

THE EFFECTS OF VARIABILITY AND COVARIABILITY ON  
PROTEIN MODULES USING SINGLE CELL MASS  
SPECTROMETRY

A DISSERTATION  
SUBMITTED TO THE DEPARTMENT OF CHEMICAL AND  
SYSTEMS BIOLOGY  
AND THE COMMITTEE ON GRADUATE STUDIES  
OF STANFORD UNIVERSITY  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

Kyle M. Kovary

August 2020

# Abstract

My abstract.

# Acknowledgements

I wish to thank my advisor, Mary, for giving me the opportunity to work in such a wonderful and inspiring research group and for giving me her invaluable support. She has been a tremendous mentor for me and more importantly, it has been an honor to be her first graduate student. I am also grateful to all my lab mates for the spectacular working atmosphere and cooperation.

I would like to extend my special appreciation to the committee members for providing timely and constructive feedback on my thesis work, without which this work could have never become what it is today.

I would like to thank the Department of Chemical and Systems Biology, my classmates, and fellow graduate students, who have created such an outstanding environment, for providing infinite amount of support. Moreover, I would like to thank everyone in the CSB administration for taking care of administrative aspects for graduate students.

Lastly, I would like to thank my family for their support during my time at Stanford. My warmest thanks belong to them for a lifetime of love, care, and support.

# Contents

<b>Abstract</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background . . . . .	1
1.1.1 Specific background . . . . .	1
<b>2 Expression variation and covariation impair analog and enable binary signaling control</b>	<b>2</b>
2.1 Abstract . . . . .	3
2.2 Introduction . . . . .	3
2.3 Results . . . . .	5
2.3.1 Computational simulations using reported levels of expression variation show a dramatic loss of analog single-cell transmission accuracy . . . . .	5
2.3.2 Development of a method to accurately measure the relative abundance of tens of proteins in a single cell . . . . .	8
2.3.3 Low variation in the relative abundance of proteins explains how cells are able to accurately control analog single-cell functions	9
2.3.4 Identification of covariance between the relative abundances of components in the ERK pathway . . . . .	11

2.3.5	Model analysis demonstrates that expression variation improves control of how many cells in a population make a binary cell-fate decision . . . . .	12
2.3.6	Understanding the role of variation in MEK and ERK expression in regulating bimodal ERK activity . . . . .	13
2.3.7	Understanding the role of covariation in MEK and ERK expression in regulating bimodal ERK activity . . . . .	15
2.3.8	Constraints on accurate control of analog and binary signaling by expression variation and covariation . . . . .	17
2.4	Discussion . . . . .	20
2.4.1	Variation in signaling protein expression between individual cells is lower than expected . . . . .	20
2.4.2	Competing roles of expression variation in enabling accurate analog single-cell signaling and controlling population-level binary signaling . . . . .	21
2.4.3	Expression variation, covariation, and number of pathway components define accuracy of analog versus binary signaling systems . . .	22
2.5	Materials and Methods . . . . .	23
2.5.1	Reagents and tools table . . . . .	23
2.5.2	<i>Xenopus laevis</i> egg collection and activation . . . . .	24
2.5.3	SRM sample preparation . . . . .	24
2.5.4	SRM data acquisition . . . . .	25
2.5.5	SRM data statistical analysis . . . . .	26
2.5.6	Cell culture . . . . .	27
2.5.7	EKAR-EV-NLS stable cell line . . . . .	27
2.5.8	Immunofluorescence . . . . .	27
2.5.9	siRNA transfection . . . . .	28
2.5.10	Image acquisition . . . . .	28
2.5.11	Image segmentation and tracking . . . . .	29
2.5.12	Signal measurement of segmented cells . . . . .	29
2.5.13	Modeling for figures 1A and 1B . . . . .	30

2.6	Main Figures . . . . .	30
2.7	Supplemental Figures . . . . .	30
<b>3</b>	<b>Concluding Remarks</b>	<b>45</b>

# List of Tables

S1	Short table caption for List of Tables . . . . .	30
----	--	----

# List of Figures

2.1	Computational simulations using reported levels of expression variation show a dramatic loss of analog single-cell transmission accuracy . . .	31
2.2	Development of a method to quantitatively measure relative abundances of tens of endogenous proteins in parallel in single <i>Xenopus</i> eggs . . .	32
2.3	Single-cell variation in relative protein abundance is typically 5–10% in <i>Xenopus</i> eggs . . . . .	33
2.4	MEK and ERK expression covary in <i>Xenopus</i> eggs and cultured human cells . . . . .	34
2.5	Using a general five-step model to understand the effect of variation on controlling the fraction of cells in the population that respond to input stimulus . . . . .	35
2.6	Live-cell imaging experiments and simulations using an established MEK/ERK signaling model show that variation between MEK and ERK expression widens the window over which input stimuli can control the fraction of cells that are activated in the population . . . . .	36
2.7	Single-cell imaging experiments and model simulations show that covariation between MEK and ERK expression facilitates control of bimodal ERK activation . . . . .	37
2.8	Competing demands on variation and covariation in the control of analog single-cell versus binary population-level signaling outputs . .	38
S1	Comparison of fIDL and mutual information (MI) analysis . . . . .	39
S2	Selected reaction monitoring mass spectrometry approach to measure tens of proteins in parallel in single <i>Xenopus</i> eggs . . . . .	40

S3	Bootstrap analysis of CVs of the relative abundance of 26 proteins using a 60-egg set collected at 60 min after egg activation . . . . .	41
S4	Comparison of technical and biological variation in the SRM mass spectrometry measurements . . . . .	42
S5	Representative images of immunohistochemistry staining of the proteins studied . . . . .	43
S6	siRNA-mediated depletion experiments . . . . .	44

# Chapter 1

## Introduction

### 1.1 Background

Broad overview.

#### 1.1.1 Specific background

Specific text.

# Chapter 2

## Expression variation and covariation impair analog and enable binary signaling control

Kyle M. Kovary, Brooks Taylor, Michael Zhao, Mary N. Teruel

The writing in this chapter contributed to the following publication:

K.M. Kovary, Brooks Taylor, M.L. Zhao, and M.N. Teruel. Expression variation and covariation impair analog and enable binary signaling control. *Molecular Systems Biology*, 14(5):e7997, 2018

## 2.1 Abstract

Due to noise in the synthesis and degradation of proteins, the concentrations of individual vertebrate signaling proteins were estimated to vary with a coefficient of variation (CV) of approximately 25% between cells. Such high variation is beneficial for population-level regulation of cell functions but abolishes accurate single-cell signal transmission. Here, we measure cell-to-cell variability of relative protein abundance using quantitative proteomics of individual *Xenopus laevis* eggs and cultured human cells and show that variation is typically much lower, in the range of 5 – 15%, compatible with accurate single-cell transmission. Focusing on bimodal ERK signaling, we show that variation and covariation in MEK and ERK expression improves controllability of the percentage of activated cells, demonstrating how variation and covariation in expression enables population-level control of binary cell-fate decisions. Together, our study argues for a control principle whereby low expression variation enables accurate control of analog single-cell signaling, while increased variation, covariation, and numbers of pathway components are required to widen the stimulus range over which external inputs regulate binary cell activation to enable precise control of the fraction of activated cells in a population.

## 2.2 Introduction

Vertebrate signaling has been shown to control both binary and analog outputs. Here, we use the term binary if the output is bimodal and the term analog if the output signal changes in parallel with the input signal without bifurcations during the transmission. Examples of binary signaling decisions include the commitment to start the cell cycle [2], cell differentiation [3, 4, 5], apoptosis [6], action potentials [7] and the explosive secretory response of mast cells when encountering an antigen [8]. Effective analog signaling in individual cells has been observed, for example, in the visual transduction system where the number of absorbed photons proportionally increases electric outputs in cone cells [9], in single-cell IP<sub>3</sub> and Ca<sup>2+</sup> regulation by GPCRs [10], as well as for CD-8 [11] and IL-2 signaling [12] in T cells. Analog

signaling is also needed to accurately regulate the timing or duration of intermediate cell processes such as in the cell cycle where the time between the start of S-phase to mitosis has only small variation between individual cells [13]. Such precise regulation of durations requires low noise in the signaling steps before mitosis [14]. Together, these examples suggest that accurate analog signaling is important for graded control of cell outputs in single cells as well as for accurate internal timing.

A main motivation for our study was the high levels of protein expression variation that have been reported in vertebrate cells with coefficient of variations (CVs) of approximately 25% [15, 6, 16]. Such high levels of expression variation are beneficial for binary signaling which is often regulated at the population level rather than single-cell level. In population-based signaling, a goal of organisms is to use different levels of input to regulate the percentage of cells in a population that make a binary decision such as whether to proliferate, differentiate, or secrete. For input stimuli to control which percentage of cells are activated, high noise in signaling is needed between cells in the population such that individual cells have different sensitivities to input stimuli [17, 18, 19, 20, 5]. However, the same high noise needed to control population-level signaling does not have any benefit for analog signaling and just serves to degrade signal transmission. These different demands on noise for analog and binary signaling suggest that there is a trade-off for noise between population-level and single-cell signaling [21]. Specifically, the reported high levels of expression variation and signaling noise in mammalian cells [15, 22, 16, 23] raise the question of how noise in a signaling system can be low enough for accurate analog signaling. It also remained unclear how the different potential internal noise sources could generate optimal conditions for analog single-cell versus binary population-level signaling.

Here, we measure cell-to-cell variation in the relative abundance of pathway components to understand the limits of analog and binary signaling accuracy. We also investigated the role of covariation of pathway components since we realized that covariation could exacerbate the analog signaling problem and/or enable the control of population-level binary signaling. We considered that previous estimates of cell-to-cell variation in protein expression might be too high due to experimental challenges in accurately measuring small differences in protein abundance between cells and accounting for

“hidden variables” such as differences in cell size and cell cycle state [24]. To determine lower limits of protein variation, we developed single-cell quantitative proteomics methods in single *Xenopus laevis* eggs and employed quantitative normalization of cultured human cells to accurately measure variations in protein abundance normalized by protein mass. We found that cell-to-cell variation in relative protein abundance is much lower than expected, with CVs of between 5 and 15%, suggesting that expression variation is less limiting than currently believed and is compatible with accurate analog signal transmission. Furthermore, our simulations show that these experimentally observed low levels of expression variation pose a challenge for cells to accurately control population-level decisions. One potential strategy to increase pathway output variation was revealed by experiments which showed significant covariation between the single-cell expression of two sequential signaling components, MEK and ERK. Our modeling showed that such increased covariation—which increases the overall noise in the signaling pathway—allows populations of cells to control the percentage of cells that activate ERK over a wider range of input stimuli, suggesting that covariation of signaling components is one strategy for populations of cells to more accurately control binary cell-fate decisions. Finally, we developed a metric to describe how systems can optimize the shared use of pathway components to control single-cell analog and population-level binary signal transmission by using different numbers of regulatory components, levels of expression variation, and degrees of covariation.

## 2.3 Results

### 2.3.1 Computational simulations using reported levels of expression variation show a dramatic loss of analog single-cell transmission accuracy

Our study was motivated by the reported high levels of expression variation and the detrimental impact that this source of noise may have on analog single-cell signaling, especially since signaling pathways typically have multiple components

which necessarily results in even higher cumulative signaling noise. To define the general control problem of how expression variation increases overall signaling noise and limits signaling output accuracy, we carried out simulations by applying a relative fold-change in input signal ( $R$ ) to a signaling pathway and stochastically varying the expression of pathway components for each simulation. To determine how accurately a multi-step signaling pathway can transmit a relative input stimulus ( $R$ ) to an analog output ( $A^*$ ), we modeled the signaling pathway shown in Fig 2.1A. Specifically, we used a five-step model where a relative change in input  $R$  acts through four intermediate steps, possibly reflecting a kinase cascade with counteracting phosphatases, to generate corresponding changes in the output  $A^*$ . The regulation of these steps can be at the level of activity or localization of pathway components. We considered five steps with 10 variable regulators to be a typical signaling pathway since it has been shown that step numbers in signaling pathways can range from very few in visual signal transduction [25] to over 10 steps in the growth-factor control of ERK kinase and cell cycle entry [26]. In our simulations, each of the parameters represents a regulatory protein that activates or inactivates one of the pathway steps. We assumed that each of these components has “expression variation,” meaning that their concentrations vary between cells with a coefficient of variation (CV) calculated as their standard deviation divided by their mean value in the cell population. We simulated this expression variation by multiplying each parameter in the model with a log-normal stochastic noise term with a CV of either 5, 10, or 25% [5]. As is apparent in the top plots in Fig 2.1B for a CV of 5%, the signaling responses of cells to threefold (red) and ninefold (blue) increases in the input stimulus,  $R$ , can be readily distinguished from the signaling responses of unstimulated cells (black traces). For a higher CV of 10%, the signaling responses to a threefold increase in  $R$  partially overlap with the unstimulated cell responses, and only the responses to a ninefold increase in  $R$  can be unequivocally distinguished from unstimulated cell responses. For a CV of 25%, even responses to a ninefold increase in input stimulus overlap with the responses of unstimulated cells, showing a dramatic loss in signaling accuracy.

One way to overcome this dramatic loss in signaling accuracy due to expression variation of pathway components is to increase the input stimulus. We reasoned that

we could use a fold-increase parameter to quantify the loss in signal accuracy. We thus defined a fold-Input Detection Limit (fIDL) as the minimal fold-stimulus needed to generate signaling responses that can, in 95% of cases, be distinguished from cell responses in unstimulated cells (see 2.5 for calculation). Figure 1C shows an example of how the fIDL is calculated by determining the minimum fold-input stimulus that is needed to have only a 5% overlap between the resulting signaling output distributions ( $A^*$ ) of unstimulated and stimulated cells (black and green histograms, respectively). In the case shown, an fIDL stimulus of 2.83 is needed to overcome the loss of signaling accuracy caused by having 10% expression variation in pathway components. We used fIDL instead of a commonly used mutual information metric since mutual information between input ( $R$ ) and output ( $A^*$ ) has a strong dependency on the dynamic range of the system output, while the fIDL is largely independent of saturation (Fig S1). As shown in the barplots in Fig 2.1D, increasing the CV of pathway components from 10 to 25% increases the fIDL from 2.83 to 14, a stimulus requirement that is likely prohibitive for analog single-cell signal transmission. Our realization that fIDLs are very high for reported expression variation levels was a main motivation for our strategy below to more accurately measure expression variation in order to understand whether and how analog signaling in single cells is limited by this noise source.

We also wanted to determine whether the expression of vertebrate proteins may covary since covariance has been shown to exist in a yeast regulatory pathway [27]. We considered that if proteins within a signaling pathway covary, the overall noise in the output response would increase. To illustrate the effect of covariance can have on a multi-step analog signaling pathway, we added covariation to the model shown in Fig 2.1A by making the positive regulators (e.g., kinases) covary together and also made the negative regulators (e.g., phosphatases) covary together. As shown in Fig 2.1E, covariance causes the error propagation to increase, and the overall noise in the signaling output is much higher compared to the case where proteins in the same pathway vary independently of each other. Given that covariation causes a marked increase in the overall noise of the signaling response, one would expect that covariation between components of the same signaling pathway should generally be

avoided in order to have accurate analog signaling.

