AWARD NUMBER: W81XWH-16-1-0085

TITLE: Using High-Precision Signaling Activity Imaging to Personalize Ras Pathway Inhibition Strategies in Neurofibromatosis

PRINCIPAL INVESTIGATOR: John Albeck

CONTRACTING ORGANIZATION: University of California, Davis
Davis, CA 95618

REPORT DATE: June 2018

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
June 2018	Annual	06/01/2017-05/31/2018

4. TITLE AND SUBTITLE
Using High-Precision Signaling Activity Imaging to Personalize Ras Pathway Inhibition Strategies in Neurofibromatosis

5a. CONTRACT NUMBER
W81XWH-16-1-0085

5b. GRANT NUMBER
NF150076

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
John Albeck, Taryn Gillies, Carolyn Teragawa, Heather Blizard, Michael Pargett

E-Mail: jgalbeck@ucdavis.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

UNIVERSITY OF CALIFORNIA, DAVIS
1850 RESEARCH PARK DR, STE 300
DAVIS CA 95618

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The goal of this project is to characterize differences in Ras pathway signaling in using live cell activity measurements in cell types and genetic contexts relevant to NF1. During this award period, we have completed data collection for single-cell measurements of ERK activity in both cell lines and in patient-derived fibroblasts. While both technical and biological variability are substantial in these experiments, we have developed improved methods of statistical comparisons to enable us to detect significant differences in signaling kinetics in NF1-deficient cells. Our results indicate that deficiency for NF1 results in elevated ERK activity in unstimulated cells, but has no measureable effect on growth-factor stimulated pathway activity. Based on initial studies of gene expression in response to altered ERK activity, we have identified characteristics of genes whose expression level is regulated by the integrated level of ERK activity over time, providing a potential mechanism by which the elevated baseline level of pathway activity in NF1 cells may alter gene expression profiles.

15. SUBJECT TERMS

None listed

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	61	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

INTRODUCTION……. 4
KEYWORDS……. 4
ACCOMPLISHMENTS……... 4
IMPACT…….10
CHANGES/PROBLEMS……. 10
PRODUCTS…….. 11
PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS…….. 12
SPECIAL REPORTING REQUIREMENTS…… 14
APPENDICES……14

INTRODUCTION: Many of the pathological effects of *NF1* mutations result from failure to regulate the activity of the Ras signaling pathway. Inhibition of this pathway is one of the major therapeutic strategies currently being explored for NF1 patients. The effectors of the Ras pathway include the Raf/MEK/ERK signaling cascade, which stimulates proliferation and controls the expression of many downstream genes. Recent cell biology studies have shown that the kinetics of ERK activation, including the intensity, duration, and probability of response, are critical for deciding how the cell responds to Ras activation and which genes are expressed. We have also found that different inhibitors of the Ras pathway have unique effects on these kinetic features and on cellular proliferation and gene expression. Therefore, an understanding of kinetic changes in kinase activity and their effects at the cellular level is important in designing effective inhibition strategies. Our preliminary studies demonstrate our ability to resolve these kinetic differences at high temporal resolution using high-throughput live-cell imaging of ERK, while similar measurements are not possible using traditional biochemical methods. The goal of this project is to develop reporter cells and perform real-time signaling activity measurements in cell lines and genetic configurations relevant to NF1, to determine whether *NF1* mutant cells differ in signaling intensity, duration, threshold, or basal activity. We also plan to investigate how ERK kinetics in wild type and *NF1* mutant cells respond to a panel of pathway inhibitors to determine which dose levels and inhibitors represent the best strategy for normalizing mutant signaling activity and cell responses.

KEYWORDS: *Provide a brief list of keywords (limit to 20 words).* Ras, ERK, Akt, live-cell imaging, FRET reporter, inhibitor, kinetics.

ACCOMPLISHMENTS: *The PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency Grants Officer whenever there are significant changes in the project or its direction.*

What were the major goals of the project?

The main goal of this research project is to characterize the kinetic differences in Ras-driven signaling between wild type and NF1-deficient cells. The tasks and objectives are as follows:

Specific Aim 1 - Identify the kinetic differences in Ras effector activity resulting from NF1 mutations	**Timeline**	**Site 1**
Major Task 1 ***Establish and validate isogenic reporter cell line models carrying NF1 mutants***	Months	
Obtain CRISPR reagents, generate NF1+/- and -/- astrocyte, Schwann cells, and C2C12s	1-3	100%
Transfect/transduce C2C12 and MPNST cell lines, select, and test by live-cell microscopy.	4-6	100%
Transfect immortalized astrocyte and Schwann cells, test by live-cell microscopy.	6-9	100%
Milestone(s) Achieved		
Stable reporter cell lines for initial test, NF1	3	100%
Optimized imaging protocol and analysis pipeline	6	100%
All stable reporter cell lines ready	9	100%

Major Task 2 ***Measure and compare kinetic modes of ERK and Akt activity in NF1 wild type and mutant cell lines.***	Months	Completion
Collect data for different growth factor stimulation conditions	6-12	100%
Extract data and perform quality control on dataset.	6-14	100%
Perform statistical comparisons	12-16	100%
Milestone(s) Achieved:	4	
Full dataset collected	15	100%
Grid of statistical comparisons complete	16	100%
Major Task 3 ***Measure ERK and Akt kinetics in patient-derived cells***	Months	Completion
Develop stable reporter cells and optimize imaging protocol	12-15	100%
Collect full dataset, extract data, perform quality control	14-18	100%
Perform statistical comparisons	16-18	80%
Milestone(s) Achieved:		
Dataset collected	18	100%
Statistical data available	19	80%
Major Task 4 ***Correlate signaling kinetics with cellular behaviors***		
Perform live-cell and immunofluorescence experiments	16-20	80%
Data analysis	20-24	50%
Milestone(s) Achieved:		
Data set collected	20	80%
Table of similarity measurements completed	24	20%

Specific Aim 2 - Identify the Ras pathway inhibitors that best normalize signaling kinetics, gene expression, and cell proliferation rates for NF1 mutant cells.	**Timeline**	**Site 1**
Major Task 1 ***Identify Ras-pathway inhibitors that reverse single-cell NF1 kinetic modes and responses in cell lines***	Months	Completion

Perform live-cell inhibitor experiments	20-24	100%
Image data extraction and data set assembly	20-26	90%
Statistical analysis and comparisons	26-30	0%
Milestone(s) Achieved		
Dataset collected	24	100%
Statistical data available	30	0%
Major Task 2 ***Compare drug responses between different patient-derived cells***		
Collect inhibitor data on patient cell lines	18-24	100%
Image data extraction and data set assembly	22-25	100%
Statistical analysis and comparisons	28-32	20%
Milestone(s) Achieved:		
Dataset collected	26	100%
Statistical data available	32	20%
Major Task 3 ***Link gene expression in response to Ras pathway inhibitors to cellular responses***		
Perform inhibitor treatment and proliferation/differentiation assays	24-30	0%
Collect mRNA samples and perform mRNAseq.	24-30	0%
Data analysis	28-32	0%
Milestone(s) Achieved		
Cell behavior data available	30	0%
mRNAseq data available	30	0%
Comparison of inhibitor data complete	32	0%
Major Task 4 ***Characterize individual cell adaptive responses to Ras pathway inhibitors in MPNST cells***		
Perform live-cell imaging of inhibitor treatment in MPNST cells	30-32	0%
Extract image data and perform analysis	32-36	0%
Milestone(s) completed		
Image data set available	34	0%
Statistical comparisons complete	36	0%

What was accomplished under these goals?

