization traces obtained from lineage experiments in 2% glucose. Five replicates of 2% glucose experiments were used for determination of this threshold. We also considered other thresholds for burst quantification, ranging between 0.6 and 1.6 standard deviation. The results of analysis of Msn2 localization trajectories in the time lapse lineage experiments were found to be robust to the choice of a specific threshold (fig. S11, S12).

Feature Selection using Lasso

Lasso is a shrinkage based regression analysis technique, popularly employed in machine learning, to ascertain the importance of or select for predictor variables, responsible for generating the observed outcome of a statistical process (*47-50*). It operates by shrinking the values of the regression coefficients through introducing a penalty on the magnitude of the coefficients. Lasso coefficients are obtained by the minimization of the following problem

$$\hat{\beta}^{lasso} = \underset{\beta}{arg\min}\left\{\frac{1}{2N}\sum_{i=1}^{N}(y_i - \beta_0 - x_i^T\beta)^2 + \lambda\sum_{j=1}^{p}|\beta_j|\right\} \quad [1]$$

y_i: observed response at i^{th} instant.
x_i: vector of predictor variables at i^{th} instant.
N: number of observations.
β: a vector of regression coefficients having size p.
λ : a positive regularization parameter that controls the amount of shrinkage.

As the value of λ is increased, the amount of penalty imposed on the coefficient sizes becomes larger, thus resulting in more and more coefficients being shrunk towards zero. By causing some of the coefficients to be exactly zero, a selection of a continuous subset of predictor

variables is generated. More mathematical details about Lasso for feature selection can be found in the literature (*51*).

In this framework, an observed transcription burst of CFP in a single cell could occur due to two factors: a) modulation of a dynamical feature of Msn2 signal, and b) noise due to inherent stochastic nature of transcription. For an individual cell, this can be generically expressed as

$$z_{t'} = g\left(u_t^{amp}, u_t^{dur}, u_t^{freq}\right) + \varepsilon \; ; \quad t' = t + d \qquad [2]$$

where, d: CFP maturation delay; $z_{t'}$: delayed burst of CFP observed at time t'; u_t^{amp}: amplitude of localization at time t; u_t^{dur}: duration of localization at time t; u_t^{freq}: frequency of localization at time t; ε = noise term approximating the stochasticity of transcription; $g(.)$ is a user defined functional form.

Since Msn2 and CFP snapshots were collected in the same cell after every 2.5 minutes and 60 minutes respectively, each cell's localization trajectory was first divided into windows of 60 minutes. A window size of 60 minute was chosen to keep the length of localization and CFP traces equal. In each such window of 60 minutes for a single cell, the average amplitude, duration and frequency of localization were quantified. Applying the above idea of Lasso specifically in context of our dataset and focusing only on the deterministic component of equation [2] then led to the representation of the following mathematical form for an individual cell

$$\hat{y}_w = \sum_{i=0}^{\tau} \alpha_i \, u_{w-i}^{amp} + \sum_{i=0}^{\tau} \beta_i \, u_{w-i}^{dur} + \sum_{i=0}^{\tau} \gamma_i \, u_{w-i}^{freq} \qquad [3]$$

where, u_w^{amp}: localization amplitude at time window w; u_w^{dur}: localization duration at time window w; u_w^{freq}: frequency of localization at time window w; $\hat{y}_w$: approximation of a burst of CFP observed at time window w, by the right hand side of equation [3]. α, β, and γ are regression coefficients governing relationship between CFP and amplitude, duration and frequency of Msn2 signal respectively. Further, we included the possibility that CFP at a time window can not only be activated by a signal feature quantified from that time window but also from two time windows preceding it. This is represented in parameter τ in equation [3], where τ=2 in this case. Incorporating this additional possibility allowed us to account for the delay between activation of CFP and the occurrence of a localization stimulus triggering it. Finally, we solved

$$\min_{\alpha,\beta,\gamma} \left\{ \frac{1}{2W} \sum_{w=1}^{W} (y_w - \hat{y}_w)^2 + \lambda \sum_{j=1}^{W} \left| \alpha_j + \beta_j + \gamma_j \right| \right\} \qquad [4]$$

where, W: total number of time windows in all cells taken together; λ: the regularization parameter,
y_w: CFP burst observed at time window w.

