DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Emergence and Transmission of Noise Created at Cell Division

Citation

Jung, Yoonseok. 2017. Emergence and Transmission of Noise Created at Cell Division. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:41141533

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

Emergence and Transmission of Noise Created at Cell Division

A dissertation presented

by

Yoonseok Jung

to

The School of Engineering and Applied Sciences

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Applied Physics

Harvard University
Cambridge, Massachusetts

April 2017

© 2017 – Yoonseok Jung

All rights reserved.

Advisor: Professor Johan Paulsson Yoonseok Jung

Emergence and Transmission of Noise Created at Cell Division

Abstract

Molecules in cells collide and react randomly, creating stochastic fluctuations in synthesis and degradation which creates heterogeneity among genetically identical cells in the identical environments. However, the random motion of cellular components also creates spatial heterogeneity which at cell division means that the two sister cells can receive different amounts of each component. Such partitioning errors could in principle explain much of the heterogeneity that now is attributed to stochastic gene expression, but has been largely unexplored. This thesis aims to study the contribution of partitioning noise at cell division on cellular variability. The effect of partitioning errors on the overall heterogeneity can also accumulate over multiple divisions, depending on how quickly deviations are corrected during the cell cycle. Thus, its effect can only be properly estimated using accurate time-lapse measurements over many single cells while keeping track of sibling cells. In the first part, we present a microfluidic device that enables high-throughput and accurate measurements under exceptionally homogeneous growth condition both in time and space. In the second part, using *Schizosaccharomyces pombe* as a model organism, we systemically measure partitioning noise for high-abundance

proteins with various localizations and reveal that a significant amount of partitioning noise composes protein noise at birth, ranging from 33% of cytoplasmic proteins to 57% of vacuolar proteins. Next, by leveraging our microfluidic device, we directly measure how partitioning noise is corrected over cell cycle. Surprisingly, all of the measured strains displayed a simple exponential correction curve with a half-life of 1 generation, as expected when cells produce the same amount regardless of their starting value. That is, the correction curve is fully described by a passive control model in which fluctuations are corrected by regression to the mean without feedback. We further use this model to identify the total contribution to protein noise from partitioning errors, and thus also to identify the noise component that appears to come from stochastic gene expression. Overall, our work demonstrates the significance of random partitioning on protein noise.

Table of Contents

Abstract……iii

Acknowledgements……vi

1. Introduction……..1

2. Microfluidic Device for High-throughput Imaging of Fission Yeasts……..5

1. Introduction……..…….…….6

2. Design of mother machine for fission yeast……..…10

3. General pipelines to analyze a time-lapse movie……..13

4. Homogenous condition over time and space…….14

5. Size control in fission yeast…… …. 19

6. Discussion and continuing works……..22

3. Effects of Random Partitioning on Protein Noise…….26

1. Introduction……27

2. Detecting sister cells that just completed cytokinesis……30

3. Significant partitioning noise generated at each cell division.…..…….……32

4. Half of partitioning noise transmits to the next cell cycle…….36

5. Passive transmission of partitioning noise…….41

6. Decomposing the effects of random partitioning from protein noise……43

7. Discussion and outlooks……49

Appendix A: Methods……..53

Appendix B: Fabrication of Microfluidic Devices……71

Appendix C: Supplemental Figures…… ..……80

Reference……..90

Acknowledgements

First and foremost, I would like to thank Johan Paulsson for being such a great adviser. Not only this work would have not been possible without his intellectual guidance, but also I would have not reached here without his continuous help. During many hard times, he was always supportive of me and often gave great insights to encourage me.

I thank Philippe Cluzel and Donhee Ham for being great committee members for my dissertation.

I am grateful to the members of the Paulsson lab. I thank Dann Huh for first introducing me to the lab and teaching many experimental techniques to a person who had never held a pipette back then. As a collaborator, it was also a great experience to have wonderful discussions with him. I thank Nathan Lord for helping me make mother machine and teaching me how to run and analyze time-lapse experiments. I thank Burak Okumus for having many great discussions and teaching me valuable microscopy techniques. I thank Dirk Langraf for providing me plasmids and giving me great feedbacks whenever experiments fail. I thank Rishi Jajoo for having an opportunity to collaborate with him. I thank Sadik Yildiz for helping me build many analysis scripts and being a good lab mate. Many overnight works with him to build a turbidostat will be not forgotten. I also thank for all the other members: Billy, Raul, Tom, Laurent, Andy, Shay, Silvia, Charles, Somenath, Scott, Emanuel for being great coworker within the lab.

Many people helped me to fabricate microfluidic devices. I thank the staffs at the center of nanoscale centers for providing me a wonderful facility. I thank Calixto Saenz for maintaining the cleanroom for microfluidic devices. I also would like to thank

Kaushik Ragunathan and Danesh Moazed for being supportive collaborators. I would like to thank Samsung Scholarship for financially supporting me during graduate years.

It was a great luck to have great friends during graduate years. Especially, I would like to thank Dongwan Kang, Gwangpil Jung, Byungmo Choi, Jaeyun Moon, Gyeongyun Kim, Junsung Yang for their supports and sharing great moments with me. I thank Hae Na Park for her patience, support, and caring me a lot.

I thank my grandparents grandparents-in-law for their warm encouragements. My grandfather-in-law had shown his integrity and modesty in his life, which I will always admire.

Most importantly, I thank my parents for their constant support, love and patience. Without them, I would have not reached here. They taught me being calm under pressure and how to appreciate every day of my life. I also thank my younger sister for her support. I am truly grateful to have these wonderful family.

This work is dedicated to my parents.

Chapter I:

Introduction

Cellular life is full of randomness [1]. Inside a cell, biomolecules such as mRNAs and proteins are synthesized and degraded by biochemical reactions triggered by random collisions between their upstream components. For example, the transcription of an mRNA begins with RNA polymerase by chance finding and binding to a promoter region in the genome [2]. The times of the birth and death events of biological components is thus inevitably probabilistic [3-5], causing stochastic variation in abundances of all components [6, 7]. The most studied example is perhaps gene expression 'noise', where genetically identical cells grown under identical homogenous environments can differ greatly in protein levels due to stochastic transcription and translation [8-10]. Gene expression noise is sometimes conceptually separated into two components, based on how it originates: 'intrinsic' noise arises from probabilistic births and deaths within the

system itself and, 'extrinsic' noise arises from variability in other components that affect the system [11, 12]. Intrinsic noise stems from low copy numbers as limited numbers of reactions involved generally do not sufficiently average out randomness. Indeed, it has been directly measured that many mRNAs and proteins exist in only a handful of copy numbers per cell in *E. coli* [13], yeasts [14, 15], and mammalian cells [16, 17]. When such a component participates in biochemical reactions of another species, its fluctuation perturbs the chemical parameters, thereby propagating its noise to the downstream as an extrinsic noise [1, 6]. For example, fluctuation in concentration of mRNA affects the synthesis rate of its cognate proteins; thus, noise in mRNA thus transmits to protein noise. Therefore, even high abundance proteins can exhibit significant fluctuations as a consequence of randomness in its upstream factors such as ribosomes, transcripts, transcriptional factors, and even the states of chromatin [15, 18]. Given that chemical networks in a cell are highly interconnected [19-21], it has been suggested that gene expression noise may be a widespread phenomenon.

Supporting this view, many experimental studies indeed revealed significant non-genetic heterogeneity in the concentrations of mRNAs and proteins under homogenous conditions, regardless of their expression levels for a variety of organisms [13, 17, 22-24]. Notably, these genome-wide studies [13, 24, 25] highlighted the significance of stochastic gene expression by successfully fitting the observed distributions to simplified toy models consisting of random birth and death processes. These models, however, often

ignored one of the most fundamental stochastic processes that potentially contribute to molecular noise; every molecule randomly ends up being in one of daughter cells at cell division because its location inside a cell is inherently governed by stochastic movements [26-28].

Stochastic partitioning process can produce sister cells having asymmetric mixture of cellular components even at symmetric division in volume, thus their molecular concentrations are randomized at each division. Indeed, many studies on organelle inheritance revealed significantly asymmetric spatial patterns of cellular compartments between sister cells at cell division [29-35], yet its effect was not linked to protein noise in a quantitative way. More critically, Huh and Paulsson theoretically showed that noise arising from random partitioning is inherently indistinguishable from gene expression noise when using snapshot measurements because very different models can fit the same data; even when noise uniquely comes from partitioning, the shapes of the resulting distributions and the fold-change in variances to changes in transcription and translation rates can still be exactly as expected from models of stochastic gene expression [36]. Hence, the question of how much of protein noise comes from gene expression noise has not yet been answered at all.

The noise from partitioning can also accumulate over divisions. Thus to appreciate its total contribution two questions must be answered; 1) how much noise is generated per division? and 2) how rapidly is the noise corrected during cell growth? To address the

latter, we must follow the differences between sister cells over cell growth in high-throughput. In Chapter II, we describe our invention of an experimental platform for high-throughput single cell imaging using microfluidics. In Chapter III, we systemically measure the noise from random partitioning in a quantitative an systematic manner and argue its significance on protein noise.

Chapter II:

Microfluidic Device for High-throughput Imaging of Fission Yeasts

Contributions

Development of a GUI program for manual correction of segmentation masks and algorithm to segment cells with phase-contrast images were done in collaboration with Sadik Yildiz. I performed all other works.

II.1. Introduction

The state of a cell reflects its past dynamics, but in a one-to-many relationship where many different paths can lead to the same current state. The same is true for statistical properties: just like many different distributions can produce the same bulk average, many different stochastic processes can produce the same overall distribution [37]. This has been a very real practical problem. Specifically, stochastic models are often evaluated by comparing predicted and measured distributions, but very different stochastic reaction mechanisms can create the exact same distributions. For example, bursty synthesis of mRNAs or proteins can produce the same distributions as many other non-bursty mechanisms that are equally plausible [18, 21], and yet they have been taken as evidence of the former.

Much more information can be gleaned by directly monitoring processes in individual cells through time-lapse imaging [38-40]. This idea has a long history in biology. For example, in the 1970s, a series of landmark papers [41, 42] by Paul Nurse revealed a surprisingly tight distribution of cell sizes at mitosis in fission yeast, demonstrating that it is achieved through negative feedback between initial size and generation time by periodically recording the images of cells on a panchromatic film. However, the molecular mechanisms that underlie the size control have only recently been established much thanks to the advent of fluorescent protein reporters [43, 44], which began to be used for quantification of protein level in the 2000s [11, 45-47].

Time-lapse fluorescence microscopy enables us to monitor intracellular dynamics with fluorescent reporters in a single cell level, and has helped to uncover many hidden aspects that cannot be appreciated in population level experiments. For example, time-lapse experiment has helped us understand how cells make fate decisions to enter alternative states using noisy genetic networks [48-50]. In addition, it is a great platform to test how synthetically engineered circuits work and to how improve them [45, 51]. Thus, time-lapse imaging of single cells has established itself as a powerful tool to study a wide variety of cellular systems.

The conventional method for time-lapse measurement is simply to grow cells on an agarose pad, which can be applied for a wide variety of model organisms [49, 52-54]. However, several major challenges remain. First, it does not allow long-term imaging. The first few generations are often hard to interpret because cells are adjusting to the change in conditions from liquid culture to colony. Because the number of cells exponentially increases, the population then quickly accumulates and often forms second layers when the population size reaches a few hundred cells, i.e., in 7-8 generations. This severely limits the duration of experiments. Even doubling the field of view only extends the time series by one generation. Given that some transitions in cellular states occur on very long time scales (for example >50 generations for spontaneous changes of histone modifications in fission yeast [55]), much longer duration is needed to study rare events. A second challenge is that agar pad experiments do not provide homogeneous growth

conditions for cells. Since cell growth depends on nutrients diffused from the agarose pad, its local concentration is likely to fluctuate over time and space depending on density of nearby cells, thus each cell can have a different history in cellular environments. This makes it hard to distinguish if cellular changes come from intracellular reactions or from environmental changes. Third, tracking cells in growing colonies is nontrivial and often laborious because each cell is very similar in their size, are densely packed, and moves almost randomly due to growth of surrounding cells. When accurate lineage tracking is required, small errors in tracking can completely change the conclusions. Fourth, agar pads generally do not provide high-throughput measurements, because images in each field of view must be taken frequently enough to reliably track the cells and because the colonies must be spaced out to avoid the risk of merging colonies. Finally, because cells are being in close proximity over time, optical signals coming from nearby cells or 'halo' interfere with its own signal, distorting quantitative measurement due to the point-spread function of light.

To address the above challenges, alternative approaches using a microfluidic device have recently emerged [56-58]. One famous example is the 'mother machine', a microfluidic device originally developed for *E. coli*, which has been shown to overcome many experimental limitations of agarose pads [59]. The mother machine has cell growth channels, in which the rod-shaped *E. coli* cells fit and grow while aligned linearly due to structural constraints. One end of the cell channel is closed off, and thus a 'mother' cell

stays at the end, while its progeny are pushed out of the open end as they grow. The growth in an organized shape makes tracking of growing cells easier. Also, the cell growth channel is connected to a larger media channel where growth medium constantly flows by, allowing nutrients to diffuse around the cells and also washing away cells coming out of the growth channel. This simple yet effective design has allowed some labs to image single cells over time with high sampling depth under homogenous environments, and provided a great platform for many valuable studies in bacteria [50, 51, 60-62]. However, for eukaryotes, there is a limited number of microfluidic devices that provide long-term imaging in a high-throughput way. For *Saccharomyces cerevisiae,* there is a similar version of mother machine that only confine mother cells taking advantage of their bigger cell size due to asymmetric division, but the duration of an experiment is limited since they eventually die because of aging effect (half life = 20 generations) [63], and also there is possibility that cells are mechanistically stressed since they become larger as they age while dimensions of the chip stay constant (personal communication with Dirk Landgraf). For *Schizosaccharomyces pombe*, which does not seem to age significantly [64], there are two reported devices to our knowledge. The first one [65] is basically a version of mother machine with wider cell growth channels, such that cells form colonies in the cell channels as on an agarose pad, hence harboring the similar disadvantages including non-trivial tracking. The second one used a clever trick to slightly open the end of the cell channel with narrower trenches, rather than to close it as in the mother machine, so that medium can actively flow along the cell channel.

However, it was only possible to trap and track ~100 mother cells because loading of cells was inefficient and mother cells were often lost over time due to unstable flow [66]. More importantly, this device completely does not provide lineage information, as one of the two daughter cells is immediately washed out upon cytokinesis. As described in Chapter III, following both sister cells over time is critical to measure how partitioning errors is corrected over cell growth. Finally, the utility of high-throughput imaging and small statistical errors is severely compromised if systematic artifacts, e.g. due to systematic spatial or temporal biases in the device, instead come to dominate the total error. Therefore, we set out to design a new microfluidic device for fission yeasts based on the mother machine for bacteria [59].

II.2. Design of mother machine for fission yeast

In terms of designing microfluidic devices, the main difference between *E. coli* and *S. pombe* is their size. *E. coli* is ~2-5μm in length on average at division, depending on growth conditions, and ~0.5-0.8μm in width, whereas *S. pombe* is ~14-18 μm in length and ~3.8μm in width at division. The larger cell size requires the longer cell channel, which makes it more difficult to provide fresh growth medium to cells passively by diffusion as in the original design. To overcome this, I developed and tested a very large number of microfluidic devices, with different designs or dimensions to make sure that there are no any signs of poor or non-uniform feeding. Most of them enabled cell

growth yet with spatial biases inside the cell channel. For example, an original design with larger dimensions showed slower growth of old-pole cells.

The final solution included shallow media feeding layer on the side, surrounding the cell channel and connected to the main media channel where medium flow (Figure II. 1A). In this way, fresh medium has much more space to flow around cells and actively flows to cells from an inlet. The height of the shallow layer (2.2μm) is lower than the main cell channel (4.2μm), thus cells can only fit the main channel when loading (Figure II.1B) and over growth (Figure II.1C). To prevent the shallow layer from collapsing into a cover glass, we also included circular pillars to support it (Figure II.1A). As a result of active flow, we could not detect any spatial bias within or between growth channels, despite the great statistical throughput. To minimize the halo effects from nearby cells, the cell channels are separated by 12μm. The height of the cell channel was carefully optimized with two considerations: loading is not efficient if too low, while cells tilt if too high which makes it difficult to quantify the cell size. On average, eight newborn cells can be imaged per the cell channel (85μm long), and nine parallel channels can fit in one field of view with 60x magnification. Cells stay in the same horizontal plane over time, thereby helping subsequent image analysis such as segmentation and tracking (Figure II. 1C). A total of 10 independent lanes were embedded in a single microfluidic chip, therefore multiple experiments are possible in a parallel way.

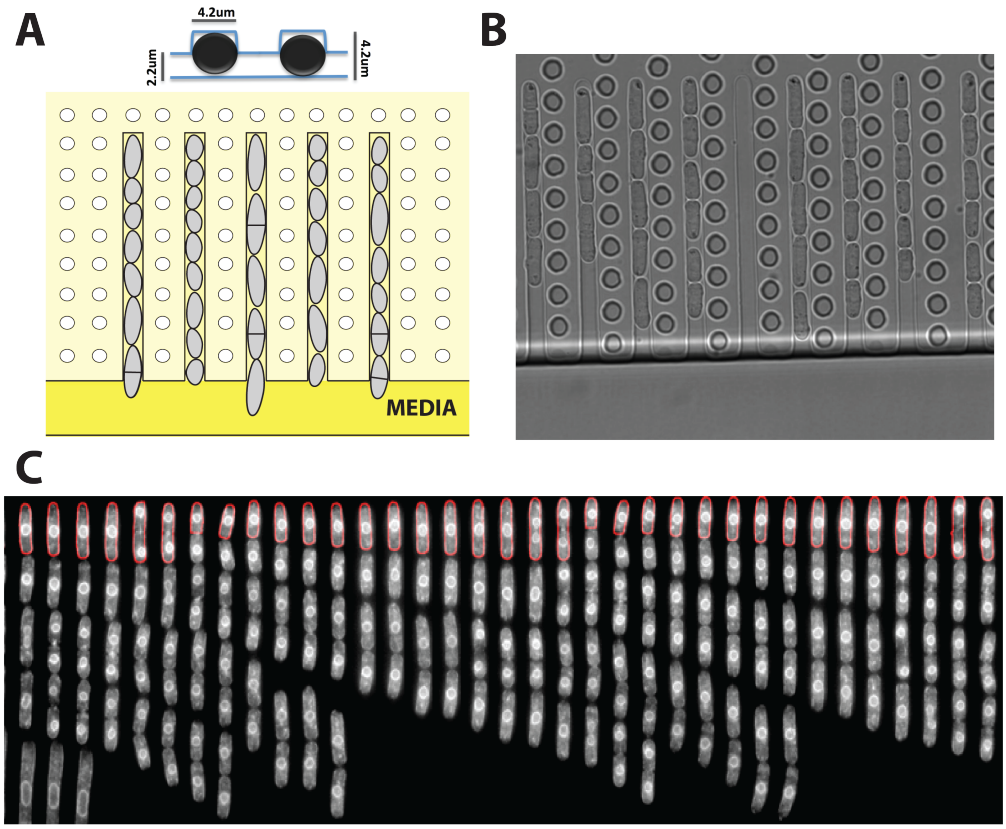


Figure II.1 | Design of the mother machine for fission yeasts. **(A)** A schematic of the modified microfluidic device for *S. pombe*. The cell channel filled with cells is designed as 4.2 um x 4.2 um rectangular shape (upper panel) to fit the shape of fission yeast with a diameter of ~3.8 um. As illustrated in a top view of the device (lower panel), each cell channel is surrounded by the shallow media layer (light yellow) and the main media layer (dark yellow), and growth medium actively flows in both of layers. The shallow media layer is supported by multiple pillars (white circles) of 2.2 um height to keep from collapsing while not allowing cells to enter. Each cell channel is 85 um long, which can accommodate four dividing cells (16 um in rich media), and is separated by 12 um to minimize the halo effects between them. **(B)** An example bright-field image of the device

Figure II.1(continued) | after loading the cells into it. With a 60x objective, a maximum of 9 independent channels per a field of view can be imaged simultaneously, and about 5 dividing cells can be accommodated within a channel. **(C)** An typical time-lapse movie within a single cell channel is represented as a kymograph. The images were taken every 15 minutes for the strain (DH115) in which mGFPmut3 is translationally fused to *cnx1* that localizes to endoplasmic reticulum. The red lines indicate the boundaries of segmentation masks for mother cells over time.