### 2.3.2 Development of a method to accurately measure the relative abundance of tens of proteins in a single cell

To probe the lower limits of protein expression variation, we selected a system with a need for analog single-cell signaling that was also suitable for parallel proteomics analysis. We chose *Xenopus laevis* eggs for three reasons. First, previous studies showed that the timing of the cell cycle during early embryogenesis is very precise with an accuracy of 5% [28], suggesting that the *Xenopus* system must have accurate analog signaling to maintain such timing. Second, eggs do not grow in size and have only minimal new synthesis and degradation of mRNA, two features which we thought would reduce protein expression variation. Third, *Xenopus laevis* eggs are well suited for single-cell proteomics analysis due to their large size (Ferrell, 1999), allowing us sufficient starting material to very sensitively measure and compare relative abundances of many proteins simultaneously in the same cell.

To accurately compare the relative abundance of tens of endogenous proteins in parallel in single cells, we used selected reaction monitoring mass spectrometry (SRM-MS), a low-noise quantitative mass spectrometry method [29, 30, 5] (Fig S2). Cytoplasmic proteins were extracted from eggs and subjected to trypsin digestion and phosphatase treatment before undergoing targeted quantification on a triple quadrupole mass spectrometer. Heavy isotope-labeled reference peptides were spiked in proportionately to a measured total protein concentration, and the ratio of the light (endogenous) peptide to the heavy (synthetic) peptide was used as a readout of relative protein abundance. Small calibration errors were further corrected for during the analysis using the median of 22 normalized peptide intensities as a correction factor similar to previous studies [29, 31, 30]. We measured relative protein abundance (abundance over total protein mass) as a measure of protein concentration since reaction rates and signaling processes depend on the concentration rather than abundance of proteins [32].

We first validated our method using bulk cell analysis at different timepoints

during the first cell cycle, a process which can be initiated by addition of calcium ionophore and takes approximately 90 min to complete [33]. We measured the abundances of a set of 26 proteins that we selected to include known regulators of signaling and cell cycle progression, as well as several control proteins (Fig 2.2A; Table EV1). Timecourse analysis over the first cell cycle further showed that we could observe the expected cycling behavior of Cyclin A and Cyclin B (Fig 2.2B). We next demonstrated that we could measure timecourses of relative protein abundances in single cells by carrying out measurements at five timepoints with five eggs each (Fig 2.2C; Table EV2). Except for a few known cell cycle-regulated genes, Cyclin A, Cyclin B, Cdc6 and Emi1, all of the measured proteins changed their abundance on average less than a few percent during the first egg cell cycle [34]. The constant average level of many of these signaling and cell cycle proteins can in part be explained by only minimal mRNA synthesis during early *Xenopus laevis* cell cycle [35].

### 2.3.3 Low variation in the relative abundance of proteins explains how cells are able to accurately control analog single-cell functions

We next focused on analyzing the extent to which protein concentrations vary between single cells. We first analyzed the set of 25 individual eggs from Fig 2.2C and determined the variation of each protein in each of the batches of five eggs collected at each of the five time-points (Fig 2.3A, left). Markedly, all CVs were much lower than expected with the median CV across all proteins and time-points being only 7% (Fig 2.3A, histogram in right panel). To independently verify these low variation measurements, we collected and analyzed a larger set of 120 individual eggs: 60 eggs collected at 60 and at 80 min after activation. To test for reproducibility of the measured variation, we divided the 60 eggs at each time-point into batches and carried out a variation analysis (Fig 3B; Table EV3). Bootstrapping analysis showed similar low variation (Fig S3). As further validation, the variations measured in the two independent experiments were similar to each other (Fig 2.3C). We also noted that most of the proteins that have high cell-to-cell variation (marked as red circles in Fig

2.3C) also change their abundance during the cell cycle (Fig 2.2C), suggesting that high CVs reflect proteins whose abundances are actively regulated. Thus, our finding of low CVs answers the question raised in Fig 2.1A–D of how cells can accurately control analog single-cell signaling outputs. Since expression variation can be as low as 5–10%, this main source of signaling noise is compatible with accurate single-cell signaling and timing control. Such low variation may also permit accurate timing in the *Xenopus laevis* embryonic cell cycle, which has been measured to be on the order of  $\pm 5\%$  between eggs [28].

It should be noted that for some proteins, the biological variation might be even lower than we were able to measure in these experiments. To test whether there is a lower limit for measuring variation, we carried out control experiments in which 30 individual eggs were lysed and mixed together to remove biological variability (Fig S4A). This mixed lysate was then pipetted into 30 individual tubes, and the sample in each tube was prepared and analyzed separately by SRM mass spectrometry. The variation between these 30 individually prepared and analyzed aliquots of the same starting lysate were compared to obtain a measure of technical variation. As shown in Fig EV4B, the technical variation is comparable to the lowest CV measurements we show in Fig 2.3A–C, suggesting that further technical improvements may reveal even lower biological variation.

Our analysis so far argues that expression variation can be much lower than previously assumed, which would enable accurate analog single-cell signaling as shown by how decreasing expression variation in Fig 2.1B allows for less overlap between unstimulated and stimulated cell responses. We next tested whether we would find the same low variation in protein expression in cultured human cells (HeLa cells) by carrying out immunocytochemistry experiments (Figs 3D and EV5). To accurately measure relative protein abundances, we first gated for cells in the same G0/G1 cell cycle state by using Hoechst DNA stain measurements (2n-peak) [2]. We further normalized the abundance of each protein to total protein mass in each cell. The latter was measured using an amine-reactive dye that stains all proteins in a cell [36]. Since total protein mass is proportional to cell volume [37], normalization by total protein mass can be used as a measure of protein concentration, analogous to the

normalization we used in the single-egg experiments. To minimize small illumination non-uniformities associated with imaging, we also confined our analysis to cells in the center area of images where the illumination and light collection is more uniform (see Materials and Methods). For comparison with the Xenopus egg data, we measured corrected CVs for the relative abundances for ERK, MEK, MCM5, and MCM7 as well as the control proteins GAPDH and ENO1. We validated the specificity of the antibodies by showing that the immunocytochemistry staining could be knocked down by the respective siRNAs (Fig S6). The resulting CVs for relative protein abundance were in the 10–15% range, lower than typically reported mammalian protein CV values [15, 38, 16].

### 2.3.4 Identification of covariance between the relative abundances of components in the ERK pathway

We next determined whether there was covariance between proteins by analyzing the same 120-egg proteomic dataset shown in Fig 2.3B. As shown in Fig 2.4A, our correlation analysis uncovered several covarying regulatory proteins. For example, there was significant co-regulation between MCM5 and MCM7 (Fig 2.4A-B), which is expected since they function as part of a stable MCM Helicase complex that can protect subunits from degradation in mammalian cells [39]. Nevertheless, we were surprised to also find significant covariation between MEK (MAP2K1) and ERK (MAPK1) (Fig 2.4A-B) because such covariance adds extra noise to the signaling pathway and would not be beneficial for accurate analog signal transmission. As further validation of the statistical significance of the covariance, the P-values for MCM5/MCM7 and the MEK/ERK covariation remained significant, even after adjustment for multiple comparison testing by using Benjamini-Hochberg corrections (Table EV4).

To determine whether the covariances we observed in *Xenopus laevis* eggs are conserved in human cells, we carried out single-cell immunohistochemistry measurements. As shown in Fig 2.4C, we found a strong covariance between MCM5 and MCM7. siRNA-mediated depletion experiments confirmed that MCM5 and MCM7 likely co-stabilize each other as both levels are reduced upon knockdown of either MCM5 or

MCM7 in HeLa cells (Fig S6). While control experiments showed weak covariation between MCM5 and the control protein GAPDH, we once again found a significant covariation between MEK and ERK, similar to the covariance we had observed in *Xenopus laevis* eggs (Fig 2.4C). This co-regulation is likely due to shared upstream expression regulation, or indirect feedbacks, as siRNA-mediated depletion of MEK and ERK showed opposing effects on ERK and MEK expression, respectively (Fig S6). The unexpected covariation between MEK and ERK in both *Xenopus laevis* eggs and human cells made us consider whether it might be beneficial for a cell to have components of the same pathway covary, possibly in the context of binary cell activation that is often associated with MEK and ERK signaling pathways.

### 2.3.5 Model analysis demonstrates that expression variation improves control of how many cells in a population make a binary cell-fate decision

As mentioned in the Introduction, previous studies showed that noise in signaling can be beneficial by widening the range of input stimuli that controls the percentage of cells in a population that are activated or not [5, 21]. We were therefore interested to understand whether and how variation and covariation of the expression of pathway components could be main sources of noise for the control of binary cell activation. We first focused on variation and carried out simulations to understand the effect of variation of pathway components on binary signaling at the population level. As shown in the schematic in Fig 2.5A, we used the model introduced in Fig 2.1A but now assumed a last regulatory step whereby a cell with a  $y_5^*$  value above 10 would trigger a switch into an active state while a cell with an output value  $y_5^*$  below the threshold of 10 would remain inactive. This last step is denoted as B\* versus B, reflecting the active and inactive binary output state, respectively. The results discussed here are largely independent of the value of the threshold (see 2.5).

We used this binary model to determine the percentage of cells in a population that will switch into the active state for different fold-increases of input stimuli and different levels of expression variation. As shown by the black circles in Fig 2.5B,

if there is no expression variation of pathway components, all cells will reach the threshold and abruptly switch from the inactive to active state within a very narrow stimulus window. As the expression variation of pathway components increases and the cells become more different from each other, the percentage of cells in a population that switch from the inactive to active state can be controlled over a wider range of input stimuli. Increasing the CV of pathway components to 40% results in a close-to-linear relationship in the five-step model between percent of cells activated and relative input stimulus amplitude.

This widening of the input stimulus control window can be quantified by fitting the fractional activation data with an apparent Hill coefficient (aHC) that measures how well the population-level output can be controlled by the input. The fitted Hill coefficients for systems with different amounts of protein expression variation are shown in the bar plot in Fig 2.5C. A system with a smaller aHC can be more accurately controlled over a wider range of input levels which would be desirable in physiological settings where external hormone input stimuli may not be precise themselves. Another consideration to take into account is that physiological responses to hormone stimulation can typically be elicited over a 10-fold or greater range of relative hormone stimulus increases ( $R$ ) [40, 41, 42]. For a five-step signaling pathway, accurately transmitting a 10-fold range of input stimuli means that there should be a nearly linear relationship between input stimulus and percent of activated cells (Fig 2.5B). Such a broad and nearly linear relationship requires that the signaling pathway has high overall variation (approximately 40%) which could originate from variation in expression of individual pathway components or from other sources of noise.

### 2.3.6 Understanding the role of variation in MEK and ERK expression in regulating bimodal ERK activity

Since the MEK/ERK signaling pathway often controls binary single-cell decisions such as whether cells divide or differentiate [43], we used MEK and ERK as examples of variable signaling components to evaluate the role of expression variation in population-level cell-fate decisions. We first validated experimentally that ERK signaling output

was bimodal or at least variable for intermediate stimuli in the same population by using EGF stimulation of human MCF10A cells. Specifically, we generated an MCF10A cell line expressing a FRET sensor of ERK activity to measure ERK activation in live cells [44, 45]. The FRET intensity of this sensor, EKAR-EV, was shown previously to faithfully report pERK levels in MCF10A cells [46]. We used EGF to activate the pathway, and after 60 min, cells were fixed and stained with antibodies to measure the abundances of MEK and ERK, so that the pathway response could be related back to the relative level of the two proteins. The MEK and ERK abundance values were normalized by an intracellular total protein stain following established protocols from [36] in order to correct for cell volume and to obtain relative protein abundances. An EGF titration showed that there was indeed bimodal ERK activation and that intermediate stimuli doses could induce heterogeneous responses (Fig 2.6A). We quantified the ERK activity in each time-course by calculating integrated ERK activity as the area under the curve after EGF stimulation. As shown in Fig 2.6A, the integrated ERK activity values showed two peaks, allowing us to define cells as active or inactive using the indicated threshold (dotted vertical black line). From the histograms in Fig 2.6A, it is apparent that the fraction of activated cells in the cell population increases as the EGF concentration increases. This relationship is more directly plotted in Fig 2.6B. Thus, in these human MCF10A cells, there is a wide range of input stimuli over which the fraction of the cell population that is activated can be controlled.

We next determined whether natural variation in MEK and ERK abundances indeed matters in determining whether individual cells have active ERK or not since it is conceivable that the level of active ERK is controlled by other factors such as variable numbers of receptors or variations in phosphatase activity. If expression variation matters for controlling activated ERK levels, the single-cell expression of MEK and ERK should on average be higher in cells with high ERK signaling compared to cells with low ERK signaling when analyzed in the same population of cells for the same intermediate input stimulus. Indeed, when we compared relative MEK and ERK abundances in cells with active or inactive ERK activity, we confirmed that activated cells have on average higher MEK and ERK concentrations, and inactive

cells have on average lower MEK and ERK concentrations (Fig 2.6C). These results argue that natural single-cell variation in the concentrations of MEK and ERK does matter in determining whether or not a cell will be activated.

The five-step model in Fig 5 conceptually showed how expression variation can broaden the range over which input stimuli can control binary cell fates. We next used an established model of the MAPK pathway to better understand whether and how natural variation in MEK and ERK expression contributes to the controllability of bimodal ERK signaling in a population over a broader range of input stimuli. The model has seven protein species: Ras, MEK, ERK, four phosphatases, and RasGTP as the input [47], and we added random log-normal noise with 10% CV to each simulation. We tested the model over a range of RasGTP input doses as a proxy for receptor input. When 15% random variation in both MEK and ERK was added to the model, the output of the model, phosphorylated ERK (pERK), which reflects ERK activity, became variable between cells and was bimodal for intermediate concentrations of EGF stimuli as shown by the time-course traces in Fig 2.6D. Figure 2.6E better illustrates the effect of adding variation to the model. The red and blue curves in Fig 2.6E show the percentage of cells with activated ERK at different doses of receptor stimulation when either 3 or 20% random variation in MEK and ERK was added. Increasing the expression variation of MEK and ERK in each simulation results in a more linear relationship between input stimulus and percent of activated cells, thus allowing for improved controllability of the percentage of activated cells in the population over a wider range of stimuli. Such a wide range over which stimuli regulate the cell function is important given that receptor stimuli have significant noise at the level of local ligand concentrations and receptor abundance.