Lasso analysis method has been applied to data from experiments where cells have been subjected to stress, thus causing activation of CFP (i.e. experiments from 0.25% glucose and

0.1% glucose concentrations), and results have been shown in Fig. 8B. CFP time trace of single cells in a given glucose stress environment were normalized by CFP expression obtained from cells in stress-free 2% glucose concentration. The mean square error (MSE) of the solution to equation [4] for each regularization parameter λ has been shown in fig. S19. For each feature, all coefficients yielding 10% to 20% of the maximum MSE were then averaged to quantify the degree of dependence between that feature and CFP.

Minimal linear state-space model of gene expression

Given that frequency and amplitudes have been found to be the key regulators for gene expression, this section describes the identification of a simple model to understand how amplitude and frequency of localization actually brings about the regulation of gene expression. Given the input measurements and the observed outcomes, identification of a system generally starts with defining the simplest possible model and incrementally adding in complexity until a model, that best describes the systems dynamics between the input and the output, is obtained.

System identification: discrete linear dynamical system:

The model, that approximates the underlying gene expression phenomenon, provides a formal description of the relation between the key localization parameters and gene expression (here CFP expression). Although the underlying system is continuous-time, a discrete-time model is a practical representation of the system that would explain what is observed in discrete time and would facilitate time-series modeling for empirical measurements over time. We found that a most simple representation of the system, defined by a first order linear state space equation was adequate to describe the CFP dynamics observed in natural stress. The mathematical details of discrete-time linear dynamical system are described below.

Let the CFP expression at any discrete-time instant k be denoted as the true system state variable y_k . The temporal dynamics of a CFP trajectory, y_k, for an individual cell is described by

$$y_{k+1} = \alpha y_k + \beta u_k + w_k \qquad [5]$$

where, equation [5] is called the system model. The system state y_{k+1} (CFP expression) at any time instant $(k+1)$ is driven by a) the state y_k at past instant , b) the vector of inputs u_k (here localization amplitude and frequency) at past instant k, and c) system noise w_k at past instant k. The noise due to variability in gene expression across cells, which is also influenced by the cell-cell variability in input localization dynamics, is lumped into the variable w_k. Specifically, for any cell: $u_k \in \mathbb{R}^2$, $u_k = [u_k^{amp}, u_k^{freq}]$ denotes the vector of input features at k^{th} instant (k being the time window of interest); u_k^{amp} denotes the localization amplitude at k^{th} instant; u_k^{freq} is localization frequency at k^{th} instant; $y_k \in \mathbb{R}$, denotes true CFP expression at k^{th} instant (obtained from the polynomial approximation of noisy CFP trajectories, further details mentioned in later sections). Further, the initial value true CFP is assumed to follow a normal distribution

with mean $\bar{y}_0$ and variance σ_0^2 , i.e. $y_0 \sim N(\bar{y}_0, \sigma_0^2)$; $w_k \in \mathbb{R}, w_k \sim N(0, \sigma_w^2)$ is the system noise at k^{th} instant. Theoretically, this should be a normal distribution having mean 0 and variance σ_w^2, and driving CFP expression; $\alpha \in \mathbb{R}$, defines the system parameter measuring the degree of dependence between the past and current value of true CFP expression; $\beta \in \mathbb{R}^2$, $\beta = [\beta_1, \beta_2]$ denotes the input parameters describing the relationship between localization amplitude and true CFP (β_1) or the localization frequency and true CFP expression (β_2).