II.3. General pipelines to analyze a time-lapse movie

Our device described above allows us to trace thousands of individual cell lineages over hundreds of consecutive hours, generating a great amount of high-resolution time series data. To reliably analyze this data with minimal manual labors, we developed an automatic pipeline to extract various quantities from the time-lapse movies. The key feature of our analysis is to track and extract data from all cells, including siblings and cousins, unlike most previous analyses that only track mother cells. Also, we take advantage of the lineage information to detect segmentation errors, thereby reducing the efforts of manual corrections. First, using a fluorescent cytoplasmic marker, the boundaries of cells are identified by applying local thresholds on the intensity in pixels with the T-point algorithm [67]. Second, cells within a single cell channel were extracted from each image, and a kymograph was made by combining trimmed images over time (Figure II.1C and Supplemental Figure C.1B). Third, tracking is performed. Briefly, two

consecutive segmentation masks are matched from the top, a division event is detected if there is a decrease in the size of segmentation mask, and an indicator of segmentation error is raised if the sum of sizes of two daughter cells are significantly deviated from the mother cell size or if the ratio of cell size between sister cells significantly deviates from one. This simple tracking algorithm is possible because of the linear growth trenches in the mother machine that conserve the order of cells. By tracking the cells over subsequent images, growth patterns of individual cells are calculated, and an another indicator of segmentation error is raised if the growth pattern is not monotonic or severely deviates from fitted trend, or if the generation time is too short or long (Supplemental Figure C. 1A), raising an alarm for further manual checks. Fourth, segmentation masks with possible errors detected in the previous steps are visualized and validated in our custom interactive module, and are manually corrected by comparing the patterns in nearby frames (Supplemental Figure C.1B). Finally, quantities such as cell size, fluorescent intensity profile inside a mask, positions in a field of view, and cell cycle positions are extracted from each mask and grouped by each lineage. Overall, this automatic pipeline enables to extract high-quality data for single-cell lineages in a fast yet reliable fashion.

II.4. Homogeneous condition over time and space

Maintaining a stable environment over time and space is critical to study cellular variability, because interpretations otherwise become harder. Specifically, if the

environment such as a local concentration of available nutrients fluctuates over time, cells to some extent track the environmental changes by modifying its physiology and intracellular genetic networks. Similarly, if the environment varies depending on location, the difference between cells can no longer be straightforwardly attributed to the internal mechanisms. In fact, even if the environment is homogeneous at the time of data acquisition, the effects from previous inhomogeneities can still remain. For example, cells are often prepared at a temperature different from where they grow, and hence need some time to adapt to the new environment. To address this issue, we measured how long it does take to reach stationarity (in the sense of stochastic processes, with steady statistical patterns) after loading cells, and optimized the protocol to start the imaging only once the adaption effects are not measurable (usually >16 hours for our conditions).

In Figure II.2, we show the stationarity of our device by providing the stable statistics of cell growth over time and space. The experiments were performed for strain DH60 that contains mCherry driven by the constitutive adh1 promoter in the leu1 locus, which serves as a segmentation marker. Cells were loaded into the mother machine and grown for ~16 hours without imaging under rich synthetic medium (SC+PMG media) at 32°C. After the adaptation period, images were taken every 5 minutes for ~3 days. After segmentation and tracking, a total of ~10,000 cell divisions were obtained and generation times and the length at division were quantified. The distributions of measured properties did not vary over time, and for presentational purposes we used two statistical quantities:

the average and coefficient of variation (CV = std(x)/mean(x) for random variable x) for each distribution. The mean generation time was ~144 minutes, which is slightly faster than the ~150 minutes measured in a bulk (likely because we run the experiments in the equivalent of extremely early exponential phase before any growth competition can set in), and virtually constant over time (orange line in Figure II.2A). The CV of generation time (~14% on average) also showed an almost flat pattern (green line in Figure II.2A). The mean cell length at division was measured at ~16.7μm, consistent with previous reports [44]. Both the average and the CV (~9.59% on average) were stable over 3 days of the experiment (orange and green lines in Figure II.2B, respectively). To show spatial uniformity, we compared the distributions conditioned on the positions of cells represented by order within a cell channel (inset in Figure II.2C). The distributions of generation times and dividing lengths for each relative positions were similar to each other, except for slightly longer division times and shorter dividing lengths at the lowest position in some versions of the device (Figure II.2C and Figure II.2D). We validated that this was due to lack of outliers that divide at >25μm at this position, suggesting that even this small anomaly was not statistically significant. Indeed because this is the first position where cells are washed out, a division event of long cells is less likely to be observed. To minimize any artifacts from this bias, we only proceed with cell lineages that the division of their sister cells is observed for further analysis in Chapter III.

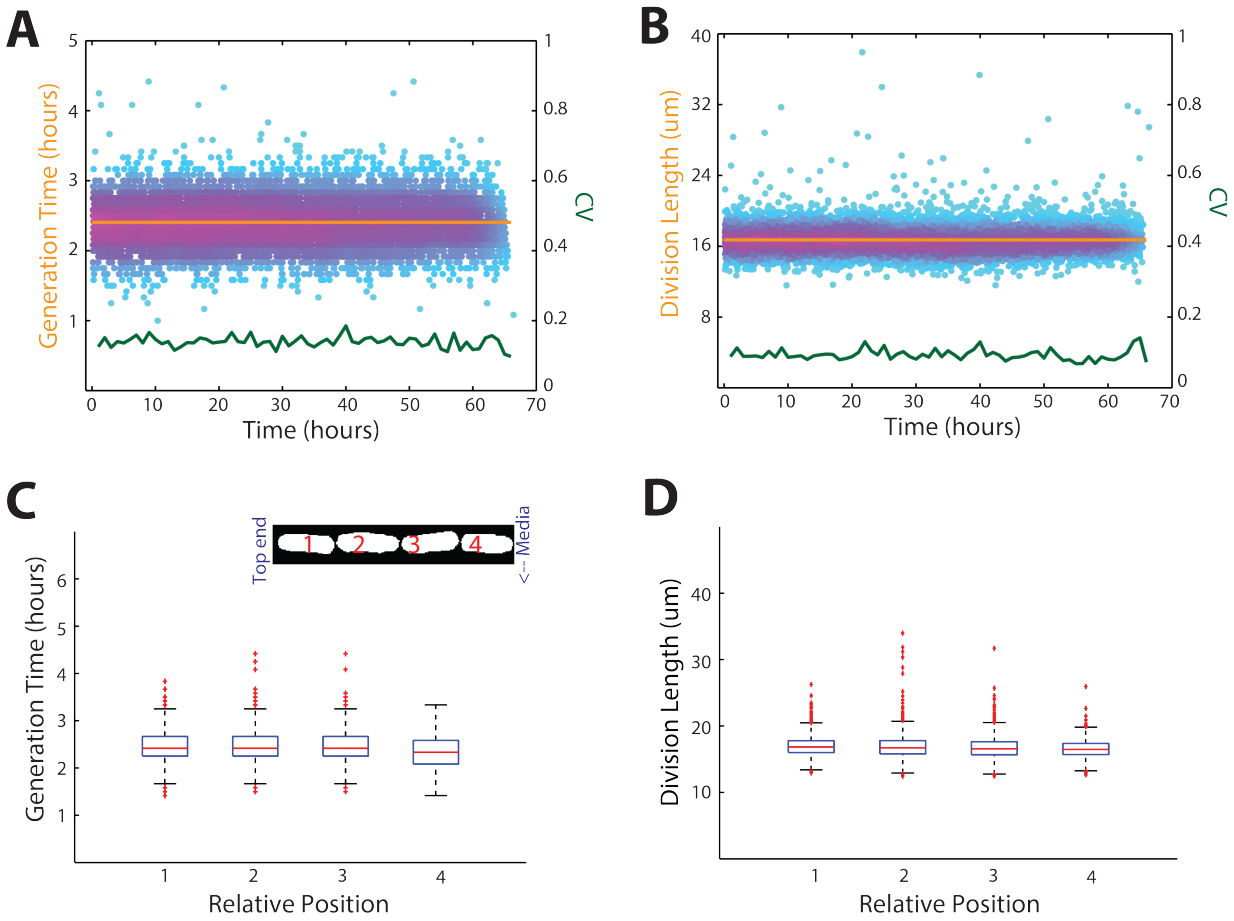


Figure II.2 | Stability of the mother machine for fission yeasts. Time-lapse experiment was performed for the strain (DH60) where Padh1-NES-mCherry is fused to the leu1 locus, which served as a segmentation reporter. Cells were grown under rich synthetic media (PMG) at 32°C. The experiment lasts for ~3 days and around 10,000 divisions were observed. **(A)** Homogenous generation time was observed over time. Each dot represents individual division event and their density was represented as a heatmap. The orange line is a moving average of generation time with a window of 1 hour, and the green line is a coefficient of variation of generation times within moving windows of 2 hours.

Figure II.2(continued) | (**B**) Homogenous division length was observed over space. Each dot represents the length of cell at division. The orange line is a moving average of division length with a window of 1 hour, and the green line is a coefficient of variation of division length within moving windows of 2 hours. (**C**) Homogenous generation time was observed over space. The inset illustrates how we represent the position of cells within the cell channel. A box plot shows the distributions of generation time for each position. Blue box measures standard deviations and the ends of dotted lines shows 25th and 75th quantiles of generation time. (**D**) Homogenous division length was observed over space. A box plot shows the distributions of division length for each position. Blue box measures standard deviations and the ends of dotted lines shows 25th and 75th quantiles of division length.

We also tested stationarity in the concentration of fluorescent proteins (Supplemental Figure C.2). Since these bleach over time, we observed a period of slight decay in their intensities (~5 generations) before eventually stabilizing once bleaching and new production balance (shaded region in Supplemental Figure C.2). Considering this complication, we exclude the data before the stabilization, and further ran experiments with different intervals between images to ensure our conclusions were not affected. We also noticed that the estimated concentration was also somewhat dependent on spatial position in a field of view (Supplemental Figure C.3). This distinct pattern (~10% monotonic decrease to the right) was reproducible day by day, and for strain by

strain, and even when analyzing different organisms with a different setup (personal communication with Nathan Lord), suggesting uneven illumination from light sources. We indeed observed that the microscope stage was slightly titled. To correct this systemic bias, we computationally calculate a moving average curve of the concentration as a proxy for a illumination pattern to improve the flat field correction (Supplemental Figure C.3 and Appendix A.1). This allowed us to accurately estimate levels in single cells. To summarize, our microfluidic device is capable of maintaining fission yeast cells under exceptionally homogenous conditions temporally and spatially, allowing us to consider stable statistical patterns for the cellular quantities of interest.

II.5. Size control in fission yeast

We first used this set-up to evaluate the control of cell size in fission yeast. We observed a negative correlation between generation time and initial cell size (Figure II. 3A, Pearson correlation coefficient = 0.6), consistent with previous studies [68]. Unlike *E. coli* [60], a "sizer model" describes the data well, i.e., a model in which the cell grows until it reaches a specific size (Figure II.3B) by modulating its division time (dotted line in Figure II.3A). In this model, 10% smaller cells have 10% longer generation time on average. However, the sizer model was not perfect under our experimental condition (SC+PMG media), since small positive correlations between the initial newborn cell size and the dividing size was observed (Pearson correlation coefficient = 0.12). That is, the

initial size of newborn is to some extent inherited to the next generation. Though small, these effects can still be important to compare different models. To demonstrate that the relationships between two variables are reproducible, we binned data by when it was measured and calculated the correlation coefficients within each bin. We confirmed that the correlations are stable over time (Figure II.3C and Figure II.3D), demonstrating that our device provides robust statistics even for small effects by enabling us to have high-quality data with enough sampling.

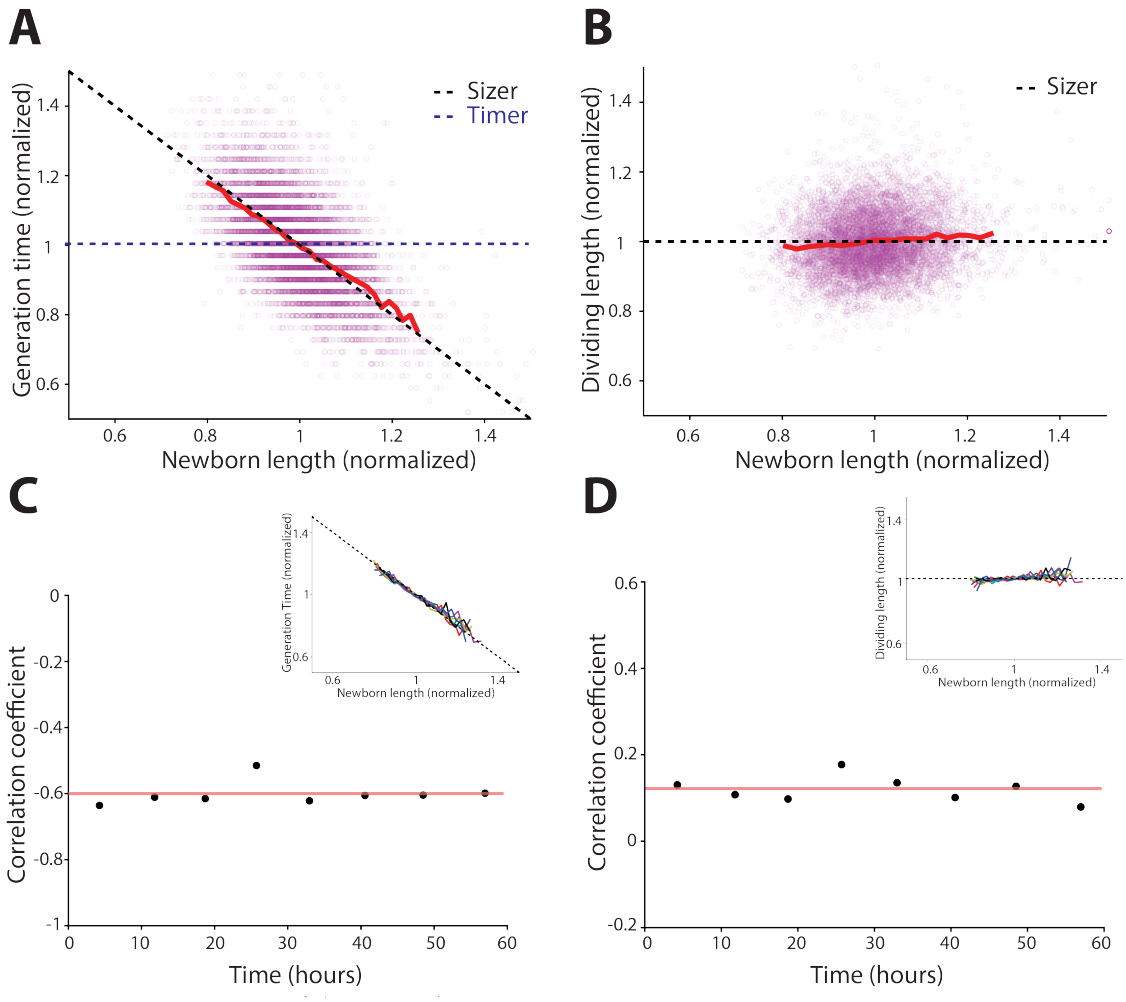


Figure II.3(previous page) | Size control in fission yeasts. Cells (DH01) were grown under rich synthetic media (PMG) at 32°C in a mother machine. The experiment lasts for ~3 days and around 10,000 divisions were observed. Images were taken every 5 minutes **(A)** Newborn length at birth and generation time are plotted after normalizing by their means. The red line indicates binned averages with windows of 1,000 events. Black dotted line represents 'sizer model', while red dotted line shows 'timer model'. **(B)** Newborn length at birth and length at division are plotted after normalizing by their means. Black dotted line represents 'sizer model' where deviations from the mean in newborn length is completely corrected over cell cycle on average. **(C)** Correlation coefficients between generation time and newborn length are plotted over time by binning data with equal widths. Inset shows newborn size versus generation time for data in each bin. **(C)** Correlation coefficients between newborn length and dividing length are plotted over time by binning data with equal widths. Inset shows length at birth versus length at division for data in each bin.

II.6. Discussion and continuing works

Tracking the dynamics of processes in single cells can address many questions that could not be easily answered with snapshot experiments. For example, to explain why genetically identical cells are different from each other, it is very helpful to directly observe the processes over time. To facilitate interpretations, it is also important that measurements are made under homogeneous conditions, so that they are not convolved with fluctuations in the environment. Furthermore, to follow slow dynamics or rare events, long time series and high sampling depth are required.

In this work, we developed a microfluidic device for fission yeast that enables high-throughput imaging under homogeneous conditions. We utilized active transport of fresh media rather than passive diffusion to accommodate larger dimensions than previous devices for *E. coli.* Thanks to our automated analysis pipeline, tracking many siblings over long periods of time became possible, which enabled us measure how partitioning errors arise and persist over time as described in Chapter III. We also used the device to measure long-term epigenetic switching dynamics in single cells, in collaboration with the Moazed group, which would not be possible with a conventional setup since the average switching time were extremely long (~14 generations on average [55]). Since the device potentially can be combined with other imaging techniques such as single molecule imaging, it may also be useful for other applications. For example, one could in principle perform single-molecule fluorescence in situ hybridization (FISH)

measurements after following cellular dynamics and lineage structure, similar to a recent report [69], but in greater throughput and correlating the FISH results to the preceding dynamics.

One caveat with the device described above is that it does not provide complete lineage structure since cell channel can contain four dividing cousins at best. To analyze relationship between cells distant in lineages (for example, cousins of cousins), the device needs to be much longer. Also, it is vulnerable to possible polarity effects since it only keep mother cells. To address these issues, we developed another device that can follow dynamics of complete lineages for 4 consecutive divisions (Figure II.4A). In this design, the cell channel is open in both ends and is surrounded by a side channel that only media can flow. Whenever there is difference in pressure between two parallel media channel, fresh media flow through side channel. This way we were able to provide homogenous environment for extremely long channels (320 um) that contains ~20 dividing cells (left panel in Figure III.4A and Figure III.4C). Our current project with this device is to study correlations in division time and cell size within lineage structures. To minimize the artificial effects from tagging fluorescent proteins, we also developed a pipeline to segment and track cells in phase images in collaboration with Sadik Yildiz in our lab (Figure II.4B). We anticipate that this new device will allow to study inheritance patterns in cell size and generation time between cells in a large lineage tree, which has been rarely measured in a controlled environment.