### 2.3.7 Understanding the role of covariation in MEK and ERK expression in regulating bimodal ERK activity

Given the need for low Hill coefficients and a broadening of the relationship between input stimulus and percent of activated cells in order to optimally control population-level responses to physiological stimuli (Fig 2.5B-C), we next determined whether

covariation could be another source of overall noise that could lower the Hill coefficient and improve controllability. Such an increase in overall noise is needed as a system with 10% expression variation may not generate sufficient signaling noise for accurate population control of binary signaling responses. We had shown in Fig 2.4C that MEK and ERK covary with each other in human HeLa cells. We now also confirmed that MEK and ERK covary with each other in the human MCF10A cells used for the FRET pERK activity measurements (correlation coefficient of 0.7; Fig 2.7A). Measurements of covariation between MCM5 and MCM7 and lack of covariation between MCM5 and GAPDH are shown as controls.

Next to understand the effect of covariation in a general multi-step signaling pathway, we added covariation to the five-step model from Fig 2.5. As shown in the model output in Fig 2.7B, when covariation is added to all components in a regulatory system that has 10% variation of pathway components, the controllability of population-level binary responses is significantly improved by reducing the relationship between input stimulus and percentage of cells activated. This improvement in controllability is demonstrated by the Hill coefficient decreasing from over five if there is no covariation to down to 2.3 if covariation is added. Thus, our five-step model demonstrates that a system with high covariation of signaling components enables population-level regulation of binary outputs over a broader range of signaling inputs.

We next tested how strong the contribution is if only a pair of pathway components covaries by using the MAPK/ERK signaling model from Fig 2.6 and assuming that only MEK and ERK covary with each other. We compared how the fraction of activated cells in a population changes if MEK and ERK expression noise was random or covaried. We were cognizant that more than two pathway components may be co-dependent in typical signaling systems. As shown in Fig 2.7C (top plot), when there is covariation of a single pair of components, there was a small but significant broadening of the relationship between the stimulus intensity and the percentage of cells in the active state that seems to be particularly significant if cells have to control the activation of small fractions of cells. When considering cell differentiation as an example of this binary signaling response, a control of only 1% of the precursor population differentiates is critical physiologically since several tissues are believed

to differentiate < 1% of precursor cells at any given time [48, 5]. The significance of the contribution of covariation in this low percent range of cell activation can be seen by using a log scale for the y-axis in Fig 2.7C (bottom panel) and testing for the effect of applying noise to the input signal R. As shown in the panel, if one wants to keep 1% of the cells in the population activated (marked by the dotted horizontal line), a 10% difference in the input signal (represented by a black arrow) would result in less error in the number of activated cells (1.4-fold versus 2.3-fold accuracy in the percent of activated cells) when comparing a system with or without covariation in a single pair of pathway elements. Thus, a system with covarying components would be significantly more accurate in this physiologically relevant regime where low percentages of activated cells need to be maintained.

The model output in Fig 2.7D also confirms our experimental results from Fig 2.6C that cells in the population with high ERK activity have on average higher MEK and ERK levels compared to cells that have low ERK activity, arguing that the concentrations of MEK and ERK are limiting in the model and thus matter in determining whether or not a cell will be activated. Together, the plots in Fig 2.7C-D show that covariation of even one pair of pathway components—MEK and ERK in this case—significantly widens the range of input stimuli over which cell-fate decisions can be controlled at the population level. Covariation of more pathway components would further widen the stimulus range and further improve controllability. This result on the importance of expression variation and covariation is particularly important if organisms need to control the activation of small fractions of cells in a population such as to enable low rates of cell differentiation [5] or apoptosis [6] in tissues.

### 2.3.8 Constraints on accurate control of analog and binary signaling by expression variation and covariation

We used our experimentally measured low CV values for relative protein abundances, together with our finding that covariation can further improve the controllability of binary signaling outputs, to explore the respective ranges of variation and covariation where single-cell and population-level signaling can be effectively controlled. As

depicted in Fig 2.8A, we employed a modification of the model from Fig 2.1A to directly compare analog and binary signaling outcomes by assuming that the same pathway drives in one case an analog single-cell output ( $A^*$ ) and, in the second case, binary cell activation if the output  $y_5^*$  reaches higher than a threshold of 10 ( $B^*$ ). We use fold-Input Detection Limits (fIDLs), as defined in Fig 2.1C-D, to quantify accurate analog single-cell signaling and aHC, as defined in Fig 2.5C, to quantify accurate controllability of population-level binary signaling. As discussed in Fig S1, the fIDL parameter is a measure of analog signaling accuracy that is inversely related to mutual information but is less dependent on the dynamic range of the output, and the Hill coefficient is an inverse measure of the input range over which the population-level output can be controlled. The equations used to calculate the fIDL and aHC are shown at the top of Fig 2.8B-C (see 2.5 for derivation).

As shown in Fig 2.8B-C, single-cell analog or population-level binary outputs can be optimally controlled if the fIDL or aHC, respectively, are small and close to 1. The conflicting constraint between the control of single-cell analog and population-level binary signaling by expression variation can be seen clearly by combining the two graphs in Fig 2.8B-C into a single competition curve (Fig 2.8D). Increasing the variation in the concentration of pathway components moves cells along this curve from optimal conditions for analog single-cell signaling (CV of 5%, right bottom) toward optimal conditions for controlling binary population-level signaling (CV of 40%, left top) with the curve staying far away from the origin at the left bottom where analog and binary signaling would both be accurate. Thus, the same signaling system with a CV of 5% that has optimal analog single-cell accuracy loses its ability to accurately control binary population-level outputs. Similarly, a system with a CV of 40% that is optimal for controlling binary population-level outputs loses its ability to accurately control analog single-cell signaling. Thus, cells cannot have a shared pathway that controls accurate analog single-cell signaling outputs and also accurately controls binary population-level signaling outputs.

As shown in Fig 2.8C-D, as well as in Fig 2.5B, a CV of 40% or greater would be optimal for controlling population-level signaling outputs. However, our study and previous work by others suggests that such high CVs of protein concentrations are

not common [15, 16], indicating that cells must use other mechanisms to generate the necessary high signaling noise to accurately control the fraction of activated cells for population-level binary outputs. We considered that changes in the number of pathway components as well as the covariance of pathway components are strategies to alter the overall signaling output noise. We used the fIDL versus aHC co-dependency curve to determine how changes in pathway component numbers control analog or binary signaling (Fig 2.8E). While our analysis so far assumed 10 regulatory elements, fewer or higher numbers of signaling steps are common in signaling systems. Notably, changing the number of signaling steps improves one signaling mode at the cost of the other. Fewer signaling steps move the system toward improved analog single-cell signal transmission and more signaling steps toward improved control of population-level binary outputs. To illustrate the effect of increasing or decreasing signaling steps with examples: since many signaling systems are complex with likely 20 or more regulators [47, 16], such complex systems must necessarily be mediating population-level signaling responses. In contrast, the visual signal transduction pathway in retinal cone cells, which transduces light intensity inputs proportionally into electrical outputs, has only a few main regulatory components [9] which benefits the control of analog single-cell signaling responses.

Our modeling and experimental data in Figs 5 and 7 showed that a potent strategy to increase noise, without adding expression variation to individual components, is based on positive covariation between pathway components. Covariation can increase accurate binary signal transmission as we show in the case of the MEK/ERK signaling pathway. Indeed, Fig 2.8F shows that adding covariation moves cells away from a state where they can accurately perform analog single-cell signaling toward a state where they can accurately control the percentage of activated cells at the population level. These results suggest that covariation is a useful strategy to improve the control of population-level binary cell functions without that the expression variation or number of pathway components themselves have to be increased. We also note that covariation can in some cases increase rather than decrease, analog single-cell accuracy if directly opposing enzymes (e.g., a kinase and a phosphatase) covary with each other [12]. Together, our analysis shows that cells have a versatile set of internal

tools to control whether a signaling pathway can accurately control single-cell analog or population-level binary signaling by changing either the expression variation of individual components, the number of pathway components, or the covariation in expression between components. Furthermore, if pathways share components, these model calculations argue that analog signals have to minimize component numbers by branching out early in a pathway, while binary population-level signal responses would optimally be transmitted through more pathway steps and with pathway components covarying with each other (Fig 8G).

## 2.4 Discussion

### 2.4.1 Variation in signaling protein expression between individual cells is lower than expected

Variation in mRNA and protein expression between individual cells is believed to be a main limitation for cells to accurately transduce receptor inputs to control analog functional outputs. In particular, studies in model organisms and cultured mammalian cells suggest that the main sources of noise are likely the small number of mRNA molecules, which is common for signaling proteins, and the frequently observed bursting behavior in gene expression which can further increase the variation in the number of mRNA molecules present in a cell at a given time. Together with the observed variation in the expression of fluorescently conjugated signaling proteins and the observed variation in antibody staining of signaling proteins, it has often been assumed that the CV of signaling proteins between cells in the same population must be quite high, with numbers of about 25% CV being frequently used.

Our study investigated signaling protein variation by measuring variation of a small set of signaling proteins in *Xenopus* eggs and a subset of proteins also in mammalian cells. Specifically, we measured variation by normalizing the expression of individual proteins by the total protein mass. Using this strategy, we found CV values for proteins between individual *Xenopus* eggs of 5–10% and between human cells in the 10–15% range, much lower than expected. We used in both cases total

protein mass of a cell for normalization since cell volume is believed to scale closely with cell protein mass. Variation of the concentration is in most cases an optimal measure of variation, as the relevant parameter for the activity of a signaling protein is its cytosolic concentration, or the abundance of a particular signaling protein in a cell divided by the volume of the cell. The low CV values of 5–15% that we measured would make it possible for sensory, hormonal, or other analog signal transduction systems to accurately transduce information about the amplitude of an input to gradually control the output, as we demonstrated using a minimal model of a typical signaling pathway with five steps. As shown in Fig 1A–D, a five-step system with low variation in pathway component expression can with high accuracy distinguish a threefold increase in an input stimulus from a one-fold increase, while a system with 25% variation can only distinguish a much larger fold-increase.

#### 2.4.2 Competing roles of expression variation in enabling accurate analog single-cell signaling and controlling population-level binary signaling

A main interest of our study was to better understand the competing requirements of analog signaling systems, that need to accurately control a gradual output response of a single cell, from binary signaling systems that need to control the percentage of cells in a population that trigger a particular cell-fate transition. In the analog single-cell case, low noise is optimal, while in the latter binary population-level case, high noise is optimal. In the binary case, the critical role of increased variation in the expression of signaling proteins is to broaden the response behavior in a population of cells so that the fraction of cells that trigger a cell-fate switch can be controlled over a broader range of input signals. If there is no noise, all cells would trigger the cell-fate change at precisely the same amplitude of an input stimulus. Given that input signals can also be noisy, one can argue that optimal binary systems should be able to control whether 10 or 50% of cells in a population are activated over a range of input stimuli of about 5, which would allow for approximately linear control of the fraction of activated cells by changes in the amplitude of typical receptor inputs.

Our model calculations showed that such a system requires much higher variation in the expression of signaling proteins in the 40% range compared to the low variation required for optimal signaling for analog systems.

These competing needs for high versus low variation for different types of signaling raised the question whether cells use alternative mechanisms to increase overall signaling noise in a system in order to still allow cells to keep the variation of individual signaling proteins relatively low if these same components are also needed in other situations for binary signaling. We show that having high numbers of signaling components involved in a binary decision is a powerful strategy to generate more noise as more components make a cumulative noise contribution that increases noise. We further showed in model calculations that covariation between signaling proteins in the same pathway reduces analog signal transmission accuracy but also found that covariation can both increase the overall noise/variation of the functional output. These considerations led us to measure protein covariation in *Xenopus* eggs and human cells, and we observed a significant covariation between MEK and ERK expression in both systems. Model calculations of the ERK pathway, together with measurements of ERK activity and ERK and MEK expression levels, showed that the increased variation due to covariation increases the range over which input EGF signals can control binary ERK activity output. Of note, the contribution from a single pair of covarying signaling proteins is relatively small, and strong effects resulting from covariation require multiple signaling proteins covarying with each other. Future studies with high quality antibodies for multiple pathway components will be needed to more generally test this covariation hypothesis in different signaling systems.

#### **2.4.3 Expression variation, covariation, and number of pathway components define accuracy of analog versus binary signaling systems**

Another goal of our study was to develop a general formalism to better understand how variation and covariation of signaling protein expression and also the number of pathway can be modulated to optimize different types of signal transmission in cells.

These factors impede the accuracy of single cell signal transmission while improving the controllability of the population-level regulation of binary cell activation. Our analysis in Fig 2.8 shows a clear competition arguing that cells should have relatively low variation of signaling proteins in the 10% range if they need to reuse these same pathway components also for population-level control of binary cell fates. This implied that other strategies are needed to increase the overall noise of the signaling pathway for the control of binary decisions. We identified that having large numbers of pathway components and having covariation between pathway components are two such strategies to increase the overall noise and to allow for population level control of cell fates over broad ranges of input stimuli. We developed a simple model that shows how many pathway components are needed and how much covariation can maximally contribute to the control of binary cell fates.

In conclusion, our study employed sensitive single-cell mass spectrometry and single-cell immunofluorescence analysis to reveal a low variation in relative protein abundances with CV values in the 5–15% range, suggesting that expression variation is not prohibitively high for analog signal transmission in single cells as was often assumed in previous studies. However, such low levels of variation make it difficult for signaling pathways to control population-level binary signaling outputs over broad ranges of input stimuli. We show that covariance of signaling components and increased numbers of pathway components can be effective mechanisms to increase overall signaling output noise and thereby allow for optimal control of binary cell-fate switches at the population level even if the variation of individual signaling components is low.

## 2.5 Materials and Methods

### 2.5.1 Reagents and tools table

Insert table

### 2.5.2 *Xenopus laevis* egg collection and activation

*Xenopus* egg extracts were prepared based on modifications of a previous protocol [28]. All of the animal protocols used in this manuscript were approved by the Stanford University Administrative Panel on Laboratory Animal Care. To induce egg laying, female *Xenopus laevis* were injected with human chorionic gonadotropin injection the night before each experiment. To collect the eggs, the frogs were subjected to pelvic massage, and the eggs were collected in 1x Marc's Modified Ringer's (MMR) buffer (0.1 M NaCl, 2 mM KCl, 1 mM MgCl<sub>2</sub>, 2 mM CaCl<sub>2</sub>, 5 mM HEPES, pH 7.8). To remove the jelly coat from the eggs, they were placed in a solution of 2% cysteine in 1x MMR buffer for 4 min and gently agitated, after which they were washed four times with 1x MMR buffer. To activate the cell cycle, eggs were placed in a solution of 0.5  $\mu$ g/ml of calcium ionophore A23187 (Sigma) and 1x MMR buffer for 3 min, after which they were washed four times with 1x MMR buffer. Single eggs were collected at their respective time-points and placed into 600  $\mu$ l tubes and snap frozen in liquid nitrogen before being stored at -80°C.