The system model [5] describes the linear dynamics of the CFP expression y_k and captures how it evolves with time depending on the localization inputs, amplitude u_k^{amp} and frequency u_k^{freq}, and the system noise w_k. Let y_k^{obs} denote the observed output of the system state y_k. The discrete-time linear observations y_k^{obs} are represented by

$$y_k^{obs} = y_k + v_k \qquad [6]$$

Equation $[6]$ is called the observation model, where the observed system output y_k^{obs} at k^{th} instant, is a measure of true system state y_k corrupted by measurement noise v_k at that instant. Specifically, for any cell: $y_k^{obs} \in \mathbb{R}$, is the observed CFP expression at k^{th} instant; $v_k \in \mathbb{R}, v_k \sim N(0, \sigma_v^2)$ is the measurement noise at k^{th} instant. Theoretically, this should be a normal distribution having mean 0 and variance σ_v^2. It represents the measurement error which may arise from technical device, low sampling rate of data collection and/or spare size of sample.

The system model [5] and the observation model [6] when combined together, represented the state-space model. The set of unknown parameters to be estimated were: $\hat{\alpha}$, $\hat{\beta}_1$, $\hat{\beta}_2$, $\bar{y}_0$, σ_0^2, σ_w^2, σ_v^2. The state-space model represented the CFP expression dynamics of all the cells in the study. However, the parameters of the model are stress concentration specific. For each stress concentration, single cell CFP time traces were fit by minimizing error from time traces of all cells (total 75 cells from 5 experiments in each glucose concentration), using the least square parameter estimation routine in MATLAB. Figure S21 shows that, as expected from theory, the noises w_k and v_k followed Gaussian distribution with mean very close to zero. The exact mean values for w_k and v_k in 0.25% glucose were 0.0026 and $6.9*10^{-17}$ respectively. The exact mean values for w_k and v_k in 0.1% glucose were 0.0034 and 0.0000342 respectively. The estimated parameters obtained from fitting under the two stress conditions are given in table S9.

Model prediction: Kalman filter:

Here, we demonstrated the goodness of the model identified in the previous section by analyzing its prediction performance. We predicted the true state y_k (CFP expression) given the inputs and the measurements. The system behavior of the Msn2 localization – CFP expression mechanism is characterized by: (a) a linear state dynamics model [5] with inputs and zero-mean Gaussian noise; and, (b) a linear observation model [6] with measurements mixed with zero-

mean Gaussian noise. For this identified linear and Gaussian model, the Kalman filter (*52-53*) is the optimal minimum mean-squared error (MMSE) estimator to perform the prediction task on the CFP expression y_k.

The predicted response $\hat{y}_{k+1}$ of the cells in each stress condition is given by the following Kalman filter equation, where all parameters in the equation have been obtained from model identification step

$$\hat{y}_{k+1} = \hat{\alpha}\hat{y}_k + \hat{\beta}u_k + \hat{\alpha}g_k(y_k^{obs} - \hat{y}_k) \quad [7]$$

The linear prediction equation [7] is similar to the system model [5] where the estimated parameters, $\hat{\alpha}$ and $\hat{\beta}$, are obtained from model identification step. In [5] the random system noise w_k drives the gene expression dynamics. The correction term $\hat{\alpha}g_k(y_k^{obs} - \hat{y}_k)$ of the Kalman filter [7] helps to track the random dynamics of the cell response due to the system noise. The gain g_k, called Kalman gain, is the design parameter for Kalman filter prediction of the gene expression using the identified model. The gain is computed iteratively in each time instant as

$$g_k = \hat{\sigma}_k^2\left(\hat{\sigma}_k^2 + \sigma_v^2\right)^{-1} \quad [8]$$

$$\hat{\sigma}_{k+1}^2 = \hat{\alpha}^2(1 - g_k)\hat{\sigma}_k^2 + \sigma_w^2 \quad [9]$$

where, $\hat{\sigma}_k^2 = \mathbb{E}[(\hat{y}_k - y_k)^2]$ denotes the prediction error variance. The system dynamics parameter $\hat{\alpha}$, the system and measurement noise variances σ_w^2 and σ_v^2 are obtained from the model identification step. The initial conditions of the cell response $\hat{y}_0$ and the prediction error variance $\hat{\sigma}_0^2$ for the Kalman filter [7] and gain computation steps [8]-[9] respectively, are