Following the strategy outlined above, we now have collected a dataset from multiple cell types that provides a consistent answer to the question of how mutant NF1 cells differ from wild type in Ras signaling kinetics. We find that NF1 deficiency results in an elevated baseline of Ras signaling, but does not alter the amplitude or duration of growth factor-stimulated signaling events.

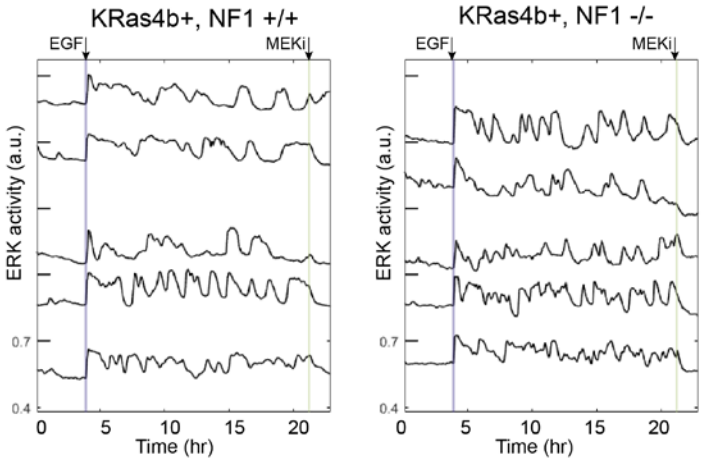


Figure 1. Example single-cell activity profiles. MEFs with the indicated genetic configuration were treated first with EGF, followed by MEK inhibitor (MEKi) during live-cell imaging. ERK activity was calculated based on FRET ratio and is shown for 5 randomly selected cells (out of several hundred collected per experiment) in each condition.

One line of evidence for this conclusion comes from an isogenic series of cell lines that we have investigated in collaboration with Frank McCormick's group. These cells are mouse embryonic fibroblasts that have been engineered to express single Ras gene isoforms. The advantage of this system is that differences in signaling due to Ras patway mutations can be definitively attributed to the genetic configuration of the Ras pathway, without ambiguities arising from other mutations, such as the ones typically found in cancer cell lines. This panel includes both KRas4b^{WT}/NF1+/+ and KRas4b KRas4b^{WT}/NF1-/- cells, enabling a direct comparison between NF1-deficient and wild type cells. Also included are a number of mutant isoforms of KRas (including G12V, G12C, and Q61R), which enabled us to compare the effects of loss-of-function mutants in NF1 (a Ras GTPase-activating protein, or GAP) to the effects of cancer-driving mutations that prevent Ras from responding to GAPs. We have expressed an ERK activity reporter in each of the cell lines of this panel and have performed extensive live-cell microscopic interrogation to assess the differences between these cells. Computational image processing was used to extract activity profiles from each individual cell. Examples of such profiles can be found in Figure 1.

To make quantitative comparisons of different kinetic aspects of ERK activity, we developed novel analysis procedures to compare the activity profiles of each cell. Each single cell trace was normalized to the minimum value in a 1 hour window following treatment with the MEK inhibitor PD0325901 (typically added at the end of the experiment). Baseline values were calculated by taking the mean of the 3 hour window prior to stimulation for each cell. The mean was calculated from a 3.5 hour window following treatment with the specified growth factor or vehicle control. Mean derivative is calculated as the sum of the absolute value of the derivative over a 3.5 hour window following stimulation. Responders were defined as cells who's post-stimulation ERK activity increased at least 5% compared to the average baseline value, had a higher maximum derivative at the time of stimulation compared to the baseline region, and had a significantly different ($p < .05$) distribution of ERK activity values following stimulation compared to the distribution in the baseline region. Response location was defined as the time point where the cell reached its maximum ERK activity value following stimulation in each responding cell. Duration was calculated as full width at half maximum for each responding cell. The mean of each metric was calculated for every replicate experiment, and the distribution of these replicate means was compared for each cell line against K-Ras wild type. Each comparison was made by t-test, with false discovery rate controlled across all comparisons via the method of Benjamini and Hochberg. For metrics where replicates were available at the single cell level, as well as across experiments, the standard deviation of the mean for each experiment was determined from single-cell samples and added to experimental variance according to a linear error model.

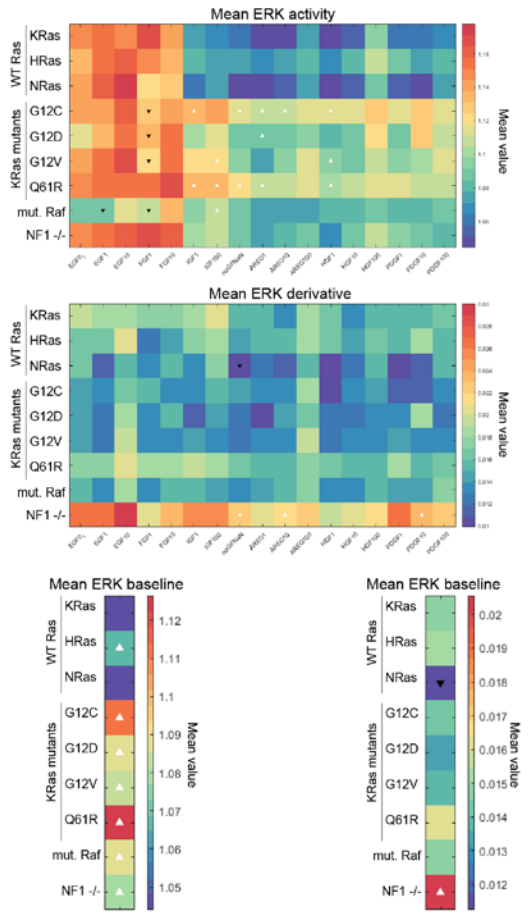


Figure 2. Comparison of quantitative kinetic metrics of ERK activity between MEFs with different Ras configurations. MEFs expressing single Ras isoforms as indicated at the left (including KRas-wt, NF1-/-, bottom row) were scored as described in the main text, across various treatment conditions (listed across the bottom). The value of each metric is indicated by color; triangles indicate a statistically significant increase (white) or decrease (black) relative to the other cell lines.