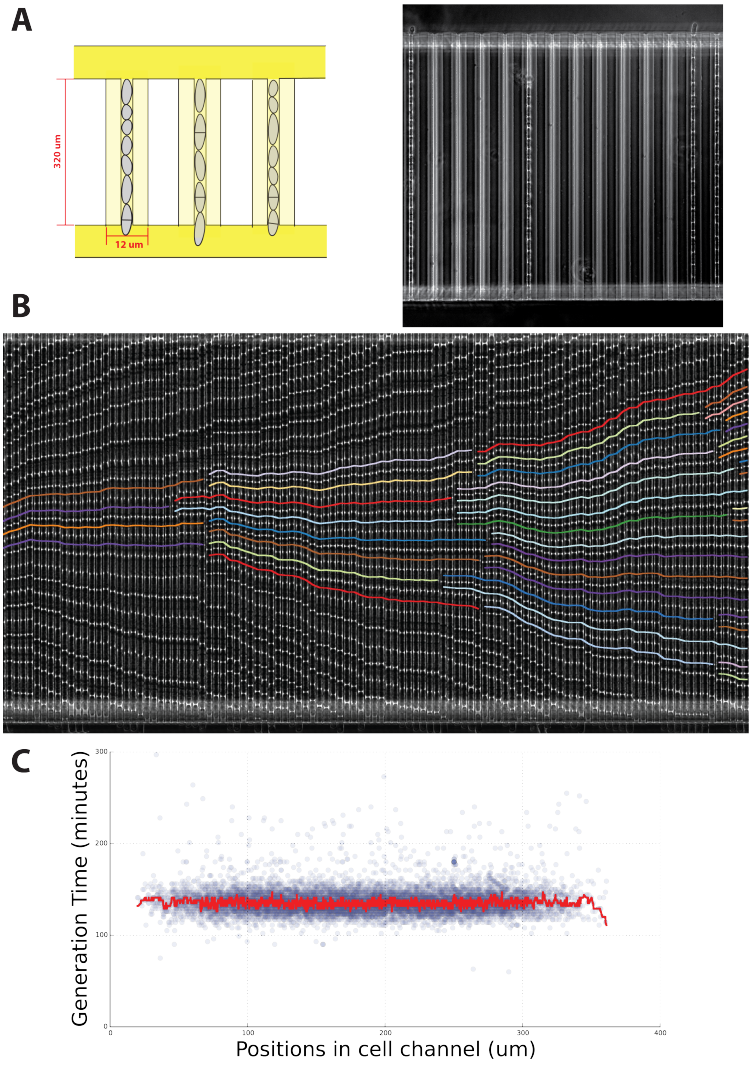


Figure II.4 | Open ended microfludic device for fission yeasts. **(A)** Right panel shows a schematic of the open ended microfluidic device for *S. pombe*. The cell channel (320 um) filled with cells is designed as 4.2 um x 4.2 um rectangular shape to accommodate cells, and is surrounded by a side channel (2.2 um) in which cells cannot enter and only medium flows. Left panel is a phase-contrast image of the device filled up with cells. Each cell channel can have >20 dividing cells. **(B)** A kymograph for a single cell channel with

Figure II.4(continued) | tracking lines. Cells (DH0) were segmented in phase contrast images and tracked. Images were taken every 4 minutes. Each line represents a trace of centroid of cells and the discontinuity between lines shows a division event. **(C)** Generation time over position inside a single cell channel is plotted. A red curve shows a moving average with a window size of 30.

Chapter III:

Effects of Random Partitioning on Molecular Noise

Contributions

This work will be submitted as

Jung, Y.*, Huh, D.*, Paulsson, J. *Noisy protein concentrations in S. pombe broadly reflect partitioning errors at cell division.*

All of the authors participated in developing general ideas and writing a manuscript. Dann Huh performed snapshot experiments with confocal microscopy. I performed time-lapse experiments in microfluidic devices. Both Dann Huh and I built strain libraries and developed imaging set-ups.

III.1. Introduction

The motion of molecules in cells is random to a large extent, whether driven by thermal fluctuations and/or energy-dissipating motors [70]. As a result, the molecular collisions leading to production or degradation of cellular components are inherently probabilistic, causing spontaneous fluctuations in concentrations [3, 21, 71, 72]. However, because the locations of those components are random, different parts of the cell will have different numbers of molecules even when the average levels are spatially homogenous. At cell division, this asymmetry can cause spontaneous differences in cellular concentrations simply because more molecules can by chance end up in one of daughter cells than another [26-28, 73]. In fact, it was mathematically demonstrated that the effects of such 'partitioning errors' can be surprisingly similar to the noise coming from random production and degradation [36]. For example, the shape of the resulting distributions of numbers of molecules per cell can be similar, and the variances can respond in the same way to perturbations in e.g. transcription or translation [36]. Furthermore, the resulting randomness can be very hard to suppress during the next cell cycle due to delays and information loss in feedback loops [71, 74].

Partitioning errors are particularly pronounced for low copy components. However, in many cases, molecules can segregate together in groups and the partitioning errors are then determined by the effective number of partitioning units rather than the number of individual molecules, suggesting that even components in high numbers can exhibit

significant errors at division if they segregate in a few compartments [33, 75]. Indeed, the intracellular space can be highly structured with various types of biomolecular groups composed of many different species. Some types of molecules naturally form clusters even in prokaryotes, for examples some types of plasmids [76, 77] or organelle-like structures such as carboxysomes in cyanobacteria [78]. Although prokaryotes have less spatial structures in cellular space, eukaryotes have a diverse range of organelles, concentrating certain proteins into compartments [79, 80]. Even the organelles themselves may be spatially structured, their locations in the cell are not necessarily quantitatively ordered. Therefore, we reasoned that protein levels in eukaryotes could be particularly sensitive to stochastic segregation, despite or even because of the the apparent spatial organization.

Several studies have indeed reported uneven patterns of localizations at cell division for a wide variety of organelles [29, 30, 32-35, 81], but the differences have not been systematic or quantitative. Furthermore, to evaluate the significance of random partitioning, the magnitude of the partitioning noise resulting from the partitioning errors should be compared to the total protein noise, to determine if the partitioning noise is overshadowed by other types of noise or if they themselves dominate the heterogeneity.

Another complication when analyzing the effect of random partitioning on protein noise is that the heterogeneity caused can remain incompletely corrected for several generations, thus the effect accumulates to some extent. Thus, even if minute partitioning

noise is generated per division, the total effect can be still significant when summed up over periodic cell divisions. Also, partitioning errors in one component can appear as production noise in another over cell cycle. For example, random segregation of mRNAs between two cells will cause random differences in protein production rates. Though partitioning errors occur almost instantaneously when cells divide, its effects on heterogeneity in cells can thus occur gradually over several cell cycles. Even if time-lapse movies revealed significant variation in production or degradation rates, it is thus possible that most heterogeneity ultimately 'originates' from uneven segregation at cell division, making it essential to measure now only the partitioning error but also how quickly the resulting partitioning noise is eliminated in subsequent cell cycles.

In this work, we set out to systemically evaluate the contribution of random partitioning on protein noise in *Schizosaccharomyces pombe*. We first build a library of non-intrusive and localization-verified reporter proteins for proteins with diverse localization patterns. We then quantitatively survey how much partitioning error is generated per division, and using a high-throughput microfluidic device, we measure how the resulting partitioning noise is corrected during subsequent cell growth by following the traces of sister cells after cell divisions. Finally, we develop analysis tools to extract the net contribution of random partitioning from the total protein noise, and compare it to gene expression noise.

III.2. Detecting sister cells that just completed a division

To systemically survey the contributions of partitioning noise to cellular heterogeneity, we chose fission yeast as a model organism because 1) it on average divides symmetrically in volume, 2) it has a wide range of organelles with different spatial patterns, and 3) its generation time is relatively fast (~140 minutes in synthetic rich media) for a eukaryote. We fused a monomeric version of GFPmut3 with the A206K mutation (mGFPmut3) [82] into 38 different proteins selected to cover various localizations including vacuole, mitochondria, nucleus, endoplasmic reticulum (ER), Golgi complex, periphery, and the cytoplasm (Figure III.1A and [91] for more details). mGFPmut3 was chosen over other fluorescence proteins evaluated because the fluorophore matures relatively quickly (half-time = 8 mins at 30°C [91], Supplemental Figure C.3) and because it is highly monomeric and thereby minimizes artifactual clusters [82].

Because the 3'UTR of the mRNA can affect protein localization [83], we used a 'seamless integration' method that does not to leave any integration scars between the insert and the endogenous 3'UTR (Figure III.1A) [84, 91]. To minimize artifactual localization pattern due to tagging fluorescent proteins, we carefully chose proteins whose localization were reported previously [80] and where we could confirm that their patterns were unaltered [91]. A cytoplasmic mCherry was also expressed as a segmentation marker in all strains, driven by a constitutive adh1 promoter (Padh1-NES-mCherry integrated in *leu1* locus). The tagged proteins are in high abundance with $>10^4$

copies per average cell [85] to ensure sufficiently strong fluorescence signals. Thus we expect the partitioning errors would mostly come from organization into larger units or organelles, rather than from spontaneous low-copy noise.

We first asked how much partitioning error is generated at cell division. Partitioning error can be defined as the added variance in newborn cells compared to their mother cell (see Appendix A.2). We therefore sought to detect pairs of sister cells that just completed cytokinesis. Exponentially growing cells were plated on an agar pad, and imaged at room temperature under a spinning disk confocal microscopy. Cells were segmented using cytoplasmic mCherry signals, and pairs of divided cells with complete septum were selected based on segmentation masks. For some components it could be important to accurately identify the time that cells divide. For example, it was shown that mitochondria move to the cell poles together during chromosome segregation, but that they subsequently re-equilibrate over the cell just before division [81]. To confirm effective separation between cells in a putative sister pair, we photo-bleached mCherry or mGFPmut3 in one of the daughter cells and monitored changes in fluorescence in the other cell [91]. This showed that our segmentation algorithm was an excellent indicator of septum completion.

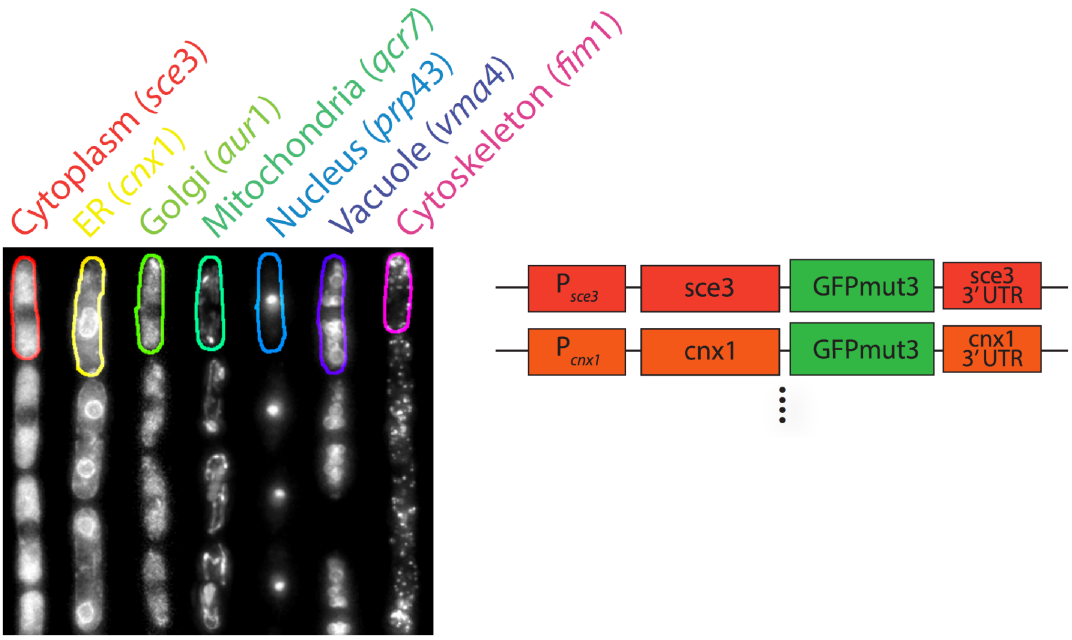


Figure III.1 | Library generation of *S. pombe* The left image represents sample images of strains with diverse localization patterns in the mother machine. The outline of cells shows segmentation masks obtained from mCherry channel images. As in the right illustration, all strains conserve 3'UTR after tagging mGFPmut3 [91].

III.3. Significant amount of partitioning noise is generated per division

By measuring hundreds of just divided pairs of sister cells, we could accurately estimate the partitioning error for each protein reporter. As cell division occurs quickly, almost instantaneously relative to the rest of the cell cycle, the volumes and molecular abundances in the two daughters sum up to the volume and abundance in the mother (for

example, see a cytokinesis event in Supplemental Figure C.6B), allowing us to calculate the increase in protein noise after cell division (Figure II.2A).

The variability in protein levels among newborn cells was significantly higher than before division for all strains, ranging from 4.5% of cytoplasmic proteins to 10.2% of vacuolar proteins (green bars in Figure II.2B). We found this increase variability is well described as a normalized difference in concentration between sibling cells (Supplemental Figure C.4), similar to previous theoretical results for total abundance [91]. We further found that the partitioning error were largely determined by where proteins localize (green dots Figure II.2B), regardless of their expression levels (Supplemental Figure C.5), indicating that random partitioning of organelles is a major source of variability. We observed a similar pattern for the total protein noise after division (orange dots Figure II.2B), suggesting that the localization in organelles may indeed play a major role in shaping the molecular variability. To assess the contribution of random partitioning, we compared the measured partitioning noise (Q) to the protein noise of newborns (CV_0) for each strain. A significant portion of the total protein noise indeed came from partitioning errors, ranging from 32.6% of cytoplasmic proteins to 56.9% of vacuolar proteins (blue bars in Figure II.2C), with similar values for different proteins in the same organelle or each localization (blue dots in Figure II.2C). Thus we confirmed theoretical speculations that partitioning errors can provide a substantial source of protein noise even for high abundance proteins.

To better understand why each organelle has a different partitioning noise, we first tested if various organelles segregated in proportion to the volume of each daughter cell. For proteins in the cytoplasm or mitochondria we observed that the abundances in each daughter cell on average were proportional to the relative volume of the daughter (dotted line in Supplemental Figure C.6A), as expected when molecules are uniformly distributed over the cell. Thus, if a particular daughter cell after division happened to have 10% larger volume compared to its sister, it on average contained 10% more of the proteins. For proteins in the endoplasmic reticulum or nucleus, total abundances are more similar between daughters of different sizes, i.e., the larger cell did not contain proportionally more of the proteins (Supplemental Figure C.6A). We noted that mitosis happens earlier than cytokinesis in fission yeast (Supplemental Figure C.6B). Since the nucleus divides symmetrically prior to cell division, proteins localizing to the nucleus therefore partitions symmetrically, regardless of any asymmetry arising at cytokinesis. Vacuoles showed the opposite pattern, i.e., the protein abundance exaggerated the volume skew: if one daughter cell was larger than the other, the abundance of vacuole proteins were more than proportionally larger. This might be explained by the fact that vacuoles are large objects, and therefore affect where the septum forms.

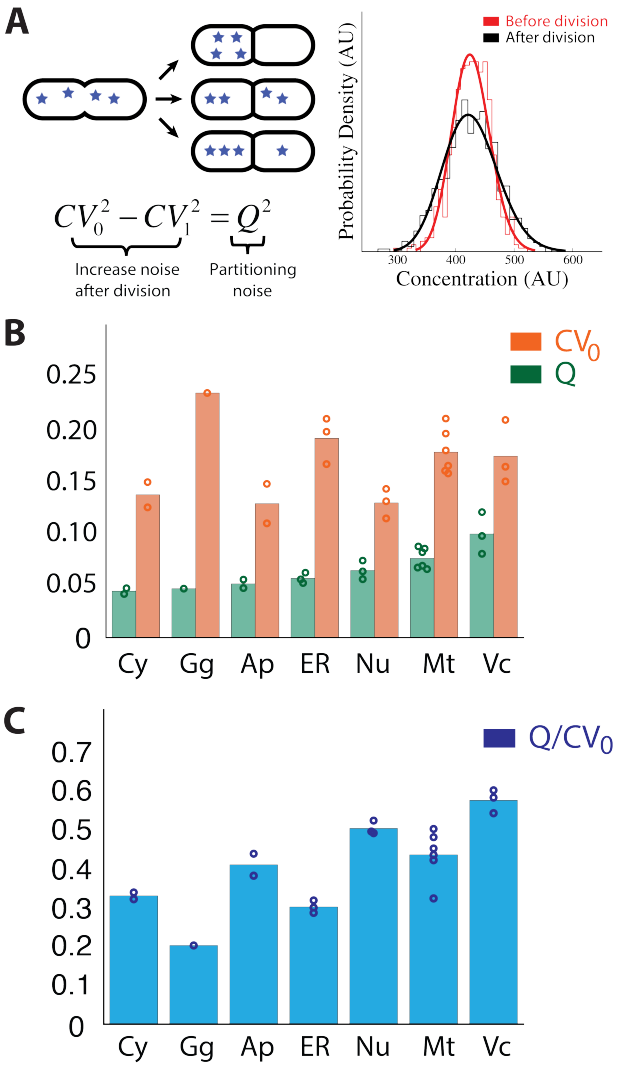


Figure III.2 | Significant partitioning noise is generated at cell division **(A)** An illustration of quantifying partitioning noise. The left schematic describes random segregation of molecules at cell division. A mother cell (left side) containing 4 molecules (blue stars) divides into two daughter cells (right side) in a probabilistic way. The daughter cells can have 0 to 4 molecules by chance even with symmetric division in volume, which elevates the variability in concentration. We define the increase noise

Figure III.2(continued) | as partitioning noise (Q) as written in the equation. The right panel describes the distributions of concentration before (red curve) and after the division (black curve) for the protein localizing into vacuoles (DH151, vma4-mGFPmut3). **(B)** The measured partitioning noises and the noise for newborns in protein concentration are plotted by localization patterns. cytoplasm (Cy), endoplasmic reticulum (ER), golgi complex (Gg), mitochondria (Mt), nucleus (Nu), vacuole (Vc), and actin patch (Ap). Each dot represents different strains, and the bars indicates the mean level of partitioning noise (green) and noise at birth (orange). **(C)** Partitioning comprises a significant portion of noise at birth in protein concentration. The ratio is calculated by dividing partitioning noise (Q) by noise at birth (CV_0).

III.4. Half of partitioning noise transmits to the next cell cycle

The very significant partitioning error that we observed for many different proteins challenges the conventional view that protein noise primarily comes from stochastic gene expression. In fact, during repeated growth and division, the errors we observe for a single division can accumulate and may even dominate the total noise. Evaluating how quickly the effects of partitioning noise resulting from instantaneous partitioning error are corrected is therefore essential.