### 2.5.3 SRM sample preparation

Single eggs were lysed mechanically by pipetting the egg in 100  $\mu$ l of lysis buffer (100 mM NaCl, 25 mM Tris pH 8.2, Complete EDTA free protease inhibitor cocktail (Sigma). The lysate was then placed in a 400  $\mu$ l natural polyethylene microcentrifuge tube (E&K Scientific #485050) and spun at 15,000 g in a right angle centrifuge (Beckman Microfuge E) at 4°C for 5 min. The lipid layer was removed by using a razor blade to cut the tube off just beneath it, and the cytoplasmic fraction was pipetted into a 1.5ml protein LoBind tube (Fisher Scientific #13 – 698 – 794), being careful to leave the yolk behind. To precipitate the proteins from the cytoplasmic fraction, 1 ml of ice cold acetone was added to each sample and placed at 20°C overnight.

To collect precipitated proteins, the samples were centrifuged at 18,000 g for 20 min at 4°C. Acetone was decanted, and the protein pellets were resolubilized in 25  $\mu$ l of 8 M urea. To fully solubilize the protein pellet, the samples were placed in a shaker

for 1 h at room temperature. The samples were then diluted to 2 M urea with 50 mM ammonium bicarbonate to a 100  $\mu$ l volume, after which protein concentration was measured in duplicate with a BCA assay by taking two 10  $\mu$ l aliquots of each sample. The proteins in the remaining 80  $\mu$ l of sample volume were reduced with 10 mM TCEP and incubated for 30 min at 37°C, then alkylated with 15 mM iodoacetamide and incubated in the dark at room temperature. Next, the samples were diluted to 1 M urea with 50 mM ammonium bicarbonate, and heavy peptides (JPT SpikeTides) were added based on BCA assay results. Trypsin (Promega #V5113) was then added at a ratio of 10 ng trypsin per 1ug protein (no < 500 ng was added to a sample). The trypsin digestion was carried out at 37°C for 12–16 h.

To stop the trypsin, formic acid (Fisher #A117 – 50) was added at a ratio of 3  $\mu$ l per 100  $\mu$ l of sample to bring the pH down to < 3. Peptides were cleaned up using an Oasis HLB uElution plate (Waters), equilibrated, and washed with 0.04% trifluoroacetic acid in water, and eluted in 80% acetonitrile with 0.2% formic acid. All solutions used are HPLC grade. Samples were then lyophilized. To remove any variance produced by phosphorylated peptides, the samples were phosphatase-treated. Peptides were resolubilized in 50  $\mu$ l of 1 NEBuffer 3 (no BSA), and calf intestinal alkaline phosphatase (NEB #M0290S) was added at a ratio 0.25 units per  $\mu$ g of peptide and incubated for 1 h at 37°C. The peptides were cleaned up again according to steps described above. Peptides were resolubilized in 2% acetonitrile and 0.1% formic acid before SRM analysis.

#### 2.5.4 SRM data acquisition

As detailed in previous publications [29, 5], 2  $\mu$ g of peptides was separated on an EASY-nLC Nano-HPLC system (Proxeon, Odense, Denmark) with a 200 x 0.075 mm diameter reverse-phase C18 capillary column (Maisch C18, 3  $\mu$ m, 120 Å) and were subjected to a linear gradient from 8 to 40% acetonitrile over 70 min at a flow rate of 300 nl/min. Peptides were introduced into a TSQ Vantage triple quadrupole mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) via a Proxeon nanospray ionization source. The transitions for the light (endogeneous) and heavy (SpikeTide)

peptides were measured using scheduled SRM-MS and analyzed using Skyline version 3.5 (MacCoss Lab, University of Washington). Relative peptide quantifications were determined by rationing the peak area sums of the transitions of the corresponding light and heavy peptides. Only transitions common between the heavy and light peptides with relative areas that were consistent across all samples were included in the quantification. Lists of transitions used for the 25-egg measurements in Figs 2C and 3A and C and for the 120-egg measurements in Figs 3B and 4A are given in Tables EV2 and EV3, respectively.

### 2.5.5 SRM data statistical analysis

To minimize sample processing differences, a maximum of 30 single eggs were prepped and analyzed at the same time by SRM mass spectrometry. While we normalized the amount of heavy reference peptides added to each egg extract to the measured single egg protein concentration, this leaves still a small measurement error between individual eggs. This is likely both a result of small errors in the measurement of protein concentration and small volume pipetting errors, causing small under- or overestimation of relative protein abundances in a sample. This small calibration error was in previous protocols corrected using a normalization factor measured as a median of a set of anchor protein peptides [29, 31, 49]. Here, we used the median of 22 normalized peptide intensities that minimally change during the cell cycle to derive a concentration correction factor for each egg (this factor was typically between 0.9 and 1.1). The lack of change in expression of these proteins during the cell cycle can be seen in Fig 2.1D. The correction we used makes the assumption that the 22 peptides are not overall co-regulated in the same direction, an assumption that is supported by both our SRM-MS and immunohistochemistry experiments (Fig 2.4). Specifically, we measured for a set of analyzed single eggs (e.g., 25 eggs in Fig 2.2A) the medians of the relative abundances for each of the 22 peptides across all eggs. To obtain a correction factor for each egg, we first normalized each peptide by the median of that particular peptide across all samples of interest (e.g., for the 25-egg analysis shown in Fig 2.1D, each peptide value was first divided by the median of that peptide across

all 25 samples). Then, we calculated the median of the 22 normalized peptide values for each egg. The resulting correction factor value was typically in the range of 0.9 to 1.1, and we divided all 26 relative protein abundances from that egg by this factor. The variation and covariation values shown in this paper use these corrected relative abundances.

### 2.5.6 Cell culture

MCF10A cells (ATCC, CRL-10317) were cultured in a growth media consisting of DMEM/F12 (Invitrogen) supplemented with 5% horse serum, 20 ng/ml EGF, 10  $\mu\text{g}/\text{ml}$  insulin, 0.5 ng/ml hydrocortisone, 100 ng/ml cholera toxin, 50 U/ml penicillin, and 50  $\mu\text{g}/\text{ml}$  streptomycin. HeLa cells were cultured in DMEM (Invitrogen) plus 10% fetal bovine serum (FBS) and penicillin–streptomycin–glutamine (PSG).

### 2.5.7 EKAR-EV-NLS stable cell line

pPBbsr2-EKAR-EV-NLS was described previously [50]. To generate stable cell lines, the construct was co-transfected with the piggybac transposase vector into human MCF10A cells using polyethylenimine. Cells with stable integration of the vector were selected for using 10  $\mu\text{g}/\text{ml}$  blasticidin (Invivogen).

### 2.5.8 Immunofluorescence

Cells were fixed by adding paraformaldehyde to the cell media for 15 min (final concentration of paraformaldehyde in media was 4%). Cells were then washed three times in PBS before they were permeabilized by adding 0.2% triton X-100 for 20 min at 4°C before being washed again with PBS. To remove cell size effects, cells were then stained with Alexa 647 NHS Ester as a marker of total protein mass and surrogate for cell volume/thickness following protocols described in [36]. The Alexa 647 NHS Ester was added at a concentration of 0.04  $\mu\text{g}/\text{ml}$  in PBS for 1 h. After washing again in PBS, a blocking buffer consisting of 10% FBS, 1% BSA, 0.1% triton X-100, and 0.01% NaN<sub>3</sub> in PBS was added, and the cells were incubated for 1 h at

room temperature. Then, primary antibodies were added overnight at 4°C, followed by incubation with secondary antibodies for 1 h at room temperature. To obtain particular protein concentrations for each cell, the mean total cell intensities of the respective antibodies were ratioed over the mean total cell intensity of the Alexa 647 NHS Ester.

### 2.5.9 siRNA transfection

siRNAs were used at a final concentration of 20 nM and are listed in the Reagents and tools table. MCF10A and Hela cells were reverse-transfected with siRNA using Lipofectamine RNAiMax according to the manufacturer's instructions. The cells were fixed 48 h after reverse transfection with siRNA.

### 2.5.10 Image acquisition

For both fixed and live-cell imaging, cells were plated in 96-well, optically clear, polystyrene plates (Costar #3904). Approximately 10,000 HeLa cells or 5,000 MCF10A cells were plated per well. For MCF10A cells, the wells were first coated with collagen (Advanced BioMatrix Cat #5005, PureCol Type I Bovine Collagen Solution) by placing 50 $\mu$ l of collagen dissolved at a ratio of 1:100 in PBS in each well, incubating for 2–3 h at room temperature, and then rinsing three times with PBS. MCF10A cells were then plated into the wells in MCF10A growth media. For assays to determine EGF responses, the media were aspirated from the cells 24 h after plating and replaced with serum starvation media for 60 h (DMEM/F12, 0.3% BSA, 0.5 ng/ml hydrocortisone, 100 ng/ml cholera toxin, PSG). For imaging, the cells were placed into an extracellular buffer consisting of 5 mM KCl, 125 mM NaCl, 20 mM Hepes, 1.5 mM MgCl<sub>2</sub>, 1.5 mM CaCl<sub>2</sub>, and 10 mM glucose. Time-lapse imaging was performed initially in 75 $\mu$ l of extracellular buffer per well to which an additional 75 $\mu$ l of extracellular buffer containing 2X EGF doses was added to stimulate the cells. Cells were imaged in a humidified 37°C chamber at 5% CO<sub>2</sub>. Images were taken every 2 min in the CFP and YFP channels using a fully automated widefield fluorescence microscope system (Intelligent Imaging Innovations, 3i), built around a Nikon Ti-E

stand, equipped with Nikon 20X/0.75 N.A. objective, an epifluorescence light source (Xcite Exacte), and an sCMOS cameras (Hammamatsu Flash 4), enclosed by an environmental chamber (Haison), and controlled by SlideBook software (3i). Five non-overlapping images were taken per well.

### 2.5.11 Image segmentation and tracking

Cell segmentation and tracking were performed using the “MACKtrack” package for MATLAB available at <http://github.com/brookstaylorjr/MACKtrack>, and described in [23]. In place of the first-step cellular identification using differential interference microscopy, the first pass whole-cell segmentation was performed here by thresholding the total protein stain image.

### 2.5.12 Signal measurement of segmented cells

Four-channel fluorescence images were taken with a 10X objective on a MicroXL microscope, and image analysis was performed using MATLAB analysis. Background subtraction was used in the Hoechst (to stain DNA and mask the nucleus), the two immunofluorescence, and the protein mass fluorescence channels. Signal intensities were corrected for non-uniformity but were still restricted to a central  $R = 350$  pixel region of 2x2 binned images (1,080 x 1,080 pixels) of the image to minimize potential spatial non-uniformities in illumination and light collection toward the corners. The Hoechst stain was used to establish a nuclear mask and to select cells in the 2N G0/G1 state based on the integrated DNA stain. The Hoechst intensity levels used to define cells in the 2N state were selected by inspection of the Hoechst histograms. The live-cell FRET measurements of nuclear ERK activity were performed on a Nikon Ti2 controlled by 3i software (Intelligent Imaging, Denver, CO). The mean nuclear intensities of the FRET and CFP channels were ratioed for each cell to obtain the normalized FRET value at each time-point. At the end of the time-courses, the cells were fixed and stained with either an ERK or MEK antibody, as well an Alexa 647 NHS Ester as an estimate of cell volume. To obtain ERK and MEK concentrations for each cell, the mean total cell intensities of the ERK and MEK antibodies were

ratioed over the mean total cell intensity of the Alexa 647 NHS Ester. The final ERK and MEK concentrations for each cell were then matched to the corresponding FRET time-course for that particular cell.

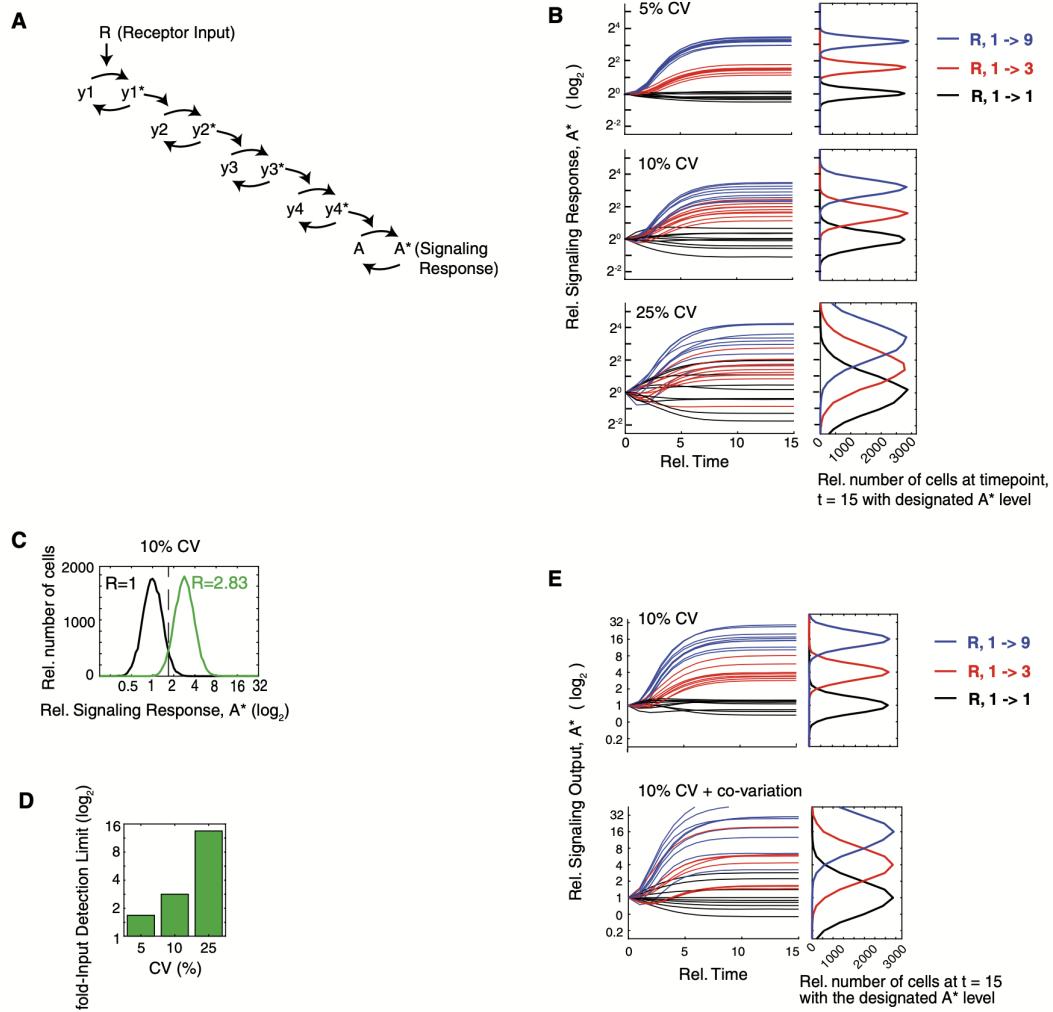
### 2.5.13 Modeling for figures 1A and 1B