$$\hat{y}_0 = \bar{y}_0\,;\ \hat{\sigma}_0^2 = \sigma_0^2$$

where, the model parameters $\bar{y}_0$ and σ_0^2 are estimated during the system identification. Since, the gain g_k is dependent on the model parameters and independent of the actual input-outputs, they can be pre-computed and saved as prediction parameters. Thus, for the prediction of the cell-response using the optimal Kalman filter, equation [7] will provide the MMSE estimates $\hat{y}_{k+1}$. More details on this is given in a separate section below, named “Kalman filter and gain calculations for prediction of CFP expression”. The prediction results for two cases have been shown in Fig. 8C. These predictions of gene expression time traces in the presence of noisy cellular response to natural stress validates the goodness of the identified model. The overall model, combining the system identification and prediction steps, is represented in a block diagram given in fig. S22.

Robustness of identification and prediction: cross-validation:

Here we demonstrated the robustness and consistency of the system identification as well as the prediction capabilities of the identified model using random cross-validation. We randomly divided our dataset into a training dataset comprising the time traces of randomly chosen 75% of the cells, and a test dataset comprising the time-traces of the remaining 25% of the cells. Then we perform system identification on the training dataset, i.e., learn the parameters of the model. Then using these parameters, we predict the response of completely new cells from the test dataset and compare it with the true responses. We repeat the above process of cross-validation 500 times; each time dividing the dataset randomly into train dataset and test dataset, learning parameters on train dataset, predicting response on test dataset, and analyzing the prediction performance. One such instances of cross validation are shown in fig. S23. The prediction results for the CFP expression levels were reasonably strong for the test datasets shown for the two stress levels.

We also calculated the mean-squared-error (MSE) for each member of the cross-validation dataset (500 in total). For MSE calculations, we averaged the squared differences between the predicted and experimental normalized-CFP levels across all experimental time points (at every 60 minutes). The mean and standard deviation of MSE from 500 cross-validation datasets in the 0.25% glucose condition was $1.9*10^{-3}\pm0.8*10^{-3}$. The mean and standard deviation of 500 cross-validation datasets in the 0.1% glucose condition was $3.3*10^{-3}\pm1*10^{-3}$. These numbers are comparable to the MSE obtained when identification and prediction was performed in the same dataset ($1.3*10^{-3}$ and $2.8*10^{-3}$ for 0.25% glucose and 0.1% glucose experiments, respectively). This shows the robustness of our analysis and that the system identification is a good representation of the actual system contributing to CFP expression.

Preprocessing CFP time traces and polynomial fitting:

As for Lasso analysis, state space modeling was also done on data from experiments in 0.25% and 0.1% glucose limited stress conditions. CFP time trace in response to stress was normalized by CFP expression obtained from cells in 2% glucose. Noisy single cell CFP trajectories were smoothed by polynomial fitting. A third degree polynomial function represented by: $y(t) = c_o + c_1 t + c_2 t^2 + c_3 t^3$ was fit to best approximate these time traces. Some sample fits for CFP trajectories are shown in fig. S24. The residuals of the fit, calculated from the mean squared error between the fit and true values, are uncorrelated and follow a normal distribution. A third degree polynomial approximation showed better quality of fits than lower order polynomials of degree one and two.

Kalman filter and gain calculations for predicting CFP expression

First, a brief overview of Kalman filter has been provided. Consider a dynamic state $x_k \in \mathbb{R}^n$ varying with discrete-time k. The linear dynamics of this stochastic process x_k is

$$x_{k+1} = Ax_k + Bu_k + w_k \quad [10]$$

where, $A \in \mathbb{R}^{n\,x\,n}$ is the system matrix, $u_k \in \mathbb{R}^p$ is the system input, $B \in \mathbb{R}^{n\,x\,p}$ is the input matrix and $w_k \in \mathbb{R}^n$ is the system noise. The measurement $y_k \in \mathbb{R}^m$ of the state follow a linear observation model