This comprehensive comparison of kinetic metrics allowed identification of metrics in which NF1-/- cells differed significantly from wild type cells. Example grids of statistical comparisons are shown in Fi. 2. As can be seen from these plots, NF1-/- cells differ in both their average baseline, which is a measure of ERK activity in the absence of stimulus, and in the mean deriviative of the baseline, which indicates that there is substantial fluctuation in the baseline level of each cell. Significant differences were not detected in growth factor-stimulated metrics such as amplitude or duration of ERK activity following EGF treatment; these were similar between wild type Ras, mutant Ras, and NF1-/- cells.

A second line of evidence for differences in Ras activity in NF1-deficient cells comes from the C8-D1A astrocyte cell line, in which we have also expressed ERK activity reporters. After isolation of a stable population of reporter cells, we introduced shRNA hairpins targeting NF1 mRNA using a lentiviral expression system. After selection, we confirmed reduction of NF1 protein by immunoblot. Extensive live-cell analysis of these NF1-kd cells showed that they responded to growth factor stimulation with similar magnitude and duration of ERK activation (Figure 3). However, when in an unstimulated state, they displayed higher level of ERK activity than the control cells, and treatment with MEK inhibitor under these conditions resulted in a larger drop in reporter signal than in the control cells. We conclude that NF1 deficiency in these model astrocyte cells results in a higher level of ERK activity in the absence of stimulation. This result is consistent with the data obtained from the genetically engineered fibroblast experiments described above. We conclude that NF1 deficiency primarily influences the baseline level of Ras/ERK activity rather than altering the kinetics of activation upon stimulation.

To understand how an elevated baseline level of Ras signaling impacts cellular behavior downstream of this pathway, we examined the expression of the Ras regulated gene Fra-1 in a model cell system, MCF-10A. Using the methodology developed in this project to quantify Ras/ERK signaling kinetics, we performed an extensive analysis of the relationship between ERK activity measurements and expression of Fra-1, which has previously been linked to invasive behavior in malignant cells. As detailed in the attached manuscript (#1), we found that Fra-1 expression increases proportionally to the integrated area of ERK activity, regardless of the amplitude or duration of activity. Such a response profile would be highly sensitive to the increased baseline activity of Ras/ERK signaling found in NF1-deficient cells. We also found evidence that other ERK-activated genes such as c-Fos and Egr-1 are less sensitive to cumulative activity and more responsive to sharp changes in activity. This finding raises the hypothesis that NF1 deficiency selectively activates genes involved in directing invasive cellular behavior, which we are currently testing further.

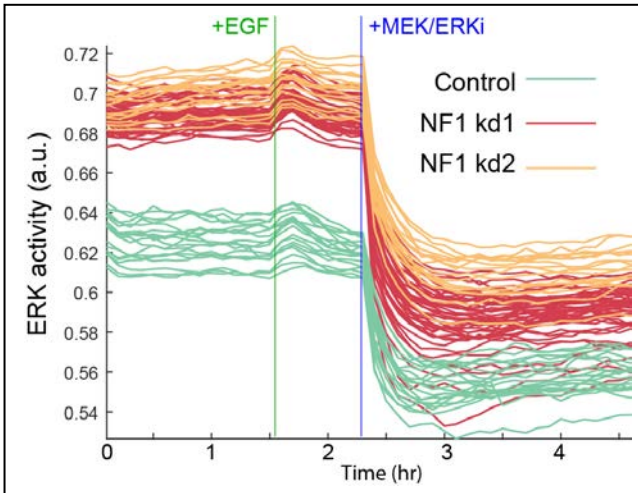


Figure 3. Measurements of ERK activity kinetics in astrocyte cell lines. C8D1A cells stably expressing an ERK activity reporter and shRNAs targeting NF1 were monitored by live cell microscopy and treated with 5 ng/ml FGF, followed by a combination of MEK and ERK inhibitors at the indicated times. Each line indicates an averaged measurement of 50-200 individual cells.

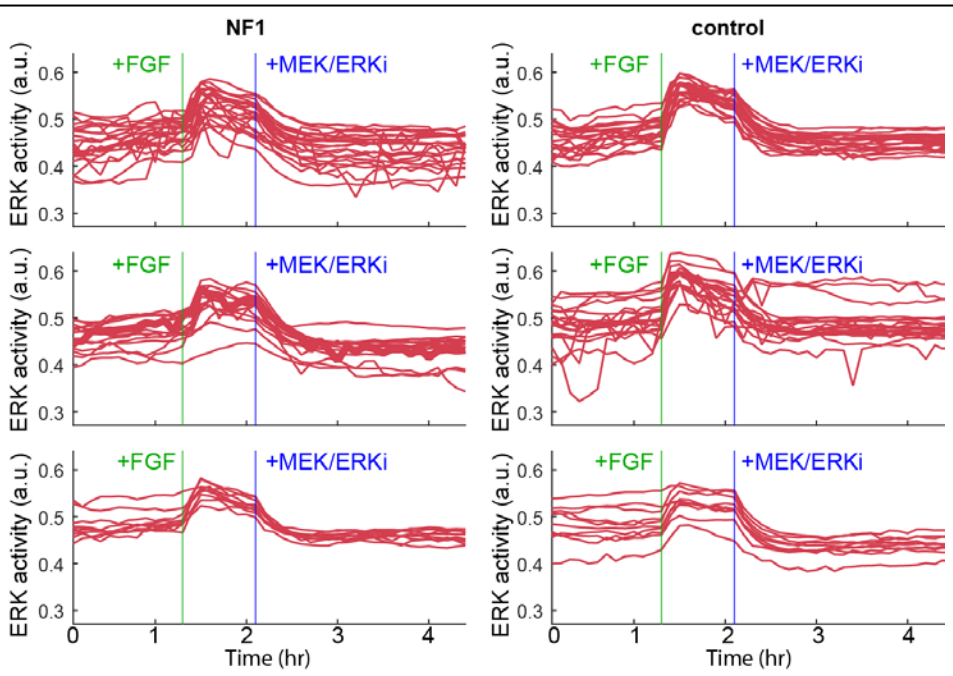


Figure 4. Reporter-based measurements of ERK activity in primary skin fibroblasts derived from 3 NF1 and 3 control individuals. After stable introduction of reporter, cells were monitored by live-cell microscopy and treated with 5 ng/ml FGF, followed by a combination of MEK and ERK inhibitors at the indicated times. Each line indicates an averaged measurement of 50-200 individual cells.