## 2.6 Main Figures

## 2.7 Supplemental Figures

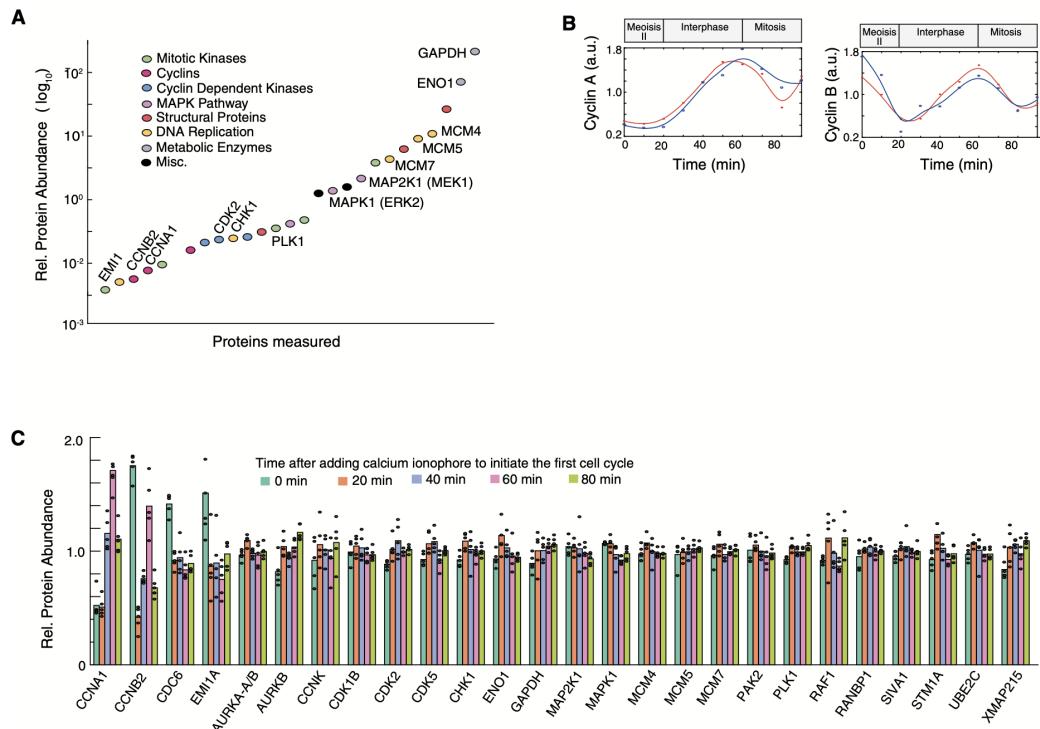
Col 1	Col 2	Col 3	Col 4
Text in Row 1a	Text in Row 1b Row 1.5b	Text in Row 1c Row 1.5c	Lots of text in Row 1d Row 1.5d
Row 2a	Row 2b	Row 2c	Row 2d
Row 3a	Row 3b	Row 3c	Row 3d

**Table S1:** Long table caption explaining everything.



**Figure 2.1:** Computational simulations using reported levels of expression variation show a dramatic loss of analog single-cell transmission accuracy

- (A) Schematic of a five-step analog signaling pathway where the asterisk represents the activated form which is assumed in this model to be a small fraction of the total.
- (B) The time course plots show how relative threefold (red) and nine fold (blue) input changes in  $R$  result in analog output responses with different degrees of noise. Random log normal expression variation was added simultaneously to each pathway component. The accuracy of analog signal transmission is dramatically reduced as the coefficient of variations (CVs) increase from 5% (top), 10% (middle), to 25% (bottom).
- (C) Example of the output response distributions of unstimulated (black) and stimulated (green) cells at the fold-Input Detection Limit (fIDL). The fIDL represents the minimal stimulus,  $R$ , needed to distinguish the output of stimulated cells from unstimulated cells with 95% accuracy, as marked by the vertical black dashed line. For the system in (A) with a 10% CV in each pathway component, the fIDL is 2.83.
- (D) Bar plot comparing the fIDL values for the system in (A) with CVs of 5, 10, and 25%.
- (E) Simulation of the pathway model in (A) but now comparing the situation in which the pathway components are all uncorrelated with each other (top) with the situation in which the activating and deactivating pathway components co-vary with each other, respectively (bottom). The overlapping output distributions in the right panels show that co-variance of components in the same pathway would introduce a marked loss in signal transmission accuracy.

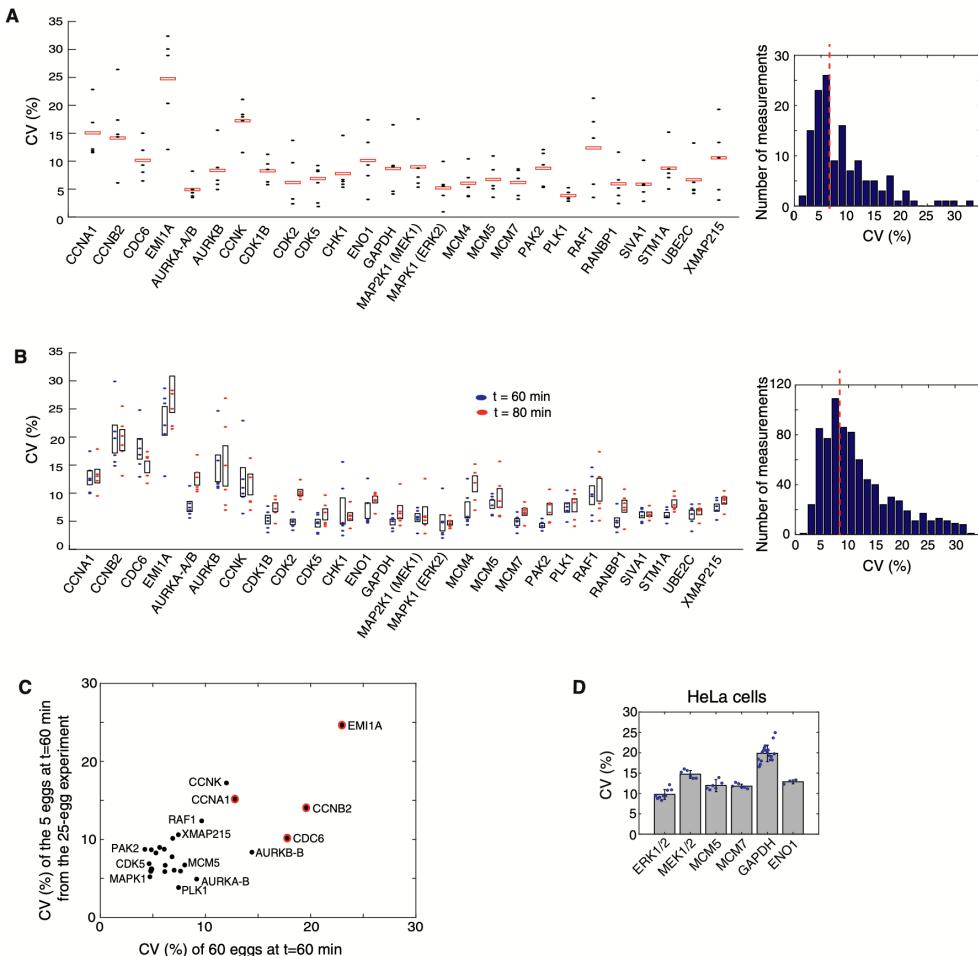


**Figure 2.2:** Development of a method to quantitatively measure relative abundances of tens of endogenous proteins in parallel in single *Xenopus* eggs

(A) Comparison of protein abundance of a set of cell cycle, signaling and control proteins in Xenopus eggs. Abundance measurements are based on SRM-MS measurements of the combined cell extracts from 5 eggs collected at the same time and before initiation of the first cell cycle. Quantitation of relative protein abundance was carried out by adding heavy isotope-labeled reference peptides to the egg extracts.

(B) Timecourse analysis of changes in Cyclin A and Cyclin B levels during the first Xenopus cell cycle measured in combined cell extracts from 5 eggs per timepoint.

(C) Five individual eggs were collected at five timepoints: 0, 20, 40, 60, and 80 min after the addition of calcium ionophore. To minimize variability due to sample handling and instrument sources, the 25 individual eggs were prepared for mass spectrometry analysis at the same time and were then analyzed in sequential runs on the same mass spectrometer. Barplot shows relative abundance changes of the 26 proteins shown in (A) tracked through the first egg cell cycle. Each black dot represents the value from an individual egg.



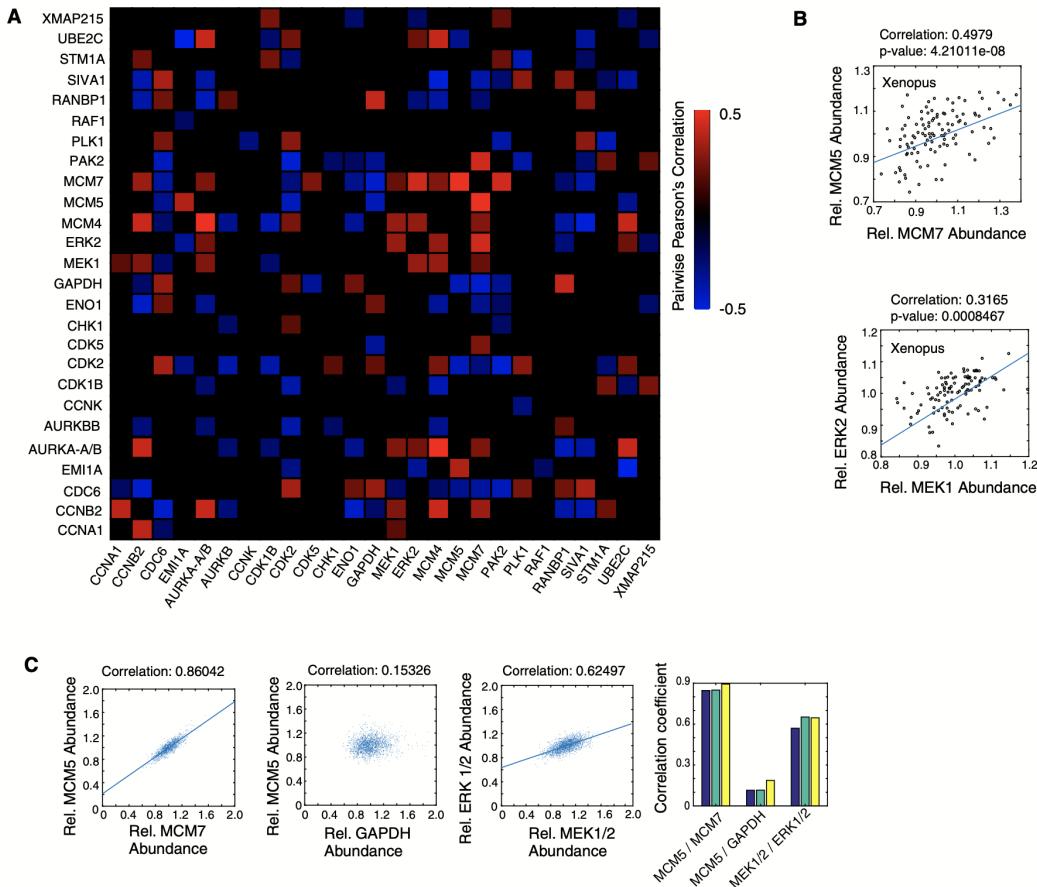
**Figure 2.3:** Single-cell variation in relative protein abundance is typically 5–10% in Xenopus eggs

**(A)** Variation analysis of the relative abundance data from Fig 2.2C. Each point represents the coefficient of variation (CV) of the relative abundance of a protein between five individual eggs in a batch collected at the same timepoint. Red boxes mark the mean CV of 5 batches, each batch collected either at 0, 20, 40, 60, and 80 minutes after cell-cycle activation. (Right) The histogram of the measured CVs for all 26 proteins at five timepoints shows that CVs typically range from 5 to 10% with a mean CV of 7%.

**(B)** Variation analysis of a second independent set of 120 eggs. Sixty individual eggs were collected at 60 (blue) and at 80 (red) minutes after the addition of calcium ionophore. The 60 eggs at each timepoint were divided into six batches of 10 eggs analyzed sequentially on the mass spectrometer to minimize technical variation. The CV of the relative abundance of each protein between 10 individual eggs in a batch was calculated and plotted as filled blue and red ovals. The black boxes mark the 25th to 75th percentile of the six batch-calculated CVs for each protein at either the 60- or 80-min timepoint. (Right) The histogram of the 312 CV measurements (6 CVs of 26 proteins at 2 timepoints) shows the mean CV is 9%.

**(C)** Control scatter plot shows that the CVs of the 26 measured proteins are similar between two independent experiments: the 25-egg experiment shown in (A) and the 60-egg, 60-min experiment shown in (B). Red circles indicate proteins that have both high CV and change their abundance during the cell cycle.

**(D)** CVs for a set of human homologs in HeLa cells. Immunocytochemistry was performed on cells plated in 96-well wells (representative images are shown in Fig EV5). Each blue dot represents the CV calculated from the ~5,000 cells in the respective well. Each barplot shows the mean CV of 3–12 wells. Error bars show standard deviation of the wells for that condition. Data shown are representative of three independent experiments.

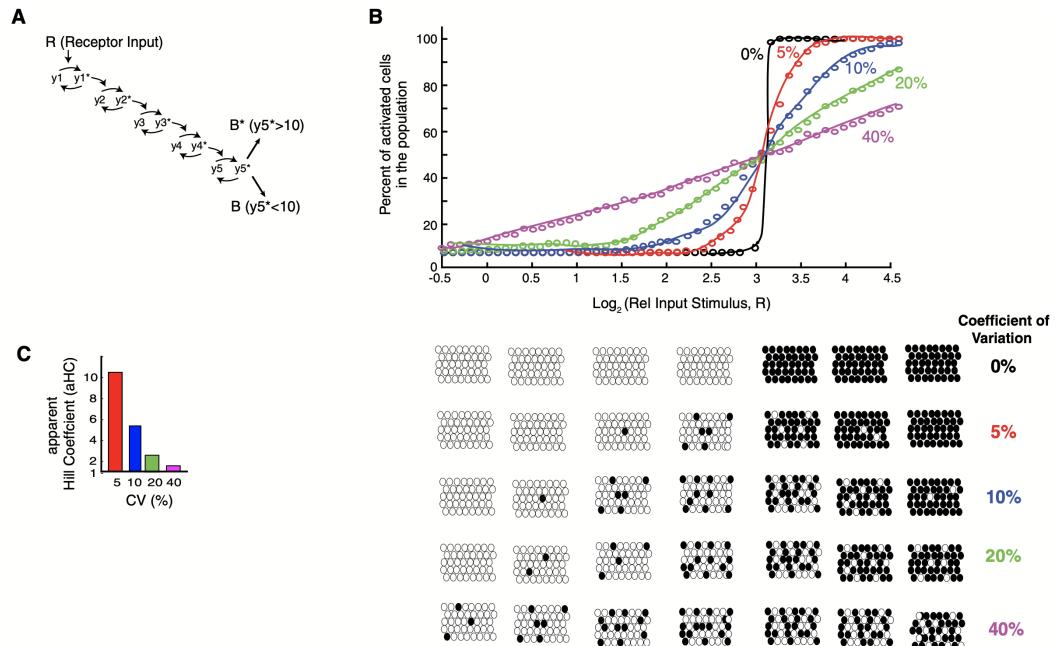


**Figure 2.4:** MEK and ERK expression covariation in Xenopus eggs and cultured human cells

(A) Heatmap of Pearson's correlation values between the respective proteins in Xenopus eggs. Twenty-six relative protein abundances were correlated pairwise in 120 single eggs. Only correlations with a P-value less than 0.05 are shown. P-values were adjusted for multiple comparison testing using Benjamini-Hochberg corrections (Table EV4).