To quantify how long partitioning noise from one division remains before it is corrected, the variability in protein concentration at certain cell cycle position needs to be decomposed into the noise originated from previous random partitioning and that from others. Separating one source of fluctuation from others is often challenging. However, our large time-series data sets allows us to bin the data, and consider the heterogeneity

observed in the small subset of individual cells that by chance had negligible partitioning errors in previous or subsequent divisions. Thus we identified the sibling pairs with insignificant partitioning errors, followed the cells over time, and compared how the protein noise changes during the cell cycle depending on whether or not there was a partitioning error. For this analysis it is important to track a large number of daughter cell pairs, which is why we modified published microfluidic devices for fission yeast, now enabling us to track the full traces of siblings and cousins for two consecutive divisions (see Chapter II). We confirmed that the growth environment was homogenous in that device as well (Supplemental Figure C.7), and that the partitioning statistics were stable over time (Supplemental Figure C.8). Thousands of cell divisions were observed for each strain, excluding strains where the signal was weak compared to the background autofluorescence, and we tracked both sister cells over at least one more cell cycle and division (Figure III.3A).

To find pairs that effectively have no partitioning errors, we quantified partitioning error for each division by calculating the discontinuity in protein concentration at division (Figure III.3A, Appendix A.2), and sorted newborn cells by it (Figure III.3B). Then, we set the maximum threshold on partitioning error that makes the closest distribution of newborns to that of mothers (Figure III.3B), and found that newborns cells with partitioning errors below the threshold have an almost identical distributions in protein concentration before and after division, meaning that the effect of partitioning error is

effectively removed upon conditioning (<4% difference between their cumulative distributions, Figure III.3C). Note that this is not a trivial result, since we compare the distribution in protein concentration but condition on the partitioning error, which is an independent variable, and yet the entire distributions became indistinguishable (Appendix A.2). We found no dependence on polarity (<3% difference in CDFs old-pole and new-pole cells), implying two daughter cells carry identical amount of partitioning noise (Supplemental Figure C.9).

Figure III.3E illustrates how vacuolar protein noise changes over the cell cycle with and without partitioning errors in the most recent division (red and purple dots, respectively), revealing a substantial difference. Even when conditioning the data on insignificant partitioning error, the noise decreases to some extent over the cell cycle. This could reflect the effects on partitioning errors from previous cell cycles.

At the end of the cell cycle, where the mass of *S. pombe* goes through a plateau phase (green shaded area in Figure III.3E and Figure III.3F) where the cell volume does not increase, both protein noise and partitioning noise stayed constant, consistent with the expectation of very little gene expression in this phase. The slight increase in fluorescence signal from GFP during this window (black curve in Figure III.3D) was indeed as expected given the fluorescence maturation time (Supplemental Figure C.10). In contrast, during growth phase, the noise coming from the random partitioning at previous division exponentially decreased to half of its value (black curve in Figure III.

3F) for all measured strains (Figure III.3G). This suggests that half of the noise created by partitioning errors are corrected during the subsequent cell cycle.

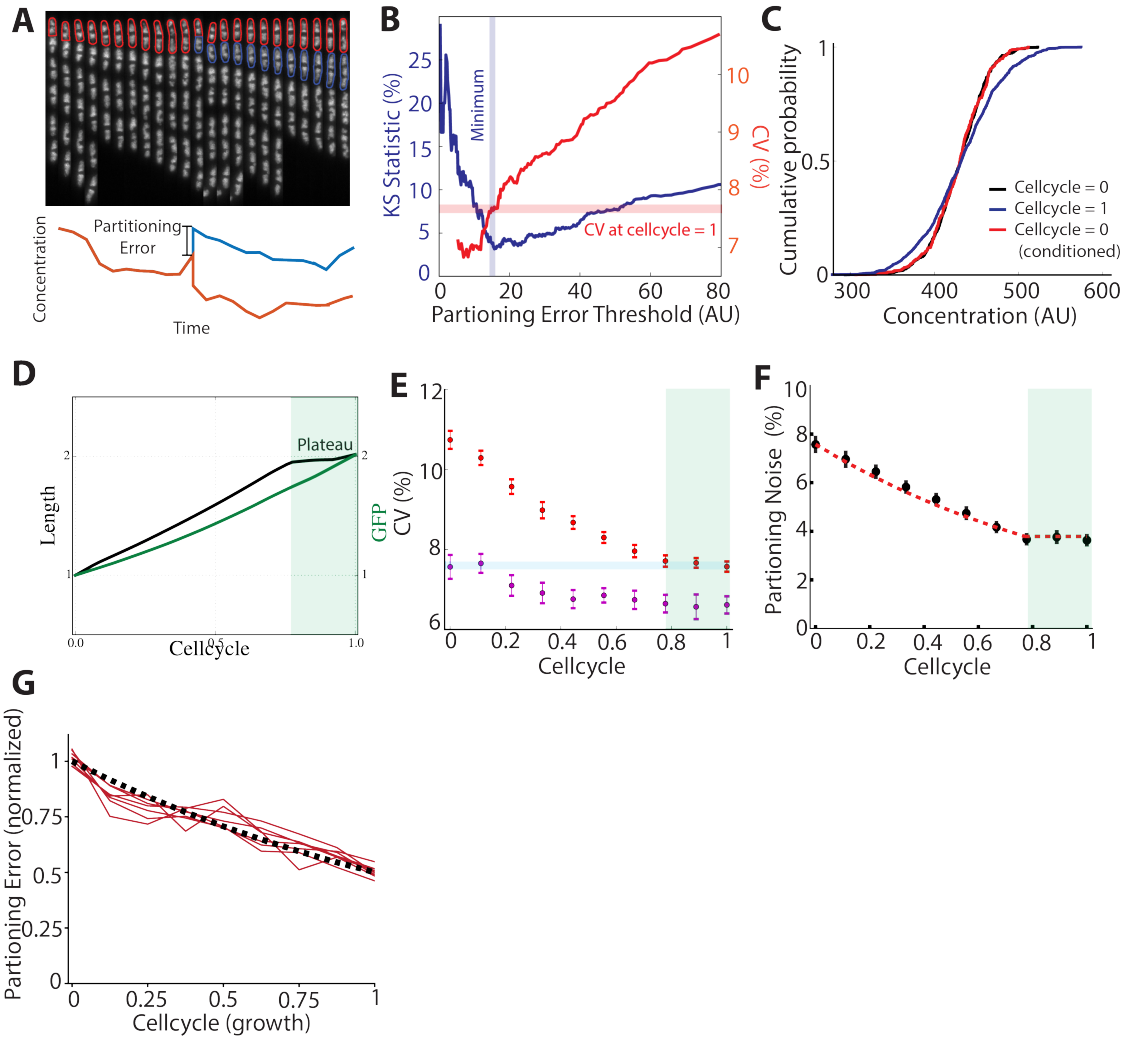


Figure III.3 | Significant partitioning noise persists over cell division Time-lapse experiment was performed for the strain (DH151, ma4-mGFPmut3) for Figure A-F, while Figure G

Figure III.3(continued) | contains the measurements for all strains. Cells were grown under rich synthetic media (PMG) at 32°C. The experiment lasts for ~3 days and images are taken every 15 minutes. **(A)** Upper panel shows a sample kymograph visualizing a mother cell and its daughter cells. The two daughter cells (red and blue) are followed after division, and their GFP concentration dynamics are plotted in the bottom panel. Partitioning error is quantified by a gap after division. **(B)** Thresholding on partitioning error removes partitioning noise. Pairs of sister cells having a partitioning error below certain threshold are collected, and Kolmogorov-Smirnov statistic (K-S statistic, blue line) and coefficient of variation (CV, red line) are calculated. The threshold with minimum of K-S statistic (blue shaded area), which makes the closest distance between cumulative distribution functions (CDF) of the conditioned pairs of newborns and mother cells, is calculated. At that threshold, the CV of conditioned newborns is similar to that of mother cell (red shaded area). **(C)** Conditioned newborns have a similar distribution to a mother cell. CDF of newborns (blue curve) show wider distribution compared to that of mother cells (black curve), indicating added variability from random partitioning, while CDF of conditioned newborns (red curve) is almost identical to that of mother cells as if there were no partitioning errors. **(D)** Average cell length (black curve) and GFP abundance (green curve) doubles over cell cycle. Plateau phase where growth stops (green shaded area) is observed for ~ 20% end of cell cycle. **(E)** Protein noise over cell cycle with and without partitioning errors (red and purple dots, respectively). **(F)** Persisting partitioning noise over cell cycle (black dots). During growth phase, it decays exponentially (red curve) to half. In contrast, during plateau phase (green shaded area), it stays constant. **(G)** Persisting partitioning noise over growth phase for all measured strains show s a similar decay pattern. The initial partitioning noise is normalized. The red curve indicates exponential curve with a half life of 0.5.

III.5. Passive transmission of partitioning noise

Deviations from an average abundance can in principle be actively corrected by compensating feedback mechanisms. For example, bacterial plasmids use negative feedback loops to correct for any deviations that arise at cell division [86], and it was reported that the initial concentration of mitochondria affects growth rate in human cells in a way that indirectly controls noise [87]. In contrast to the strong feedback we observe in the cell size control (inset in Figure III.4A), the initial protein concentrations were independent of the individual generation time of the cells, for all measured strains (Figure III.4A). Furthermore, the synthesis rate was independent of the current protein concentration (Figure III.4B). This suggests a lack of active control, i.e., the cell does not appear to sense and compensate for fluctuating abundances.

To demonstrate this we analyzed how much of the difference between newborn sister cells is corrected during cell cycle on average. Consistent with our previous observations, for all measured proteins, we found that this difference were corrected by half over cell cycle (Figure III.4C, see Appendix A.2 for more details) and that the amount added during the cell cycle was independent of the starting value. This simple mechanism for correcting noise by simply passively regressing back towards the average – which was shared by all proteins we studied with various localizations, expression level, and noise level – perhaps makes sense in the light of recent theoretical work

suggesting how very challenging and energy consuming it is to suppress noise by feedback [71], and how negative feedback control easily can have the opposite effect.

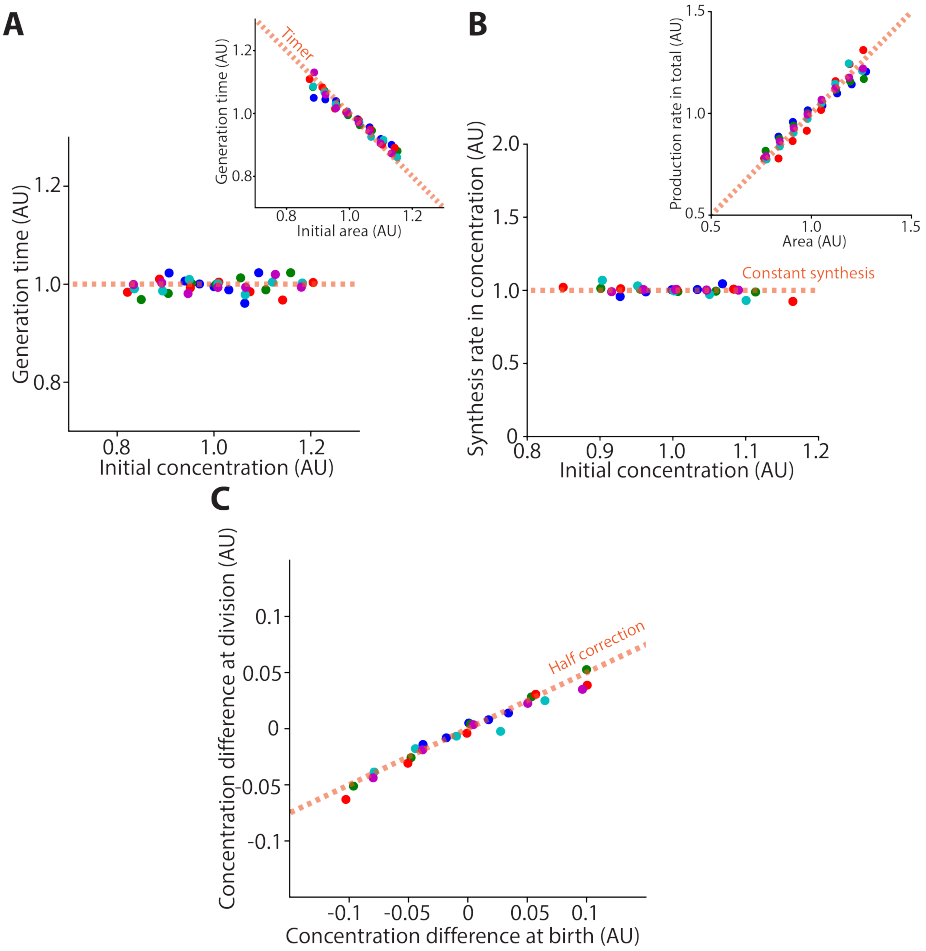


Figure III.4 | Passive transmission of partitioning noise Each dot has a unique color represented by localization patterns: cytoplasm (blue), mitochondria (red), ER (cyan), nucleus (green), vacuole (purple) **(A)** Generation time is independent of initial protein concentration. For each trace is binned with uniform width based on the distribution of initial protein concentration, and average generation

Figure III.4(continued) | times and mean initial concentrations are calculated for each bin and normalized. Inset shows similar way of plotting with initial size and generation time with orange dotted line indicating 'timer'. **(B)** Synthesis rate is independent of current protein concentration. Plots are made in a similar way to Figure (A). Inset indicates the production rate in total abundance is proportional to current cell size on average; 10% larger cell makes 10% more proteins on average. **(C)** Half of difference between sister cell is corrected during cell cycle on average. For each pair of daughter cells, the differences between them after the current division and before the next division are calculated, and normalized by mean concentration at birth. The plot is generated in a similar way as in previous graph.

III.6. Decomposing partitioning noise reveals gene expression noise

So far, we only addressed the effect of random partitioning at the most recent division, but as mentioned above we expect that partitioning noises from previous divisions also contribute to the total. To evaluate the total noise from partitioning errors, that is, the summed contributions from all previous random partitioning, one needs to know how partitioning noise decays beyond one generation. However, as observed above, the decay pattern is fully described by passive control, where the amount cells make is independent of how much they contain. For that mechanism, the noise (as measured by normalized variances) is reduced by a factor of two each cell cycle, i.e., in a geometric series (left panel in Figure III.5A). This means that the net partitioning noise can be quantified by summing up these contributions (right panel in Figure III.5A).

After computing the total noise from partitioning errors (left panel in Figure III.5B), we observed that the remaining noise is almost constant during the cell cycle, for all measured strains (Figure III.5D). In principle it is possible that this remaining noise could also come from partitioning errors in other components, e.g. mRNAs, transcription factors, polymerases etc. However, if that were the case, the noise should decrease during the cell cycle, and we observe that it is constant during the cell cycle, suggesting other mechanisms such as gene expression etc. We confirmed this result by iterative conditioned sampling, where the traces are repeatedly sampled as if the distributions are identical before and after division, thus the effects of random partitioning being iteratively washed out (Figure III.5C). This result demonstrates that random partitioning can create significant dependency of protein noise on cell cycle even if gene expression is constant over the cell cycle.

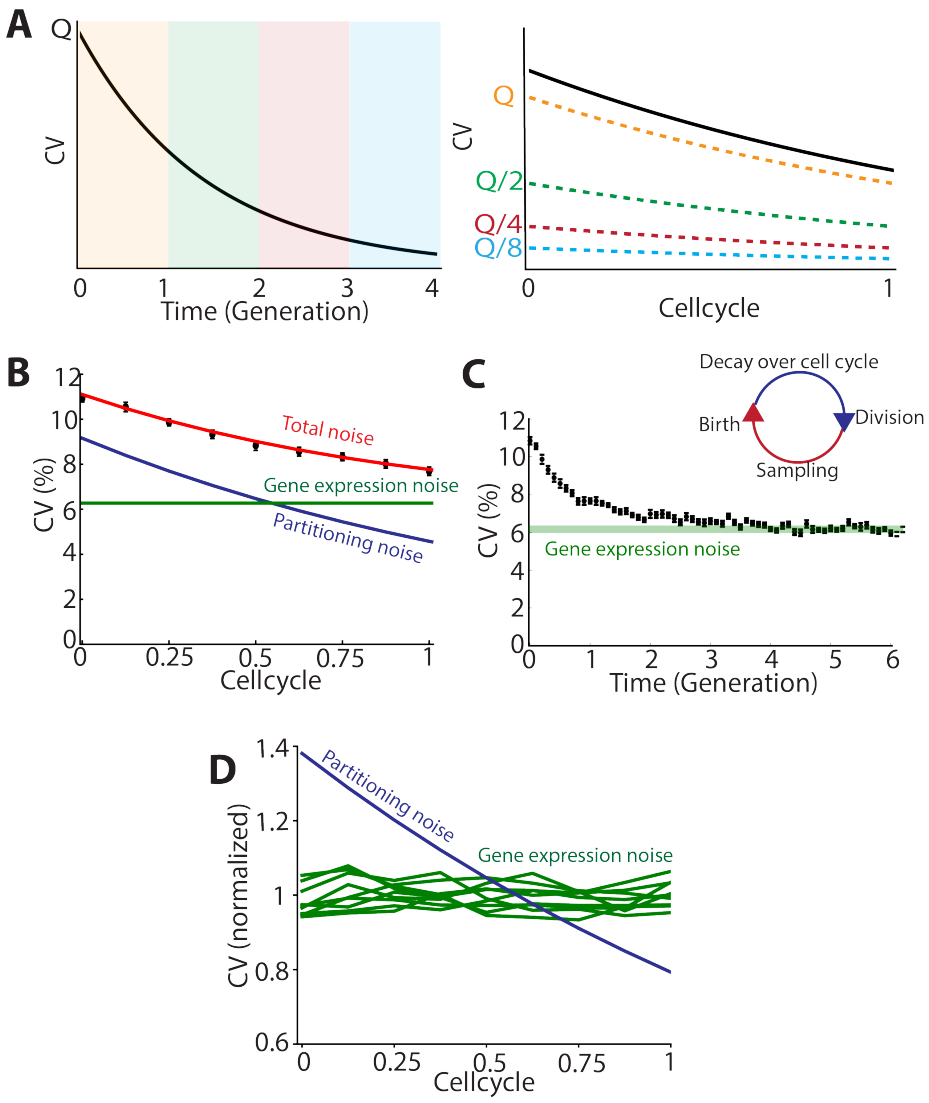


Figure III.5 | Decomposing partitioning noise reveals flat gene expression noise (A) An illustration of transmission of partitioning noise over time in a linear system. In linear system, fluctuations are corrected with same rate over time (left panel), the net contribution of partitioning noise (black curve in right panel) can be calculated by summing up the previous partitioning noises (dotted line in right panel). **(B)** An illustration of how decomposing the net partitioning noise from total noise is

Figure III.5(continued) | performed. The data used here is for vacuolar protein (DH151, ma4-mGFPmut3). The measured total protein noise (black dots) are subtracted with the net partitioning noise (blue curve), which is calculated by multiplying 4/3 to the persisting partitioning noise in Figure III.3F (Appendix A). The resulting remained noise (green dots) is almost flat over cell cycle. **(C)** Iterative conditioning reveals gene expression noise. Starting from randomly sampled 1,000 individual traces, the distribution of protein concentration at the end of cell cycle is obtained, and another 1,000 newborn cells at the beginning of cell cycle are sampled with replacement to have a similar distribution with previous distribution at the end of cell cycle, these cells are followed over cell cycle again. This procedures goes on repeatedly for 6 generations. The error bar represents standard deviation obtained from 1,000 independent simulations starting with different newborn population. **(D)** Green curves represent the remained protein noise after decomposition, normalized by its mean. The partitioning noise shows how the net partitioning nose normalized by its mean decays over cell cycle.