Finally, to examine how Ras signaling kinetics are altered in NF1-affected relative to healthy individuals, we used our methodology of single-cell measurement to compare signals in skin fibroblasts derived from patients and volunteers. Using gene delivery methods refined in the previous reporting period, we generated populations of primary fibroblasts expressing a FRET-based ERK reporter. These cells were obtained from our collaborators Katherine Rauen (UC Davis MIND Institute) and Lauren Weiss (UCSF) and included both NF1-affected individuals (*NF1* +/-) and healthy (*NF1*+/+) individuals. We have collected a full dataset from these cells and are currently completing the statistical comparisons. Our observations to date indicate that there is little difference between NF1 and control individuals at the level of signaling. A potential explanation for this similarity is that fibroblasts from NF1 individuals are heterozygous for *NF1*, and detectable changes in ERK activity may only manifest upon loss of the second allele. Evidence from mouse models has suggested that +/- cells are not equivalent to -+/+ cells in supporting fibroma development, but this difference may be too small to detect against the background of genetic variability between different individuals. These results further support the concept that the changes in Ras activity in cells with reduced NF1 activity are relatively subtle at the level of ERK activity.

What opportunities for training and professional development has the project provided?

Nothing to report.

How were the results disseminated to communities of interest?

In addition to publications, results from this project were presented in at the 2017 International RASopathies Symposium (Orlando, FL), attended by scientists, patients, and families, and at the 2017 Ras Initiative Symposium (Frederick, MD).

What do you plan to do during the next reporting period to accomplish the goals?

Our reporter measurements to date have shown that ERK signaling in NF1-deficient cells is only moderately different from wild type cells. These findings raise the question of how these relatively small changes result in changes in cell behavior that ultimately have a large impact on tissue physiology. One possibility, suggested by publication #1, is that small changes in signaling activity over time lead to a large cumulative change at the level of gene expression. Thus, while we plan to continue to follow the original research strategy, we will place particular emphasis on Aim 2, Task 3, in which gene expression profiles will be linked to changes in signaling activity. Our plan is to carry out this aim using new technologies now available to our lab, including single-cell mRNA sequencing by drop-Seq and live-cell reporters of gene expression as developed in publication #1.

IMPACT: *Describe distinctive contributions, major accomplishments, innovations, successes, or any change in practice or behavior that has come about as a result of the project relative to:*

What was the impact on the development of the principal discipline(s) of the project?

Our work to date has shown how the behavior of a signaling pathway can be measured quantitatively at the single-cell level to understand the significance of changes in biochemical kinetics. A published commentary on publication #1 stated, “These studies are a potent demonstration of the power of new technologies to answer long-standing questions in systems biology.” (Wiley, H.S., “Decoding Signal Processing at the Single-Cell Level.” (2017) *Cell Syst*. 5(6):542-543)

What was the impact on other disciplines?

Consideration of the broader impact of our work has found parallels to other biological processes in which quantitative outputs vary, such as seed germination, and has highlighted the potential utility of our data for developing quantitative models for such systems (Bradford, K.J. “Interpreting biological variation: seeds, populations and sensitivity thresholds” (2018) *Seed Science Research* 1–10. https://doi.org/10.1017/S0960258518000156)

What was the impact on technology transfer?

Nothing to report.

What was the impact on society beyond science and technology?

Nothing to report.

CHANGES/PROBLEMS:

Changes in approach and reasons for change

Nothing to report.

Actual or anticipated problems or delays and actions or plans to resolve them

Some aspects of the original research plan remain technically challenging. These include measurements of Akt activity, which we originally planned to monitor using a reporter based on nucleus-to-cytosol translocation of a fluorescently tagged substrate of Akt (FoxO3a). However, such translocation-based reporters are problematic in cell lines with highly irregular cytosolic shapes, including fibroblasts and astrocytes. Accurate quantification of Akt activity in individual cells remains obstructed by cell-to-cell variability, despite our attempts to improve the methodology. However, given our established methodology to measure ERK activity, this difficulty has not limited our ability to track Ras pathway activity. We are continuing to attempt to improve the methodology for Akt pathway activity.

Another challenge has been working with MPNST cell lines, which display heterogeneous morphology, have inconsistent growth and differentiation behavior in culture, and are resistant to stable gene expression. We are still assessing whether we will be able to complete the objectives planned for these cells, or if an alternative model will need to be used.

We have encountered another moderate delay in Aim 1, Task 4, due to the time-consuming nature of accurately identifying cell fates in time lapse images. However, we have been able to develop a software algorithm that makes it possible to automatically identify cell fates, enabling us to associate them with ERK reporter measurements in the same cell more efficiently. We anticipate that this computational method will enable us to complete this task within the next 2-3 months.

Changes that had a significant impact on expenditures

Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Nothing to report.

Significant changes in use or care of human subjects

Nothing to report.

Significant changes in use or care of vertebrate animals.

Nothing to report.

Significant changes in use of biohazards and/or select agents

PRODUCTS:

Publications, conference papers, and presentations
Nothing to report.

Journal publications.

1. “Linear integration of ERK/MAPK activity and lack of persistence detection by Fra-1”
(2017) Taryn E. Gillies, Michael Pargett, Marta Minguet, Alex E. Davies, John G. Albeck

Cell Systems 5(6):549-563.
Federal support acknowledged.

2. “Live-Cell Imaging and Analysis with Multiple Genetically Encoded Reporters”
(2017) Michael Pargett John G. Albeck
Current Protocols in Cell Biology 78(1): 4.36.1-4.36.19.
Federal support acknowledged.

Books or other non-periodical, one-time publications.

Nothing to report.

Other publications, conference papers, and presentations.

Nothing to report.

Website(s) or other Internet site(s)
Nothing to report.

Technologies or techniques
Our detailed methodology for image and reporter signal analysis has been published in Publication #2. Updates to these techniques will be included with forthcoming publications.

Inventions, patent applications, and/or licenses
Nothing to report.

Other Products
Reporter cell lines were developed as described above, and are made available upon publication. Signal analysis software has been developed as described in our publications and is made available upon request.

PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Provide the following information for: (1) PDs/PIs; and (2) each person who has worked at least one person month per year on the project during the reporting period, regardless of the source of compensation (a person month equals approximately 160 hours of effort). If information is unchanged from a previous submission, provide the name only and indicate "no change."