$$y_k = Cx_k + v_k \quad [11]$$

where, $C \in \mathbb{R}^{m\,x\,n}$ is the observation matrix and $v_k \in \mathbb{R}^m$ is the observation noise. The system and observation noise are independent of each other and are zero-mean Gaussian, i.e., $w_k \sim N(0, \Sigma_w)$ and $v_k \sim N(0, \Sigma_v)$. The initial condition of the state is random and follows Gaussian, $x_0 \sim N(\bar{x}_0\,, \Sigma_0)$. The matrices Σ_w and Σ_v are noise covariance matrices whereas Σ_0 is the covariance of the initial condition.
Let $\hat{x}_{k+1}$ denote the predicted estimate of the state x_{k+1} at time instant k. Then the prediction error e_k and prediction error covariance $\hat{\Sigma}_k$ are

$$e_k = x_k - \hat{x}_k \quad [12]$$

$$\hat{\Sigma}_k = \mathbb{E}[(e_k - \bar{e}_k)\,(e_k - \bar{e}_k)^T] \quad [13]$$

where, $\bar{e}_k = \mathbb{E}[e_k]$ is the bias of prediction. The Kalman filter provides unbiased and minimum mean squared error (MMSE) optimal prediction estimates $\hat{x}_{k+1}$. The algorithm is formally stated below:

Algorithm 1: Kalman filter:

Input: Model parameters $A, B, C, \Sigma_w, \Sigma_v, \bar{x}_0, \Sigma_0$.
Initialize: $\hat{x}_0 = \bar{x}_0$, $\hat{\Sigma}_0 = \Sigma_0$.
for $k = 0\ to\ N-1$ **do**

$$G_k = \hat{\Sigma}_k C^T \left(C\hat{\Sigma}_k C^T + \Sigma_v\right)^{-1} \quad [14]$$

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + AG_k(y_k - C\hat{x}_k) \quad [15]$$

$$\hat{\Sigma}_{k+1} = A(I - G_k)\hat{\Sigma}_k A^T + \Sigma_w \quad [16]$$

end for

---

where, $G_k \in \mathbb{R}^{n\,x\,m}$ is the Kalman gain matrix that incorporates the correction term in to the state prediction updates and N is the total number of time points.

The discrete-time linear state-space model of the dynamics of CFP expression that we have developed, is described in Supplementary Materials and Methods. In this model

$$n = 1, \qquad m = 1, \qquad p = 2$$

$$x_k = y_k \in \mathbb{R}, \qquad y_k = y_k^{obs} \in \mathbb{R}, \qquad u_k \in \mathbb{R}^2$$

$$A = \alpha \in \mathbb{R}, \qquad B = \beta \in \mathbb{R}^{1\,x\,2}, \qquad C = 1$$

$$\Sigma_w = \sigma_w^2 \in \mathbb{R}, \qquad \Sigma_v = \sigma_v^2 \in \mathbb{R}$$

$$\bar{x}_0 = \bar{y}_0, \qquad \Sigma_0 = \sigma_0^2$$

Let the prediction estimate of CFP expression state y_{k+1} at time instant k be denoted by $\hat{y}_{k+1}$ and the prediction error variance by $\hat{\sigma}_k$. For the CFP expression first order model, we do not have the true values of the system and input parameters, α and β. Using our model and the Kalman filter equations [14]-[16], the algorithm to predict CFP is as follows:

Algorithm 2: CFP expression prediction using Kalman filter:

Input: Model parameters $\hat{\alpha}, \hat{\beta}\ \sigma_w^2, \sigma_v^2, \bar{y}_0, \sigma_0^2$.
Initialize: $\hat{y}_0 = \bar{y}_0,\ \hat{\sigma}_0^2 = \sigma_0^2$.
for $k = 0\ to\ N-1$ **do**

$$g_k = \hat{\sigma}_k^2\left(\hat{\sigma}_k^2 + \sigma_v^2\right)^{-1} \qquad [17]$$

$$\hat{y}_{k+1} = \hat{\alpha}\hat{y}_k + \hat{\beta}u_k + \hat{\alpha}g_k(y_k^{obs} - \hat{y}_k) \qquad [18]$$

$$\hat{\sigma}_{k+1}^2 = \hat{\alpha}^2(1 - g_k)\hat{\sigma}_k^2 + \sigma_w^2 \qquad [19]$$

end for

---

where, g_k is the Kalman gain and N is the total number of time points.