To further test the decomposition above, we asked if gene expression noise and partitioning noise are altered differently upon perturbation of cellular system. For each organelle except for endoplasmic reticulum which seemed lethal after cloning, we generated a *Δpom1* mutant where cells divide asymmetrically in volume, and performed a time-lapse experiment under the same conditions as above. What we inferred to be the total partitioning noise, except for cytoplasmic proteins, was indeed significantly elevated (orange curve in Figure III.6A) compared to wild type (blue curve in Figure III.6A), consistent with previous observations of non-volumetric division (Supplemental Figure C.6A), whereas what we inferred to be gene expression noise (orange shaded area in

Figure III.6A) indeed was similar to the wild type (blue shaded area in Figure III.6A), suggesting that major sources for protein noise acts in a unchanged way. Next, to see if we can perturb gene expression noise, we grew wild type cells under non-optimal temperature (25°C compared to an optimum of 32°C). In this case, partitioning noise were unchanged (green curve for 25°C and blue curve for 32°C in Figure III.6B), suggesting the number of partitioning units and partitioning mechanisms are not severely perturbed as we also directly observed, while the gene expression noise indeed was significantly increased (green shaded area for 25°C and blue shaded area for 32°C in Figure III.6B), implying altered biochemical parameters due to temperature change. Chemical parameters such as a transcription rate are often inferred with fitting a distribution in snapshot measurement to a toy model [6, 13, 23, 24]. Our results show that snapshot measurements cannot be used for such fitting without first decomposing the noise to identify how much comes from partitioning, which requires a time-lapse measurement.

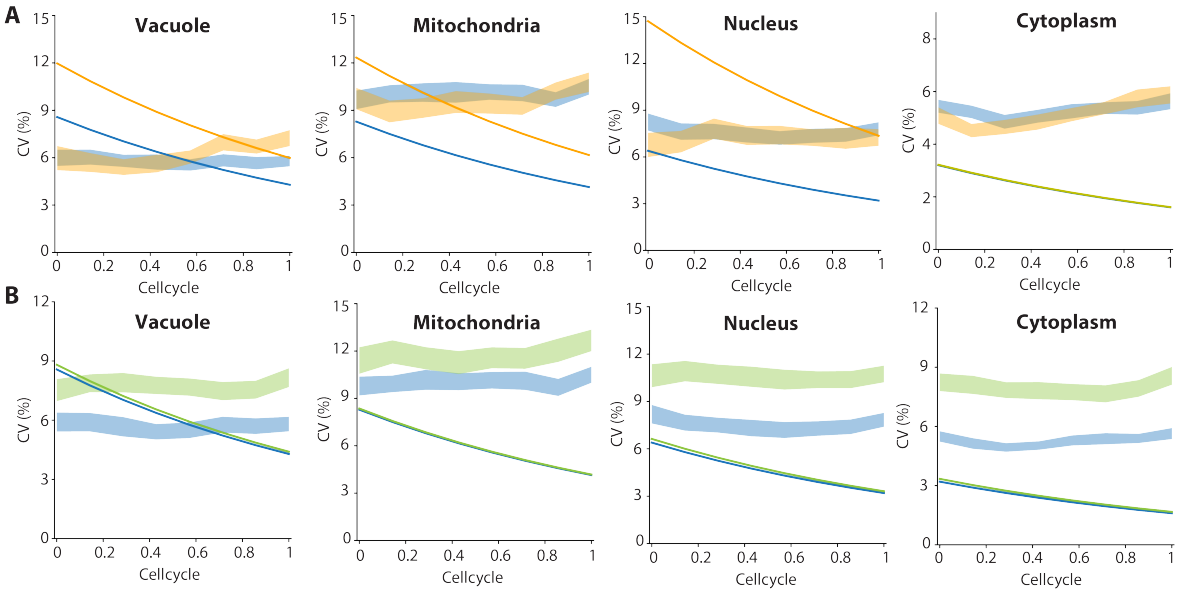


Figure III.6 | Gene expression noise and partitioning noise react differently to perturbation See Appendix A for details about the strains. Net partitioning nose and gene expression are obtained as described in Figure III.5 **(A)** Gene expression noise and partitioning noise decomposed from the protein noise for wild type and *Δpom1* mutant are compared to each other. For *Δpom1* mutants, a time-lapse measurement was performed with a identical condition to wild type cells. Solid line represents net partitioning noise, and shaded area represents gene expression noise (orange for mutants and blue for wild types). **(B)** Gene expression noise and partitioning noise at different temperature noise are compared. At 25°C, images were taken every 30 minutes (15 minutes at 32°C) since cells were grown slower. Solid line represents net partitioning noise, and shaded area represents gene expression noise (blue for 32°C and green for 25°C).

III.7. Discussion and outlooks

Protein concentrations in cells can exhibit significant variability even under homogenous environments [1, 11, 13, 23, 24, 88]. One of the main goals in this field has been to unravel the major sources of the heterogeneity [48, 49, 89], as well as to understand control circuits that exploits or compensates such 'noise' [45, 50, 51, 58, 62, 90]. Many previous studies have attributed gene expression noise to inherent transcriptional and translational process, with models that ignore the effect of random partitioning at cell division. When testing such models, the most common way to measure protein noise has been to quantify protein levels in a snapshot of an asynchronous cell population [13, 24, 25]. However, such assumption is rarely validated with a time-lapse measurement by showing the protein noise indeed stays constant over cell cycle. Furthermore, while this assumption might intuitively hold for bacterial cells where cellular components are efficiently homogeneous, this might be not true for eukaryotes with diverse cellular compartments since even high copy proteins can have a low effective numbers of partitioning units at cell division.

In this study, we challenge this view. We systemically measured the difference between in protein noise level as a function of the cell cycle and explicitly measured partitioning errors at cell division. The results showed a significant increase in protein noise immediately after division even for high copy proteins. By comparing this increase in noise or partitioning noise to the total protein noise after division, we showed

partitioning noise at cell division indeed contribute significant portions of protein noise, ranging from 33% of cytoplasmic proteins to 57% of vacuolar proteins. We further found that both partitioning noise and total protein noise was very similar between proteins that localize in the same organelles, but different between organelles, suggesting that localization pattern could be a major determinant of protein noise.

We then asked how quickly partitioning noise decays over the cell cycle, motivated by the fact that random partitioning from previous divisions could accumulate. To this end, we developed a mother machine for fission yeast that enables us to follow the protein dynamics of sister cells over time in a high-throughput manner. By following traces of newborns, we first observed that the protein noise decreases significantly over cell cycle. This result suggests that protein noise measured in a snapshot experiment is contingent on the distribution of cell cycle in a population because the level of protein noise depends on when it is measured during cell growth. Considering that exponentially growing cells contain twice as many newborn as dividing cells, partitioning noise can be more weighted compared to constant gene expression noise.

By sorting pairs of sister cells according to their partitioning errors, particularly comparing the small sub-population of such pairs that by chance did not display significant errors in the last division to the population overall, we found that the impact of partitioning errors on cell heterogeneity, as measured by the normalized variance, decays exponentially to half of its starting value over cell cycle, for all measured proteins with

various protein levels and different localizations. This is precisely as expected for passive noise control mechanisms, where errors are eliminated by regression to the mean without active compensation or feedback. This in turn enables us to determine how the partitioning noises from previous divisions accumulate at each cell cycle time-points, to account for the full effect of partitioning errors rather than just the most recent event. After decomposition, we also observed that the remaining noise from other processes in the cell was stable over the cell cycle, suggesting it is mainly from random gene expression as opposed to partitioning errors in other components.

This observation suggests that noise in upstream components either does not efficiently transmit to the protein, or that the noise in those components are not dominated by partitioning noise. Our lab observed a similar phenomenon in *E. coli*, so this might stem from general principle. One possibility is the remaining noise comes from noisy but short-lived mRNA fluctuations, as suggested in the literature. Our current effort is to understand this issue by combining datasets from two distinct organisms, bacteria and yeast.

We further showed that only the gene expression part of the noise was elevated when perturbing chemical parameters in a cell by temperature shifts. In contrast, only the partitioning noise was changed when making cell divide asymmetrically with a division-control mutant. Without the decomposition, these effects would be harder to interpret.

Overall, this work reveals the great significance of random partitioning on protein noise, which has been largely ignored as a major source for molecular noise in the numerous studies that instead attributed expression noise to stochastic production and degradation.

Appendix A: Methods

A.1 Chapter II experimental methods

Strain table

Strain	Genotype	Parent
DH0	*ura4-D18, leu1-32 h-*	FWP172
DH60	*leu1::Padh1-GST-NES-mCherry-leu4+ leu1-32 ura4-D18 h-*	DH0

Mother machine chip preparation

The silicon master wafer was briefly cleaned with a pressured airgun and placed flat on the top of folded aluminum foils. Dimethylsiloxane monomer (Sylgard 184, Dow Corning) was mixed in a 5:1 ratio with a thermal curing agent with PDMS mixer, and ~30 ml was poured onto the wafer. To remove bubbles within mixture, degassing was performed with a vacuum pump for ~1 hour. Mixture was then cured at 65°C for 1 day, peeled off, and cut into individual devices. Holes to inlets and outlets were made using a biopsy punch, and residual debris in the device and a prepared cover slip were removed with scotch tape. On the day of experiment, each device was then bonded to cover slip using oxygen plasma treatment (45 seconds at 75W with O2 at ~170mTorr) and placing on a plate heater at 95°C for 3 minutes. We noticed much higher efficient loading of cells when chips were used immediately after bonding.

Open ended chip preparation

We noticed the cell channel in this device is prone to collapsing to cover slips when following the above protocol because it has much longer structure to support. To address this issue, we changed the protocol to make PDMS stiffer so that it is less vulnerable to structural deformation. The changes made were that PDMS mixture was made in 3:1 ratio rather than 5:1 and that it underwent the second round of incubation at 80°C for 3 days after peeling and punching steps. Also, the plasma was treated longer (90 seconds). Other steps were identical.

Cell preparation

10 ml of *Schizosaccharomyces pombe* cells were grown until reaching early exponential phase (OD~0.2) in SC+PMG media or in YES media at 32°C, pelleted by centrifuging at 3,000g for 3 minutes, and resuspended in 500 ul of fresh medium. In the meantime, the device was filled with fresh media and incubated at 32°C for 10 minutes. We noticed passivating with bovine serum albumin (BSA) made cells look sick after loading, so did not use it. Cells were then injected through inlets with slow pipetting after putting another pipette tips in outlet to prevent cross-contamination between different lanes. The chip was then mounted on a custom mounting platform that can fit into a tabletop centrifuge, centrifuged at 3,000g for 10 minutes, and briefly imaged to check if cells were well loaded under a microscopy. Tygon tubings (VWR) were inserted to inlets and outlets.

Then, syringes containing medium was connected to tubings and mounted on a syringe pumps (Harvard Apparatus). Initially, flow rate was ~ 50-80 ul/min for ~1 hour to clean unloaded cells, and changed to 15-20 ul/min after confirming that inlets and outlets were clean.

Microscopy and image acquisition

Imaging was performed using a Nikon Eclipse Ti inverted microscope equipped with an Orca R2 (Hamamatsu) camera, a 60X Plan Apo oil objective (NA 1.4, Nikon), an automated stage (Ludl), a Lumencor SOLA fluorescent system, and a custom incubator that makes the imaging components stay at 32°C. We used a custom MATLAB program to control the equipment through microManager (*A. Edelstein 2010*). For fluorescence image acquisition, we used the following filter sets: GFP (Semrock GFP-1828A), mCherry (Semrock mCherry-B). The experiment for open ended device was performed with a 20X Cfi Super Fluor 20x dry objective (NA 0.75), and phase contrast images were obtained with with Te-C Lwd Ph2 Module. Before starting image acquisition, cells were allowed to grow within the imaging setup for >14 hours to make sure that they adapt to a new environment. The typical exposure time were 80-200ms and z-stacking (5 slices with 0.8 um intervals for GFP channel and 3 slices with 1 um for RFP channel) was performed to cover the total signal inside the cell. For GFP channel, 2x2 binning was used to increase the signal to noise ratio. To correct focal drift over time, the initial image in RFP channel for each lane underwent z-stacking (16 slices with 0.5 um intervals) and the z-

position with highest summed signals was selected. This position was not used for subsequent analysis. We set one sacrificial for the last position near the edge of coverglass, made it stay there between consecutive rounds of image acquisition, and provided oil to an objective if needed.

Image processing

Images in RFP channel, which served as a segmentation marker, were summed over z axis. We applied a threshold obtained with T-point method to z-summed images, filtered out masks by their sizes and morphologies, and the resulting segmentation masks were refined with local profiles of intensities. Images was then again summed over y axis to find the locations of the cell channels, and cut into sub-images that contain each cell channel with segmented cells. For each segmented mask, the following quantities were measured: frame number, area, length in major axis, length in minor axis, the orientation, the eccentricity, the xy centroid positions in a field of view, and the number of masks above it (relative position). Images in GFP channel were then overlaid with binary segmentation masks, and the following quantities were calculated: the total intensity inside a mask (a proxy for total abundance), the mean intensity inside a mask (a proxy for concentration), and the z-position with a maximum total intensity (focal plane). The masks with different focal planes were disregarded later but it was minimal (<0.5%). We calculated the moving average curve of GFP concentrations over frames, and only proceed with dataset after the bleaching effect was stabilized. For open ended

microfluidic device, the phase contrast images were subtracted by the images without cells to remove signal comes form edges in a device, then cell tips were identified by similar method. In this device, we observed cell grow linearly without tilting in the cell channel, hence the width of cell channel was used as a proxy of cell width.

Lineage tracking

After having segmentation masks for each cell channel, masks in consecutive frames were compare from the top. Masks were iteratively matched from the top end with comparing the cell size between them. Cell division was identified if observing a significant decrease in cell size. For each newborn cell, the ratio of its size to that of a mother was calculated and used as an potential indicator of segmentation error if it is not within the range of (0.4, 0.6). When a new division events was detected, each daughter cell was assigned with a unique lineage ID, a lineage ID of its sister, a lineage ID of its mother, an indicator whether it has old pole or not, and an indicator if it completed a next division. We only proceeded with lineages that completed a division and a division time was calculated, which was also used as an indicator of errors. After completing a first round of lineage tracking procedures, a growth curve of each lineage was regressed with two linear lines (growth phase and plateau phase) and an indicator for segmentation error was assigned if there is a significant deviation from the regression lines. All above steps were automated with a custom MATLAB scripts. Kymographs were then generated, and potential errors detected in the above steps were manually inspected with a custom GUI.

Flat field correction

We observed a spatial patterns of mean fluorescence intensities. Since this was reproducible independent experiments, we concluded this is due to uneven illumination of light sources. We computationally correct this flat field by assuming every cells should be statistically identical over space. To this end, all segmented masks including lineages that did not complete divisions were collected and mean fluorescences within each mask were plotted over x and y axis. We found a linear regression fits well the mean curve over space. Indeed, we found our stage holder is tilted in a linear way. Hence, we divided each fluorescence concentration by the linear curve for each segmentation mask.

A.2 Chapter III experimental methods

Strain table

Strain	Genotype	Localization	Time-lapse
DH110	*eri1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Cytoplasm	X
DH111	*int6-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Cytoplasm	O
DH112	*sce3-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Cytoplasm	O
DH113	*tif452-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Cytoplasm	X
DH114	*wis1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Cytoplasm	X
DH115	*cnx1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Endoplasmic reticulum	O
DH116	*ero12-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Endoplasmic reticulum	X
DH117	*ogm2-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Endoplasmic reticulum	X
DH118	*ogm4-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Endoplasmic reticulum	X
DH119	*ost1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Endoplasmic reticulum	O
DH120	*aur1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Golgi	X
DH125	*coq7-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Mitochondria	X
DH126	*gut2-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Mitochondria	X
DH127	*hsp10-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Mitochondria	X
DH128	*isu1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Mitochondria	O
DH129	*lon1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Mitochondria	X
DH130	*mrpl8-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Mitochondria	X
DH131	*mrps5-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Mitochondria	X
DH132	*qcr7-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Mitochondria	X
DH133	*shy1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Mitochondria	X
DH134	*sod2-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Mitochondria	O
DH137	*cut15-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Nucleus	X

DH139	*nup132-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Nucleus	X
DH140	*pla1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Nucleus	O
DH141	*prp43-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Nucleus	O
DH142	*skb15-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Nucleus	X
DH149	*abc2-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Vacuole	O
DH150	*vac8-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Vacuole	X
DH151	*vma4-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Vacuole	O
DH153	*arc3-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Actin patch	X
DH156	*fim1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 h-*	Actin patch	X
YJ052	*leu1-32 ura4-D18 pom1△::KanMX h-*	N/A	O
YJ067	*sce3-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 pom1△::KanMX h-*	Cytoplasm	O
YJ068	*isu1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 pom1△::KanMX h-*	Mitochondria	O
YJ069	*qcr7-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 pom1△::KanMX h-*	Mitochondria	O
YJ070	*prp43-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 pom1△::KanMX h-*	Nucleus	O
YJ071	*vma4-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18 pom1△::KanMX h-*	Vacuole	O
YJ078	*shm1-mGFPmut3 leu1::Padh1-GST-NES-mCherry-leu1+ leu1-32 ura4-D18*	Mitochondria	

Transformation protocol

We followed the protocol from Bahler lab (*Bahler, J 1998*). Briefly, cells were grown until OD~0.5 in YES media, 5ml of culture was pelleted by centrifuging at 3,000g for 3 minutes. Pellets were washed twice with 1ml of LiAc/TE buffer and resuspended in 100ul of LiAc/TE buffer. The mixture was combined with boiled 2ul of salmon sperm DNA and 10ul of linearized plasmids, incubated at room temperature for 10 minutes, mixed with 260ul of PEG4000/LiAc/TE buffer with gentle pipetting, incubated at 32˚C

for 60 minutes. We noticed using fresh PEG4000 significantly improves the efficiency of transformation. Then, 43 ul of DMSO was added and incubated at 42˚C for 5 minutes in pre-heated water bath. Pellets were harvested by centrifuging for 3 minutes at 3,000g, resuspended in water, and plated on adequate selection plate.

Seamless integration

Refer to a published protocol for more details (D Landgraf, 2016). Briefly, ura4 cassettes flanked by homologous regions for target genes were first transformed and selected on SC+PMG plate. Then, mGFPmut3 cassettes flanked by the same homologous regions are transformed again and selected on SC+PMG plate with 2% 5FOA plate.

Pom1 mutant construction

We amplified pom1*Δ*::KanMX cassettes in YJ052 (a gift from Fred Winston) with using colony pcr, transformed it, and selected on YES+Kan plate to generate DH112, DH128, DH132, DH141, DH151 to generate YJ067, YJ068, YJ069, YJ070, YJ071, respectively.