Name:	*John Albeck*
Project Role:	*PI*
Researcher Identifier (e.g. ORCID ID):	*1234567*
Nearest person month worked:	4
Contribution to Project:	*Supervised all areas of research on the project, participated in data generation and analysis.*

Funding Support:	*N/A*

Name:	*Taryn Gillies*
Project Role:	*Graduate Student*
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	*3*
Contribution to Project:	*Developed signal analysis methods for application to ERK and Akt signals; constructed cell lines and performed microscopy experiments*
Funding Support:	*This award*

Name:	*Heather Blizard*
Project Role:	*PI*
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	*12*
Contribution to Project:	*Maintained cell cultures, generated DNA constructs, developed reporter cells and cell lines; prepared frozen cell stocks*
Funding Support:	*This award*

Name:	*Michael Pargett*
Project Role:	*Postdoctoral Associate*
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	*4*
Contribution to Project:	*Developed, maintained, and debugged image analysis software; developed signal analysis techniques; performed data analysis*
Funding Support:	*This award and National Institutes of Health*

Name:	*Carolyn Teragawa*
Project Role:	*Research Scientist*

Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	*3*
Contribution to Project:	*Performed cell culture and microscopy experiments*
Funding Support:	*This award and National Institutes of Health*

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

No

What other organizations were involved as partners?

- **Organization Name:** University of California, San Francisco
- **Location of Organization:** San Francisco, CA
- **Partner's contribution to the projec**t:

  Collaboration NF patient and control fibroblasts were obtained from Dr. Lauren Weiss; NF1-deficient mouse fibroblasts were obtained from Dr. Frank McCormick. Both Drs. Weiss and McCormick contributed to discussions of experimental planning and preliminary data.

SPECIAL REPORTING REQUIREMENTS

Nothing to report.

APPENDICES

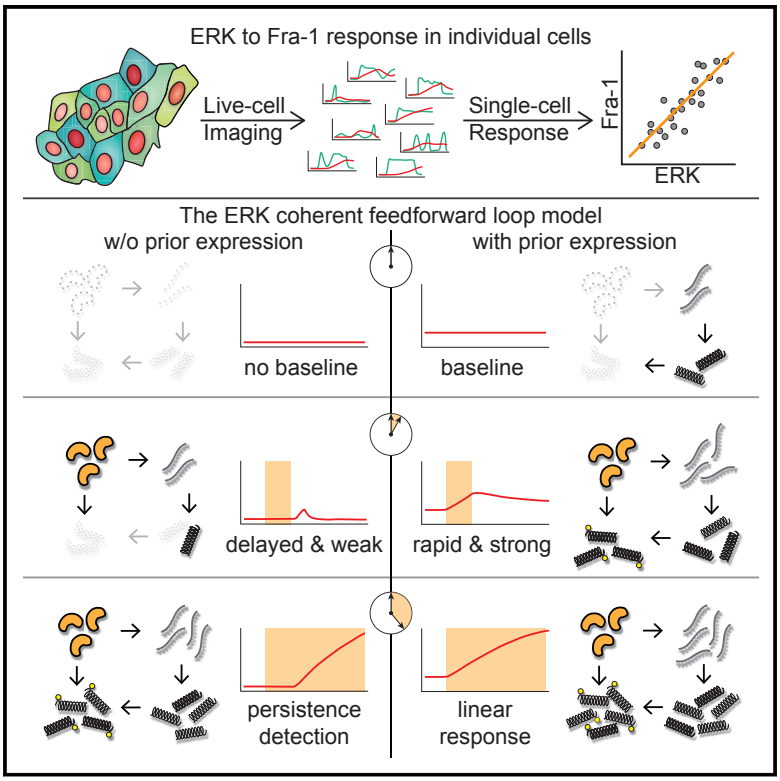
Publications #1 and #2, referenced above, are attached .

Article

Cell Systems

Linear Integration of ERK Activity Predominates over Persistence Detection in Fra-1 Regulation

Graphical Abstract



Highlights

- Multiple reporters are needed to capture the full dynamic range of ERK activity
- Fra-1 induction is related linearly to the integrated activity of ERK
- Basal transcription rate determines whether a gene may act as a persistence detector
- ERK-independent factors contribute substantially to Fra-1 variance in single cells

Authors

Taryn E. Gillies, Michael Pargett, Marta Minguet, Alex E. Davies, John G. Albeck

Correspondence

jgalbeck@ucdavis.edu

In Brief

Using a combination of live-cell ERK activity reporters and mathematical analysis, Gillies et al. determine that the immediate-early gene, Fra-1, is produced in linear relation to total ERK activity. This demonstrates that the network functions as a linear integrator, not a selective filter, of ERK activity patterns.

Gillies et al., 2017, Cell Systems *5*, 1–15
December 27, 2017 © 2017 Elsevier Inc.
https://doi.org/10.1016/j.cels.2017.10.019

CellPress

Please cite this article in press as: Gillies et al., Linear Integration of ERK Activity Predominates over Persistence Detection in Fra-1 Regulation, Cell Systems (2017), https://doi.org/10.1016/j.cels.2017.10.019

Cell Systems
Article

CellPress

Linear Integration of ERK Activity Predominates over Persistence Detection in Fra-1 Regulation

Taryn E. Gillies,[1,3] Michael Pargett,[1,3] Marta Minguet,[1] Alex E. Davies,[1,2] and John G. Albeck[1,4,*]
[1]Department of Molecular and Cellular Biology, University of California, Davis, CA 95616, USA
[2]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[3]These authors contributed equally
[4]Lead Contact
*Correspondence: jgalbeck@ucdavis.edu
https://doi.org/10.1016/j.cels.2017.10.019