(B) Two examples of pairwise correlations are shown between MCM5 and MCM7 and between MEK and ERK in Xenopus eggs.

(C) Pairwise correlation analysis in HeLa cells, using MCM5 versus MCM7 as a positive control and MCM5 versus GAPDH as an uncorrelated control. Correlations between MEK and ERK concentrations are shown. Each scatter plot shows values from ~15,000 cells. The bar graphs on the right show correlation coefficients for three separate wells, containing ~5,000 cells each, for the same three correlation pairs.

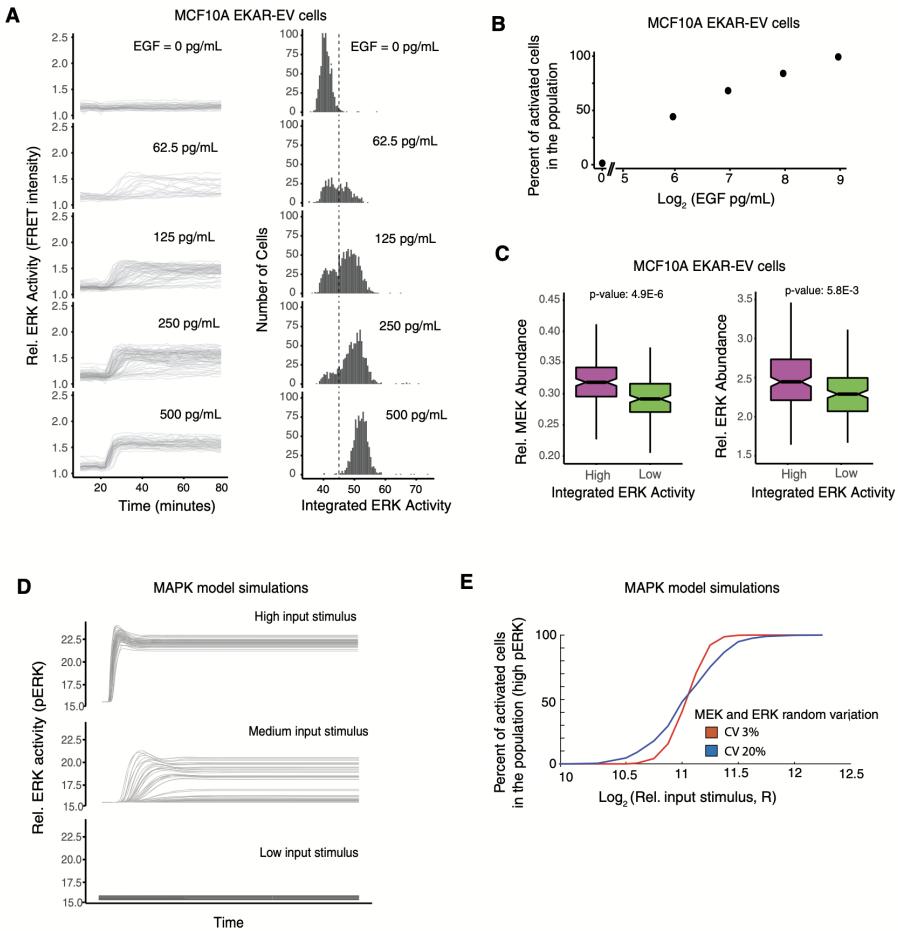


**Figure 2.5:** Using a general five-step model to understand the effect of variation on controlling the fraction of cells in the population that respond to input stimulus

(A) A binary output step was added to the model from Fig 2.1A. A threshold of 10 was used in each simulation to determine whether a cell was activated or not ( $y_5^* > 10$ ).

(B) Plot of how increasing the CVs in expression of the pathway components in this binary model from 0 to 40% increases the range over which changes in the input stimuli can change the fraction of cells in the population which trigger the binary switch and become activated.

(C) Hill coefficients were fit to the data in (B) to quantify the steepness in the curves. The steepness is an inverse measure of how wide the input range is that controls the output.



**Figure 2.6:** Live-cell imaging experiments and simulations using an established MEK/ERK signaling model show that variation between MEK and ERK expression widens the window over which input stimuli can control the fraction of cells that are activated in the population

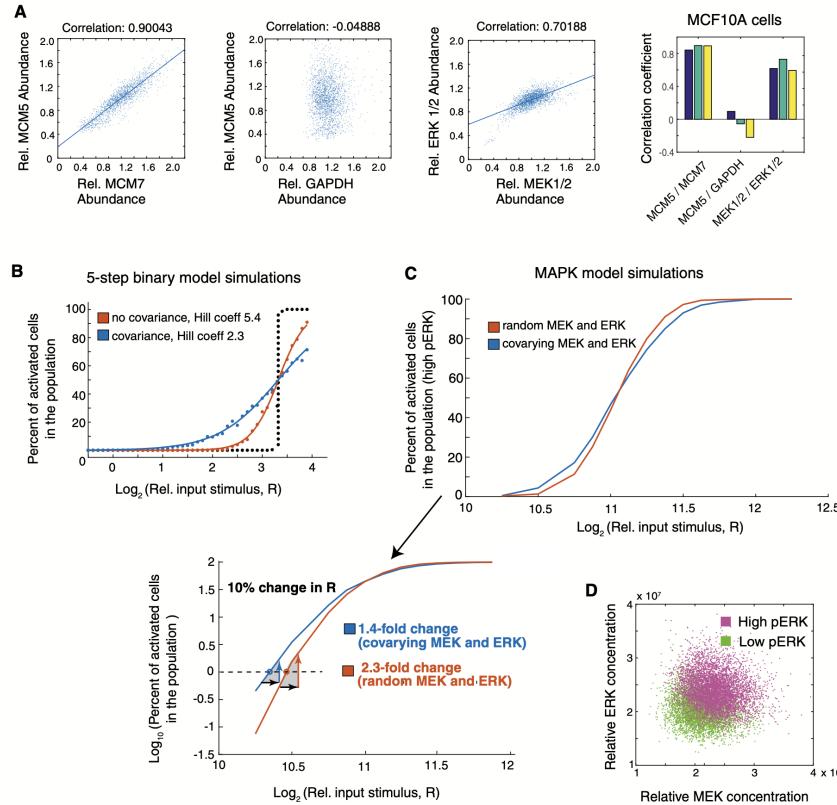
(A) MCF10A cells stably expressing the EKAR-EV FRET sensor were activated with varying concentrations of EGF after being serum starved for 48 h. (left) The plots at EGF doses show FRET intensity timecourses from > 500 individual cells. (right) Histograms show the corresponding integrated ERK activity of individual cells. The dashed line shows the threshold used to distinguish cells with active versus inactive ERK.

(B) Plot showing percentage of activated cells (cells to the right of the threshold plotted in (A)) in response to different EGF concentrations.

(C) Box-and-whisker plots of MEK (left) and ERK (right) concentrations in cells with high (top 15%, magenta) or low (bottom 15%, green) integrated ERK activity in response to EGF stimulation. The high and low conditions represent 162 and 161 cells respectively, out of a total of 1,073 cells, stimulated with 3,000 pg/ml of EGF (MEK plots), and 198 and 197 cells respectively, out of a total of 1,316 cells stimulated with 125 pg/ml of EGF (ERK plots). The non-overlapping notches between the high and low populations, as well as the low P-values, indicate that the differences between the two populations are statistically significant.

(D) Timecourse output from an established MEK/ERK model [47] in response to high, medium, and low concentrations of input (RasGTP) stimulus shows that the output for intermediate stimuli is bimodal with mainly either high pERK or low pERK cells separated by a threshold pERK intensity of approximately 17. Random log-normal noise with 15% CV was applied to MEK and ERK and 10% CV to the input stimulus (RasGTP).

(E) Model simulations resulting from applying random log-normal noise with different CVs to MEK and ERK. In all cases, random log-normal noise with 10% CV was applied to the input stimulus (RasGTP).



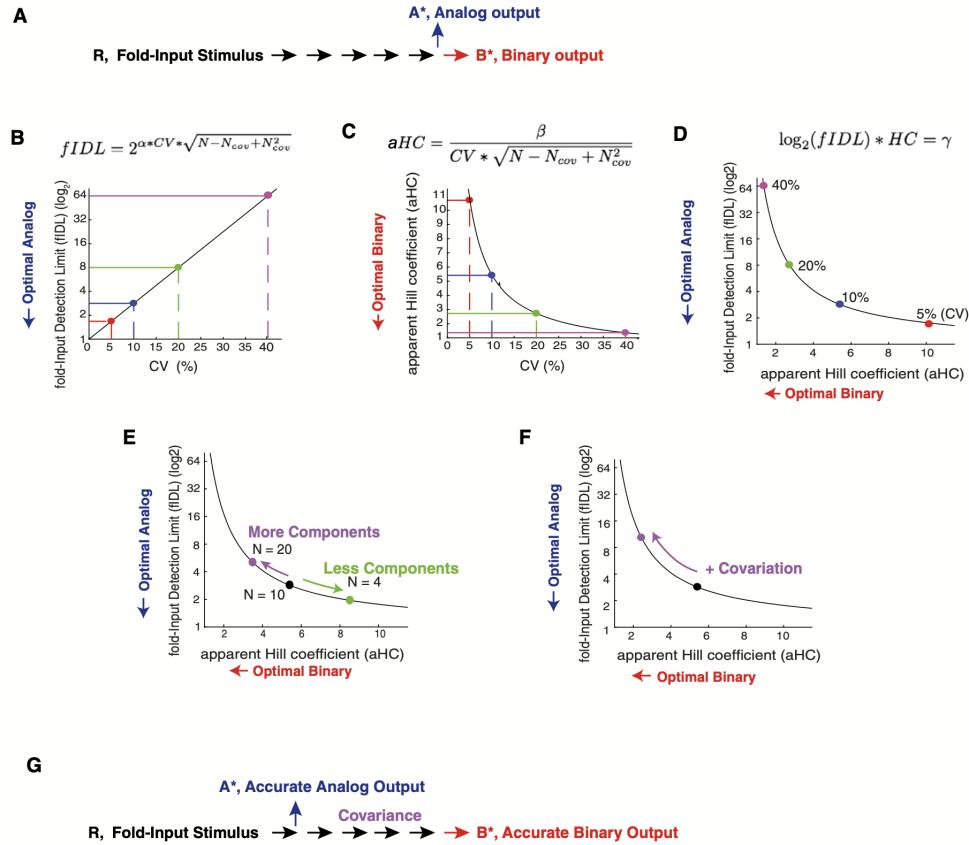
**Figure 2.7:** Single-cell imaging experiments and model simulations show that covariation between MEK and ERK expression facilitates control of bimodal ERK activation

**(A)** Immunohistochemistry experiments in MCF10A cells and pairwise correlation analysis show covariance of MEK and ERK. MCM5 versus MCM7 is used as a positive control and MCM5 versus GAPDH as an uncorrelated control. Each scatter plot shows values from ~15,000 cells. The bar graphs on the right show correlation coefficients for 3 separate wells of each correlation pair, containing 5,000 cells each.

**(B)** Using the 5-step binary model from Fig 2.5A to now look at the effect of covariance in the pathway. The same type of plot as in Fig 2.5B is shown to compare the output of the binary model if the pathway components vary randomly or covary with each other. The population response when uncorrelated CVs of 10% were applied to the pathway components is shown in red. The blue curve shows the population response when covariation was added to the model. To obtain a maximal effect, the CVs of 10% were applied to all positive and all negative regulators, respectively, such that the positive regulators covaried together and the negative regulators covaried together. Covariation in the pathway broadens the range by which input stimuli can regulate the percent of activated cells, as shown by the decrease in the apparent Hill coefficient from 5.4 to 2.3 and less steep sigmoidal response.

**(C)** Using an established MAPK model [47] to compare the effect of covarying MEK and ERK concentrations. The red curve show the results of simulations in which random log-normal noise with 15% CV was applied independently to the MEK and ERK concentrations. The blue curve shows the results of simulations in which MEK and ERK concentrations were made to covary by applying the same 15% CV lognormal noise term to both MEK and ERK in each simulation. In all cases, lognormal noise with 10% CV was applied to the input stimulus (RasGTP). The shallower slope of the blue curve show that the percent of activated cells can be regulated over a wider range of input stimuli if there is covariance between MEK and ERK.

**(D)** Output of simulations using same MAPK model as in (C). Scatter plot shows output of simulations (cells) colored by whether they had high (magenta) or low (green) ERK activity at the end of the timecourse. Cells shown were stimulated with input doses between  $2^{10.5}$  to  $2^{12}$ , a range which results in both active and inactive cells in the population as shown in (C).



**Figure 2.8:** Competing demands on variation and covariation in the control of analog single-cell versus binary population-level signaling outputs

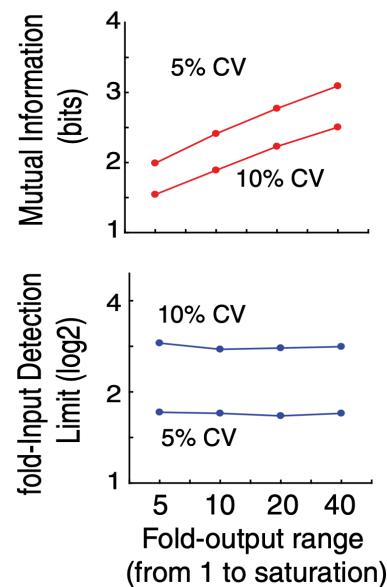
(A) Schematic of a signaling pathway that splits into an analog or binary output.

(B–D) Quantification of the competing constraints on expression variation for accurate control of single-cell analog versus population-level binary signaling outputs. Plots showing the development of a metric that quantitatively relates expression variation, analog single-cell signaling accuracy, and binary signaling accuracy. (B) Relationship between expression variation and fold-Input Detection Limit (fIDL), highlighting the physiological range of 5, 10, and 20% expression variation. (C) Relationship between expression variation and apparent Hill Coefficient (aHC). 40% expression variation enables accurate control of population-level binary outputs. (D) Integration of both relationships into a single co-dependency curve relating optimal analog single-cell and binary population-level signaling. Terms in equations: CV: expression variation; N: Total number of pathway components;  $N_{cov}$ : Number of covariant pathway components;  $\alpha = 3.3$ ;  $\beta = 1.4$ ;  $\gamma = 6.3$ .