Confocal microscopy imaging protocol

A confocal microscopy (Nikon TE2000U) at Nikon imaging center at Harvard Medical school was used. It was equipped with Yokagawa CSU-10 spinning disk head, Coherent 3W water-chilled Krypton Argon laser, Hamamatsu ORCA-AG cooled CCD camera, and Prior Proscan II motorized stage. Cells were grown in SC+PMG media to OD = 0.3~0.5, and 60 ul was placed on an agarose pad containing SC+PMG. After the media is dried

out, clean coverslip was mounted and sealed. Measurement was performed at room temperature. 200 ms and 500 ms exposure times were applied to RFP and GFP channels, respectively. Both channels were measured with z-stacking (13 slices with 0.5 um intervals) with help of autofocus algorithm in DIC channel (Metamorph). To reduce the noise level, 2x2 binning was applied for both channels. Also, a fluorescent flat plastic was used to infer flat-field profile.

Mother machine imaging protocol

It was done in a similar way as described in Appendix A.1. Images were taken every 15 minutes for wild type cells, 10 minutes for pom1 mutant cells, and 30 minutes for the measurement performed at 25˚C. For all measurements, there were >8 measurement per cell cycle on average.

Image processing

It was done in a similar way as described in Appendix A.1. In pom1 mutants, cells can be in smaller size compared to wild type due to asymmetric division. Accordingly, we changed the parameters for segmentation and detecting segmentation failure. Similarly, parameters changed for the measurement at 25˚C considering the different cell size and generation time.

Calculating partitioning error at cell division

With detected a pair of just-divided sister cells, we calculated the coefficient of variation ($CV = std(x)/mean(x)$) in protein concentration for newborns ($CV_{divided}$). Then, we summed up areas and total GFP abundances of sister cells to represent those of a mother cell, and calculated $CV_{dividing}$. Then, the partitioning error (Q) was calculated using the equation $CV^2_{divided} = CV^2_{dividing} + Q^2$. Note that Q = 0 if there were no randomness at cell division. The portion of partitioning noise to the noise of newborns were calculated as $Q/CV_{divided}$ for each strain.

Data processing

We only proceeded with a pair of sibling lineages that completed a division. For each lineage, plateau phase and growth phase were detected with fitting with two lines to the trace of cell length. To plot the traces of quantities over cell cycle in a normalized time scale, the following steps were taken. First, to separately align the two distinctive phases, the transition point from growth to plateau phase was detected with fitting. Then, data points in each phase were replotted in a time scale normalized by its mean duration, and two phases were reconnected. Finally, we resampled data (10 data points per cell cycle) by interpolation between two nearby points.

Conditioning newborns with minimal partitioning error

For each lineage, we calculated concentration discontinuity between its mother and a daughter at cell division. We define this quantity as Δ. Now let c_0 be the concentration right after the division (cellcycle = 0), and c_1 just before the division (cellcyle = 1), hence $\Delta = c_0 - c_1$. Then, our goal is to find Δ that effectively makes $c_1 \sim c_0$ where the contribution of random partitioning is minimal. To this end, we calculated the following Kolmogorov-Smirnov statistic, which measures the difference between two distributions, for certain threshold δ: $\sup | F_1(x) - F_0(x \ given \ | \Delta | < \delta) |$, where F(x) is cumulative distribution function (CDF) for protein concentration x. The former represents CDF for c_1 (for all mothers) and the latter is for c_0 for newborns whose absolute partitioning error Δ is less than δ, thus this value should be minimal when two distributions is closest to each other. We plotted the above statistic over a range of δ, picked the threshold that resulted in the minimum, and used it to sample lineages. We confirmed the mean and CV of distributions for newborns after conditioning matches to those for dividing cells.

Calculating persisting partitioning noise over cell cycle

For both of unconditioned lineage traces and conditioned traces, 500 traces were randomly sampled and the coefficient of variations over cel cycle were calculated. Sampling procedures were repeated for 100 time to have standard error of the mean. For

each sampled traces, the persisting partitioning noise were calculated with using the following equation: $\sqrt{CV_{unconditioned}{}^2 - CV_{conditioned}{}^2}$.

Estimating synthesis rate

Let X be the total abundance of some species and V be the cell volume, then the concentration $c = X / V$. Taking a time derivative of c, the following is obtained.

$$\frac{d}{dt}c = \frac{1}{V}\frac{dX}{dt} - \frac{1}{V^2}\frac{dV}{dt}$$

The first term measure the synthesis rate in concentration and the latter represents a dilution rate by cell growth. As an estimate of the first term, we calculated

$$synthesis\,rate = \frac{2}{V(f)+V(f+1)}[X(f+1)-X(f)]$$

where f is a certain frame number and X(f) is the total GFP abundance at frame f.

Partitioning error at cell division

Consider two daughter cells right after cell division with the volume of L and R, and having the total amount of some species of X and Y, respectively. In other words, each daughter cell have the concentration of $c_L = X / L$ and $c_R = Y / R$, respectively. Then, the concentration discontinuity between a mother and daughter cell, which we refer to as partitioning error (Δ) can be calculated as: $\Delta = \frac{X}{L} - \frac{X+Y}{L+R} = \frac{R}{L+R}(\frac{X}{L} - \frac{Y}{R}) = p_R(c_L - c_R)$

where $p_R = \frac{R}{L+R}$ measures the ratio in volume. Thus, partitioning error is proportional to the difference in concentration between sister cells. Note that partitioning error is zero when there is no difference between sister cells, which is a typical assumption of models that does not account for random partitioning. In such models, protein concentration is in stationary system because there is no perturbation at cell division.

Error correction in passive dilution model

A passive dilution model can be described as follows.

$$x \xrightarrow{\lambda} x+1$$

$$x \xrightarrow{\frac{\ln 2}{T}x} x-1$$

where x is the number of molecules per given area, λ is a constant synthesis rate, and T is a generation time. On average, this system is governed by the following equation:

$$\frac{d}{dt} < x(t) >= \lambda - \frac{\ln 2}{T} < x(t) >$$

Then, for a deviation from the mean $\delta = x - < x >$, it can be shown that

$$< \delta(t) \mid \delta(0) >= \delta(0) 2^{-\frac{t}{T}}$$

Thus, any deviation is corrected by half on average in one generation. Now consider a pair of siblings whose initial concentration at birth are $c_L(0)$ and $c_R(0)$, respectively, and define the difference between them as $\Delta = c_L - c_R$. Then, it can be shown

$$< \Delta(t = T) \mid \Delta(t = 0) >= \frac{1}{2} * \Delta(t = 0)$$

which was used in Figure III.4C.

Transmission of partitioning noise in passive dilution model

$$x \xrightarrow{\lambda} x+1$$
$$x \xrightarrow{(\ln 2)x} x-1$$

While the above simplified model (for simplicity, we assumed T = 1) is one of the most popular models describing gene expression noise, the effect of random partitioning is often ignored. Here, we consider the effect of random partitioning by including periodic perturbations at each division and compare the result to a conventional model. Specifically, the discontinuity in protein concentration arises at each division as follows:

$$x_{cellcycle=0} = x_{cellcycle=1} + \Delta$$

where Δ represents a random partitioning error. Over repeated random processes, the system reaches a cyclostationary state where the statistics are stationary at a given time in the cell cycle, such that

$$V_{cellcycle=0} = V_{cellcycle=1} + Q^2$$

where V is a variance function and Q is a partitioning noise. Note that Q = 0 when ignoring random partitioning and the system reaches stationarity where all the statistics are identical for any cell cycle positions.

Consider we measure V(t) for current cell cycle position $t \in [0,1]$ and let Δ_{-i} be the partitioning error at -ith previous division. Then, we define 'gene expression noise' as $V_{stationary} = V(t \mid \Delta_0 = \Delta_{-1} = \cdots = \Delta_{-\infty} = 0)$, which measures the remained noise given that all of previous divisions were perfect; the only remained randomness is stochasticity in synthesis and degradation. This gene expression noise is a stationary state for both of models with and without random partitioning since the underlying systems are identical, hence any deviations in variance moves towards gene expression noise. Since any deviation from a mean of the stationary state is corrected with a rate function of $f_{mean}(t) = 2^{-t}$ as shown in the previous section, the deviation from the variance of stationary state is corrected with a function of $f_V(t) = (2^{-t})^2 = 4^{-t}$. Then, the persisting partitioning noise from 0th division can be calculated as follows.

$$
\begin{aligned}
&\textit{Persisting partitioning noise from } 0\textit{th division} \\
&= V(t) - V(t \mid \Delta_0 = 0) \\
&= [V(t) - V_{stationary}] - [V(t \mid \Delta_0 = 0) - V_{stationary}] \\
&= [V(0) - V_{stationary}] * f_V(t) - [V(0 \mid \Delta_0 = 0) - V_{stationary}] * f_V(t) \\
&= [V(0) - V(0 \mid \Delta_0 = 0)] * f_V(t) \\
&= Q^2 4^{-t}
\end{aligned}
$$

Noe that the coefficient of variation (CV) measures standard deviation rather than variance, hence the above result should be square-rooted to compare with the results in Figure III.3C.

Separating gene expression noise from partitioning noise

Continuing with the above section, we further calculate the persisting partitioning noise from -ith division to calculate the net contribution of partitioning noises from previous divisions. With similar procedures, the below can be obtained.

$$
\begin{aligned}
&\textit{Persisting partitioning noise from } -i\textit{th division} \\
&= V(t \mid \Delta_0 = \cdots = \Delta_{-i+1} = 0) - V(t \mid \Delta_0 = \cdots = \Delta_{-i} = 0) \\
&= [V(t \mid \Delta_0 = \cdots = \Delta_{-i+1} = 0) - V_{stationary}] - [V(t \mid \Delta_0 = \cdots = \Delta_{-i} = 0) - V_{stationary}] \\
&= [V(t \mid \Delta_0 = \cdots = \Delta_{-i+1} = 0) - V_{stationary}] * f_V(t+i) - [V(t \mid \Delta_0 = \cdots = \Delta_{-i} = 0) - V_{stationary}] * f_V(t+i) \\
&= [V(t \mid \Delta_0 = \cdots = \Delta_{-i+1} = 0) - V(t \mid \Delta_0 = \cdots = \Delta_{-i} = 0)] * f_V(t+i) \\
&= Q^2 4^{-(t+i)}
\end{aligned}
$$

The last equality holds because of cyclostationarity assumption. To obtain the net contribution, we sum up all persisting partitioning noises from previous divisions.

$$
\begin{aligned}
&\textit{Net partitioning noise} \\
&= \sum_{i=-\infty}^{0} [V(t \mid \Delta_0 = \cdots = \Delta_{-i+1} = 0) - V(t \mid \Delta_0 = \cdots = \Delta_{-i} = 0)] \\
&= V(t) - V_{stationary} \\
&= \sum_{i=-\infty}^{0} Q^2 4^{-(t+i)} \\
&= \frac{4}{3} Q^2 * 4^{-t}
\end{aligned}
$$

Rewriting the above equation yields decomposition equation of noise at cellcycle t as follows: $V(t) = V_{stationary} + \frac{4}{3} Q^2 * 4^{-t}$, where the former term is gene expression noise while the latter term represents net partitioning noise. Figure III.5B shows these terms as a coefficient of variation. In Figure III.5D, we calculated gene expression noise by subtracting the inferred net partitioning noise from the observed total protein noise.

Detailed explanation of Figure III.5C

The alternative way to have gene expression noise is by conditioning zero partitioning errors forward because gene expression noise is a variance at the stationary state. In other words, $V_{stationary} = \lim_{t\to\infty} V(t \mid \Delta_0 = \Delta_1 = \cdots = \Delta_\infty = 0)$. To this end, the following procedures were repeated:

Step I: Start with randomly sampled 1,000 traces.

Step II: Calculate the noise over cell cycle.

Step III: For each trace in the previous step, sample a trace from the entire traces that minimizes the discontinuity between two traces and connect them. This way, we artificially remove partitioning errors. Go to Step II.

The above procedures were performed 1,000 times independently.

Appendix B: Fabrication of Microfluidic Devices

The devices were fabricated at the Center for Nanoscale Systems (CNS) at Harvard University. The autocad files containing the detailed designs for photolithography masks can be provided upon request. Although the designs of layers are different between mother machine and open-ended device, their vertical dimensions in z-axis are identical, hence the following protocol can be used either cases. The device contains four layers for alignment marker, a shallow media layer (a side channel for open-ended device) that provides nutrients around cells, cell channels where cell grows, and the main media layer where inlets and outlets are.

We adapted a standard protocol that utilizes repetitive patterning of photo-curable polymers with UV photolithography. The typical procedures for fabricating three dimensional structures are 1) coating photo-curable polymers on a wafer with a desired height by spinning, 2) applying patterned UV on it, 3) developing uncured polymers, and 4) repeat the steps with another patterns. However, we noticed coating photoresist on structured surface made from a previous step often produced nonuniform coating and was more sensitive to previous steps such as the duration of hard-baking and development time. This problem was particularly severe for our device because it has multiple layers and because the success of experiments requires precise and uniform dimensions. To address this issue, we developed the protocol with skipping development steps when

stacking up multiple layers so that spinning can be performed on flat surface. Since the spinning speed settings are optimized on the surface of a wafer rather than photoresist, we accordingly optimized these parameters for our purpose.

The following abbreviation was used to describe program settings in a spin coater:

speed (rpm)/acceleration (rpm/sec)/time (seconds).

Layer for alignment marker

The purpose of this step is to have alignment patterns by depositing copper to a wafer so that it is visible when aligning multiple masks through photoresist. Proceeding with multiple wafers for this step is recommended since it is a time-consuming step but can be performed in a parallel way.

1) Place a new 3” Si wafer (380 um TEST grade wafers from University Wafer) on a spinner and clean it by applying acetone and then isopropanol while spinning (3000/300/60)

2) Dehydrate the wafer on a hot plater at 150°C for 10 minutes.

3) Place the wafer onto the spin coater and wait until it cooled down.

4) Pour S1813 photoresist (Shipley) to cover ~2/3 of the wafer and spin with 3000/300/60.

5) Bake the wafer on a hot plater at 115°C for 1 minutes. (Be careful not to touch the area where alignment marker will be placed while placing it to different places.)

6) Place the wafer on mask aligner (Suss MJB4) and apply a total power of 80 mJ/cm^2 with the vacuum contact.

7) Develop the wafer for ~60 seconds in CD26 developer (Shipley), rinse it with deionized water, and remove residual water with an air gun.

8) Check if the alignment marker are thoroughly developed under an inverted microscopy. If not, go back to 7).

9) To completely remove residual uncured photoresists, place the wafer in oxygen plasma barrel (Anatech LTD SCE106), and run the program with the following settings: 45 seconds at 30 sccm O_2 with a power of 80W.

10) To deposit a thin layer (~300nm) of Cu layer, we used thermal evaporator (Sharon) with the following settings: <10-6 torr of base pressure, <3*10-6 torr of working pressure, ~4 Å/sec of deposition speed.

11) To peel off S1813 photoresist, place the wafer in ~50 ml of remover PG (MicroChem) on top of a hot heater at 80°C and leave it overnight.

12) Next day, replace the remover with fresh one and incubate again for >2 hours. Most of Cu layer should be peeled off this time.

13) Place the dish containing the wafer to a bath sonicator, and sonicate for 5 minutes. If remained Cu other than alignment marker is visible, go back to 12)

14) Rinse the wafer with isopropanol thoroughly and blow dry with an air gun.

15) The resulting wafers can be stored at room temperature before proceeding to the next steps.

Shallow media channel

The desired height in this step is 4.2 um. We recommend to use fresh SU-8 photoresist since the old one was observed to have different viscosity, probably due to evaporations. Before starting this step, testing the speed setting for desired height with using a blank wafer is highly recommended since preparing the wafers with alignment marker takes >2 days.

1) Place the wafer with alignment marker on a spinner and clean it by applying acetone and then isopropanol while spinning (3000/300/60)

2) Dehydrate the wafer on a hot plater at 150°C for 10 minutes.

3) Place the wafer onto the spin coater and wait until it cooled down.

4) Pour SU-8 2002 photoresist (Shipley) to cover ~2/3 of the wafer and spin with 1800/300/45 with an initial acceleration of 500/100/5.

5) Bake the wafer (in this order), 1 minute at 65°C, 3 minute at 95°C, and 1 minute at 65°C.

6) Place the wafer on mask aligner (Suss MJB4), align the photoresist mask carefully, check if the alignment error is less than 1um after applying vacuum contact, and apply a total power of 85 mJ/cm^2 with the vacuum contact.

7) Bake the wafer (in this order), 1 minute at 65°C, 10 minute at 95°C, and 1 minute at 65°C. Note that we intentionally extended the duration of this step so that the surface is not sticky when coating another layer of photoresist.

8) Slowly cool the wafer down.

9) Apply a small amount of SU-8 developer with a swab to the edge of wafer where there are no features.

10) For developed region, measure the height using a profilometer. The expected height is ~2.2 um.

Cell channel

The desired height in this step is 2.0 um. We noticed the height was generally lower on the surface of photoresist compared to on a wafer, probably because of different friction

parameters. Slowing down the spinning speed can be tried below 1,500 rpm/sec is not good for uniform coating. Here, we instead used a custom mixture of two photoresists.

1) Place the wafer onto the spin coater.

2) Prepare a mixture of SU-8 2002 (Shipley) and SU-8 2005 (Shipley) in the volume ratio of 3:1 and throughly mix with a clean swab.

3) Pour the mixture to cover ~2/3 of the wafer and spin with 1800/300/45 with an initial acceleration of 3000/100/5.

4) Bake the wafer (in this order), 1 minute at 65°C, 5 minute at 95°C, and 1 minute at 65°C.

5) Place the wafer on mask aligner (Suss MJB4), align the photoresist mask carefully, check if the alignment error is less than 1um after applying vacuum contact, and apply a total power of 75 mJ/cm^2 with the vacuum contact.

6) Bake the wafer (in this order), 1 minute at 65°C, 15 minute at 95°C, and 1 minute at 65°C.

7) Slowly cool the wafer down.

8) Place in a bath filled with SU-8 developer and leave it there for ~10 minutes with gentle manual agitation.

9) Rinse the wafer with isopropanol throughly and blow dry with an air gun. If white undeveloped photoresist was observed, go back to 8)

10) Measure the height using a profilometer. The expected height of this layer is ~2 um.

11) Hard bake the wafer for 10 minutes by gradually increasing the temperature from 65°C to 150°C on a hot plater.

Main media channel

The desired height in this step is ~30 um, which is not critical compared to the previous steps.

1) Dehydrate the wafer on a hot plater at 150°C for 10 minutes.

2) Place the wafer onto the spin coater and wait until it cooled down.

3) Pour SU-8 2025 photoresist (Shipley) to cover ~2/3 of the wafer and spin with 2500/300/60 with an initial acceleration of 500/100/5.

4) Bake the wafer (in this order), 1 minute at 65°C, 6 minute at 95°C, and 1 minute at 65°C.

5) Place the wafer on mask aligner (Suss MJB4), align the photoresist mask carefully, check if the alignment error is less than 1um after applying vacuum contact, and

apply a total power of 150 mJ/cm^2 twice with a waiting time of 30 seconds with the vacuum contact.

6) Bake the wafer (in this order), 1 minute at 65°C, 10 minute at 95°C, and 1 minute at 65°C.

7) Place in a bath filled with SU-8 developer and leave it there for ~10 minutes with gentle manual agitation.