SUMMARY

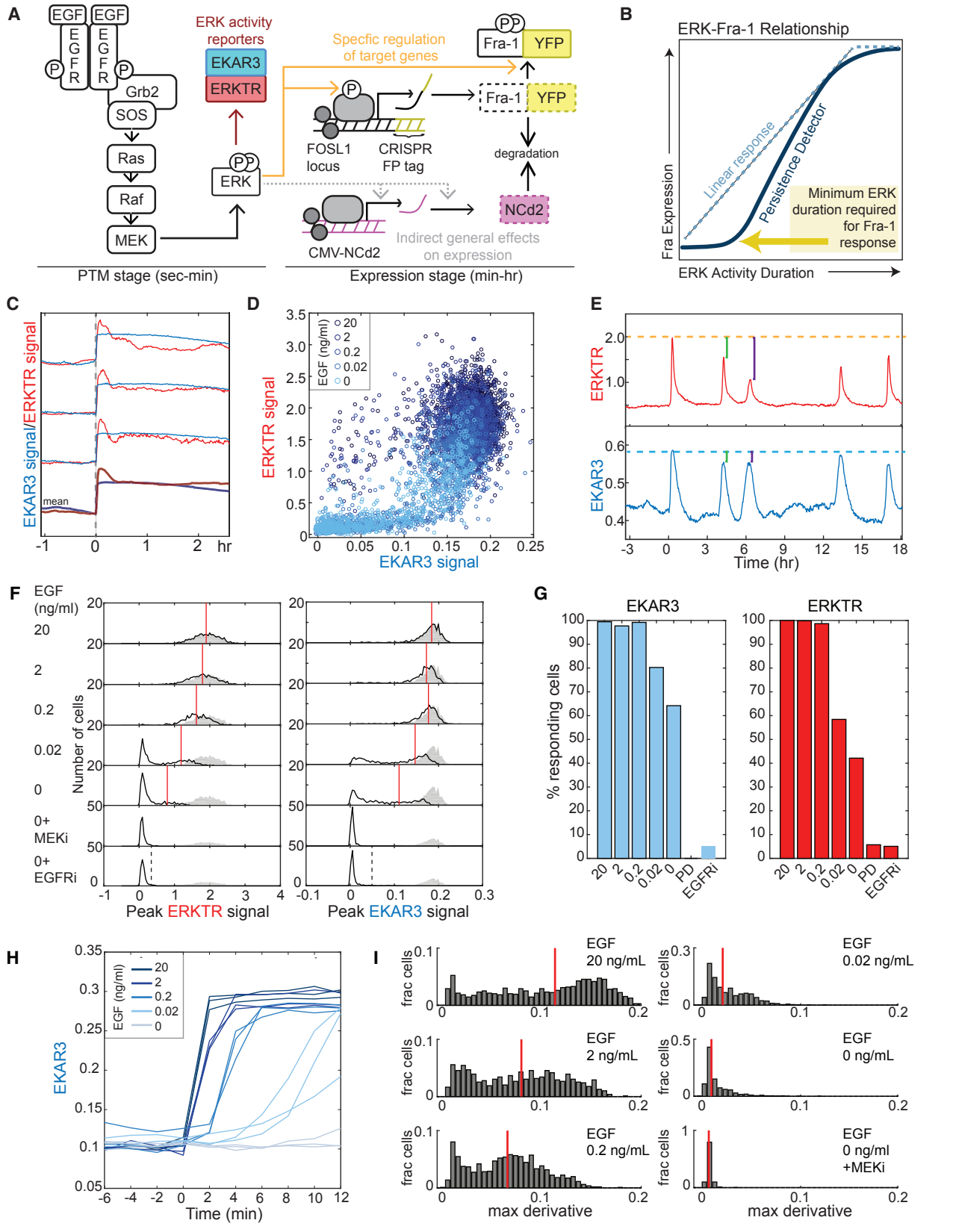
ERK signaling regulates the expression of target genes, but it is unclear how ERK activity dynamics are interpreted. Here, we investigate this question using simultaneous, live, single-cell imaging of two ERK activity reporters and expression of Fra-1, a target gene controlling epithelial cell identity. We find that Fra-1 is expressed in proportion to the amplitude and duration of ERK activity. In contrast to previous "persistence detector" and "selective filter" models in which Fra-1 expression only occurs when ERK activity persists beyond a threshold duration, our observations demonstrate that the network regulating Fra-1 expression integrates total ERK activity and responds to it linearly. However, exploration of a generalized mathematical model of the Fra-1 coherent feedforward loop demonstrates that it can perform either linear integration or persistence detection, depending on the basal mRNA production rate and protein production delays. Our data indicate that significant basal expression and short delays cause Fra-1 to respond linearly to integrated ERK activity.

INTRODUCTION

Most mammalian signaling pathways induce post-translational protein modifications (PTMs) within minutes following stimulation, followed by changes in gene expression over the course of hours or days. Because gene expression ultimately causes changes in cellular phenotype, the transduction of immediate PTM events to transcriptional machinery is critical and has been the subject of a number of quantitative studies (Hansen et al., 2015; Hao and O'Shea, 2011; Lee et al., 2014; Tay et al., 2010). However, testing models of these systems is difficult because of the very different time scales of these processes. In principle, if all cells within a population responded homogenously to stimuli, then traditional techniques, which report the average protein activity or transcript number across many cells, would be sufficient. However, recent studies have revealed that signaling responses at both the PTM and gene expression levels display a remarkable degree of cell-to-cell variability (Albeck et al., 2013; Aoki et al., 2013), and it is now clear that population-level time-course measurements distort the underlying quantitative relationship. For example, averaging across a population containing both responding and non-responding cells may give the appearance of an analog response, when the underlying cellular behavior is in fact binary (Purvis and Lahav, 2013). Furthermore, "snapshot" approaches that collect single-cell data at fixed time points can fail to reveal dynamic behaviors or artificially indicate a bimodal response (Birtwistle et al., 2012). A more suitable dataset would provide time-course data for both PTM and gene expression events within the same individual cell, so that the dynamics of each can be cross-correlated without assumption or inference.

The Ras/MAPK pathway plays a central role in many cell types in controlling cellular proliferation and differentiation. Within this pathway, transduction begins with activation of receptor tyrosine kinases such as epidermal growth factor receptor (EGFR), and culminates with activation of the kinase ERK (Figure 1A). ERK activation has been observed as early as 30 s following exposure to growth factors and typically peaks 10–15 min later (Nakakuki et al., 2010; Saito et al., 2013). Depending on the cell type or the specific growth factor ligand, ERK activity may return rapidly to baseline within 30 min or remain elevated for hours (Niepel et al., 2014). ERK phosphorylates hundreds of substrates, including transcription factors (TFs) such as SRF and Elk-1 (Gille et al., 1995), which increase transcription rates for ERK target genes (ETGs) including c-Fos, c-Myc, Egr-1, and Fra-1 (Amit et al., 2007; Treisman, 1985). The resulting increases in mRNA and protein numbers are detectable within minutes, but typically take >30 min to reach their peak level. Several ETGs, including c-Fos, c-Myc, and Fra-1, are subject to an additional layer of regulation: the protein products of these genes are initially unstable, with a half-life of ~30 min, but are protected from degradation upon phosphorylation by ERK and RSK (a kinase also activated by ERK) (Okazaki and Sagata, 1995). Of these genes, Fra-1 is of particular importance within mammary epithelial cells, where it controls epithelial-to-mesenchymal transitions (Bakiri et al., 2015) and plays a role in the development of invasive characteristics during breast cancer progression (Tam et al., 2013).