(E) increasing or decreasing the number of regulators in a pathway increases or decreases the overall noise in the pathway, respectively, and thus can be used as a way to more accurately control either binary population-level or analog single-cell functions, respectively.

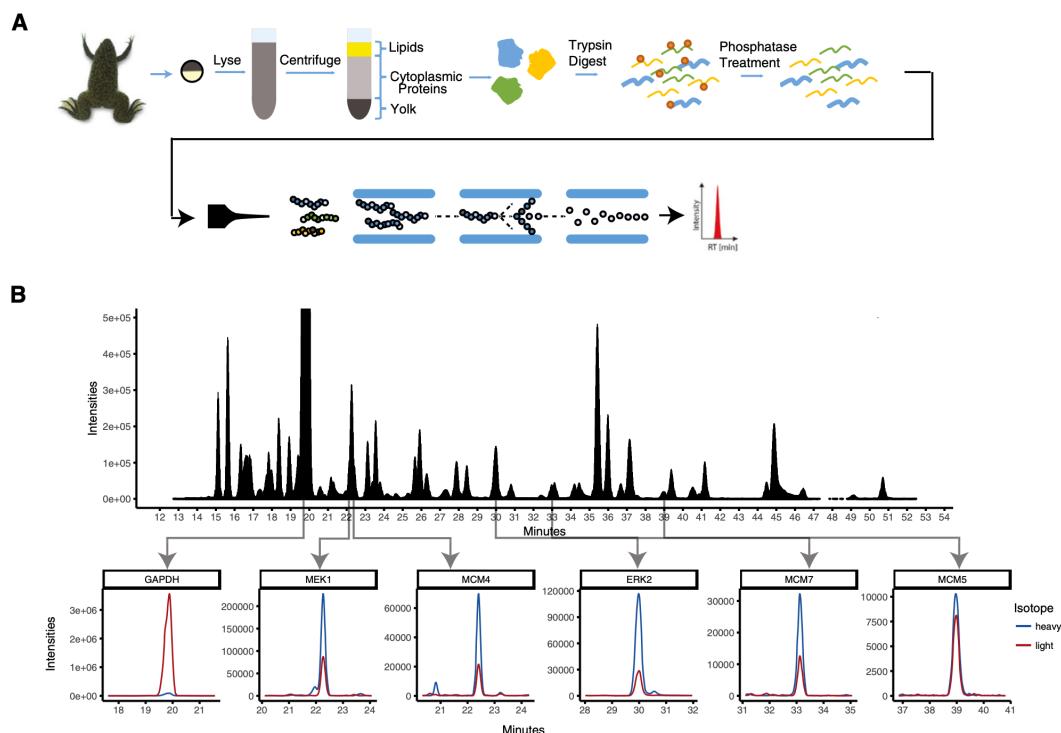
(F) Covariation between pathway components such as MEK and ERK is an effective means to increase overall noise in the pathway and thereby improve the controllability of binary population-level signaling responses while reducing single-cell analog signaling accuracy. A system with covariation can accurately control binary population level signaling without needing 40% expression variation which is likely not common.

(G) The analysis of (B–F) suggests that the same pathway components can only be shared between analog and binary signaling systems if the analog pathway branches off early after receptor stimulation. Covariation in a branch of a signaling pathway is an indication that the output is regulated by a binary output at the population level.



**Figure S1:** Comparison of fIDL and mutual information (MI) analysis

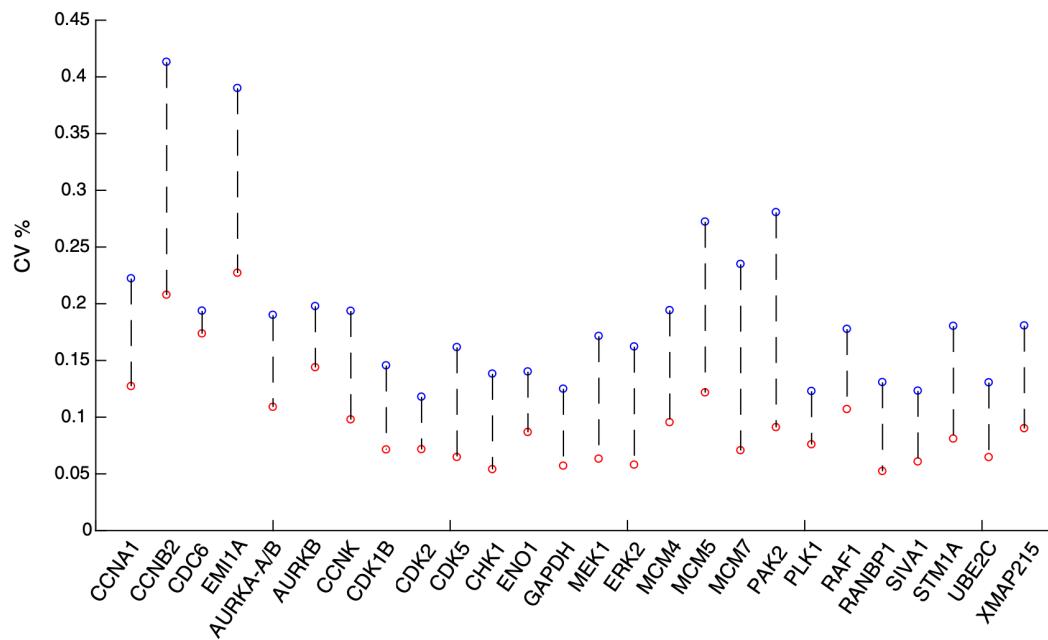
MI analysis requires a fold-output range which we added to the model in Fig 2.1A by using a saturation term for  $y_4$  (see Materials and Methods). As shown for CVs of 5 and 10%, in contrast to MI analysis (top), fIDL analysis (bottom) is largely independent of the fold-output range.



**Figure S2:** Selected reaction monitoring mass spectrometry approach to measure tens of proteins in parallel in single *Xenopus* eggs

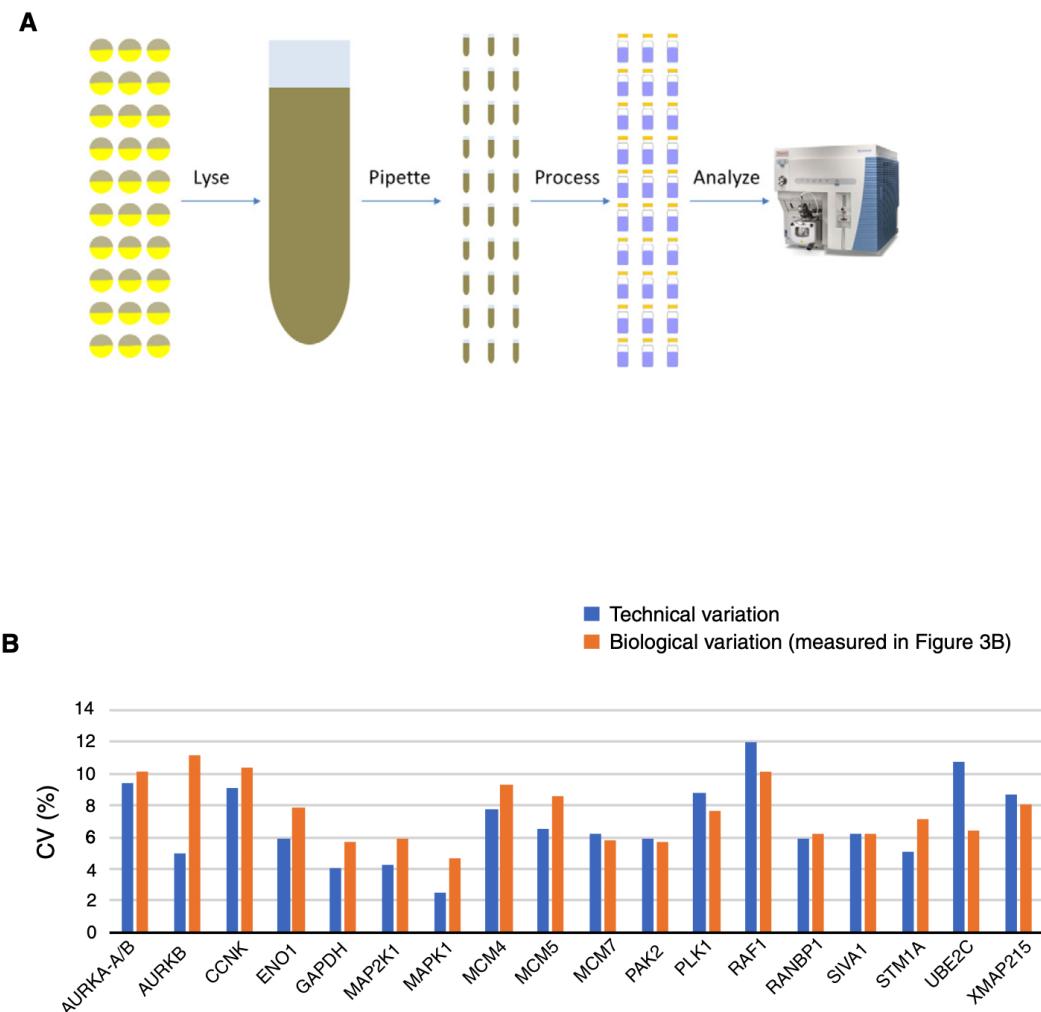
(A) Schematic of protocol to quantify the abundance of tens of endogenous proteins in parallel in a single *Xenopus* egg.

(B) Typical selected reaction monitoring mass spectrometry (SRM-MS) chromatogram.



**Figure S3:** Bootstrap analysis of CVs of the relative abundance of 26 proteins using a 60-egg set collected at 60 min after egg activation

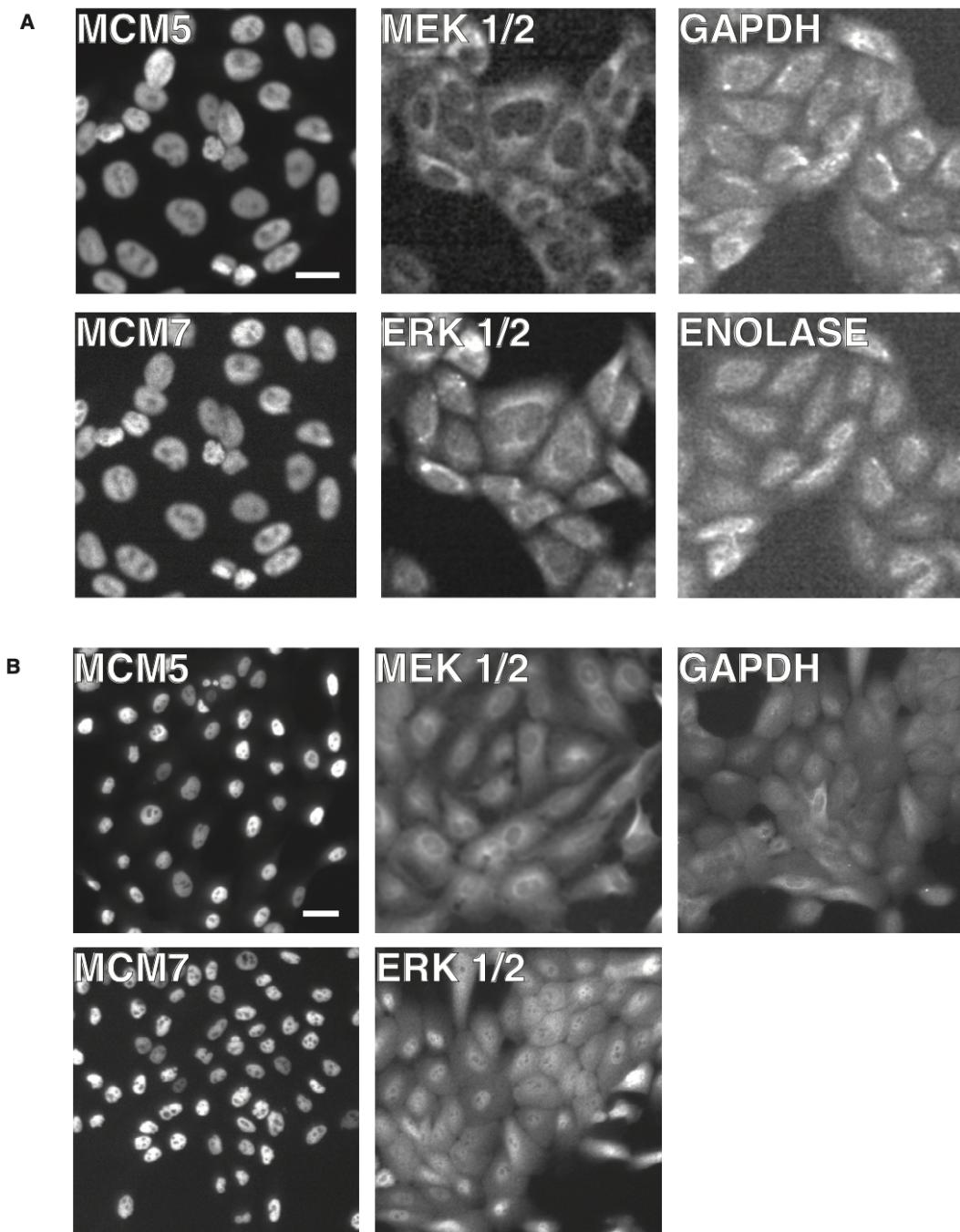
Bootstrapping of random samples was performed 2,000 times with replacement. We used the bootstrap analysis to determine CVs for the entire 60-egg dataset (blue circles) or on six batches of 10 eggs that were sequentially analyzed on the mass spectrometer (red circles). The lower CVs for batches of sequentially analyzed cells (median CV of 9% for the 26 proteins) argues that accurate concentration comparison using SRM analysis is optimally performed in batches of samples analyzed sequentially.