8) Rinse the wafer with isopropanol throughly and blow dry with an air gun. If white undeveloped photoresist was observed, go back to 8)

9) Measure the height using a profilometer. The expected height of this layer is ~30 um.

10) Hard bake the wafer for 20 minutes by gradually increasing the temperature from 65°C to 180°C on a hot plater.

Silanizing of master mold

The aim of silanizing is to produce a passivation of the surfaces to aid release from PDMS and prevents the PDMS form adhering to the master. We recommend to make >5 devices after silanizing to remove possible toxicity. A silinazing agent is toxic, so it is highly recommended to perform this step in a appropriate environment.

1) Blow dry the wafer with an air gun

2) Prepare an aluminum foil cap and place it inside a vacuum desiccator.

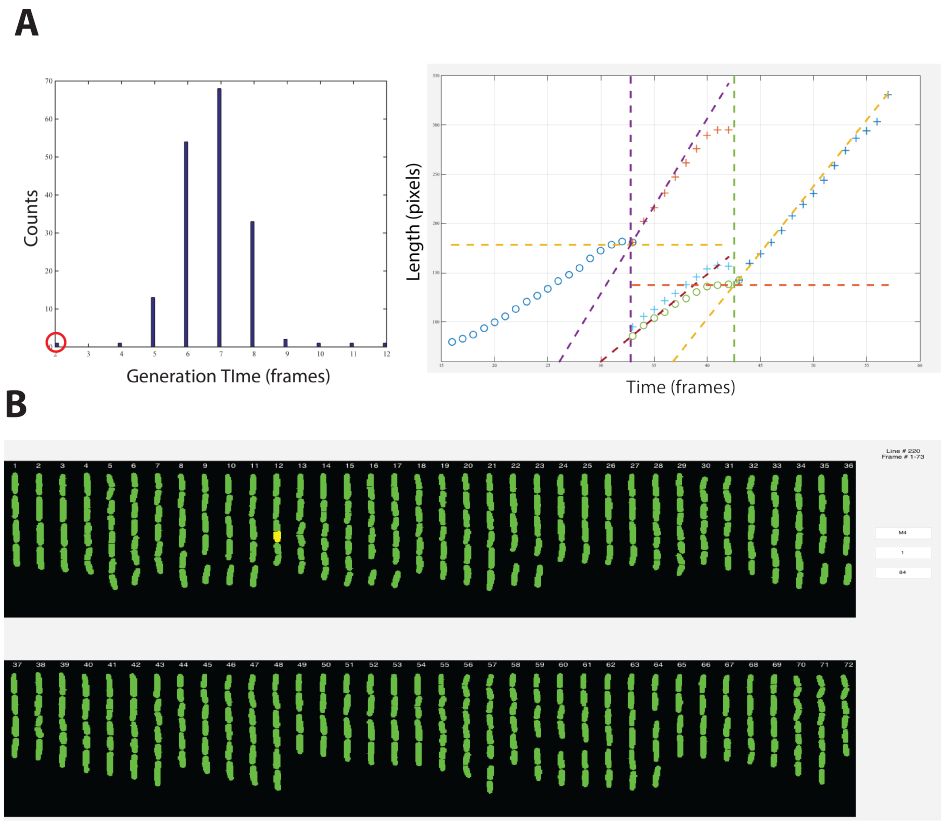
3) Place the petri dish in a vacuum desiccator and put the wafer in it with a slight tilt.

4) Put two drops of silanizing agent (tridecafluoro-1,1,2,2-tetrahydrooctyl trichlorosilane , Sigma) with pipettes.

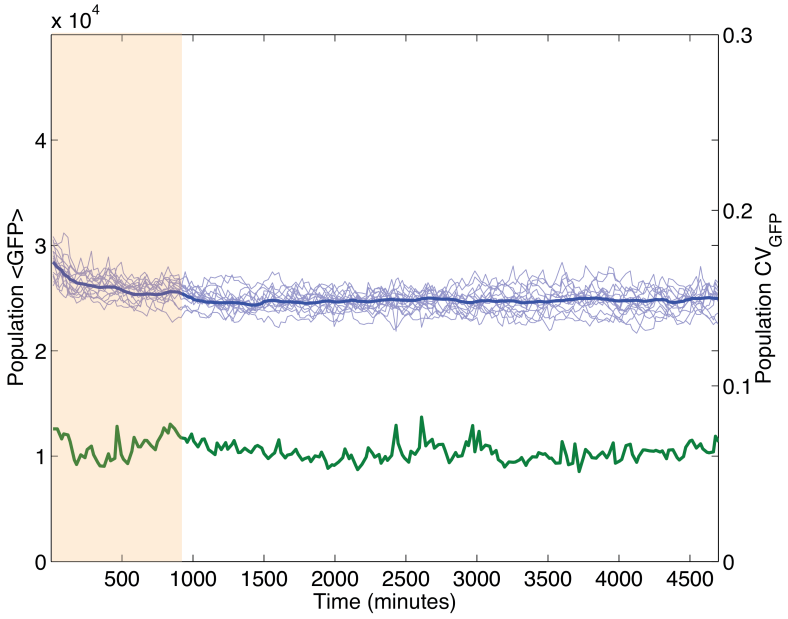
5) Run vacuum pump for ~10 minutes.

6) Place the wafer on a hot plater at 150°C for 10 minutes to evaporate the excessive silane.

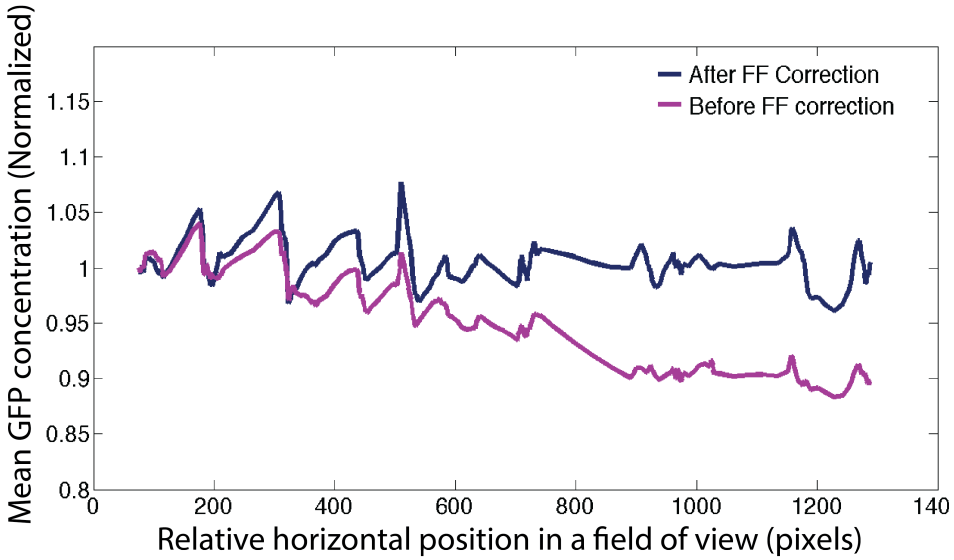
Appendix C: Supplemental figures



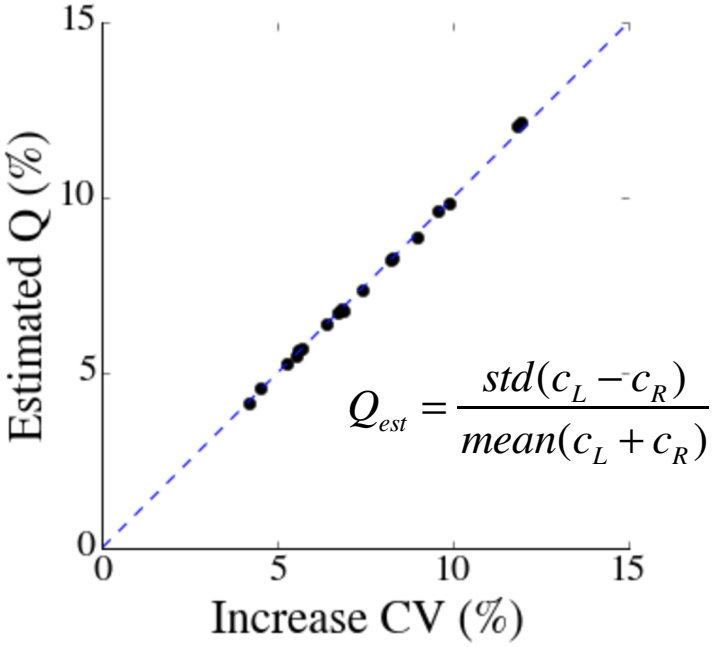
Supplemental Figure C.1 | Pipeline to analyze time-lapse movie **(A)** The left panel shows the distribution of generation time for cells within a single cel channel. The red circle indicates a possible segmentation error. The right panel plots a growth in cell size over time. Dots in different colors represent different lineages. Dotted line shows the result of fitting with two phases (growth and plateau). Errors are also detected in this procedure if certain cell size greatly deviates from the fitting. **(B)** An example of our GUI platform that enables us to correct the segmentation error. The yellow segmentation mask represents a potential error.



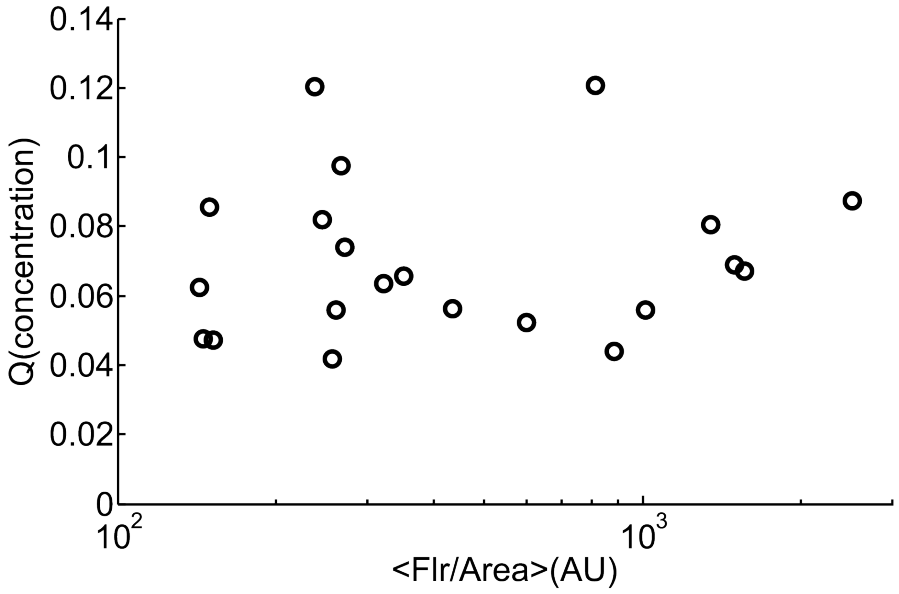
Supplemental Figure C.2 | Stable expression of GFP in the mother machine Narrow blue lines show GFP concentration traces of randomly picked mother cells over time. Wide blue line represents a moving average of GFP concentration and the green line represents the coefficient of variation within the same moving windows. We noticed the stabilization period (shaded orange area) is required to reach stationary state due to bleaching effect.



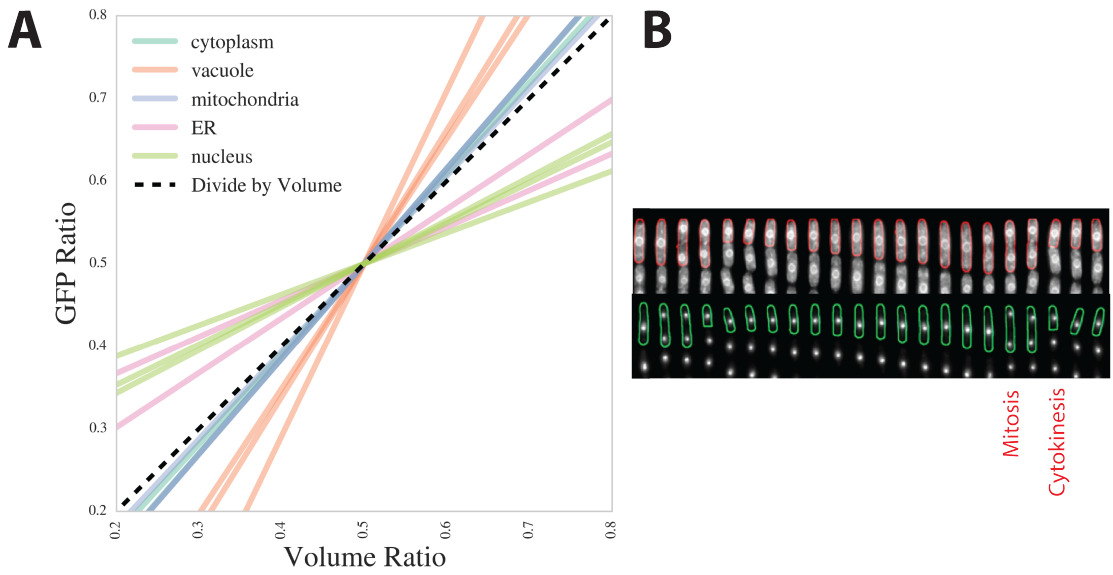
Supplemental Figure C.3 | Flat field correction We noticed uneven illumination over the spatial position. Purple curve shows an example a flat field effect for cells within a single field of view, showing 10% decrease to the right. After gathering all cells over different field of view, the mean illumination profile was calculated computationally and was applied to correct the flat-field effect (blue curve shows the corrected pattern).



Supplemental Figure C.4 | Approximating partitioning error as a normalized difference between sister cells For each strain, the increased CV after cell division versus estimated partitioning error defined as in the equation above is plotted. The partitioning error in concentration can be similarly defined as in [91], which represents a normalized difference between siblings at cell division.

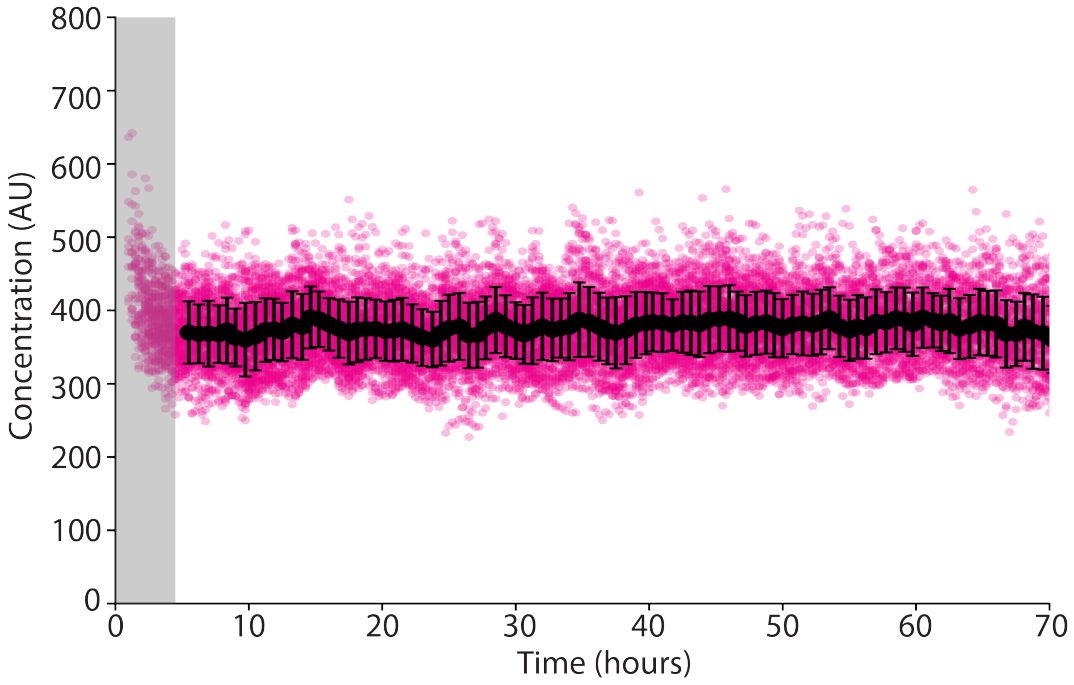


Supplemental Figure C.5 | Partitioning noise does not depend on protein expression level For each strain, their mean protein concentration and partitioning noise (Q) is plotted, showing no significant patterns.



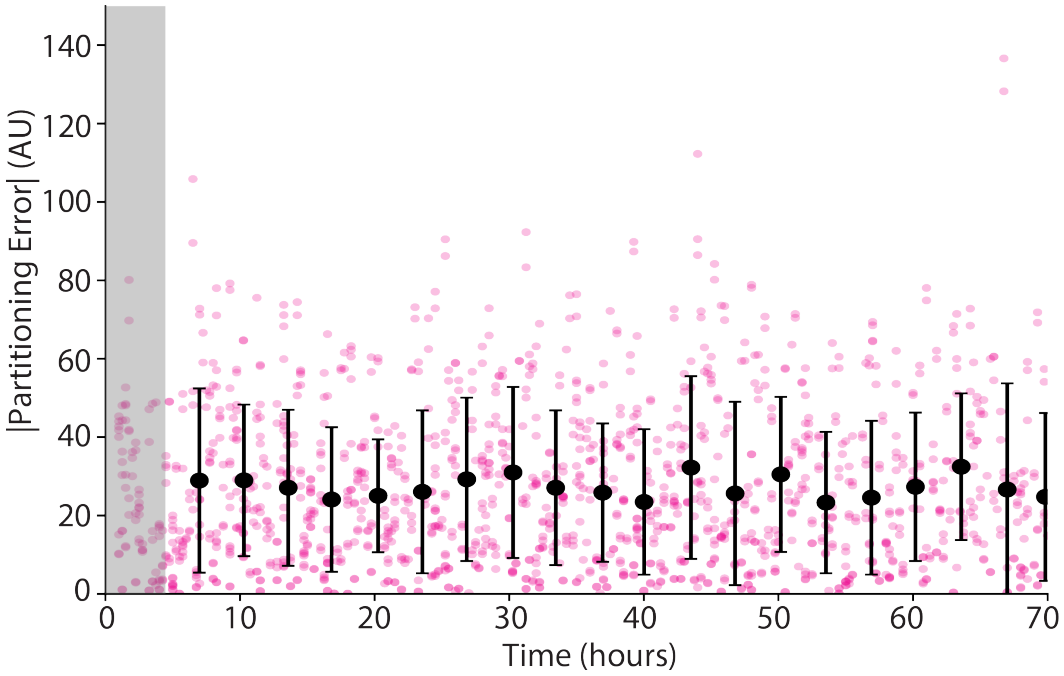
Supplemental Figure C.6 | Non-volumetric and volumetric divisions for organelles

(A) For each strain, we calculated volume ratio and total GFP ratio between siblings at birth and fitted with a regression line. In case of volumetric division, the curve should be on y=x (dotted black line). We observed the fit was reproducible since the curve largely cluster by localization patterns. Proteins in cytoplasm and mitochondria was close to volumetric division. However, proteins localizing into vacuoles showed over-volumetric division (larger cell takes more than expected), while those in endoplasmic reticulum or nucleus displayed under-volumetric division (larger cell takes less than expected). **(B)** A kymograph showing that mitosis happens earlier than cytokinesis in fission yeast.