Because ERK acts to increase the abundance of ETG proteins at the levels of both synthesis and turnover, the system forms a coherent feedforward loop (cFFL) (Mangan et al., 2003). This motif can potentially filter ERK inputs by responding to pulses of long, but not short, duration and driving approximately binary (all-or-none) expression of ETG proteins (Mackeigan et al., 2005; Murphy et al., 2002, 2004; Nakakuki et al., 2010), a behavior termed "persistence detection" (Purvis and Lahav, 2013). Other



(legend on next page)

studies have concluded that the amplitude (or peak response) of ERK activity qualitatively changes gene induction levels (Kwong et al., 2012; Woods et al., 1997). Thus, the nature of ETG regulation remains unclear, in part because studies to date have only tested a small number of conditions and used population-average or fixed-cell methods. While a common interpretation of the existing data is that ETG expression is preferentially (that is, non-linearly) sensitive to the duration of ERK activity (Legewie et al., 2005), it remains possible that ETGs respond more consistently to other patterns or to the integrated or cumulative ERK activity profile (Figure 1B). A more exact understanding of the input-output behavior of ETG expression is needed because these characteristics determine whether modulation of ETG expression by growth factors or drugs is sufficient to alter cell function.

Here, we describe a multiplexed combination of live-cell reporters to simultaneously track dynamic activity of ERK signaling and induction of Fra-1. This combination allows us to directly measure the correspondence between the kinetics of ERK activity and the rate of Fra-1 protein production in the same cell, which was not possible with previous methods. We describe an extensive analysis of how the activation of Fra-1 is affected by both the amplitude and duration of ERK activity, as controlled by both growth factor availability and small-molecule inhibitors. At both the population and the single-cell level, ERK amplitude and duration both contribute to the induction of Fra-1, such that the expression level of Fra-1 protein linearly reflects the integral of ERK activity. We incorporate this behavior into a generalized mathematical model that explains how the same gene expression motif can generate either rapid, high-gain linear responses or non-linear persistence detection of ERK activity duration.

RESULTS

ERK Response to Growth Factors Combines Digital and Analog Features

To assess how downstream effectors respond to ERK signals, it is first necessary to accurately measure ERK kinase activity. We used two genetically encoded live-cell reporters, a Förster resonance energy transfer (FRET)-based Erk activity reporter (EKAR3) and a translocation-based reporter (ERKTR) (Regot et al., 2014) (Figure 1A). Both reporters produce a ratiometric signal that is related to ERK activity (see STAR Methods for details of ratio calculations). EKAR3 and ERKTR detect sporadic pulses of ERK activity with similar but distinct kinetics; EKAR3 rises faster and attenuates more slowly than ERKTR (Sparta et al., 2015). To reliably detect ERK activity, we measured the ERK response to increasing EGF concentrations following a period of growth factor withdrawal in cells expressing both EKAR3 and ERKTR (Figures 1C and S1). Both reporters responded rapidly to stimulation (rising by 2–4 min), but ERKTR typically peaked at 6–20 min and then rapidly declined to 50%–75% of its peak value, while EKAR3 peaked immediately and maintained a plateau for 3 or more hours, suggesting it may have reached saturation. Comparing the initial response amplitudes across various doses of EGF revealed that EKAR3 and ERKTR respond to different segments of the ERK dynamic range, with EKAR3 sensitive at lower activity levels and ERKTR at higher levels (Figure 1D). Moreover, at lower EGF doses ERKTR, but not EKAR3, measured sporadic pulses as lower in amplitude than the first pulse (Figure 1E, cf. green and purple bars). While sporadic pulses of ERK activity were previously reported to be quantal in size (Albeck et al., 2013), this uniformity is likely an artifact of the saturation of FRET-based reporters. Alternatively, it is possible that nuclear phosphatases induced by ERK activity attenuate ERKTR preferentially, due to its differential localization (Shah et al., 2017).

It has remained ambiguous whether ERK responses are digital/bimodal (Altan-Bonnet and Germain, 2005; Das et al., 2009; Ferrell and Machleder, 1998) or analog (O'Shaughnessy et al., 2011; Sturm et al., 2010) with respect to EGF dose. To address this question using the advantages of both reporters, we built dose-response distributions by resolving the peak signal in each cell. At doses of EGF >2 ng/mL, both reporters indicate that essentially 100% of cells respond with a uniformly high

Figure 1. Different ERK Reporters Measure Activity in Different Ranges

(A) Schematic of the pathway and reporters used in this study. EKAR3 and ERKTR indicate ERK activity with minute-scale resolution, providing a readout of the PTM stage of EGF response. Genomic knockin of a YFP fusion tag at the *FOSL1* locus results in production of a fluorescent signal that responds to ERK activity at the level of both transcription and protein stabilization (orange arrows). NCd2 expressed from a viral promoter provides a control for general changes in gene expression capacity.

(B) Expected behavior of a persistence detector. As the duration of ERK activity increases, Fra-1 expression responds selectively to longer periods of stimulation. In the absence of mechanisms imposing this selectivity, a saturating linear response would be expected (blue dotted line).

(C) Sample EKAR3 and ERKTR measurements in single cells starved of growth factors and then stimulated at time 0 with 20 ng/mL EGF. Paired measurements from three individual cells are shown, and the mean for 489 cells below.

(D) Scatterplot of initial peak amplitude on EGF stimulation, for 5,935 single cells, stimulated with different EGF concentrations (indicated by color). Each point represents the peak EKAR3 and ERKTR signals for an individual cell within a 1.5-hr window after stimulation.

(E) Spontaneous ERK activity pulses occurring in a single cell, stimulated at time 0 with 0.02 ng/mL EGF, reported by both ERKTR and EKAR3. Dotted orange and blue lines indicate the peak signal immediately following stimulation for ERKTR and EKAR3, respectively. Green and purple bars indicate the differences of subsequent peaks relative to the initial response.

(F) Histograms of the peak ERKTR (left) and EKAR3 (right) initial pulse amplitude reached in single cells, when stimulated with various concentrations of EGF. For comparison, shaded gray regions show the fully stimulated distribution (20 ng/mL EGF) in all panels. Red lines indicate the mean values of the responding population in each condition; responders are defined as cells whose peak ERK activity is above the 95th percentile of the EGFRi (gefitinib)-treated population (dashed lines). See also Figure S1.

(G) Quantification of responding cells as detected by EKAR3 and ERKTR for multiple doses of EGF (as defined in F).

(H) Individual cell traces of EKAR3 following EGF stimulation, imaged at higher time resolution (2 min), demonstrating the increase in initial slope of EKAR3 with EGF concentration.

(I) Distributions of initial slope (maximum derivative within 30 min following stimulation) in EKAR3 following stimulation by EGF at different concentrations. Red bars indicate medians.