**Figure S4:** Comparison of technical and biological variation in the SRM mass spectrometry measurements

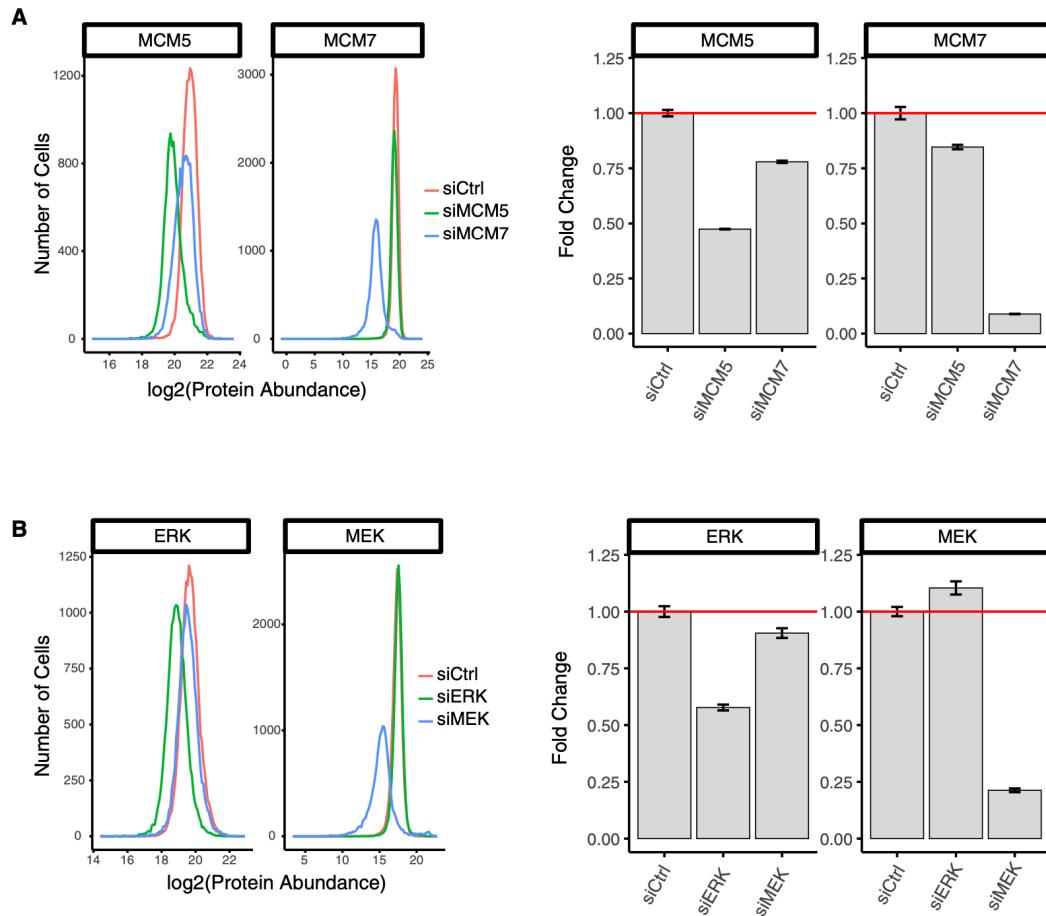
(A) Schematic of the experimental design. To measure technical variation, 30 individual eggs were lysed and mixed together to collapse any biological variability. The lysate mixture was then pipetted into 30 individual tubes and processed individually before carrying out SRM analysis to quantify sample handling variation.

(B) The CVs due to technical variation were compared to the CVs measured in Fig 2.3B.



**Figure S5:** Representative images of immunohistochemistry staining of the proteins studied

(A) Images from HeLa cells. Scale bar is  $20\mu m$ .  
(B) Images from MCF10A cells. Scale bar is  $20\mu m$ .

**Figure S6:** siRNA-mediated depletion experiments

**(A-B)** These experiments were carried out to validate the specificity of the respective antibodies and to test for co-regulation of protein expression of (A) MCM5/MC7M7 and (B) MEK/ERK. HeLa cells were transfected with the respective siRNA. Forty-eight hours later, the cells were fixed, stained, and imaged to quantify expression of the respective proteins. Approximately 1,000 cells were quantified in each histogram and barplot. Error bars show SEM. Results shown are representative of three independent experiments.

# Chapter 3

## Concluding Remarks

I conclude...

# Bibliography

- [1] K.M. Kovary, Brooks Taylor, M.L. Zhao, and M.N. Teruel. Expression variation and covariation impair analog and enable binary signaling control. *Molecular Systems Biology*, 14(5):e7997, 2018.
- [2] Steven D Cappell, Mingyu Chung, Ariel Jaimovich, Sabrina L Spencer, and Tobias Meyer. Irreversible APCCdh1 Inactivation Underlies the Point of No Return for Cell-Cycle Entry. *Cell*, 2016.
- [3] Hannah H. Chang, Martin Hemberg, Mauricio Barahona, Donald E. Ingber, and Sui Huang. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature*, 2008.
- [4] David Jukam and Claude Desplan. Binary fate decisions in differentiating neurons, 2010.
- [5] R. Ahrends, A. Ota, K. M. Kovary, T. Kudo, B. O. Park, and M. N. Teruel. Controlling low rates of cell differentiation through noise and ultrahigh feedback. *Science*, 344(6190):1384–1389, jun 2014.
- [6] Sabrina L. Spencer, Suzanne Gaudet, John G. Albeck, John M. Burke, and Peter K. Sorger. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature*, 2009.
- [7] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 1952.

- [8] I. Hide, J. P. Bennett, A. Pizzey, G. Boonen, D. Bar-Sagi, B. D. Gomperts, and P. E.R. Tatham. Degranulation of individual mast cells in response to Ca<sup>2+</sup> and guanine nucleotides: An all-or-none event. *Journal of Cell Biology*, 1993.
- [9] Vadim Y. Arshavsky, Trevor D. Lamb, and Edward N. Pugh. G proteins and phototransduction, 2002.
- [10] Mark S. Nash, Kenneth W. Young, Gary B. Willars, R. A.John Challiss, and Stefan R. Nahorski. Single-cell imaging of graded ins(1,4,5)P<sub>3</sub> production following G-protein-coupled-receptor activation. *Biochemical Journal*, 2001.
- [11] Karen E. Tkach, Debasish Barik, Guillaume Voisinne, Nicole Malandro, Matthew M. Hathorn, Jesse W. Cotari, Robert Vogel, Taha Merghoub, Jedd Wolchok, Oleg Krichevsky, and Grégoire Altan-Bonnet. T cells translate individual, quantal activation into collective, analog cytokine responses via time-integrated feedbacks. *eLife*, 3, apr 2014.
- [12] Ofer Feinerman, Joël Veiga, Jeffrey R. Dorfman, Ronald N. Germain, and Grégoire Altan-Bonnet. Variability and robustness in T cell activation from regulated heterogeneity in protein levels. *Science*, 2008.
- [13] Sabrina L. Spencer, Steven D. Cappell, Feng-Chiao Tsai, K. Wesley Overton, Clifford L. Wang, and Tobias Meyer. The Proliferation-Quiescence Decision Is Controlled by a Bifurcation in CDK2 Activity at Mitotic Exit. *Cell*, 155(2):369–383, oct 2013.
- [14] Sandip Kar, William T. Baumann, Mark R. Paul, and John J. Tyson. Exploring the roles of noise in the eukaryotic cell cycle. *Proceedings of the National Academy of Sciences of the United States of America*, 2009.
- [15] Alex Sigal, Ron Milo, Ariel Cohen, Naama Geva-Zatorsky, Yael Klein, Yuvalal Liron, Nitzan Rosenfeld, Tamar Danon, Natalie Perzov, and Uri Alon. Variability and memory of protein levels in human cells. *Nature*, 2006.

- [16] Suzanne Gaudet, Sabrina L. Spencer, William W. Chen, and Peter K. Sorger. Exploring the contextual sensitivity of factors that determine cell-to-cell variability in receptor-mediated apoptosis. *PLoS Computational Biology*, 2012.
- [17] Gürol M. Süel, Rajan P. Kulkarni, Jonathan Dworkin, Jordi Garcia-Ojalvo, and Michael B. Elowitz. Tunability and noise dependence in differentiation dynamics. *Science*, 2007.
- [18] Arjun Raj and Alexander van Oudenaarden. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences, 2008.
- [19] Tibor Kalmar, Chea Lim, Penelope Hayward, Silvia Muñoz-Descalzo, Jennifer Nichols, Jordi Garcia-Ojalvo, and Alfonso Martinez Arias. Regulated fluctuations in Nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS Biology*, 2009.
- [20] Avigdor Eldar and Michael B. Elowitz. Functional roles for noise in genetic circuits, 2010.
- [21] Ryan Suderman, John A. Bachman, Adam Smith, Peter K. Sorger, and Eric J. Deeds. Fundamental trade-offs between information flow in single cells and cellular populations. *Proceedings of the National Academy of Sciences of the United States of America*, 2017.
- [22] Raymond Cheong, Alex Rhee, Chiaochun Joanne Wang, Ilya Nemenman, and Andre Levchenko. Information transduction capacity of noisy biochemical signaling networks. *Science*, 2011.
- [23] Jangir Selimkhanov, Brooks Taylor, Jason Yao, Anna Pilko, John Albeck, Alexander Hoffmann, Lev Tsimring, and Roy Wollman. Accurate information transmission through dynamic biochemical signaling networks. *Science*, 2014.
- [24] Orsolya Symmons and Arjun Raj. What's Luck Got to Do with It: Single Cells, Multiple Fates, and Biological Nondeterminism, 2016.

- [25] L. Stryer. Visual excitation and recovery, 1991.
- [26] Gary L. Johnson and Razvan Lapadat. Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases, 2002.
- [27] Jacob Stewart-Ornstein, Jonathan S. Weissman, and Hana El-Samad. Cellular Noise Regulons Underlie Fluctuations in *Saccharomyces cerevisiae*. *Molecular Cell*, 2012.
- [28] Tony Y.C. Tsai, Julie A. Theriot, and James E. Ferrell. Changes in Oscillatory Dynamics in the Cell Cycle of Early *Xenopus laevis* Embryos. *PLoS Biology*, 2014.
- [29] Ellen Abell, Robert Ahrends, Samuel Bandara, Byung Ouk Park, and Mary N Teruel. Parallel adaptive feedback enhances reliability of the Ca<sup>2+</sup> signaling system. *Proceedings of the National Academy of Sciences of the United States of America*, 108(35):14485–14490, aug 2011.
- [30] Paola Picotti and Ruedi Aebersold. Selected reaction monitoring-based proteomics: Workflows, potential, pitfalls and future directions, 2012.
- [31] Christina Ludwig, Manfred Claassen, Alexander Schmidt, and Ruedi Aebersold. Estimation of absolute protein quantities of unlabeled samples by selected reaction monitoring mass spectrometry. *Molecular and Cellular Proteomics*, 2012.
- [32] Olivia Padovan-Merhar, Gautham P. Nair, Andrew G. Biaesch, Andreas Mayer, Steven Scarfone, Shawn W. Foley, Angela R. Wu, L. Stirling Churchman, Abhyudai Singh, and Arjun Raj. Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcriptional Mechanisms. *Molecular Cell*, 2015.
- [33] Susannah Rankin and Marc W. Kirschner. The surface contraction waves of *Xenopus* eggs reflect the metachronous cell-cycle state of the cytoplasm. *Current Biology*, 1997.

- [34] Leonid Peshkin, Martin Wühr, Esther Pearl, Wilhelm Haas, Robert M. Freeman, John C. Gerhart, Allon M. Klein, Marko Horb, Steven P. Gygi, and Marc W. Kirschner. On the Relationship of Protein and mRNA Dynamics in Vertebrate Embryonic Development. *Developmental Cell*, 2015.
- [35] Amber R. Krauchunas and Mariana F. Wolfner. Molecular Changes During Egg Activation. In *Current Topics in Developmental Biology*. 2013.
- [36] Ran Kafri, Jason Levy, Miriam B. Ginzberg, Seungeun Oh, Galit Lahav, and Marc W. Kirschner. Dynamics extracted from fixed cells reveal feedback linking cell growth to cell cycle. *Nature*, 2013.
- [37] W. H. Grover, A. K. Bryan, M. Diez-Silva, S. Suresh, J. M. Higgins, and S. R. Manalis. Measuring single-cell density. In *15th International Conference on Miniaturized Systems for Chemistry and Life Sciences 2011, MicroTAS 2011*, 2011.
- [38] Mario Niepel, Sabrina L. Spencer, and Peter K. Sorger. Non-genetic cell-to-cell variability and the consequences for pharmacology, 2009.
- [39] Erik McShane, Celine Sin, Henrik Zauber, Jonathan N. Wells, Neysan Donnelly, Xi Wang, Jingyi Hou, Wei Chen, Zuzana Storchova, Joseph A. Marsh, Angelo Valleriani, and Matthias Selbach. Kinetic Analysis of Protein Stability Reveals Age-Dependent Degradation. *Cell*, 2016.
- [40] C Atgie, F D'Allaire, and L J Bukowiecki. Role of beta1- and beta3-adrenoceptors in the regulation of lipolysis and thermogenesis in rat brown adipocytes. *Am J Physiol.*, 1997.
- [41] Prasad V.G. Katakam, Jennifer S. Pollock, David M. Pollock, Michael R. Ujhelyi, and Allison W. Miller. Enhanced endothelin-1 response and receptor expression in small mesenteric arteries of insulin-resistant rats. *American Journal of Physiology - Heart and Circulatory Physiology*, 2001.

- [42] Kenichi Kimura, David A. Low, David M. Keller, Scott L. Davis, and Craig G. Crandall. Cutaneous blood flow and sweat rate responses to exogenous administration of acetylcholine and methacholine. *Journal of Applied Physiology*, 2007.
- [43] Rony Seger and Edwin G. Krebs. The MAPK signaling cascade. *The FASEB Journal*, 1995.
- [44] John G. Albeck, Gordon B. Mills, and Joan S. Brugge. Frequency-Modulated Pulses of ERK Activity Transmit Quantitative Proliferation Signals. *Molecular Cell*, 2013.
- [45] Kazuhiro Aoki, Yuka Kumagai, Atsuro Sakurai, Naoki Komatsu, Yoshihisa Fujita, Clara Shionyu, and Michiyuki Matsuda. Stochastic ERK activation induced by noise and cell-to-cell propagation regulates cell density-dependent proliferation. *Molecular Cell*, 2013.
- [46] Hee Won Yang, Mingyu Chung, Takamasa Kudo, and Tobias Meyer. Competing memories of mitogen and p53 signalling control cell-cycle entry. *Nature*, 549(7672):404–408, sep 2017.
- [47] Oliver E. Sturm, Richard Orton, Joan Grindlay, Marc Birtwistle, Vladislav Vyshemirsky, David Gilbert, Muffy Calder, Andrew Pitt, Boris Kholodenko, and Walter Kolch. The mammalian MAPK/ERK pathway exhibits properties of a negative feedback amplifier. *Science Signaling*, 2010.
- [48] Kirsty L. Spalding, Erik Arner, Pål O. Westermark, Samuel Bernard, Bruce A. Buchholz, Olaf Bergmann, Lennart Blomqvist, Johan Hoffstedt, Erik Näslund, Tom Britton, Hernan Concha, Moustapha Hassan, Mikael Rydén, Jonas Frisén, and Peter Arner. Dynamics of fat cell turnover in humans. *Nature*, 2008.
- [49] Yuehan Feng and Paola Picotti. Selected reaction monitoring to measure proteins of interest in complex samples: A practical guide. In *Methods in Molecular Biology*. 2016.

- [50] Naoki Komatsu, Kazuhiro Aoki, Masashi Yamada, Hiroko Yukinaga, Yoshihisa Fujita, Yuji Kamioka, and Michiyuki Matsuda. Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Molecular Biology of the Cell*, 2011.

Kyle M. Kovary

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

---

(James Ferrell) Principal Advisor

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

---

(Mary N. Teruel)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

---

(James Chen)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

---

(Parag Mallick)

Kyle M. Kovary

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

---

(Lacramioara Bintu)

Approved for the Stanford University Committee on Graduate Studies.

---