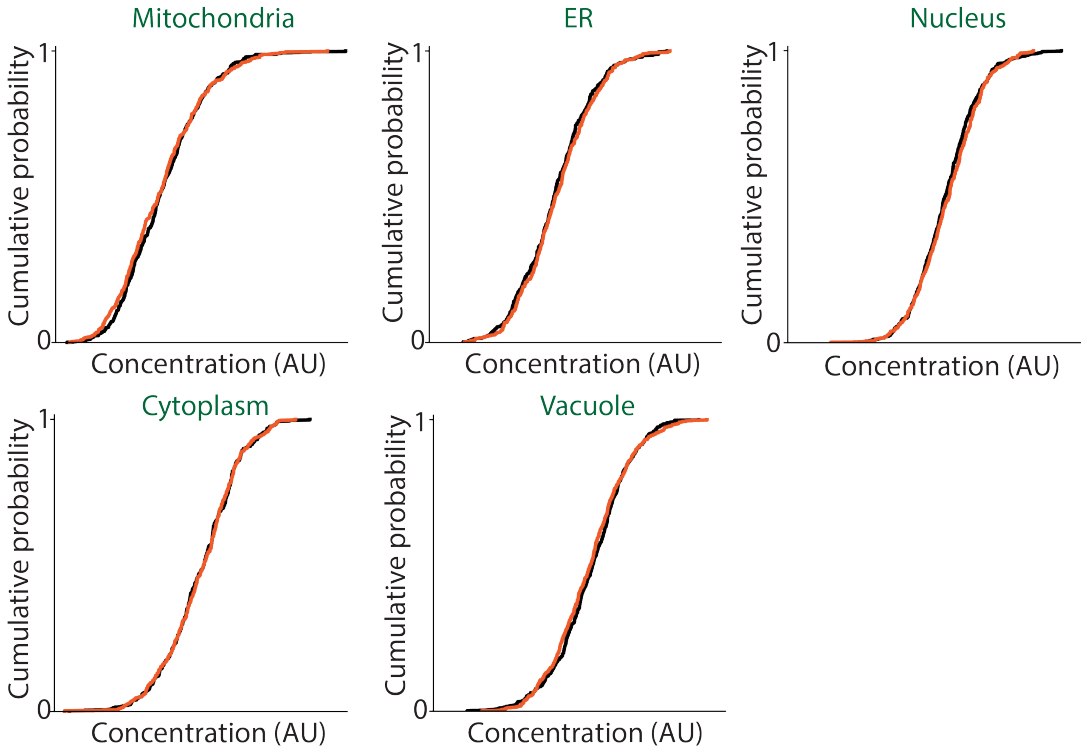


Supplemental Figure C.7 | Concentration was stable over time in mother machine.

For typical experiments in Chapter III, the stable concentration profile was observed. Black dots shows binned error bars (standard errors).

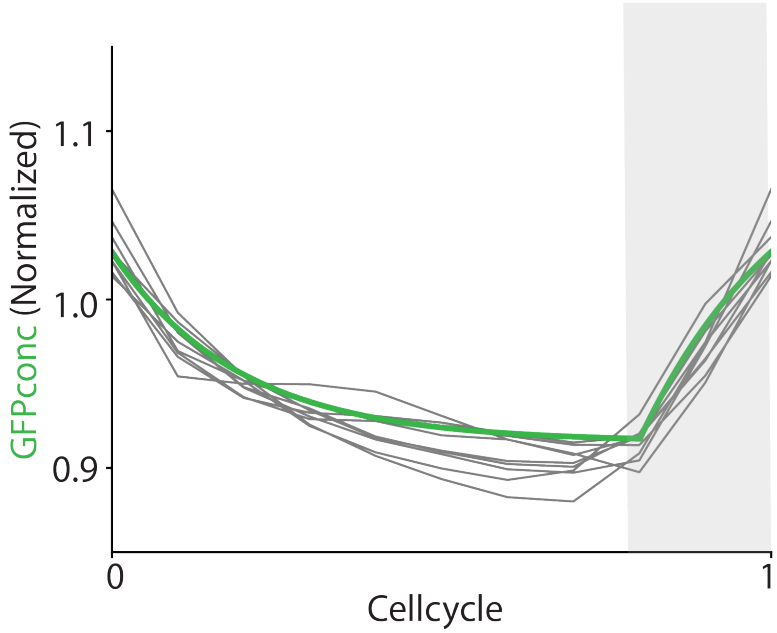


Supplemental Figure C.8 | Stable partitioning error over the course of experiment An absolute value of partitioning error or a discontinuity in protein concentration upon a division (see Figure III.3A) was plotted over time for a typical experiment. Black dots shows binned error bars (standard errors).



Supplemental Figure C.9 | Old pole cell is statistically identical to new pole cell

Cumulative distribution function (CDF) in protein concentration for old pole (red curve) and new pole cells (black curve) at birth are compared. For all measured strains, the Komologov-Smirnov statistic between them was less than 3%.



Supplemental Figure C.10 | Concentration profile over cell cycle can be described with maturation process The mean normalized GFP concentration profile for 10 strains were calculated over normalized cell cycle. The gray shaded area represents plateau phase. The green curve represents a simulation result including a maturation step with half life of 20 minutes.

Reference

1. Paulsson, J., *Summing up the noise in gene networks.* Nature, 2004. **427**: p. 415-8.

2. Elf, J., G.-W. Li, and X.S. Xie, *Probing transcription factor dynamics at the single-molecule level in a living cell.* Science (New York, N.Y.), 2007. **316**: p. 1191-4.

3. Golding, I., et al., *Real-time kinetics of gene activity in individual bacteria.* Cell, 2005. **123**: p. 1025-1036.

4. Raj, A., et al., *Stochastic mRNA synthesis in mammalian cells.* PLoS Biology, 2006. **4**: p. 1707-1719.

5. Cai, L., N. Friedman, and X.S. Xie, *Stochastic protein expression in individual cells at the single molecule level.* Nature, 2006. **440**: p. 358-362.

6. Pedraza, J.M., *Noise Propagation in Gene Networks.* Science, 2005. **307**: p. 1965-1969.

7. Pedraza, J.M. and J. Paulsson, *Effects of molecular memory and bursting on fluctuations in gene expression.* Science, 2008. **319**: p. 339-343.

8. Ozbudak, E.M., et al., *Regulation of noise in the expression of a single gene.* Nature genetics, 2002. **31**: p. 69-73.

9. Thattai, M. and A. van Oudenaarden, *Intrinsic noise in gene regulatory networks.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**: p. 8614-9.

10. Raser, J.M. and E.K.O. Shea, *Control of Stochasticity in Eukaryotic Gene Expression Jonathan.* Science, 2006. **304**: p. 1811-1814.

11. Swain, P.S., M.B. Elowitz, and E.D. Siggia, *Intrinsic and extrinsic contributions to stochasticity in gene expression.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**: p. 12795-800.

12. Hilfinger, A. and J. Paulsson, *Separating intrinsic from extrinsic fluctuations in dynamic biological systems.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**: p. 12167-12172.

13. Taniguchi, Y., et al., *Quantifying E. coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells.* Science (New York, NY), 2010. **329**: p. 533-538.

14. Ghaemmaghami, S., et al., *Global analysis of protein expression in yeast.* Nature, 2003. **425**: p. 737-41.

15. Lipson, D., et al., *Quantification of the yeast transcriptome by single-molecule sequencing.* Nature biotechnology, 2009. **27**: p. 652-658.

16. Chen, K.H., et al., *RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells.* Science (New York, N.Y.), 2015. **348**: p. aaa6090.

17. Albayrak, C., et al., *Digital Quantification of Proteins and mRNA in Single Mammalian Cells.* Molecular Cell, 2016. **61**: p. 914-924.

18. Paulsson, J., *Models of stochastic gene expression.* Physics of Life Reviews, 2005. **2**: p. 157-175.

19. Hilfinger, A., et al., *Constraints on Fluctuations in Sparsely Characterized Biological Systems.* Physical Review Letters, 2016. **116**: p. 1-5.

20. Hilfinger, A., T.M. Norman, and J. Paulsson, *Exploiting Natural Fluctuations to Identify Kinetic Mechanisms in Sparsely Characterized Systems.* Cell Systems, 2016. **2**: p. 251-259.

21. Paulsson, J., *Prime movers of noisy gene expression.* Nature genetics, 2005. **37**: p. 925-926.

22. Ras, W., et al., *Noise in eukaryotic gene expression.* 2003. **249**: p. 247-249.

23. Newman, J.R.S., et al., *Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise TL - 441.* Nature, 2006. **441 VN -** p. 840-846.

24. Bar-Even, A., et al., *Noise in protein expression scales with natural protein abundance.* Nature Genetics, 2006. **38**: p. 636-643.

25. Newman, J.R.S., et al., *Single-cell proteomic analysis of S . cerevisiae reveals the architecture of biological noise.* Nature, 2006. **441**.

26. Birky Jr, C.W., *The partitioning of cytoplasmic organelles at cell division.* International review of cytology. Supplement, 1982. **15**: p. 49-89.

27. Birky, C.W., *Relaxed cellular controls and organelle heredity.* Science, 1983. **222**: p. 468-476.

28. Hennis, A.S. and C. Birky, *Stochastic partitioning of chloroplasts at cell division in the alga Olisthodiscus, and compensating control of chloroplast replication.* Journal of cell science, 1984. **70**(1): p. 1-15.

29. Bergeland, T., et al., *Mitotic partitioning of endosomes and lysosomes.* Current Biology, 2001. **11**(9): p. 644-651.

30. Sheahan, M.B., R.J. Rose, and D.W. McCurdy, *Organelle inheritance in plant cell division: The actin cytoskeleton is required for unbiased inheritance of chloroplasts, mitochondria and endoplasmic reticulum in dividing protoplasts.* Plant Journal, 2004. **37**: p. 379-390.

31. Mukherji, S. and E.K. O'Shea, *Mechanisms of organelle biogenesis govern stochastic fluctuations in organelle abundance.* eLife, 2014. **2014**: p. 1-17.

32. Rafelski, S.M., et al., *Mitochondrial network size scaling in budding yeast.* Science, 2012. **338**(6108): p. 822-824.

33. Fagarasanu, A., et al., *Molecular mechanisms of organelle inheritance: lessons from peroxisomes in yeast.* Nature reviews. Molecular cell biology, 2010. **11**: p. 644-654.

34. Shima, D.T., et al., *Partitioning of the Golgi apparatus during mitosis in living HeLa cells.* Journal of Cell Biology, 1997. **137**: p. 1211-1228.

35. Conradt, B., et al., *In vitro reactions of vacuole inheritance in Saccharomyces cerevisiae.* Journal of Cell Biology, 1992. **119**: p. 1469-1479.

36. Huh, D. and J. Paulsson, *Non-genetic heterogeneity from stochastic partitioning at cell division.* Nature Genetics, 2011. **43**: p. 95-100.

37. Locke, J.C.W. and M.B. Elowitz, *Using movies to analyse gene circuit dynamics in single cells.* Nature Reviews Microbiology, 2009. **7**: p. 383-92.

38. Trapnell, C., *Defining cell types and states with single-cell genomics.* Genome Research, 2015. **25**: p. 1491-1498.

39. Spiller, D.G., et al., *Measurement of single-cell dynamics.* 2010. **465**.

40. Longo, D. and J. Hasty, *Dynamics of single-cell gene expression.* Molecular Systems Biology, 2006. **2**.

41. Nurse, P., *Genetic control of cell size at cell division in yeast.* Nature, 1975. **256**: p. 547-551.

42. Nurse, P., *Growth in Cell Length in the Fission Yeast Schizosaccharomyces Pom Be.* Journal of Cell Science, 1985. **376**: p. 357-376.

43. Martin, S.G. and M. Berthelot-Grosjean, *Polar gradients of the DYRK-family kinase Pom1 couple cell length with the cell cycle.* Nature, 2009. **459**: p. 852-856.

44. Wood, E. and P. Nurse, *Pom1 and cell size homeostasis in fission yeast.* Cell Cycle, 2013. **12**: p. 3228-3236.

45. Elowitz, M.B. and S. Leibler, *A synthetic oscillatory network of transcriptional regulators.* Nature, 2000. **403**: p. 335-338.

46. Wan-Ling, C., et al., *Engineered GFP as a vital reporter in plants.* Current Biology, 1996. **6**: p. 325-330.

47. Chalfie, M., et al., *Green fluorescent protein as a marker for gene expression.* Science, 1994. **263**: p. 802-805.

48. Süel, G.M., et al., *An excitable gene regulatory circuit induces transient cellular differentiation.* Nature, 2006. **440**: p. 545-550.

49. Süel, G.M., et al., *Tunability and noise dependence in differentiation dynamics.* Science (New York, N.Y.), 2007. **315**: p. 1716-9.

50. Norman, T.M., et al., *Memory and modularity in cell-fate decision making.* Nature, 2013. **503**: p. 481-6.

51. Nathan, L.P.-t., D.L. Glenn, and V. Johan, *Synchronous long-term oscillations in a synthetic gene circuit.* Nature, 2016. **538**: p. 1-4.

52. Locke, J.C.W., et al., *Stochastic Pulse Regulation in Bacterial Stress Response.* Science, 2011. **334**: p. 366-369.

53. Young, J.W., et al., *Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy.* Nature protocols, 2012. **7**: p. 80-88.

54. Sprinzak, D., et al., *Cis-interactions between Notch and Delta generate mutually exclusive signalling states.* Nature, 2010. **465**: p. 86-90.

55. Ragunathan, K., G. Jih, and D. Moazed, *Epigenetic inheritance uncoupled from sequence-specific recruitment.* Science, 2014. **348**: p. science.1258699-.

56. Moffitt, J.R., J.B. Lee, and P. Cluzel, *The single-cell chemostat: an agarose-based, microfluidic device for high-throughput, single-cell studies of bacteria and bacterial communities.* Lab on a Chip, 2012. **12**: p. 1487.

57. Ullman, G., et al., *High-throughput gene expression analysis at the level of single proteins using a microfluidic turbidostat and automated cell tracking.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**: p. 20120025.

58. Balaban, N.Q., et al., *Bacterial Persistenceas a Phenotypic Switch.* Science, 2004. **305**: p. 1622-1625.

59. Wang, P., et al., *Robust growth of escherichia coli.* Current Biology, 2010. **20**: p. 1099-1103.

60. Taheri-Araghi, S., et al., *Cell-Size Control and Homeostasis in Bacteria.* Curr. Biol., 2014: p. 385-391.

61. Campos, M., et al., *A constant size extension drives bacterial cell size homeostasis.* Cell, 2014. **159**(6): p. 1433-1446.

62. Uphoff, S., et al., *Stochastic activation of a DNA damage response causes cell-to-cell mutation rate variation.* Science, 2016. **351**(6277): p. 1094-1097.

63. Lee, S., et al., *Whole lifespan microscopic observation of budding yeast aging through a micro fl uidic dissection platform.* Pnas, 2012. **109**: p. 4-8.

64. Coelho, M., et al., *Fission yeast does not age under favorable conditions, but does so after stress.* Current Biology, 2013. **23**: p. 1844-1852.

65. Nobs, J.-B. and S.J. Maerkl, *Long-term single cell analysis of S. pombe on a microfluidic microchemostat array.* PloS one, 2014. **9**: p. e93466.

66. Spivey, E.C., et al., *3D-printed microfluidic microdissector for high-throughput studies of cellular aging.* Analytical Chemistry, 2014. **86**: p. 7406-7412.

67. Hoover, A., et al., *An experimental comparison of range image segmentation algorithms.* IEEE transactions on pattern analysis and machine intelligence, 1996. **18**(7): p. 673-689.

68. Sveiczer, A., B. Novak, and J.M. Mitchison, *The size control of fission yeast revisited.* Journal of Cell Science, 1996. **2957**: p. 2947-2957.

69. Hormoz, S., et al., *Inferring Cell-State Transition Dynamics from Lineage Trees and Endpoint Single-Cell Measurements.* Cell Systems, 2016. **3**: p. 419-433.e8.

70. Brangwynne, C.P., et al., *Cytoplasmic diffusion: Molecular motors mix it up.* Journal of Cell Biology, 2008. **183**: p. 583-587.

71. Lestas, I., G. Vinnicombe, and J. Paulsson, *Fundamental limits on the suppression of molecular fluctuations.* Nature, 2010. **467**: p. 174-178.

72. Yu, J., et al., *Probing gene expression in live cells, one protein molecule at a time.* Science (New York, N.Y.), 2006. **311**: p. 1600-3.

73. Birky, C.W. and R.V. Skavaril, *Random partitioning of cytoplasmic organelles at cell division: the effect of organelle and cell volume.* Journal of theoretical biology, 1984. **106**(4): p. 441-447.

74. Huh, D. and J. Paulsson, *Random partitioning of molecules at cell division.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**: p. 15004-9.

75. Marshall, W.F., *Stability and robustness of an organelle number control system: modeling and measuring homeostatic regulation of centriole abundance.* Biophysical journal, 2007. **93**(5): p. 1818-1833.

76. Summers, D.K. and D.J. Sherratt, *Multimerization of high copy number plasmids causes instability: Cole 1 encodes a determinant essential for plasmid monomerization and stability.* Cell, 1984. **36**: p. 1097-1103.

77. Bedbrook, J.R. and F.M. Ausubel, *Recombination between bacterial plasmids leading to the formation of plasmid multimers.* Cell, 1976. **9**(4): p. 707-716.

78. Savage, D.F., et al., *Spatially ordered dynamics of the bacterial carbon fixation machinery.* Science, 2010. **327**(5970): p. 1258-1261.

79. Huh, W.-K., et al., *Global analysis of protein localization in budding yeast.* Nature, 2003. **425**: p. 686-691.

80. Matsuyama, A., et al., *ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe.* Nature biotechnology, 2006. **24**: p. 841-7.

81. Jajoo, R., et al., *Accurate concentration control of mitochondria and nucleoids.* Science, 2016. **351**(6269): p. 169-172.

82. Landgraf, D., *Quantifying localizations and dynamics in single bacterial cells.* 2012.

83. Haim, L., et al., *A genomic integration method to visualize localization of endogenous mRNAs in living yeast.* Nature methods, 2007. **4**: p. 409-412.

84. Landgraf, D., et al., *Scarless gene tagging with one-step transformation and two-step selection in Saccharomyces cerevisiae and Schizosaccharomyces pombe.* PloS one, 2016. **11**(10): p. e0163950.

85. Wu, J.-Q. and T.D. Pollard, *Counting cytokinesis proteins globally and locally in fission yeast.* Science, 2005. **310**(5746): p. 310-314.

86. Paulsson, J. and M. Ehrenberg, *Noise in a minimal regulatory network: plasmid copy number control.* Quarterly reviews of biophysics, 2001. **34**(01): p. 1-59.

87. das Neves, R.P., et al., *Connecting variability in global transcription rate to mitochondrial variability.* PLoS Biology, 2010. **8**.

88. Sigal, A., et al., *Variability and memory of protein levels in human cells.* Nature, 2006. **444**: p. 643-646.

89. Spencer, S.L., et al., *Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis.* Nature, 2009. **459**(7245): p. 428-432.

90. Eldar, A., et al., *Partial penetrance facilitates developmental evolution in bacteria.* Nature, 2009. **460**: p. 510-514.

91. Huh, Dann. *Non-genetic Heterogeneity Originating at Cell Division*. Harvard University, 2011.