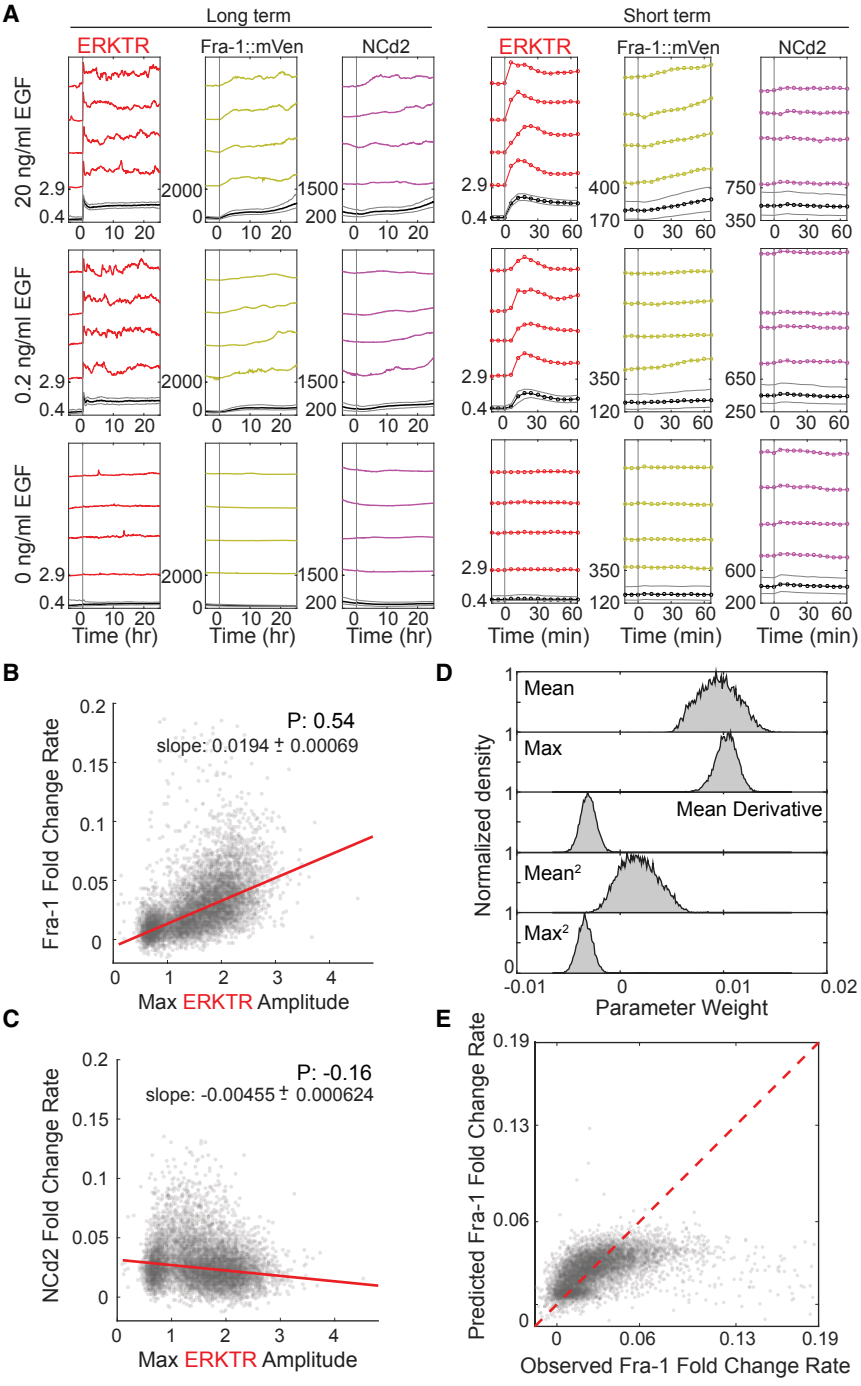


Figure 2. Measuring Fra-1 Production as a Function of ERK Activity

(A) Mean and representative single-cell time-course measurements of ERKTR, Fra-1::Venus, and NCd2 in the same cells following stimulation with 20, 0.2, or 0 ng/mL EGF. Prior to imaging, cells were cultured in the absence of growth factors for 48 hr and then stimulated with EGF at different concentrations. Four single cells and the mean (with 25th–75th percentile range as gray lines) are shown for each concentration of EGF, on both long and short time scales.

(B) Correlation between stimulated ERK activity and Fra-1 production rate. The maximum amplitude of ERKTR is plotted against the maximum slope of Fra-1::mVenus fold change within the first 60 min following EGF stimulation for each cell. Cells from multiple conditions are combined on a single plot. Red line represents a linear regression fit. P, Pearson's correlation; slope is indicated with 95% confidence intervals.

(C) Correlation between stimulated ERK activity and NCd2 production rate. Plot was constructed as in (B), using NCd2 instead of Fra-1::mVenus.

(D) Contributions of ERK parameters to a PLSR model of Fra-1 expression. A PLSR model was constructed using the mean, maximum amplitude, mean derivative, and squared terms of ERKTR over the first 60 min following EGF treatment as inputs, and maximum slope of Fra-1 intensity as an output. Parameter weight distributions were generated by bootstrapping 8,629 cells with replacement 10,000 times. ERKTR was used because of its broader amplitude response.

(E) Prediction of Fra-1 rate by linear modeling of ERK signal terms, plotted against measured Fra-1 value. Values falling on the red line are predicted exactly by the model.

amplitude signal (Figures 1F and 1G). Conversely, treatment with inhibitors of EGFR or MEK reduced all cells to a uniformly low ERK reporter signal, allowing us to set threshold criteria above which signals were considered specific ERK responses (Figure 1F). At 0.02 ng/mL EGF the fraction of responding cells fell to 80%, assessed by the more sensitive EKAR3 (58% when using ERKTR; Figure 1G). For the same EGF stimulus, we used ERKTR to distinguish peak ERK activity and found that the median amplitude for responding cells was 38% lower at 0.02 compared with 20 ng/mL EGF (20% lower for EKAR3; Figure 1F).

When stimulated only with unsupplemented medium, 65% of cells displayed a detectable ERK response, with an even lower median (58% lower than with 20 ng/mL EGF). The rate of reporter signal increase also rose with increasing EGF dose (measured by EKAR3; Figures 1H and 1I). Another EGFR ligand, amphiregulin, displayed essentially identical dose-response characteristics (Figure S1). These data indicate that ERK activity in individual cells does not respond in a strictly digital manner to EGF and similar ligands. Instead, ERK activity reflects the stimulus strength in its amplitude and rate of activation, but with an increasing fraction of refractory cells at lower concentrations of EGF.

Fra-1 Production Rate Responds Linearly to ERK Activity

To enable dynamic measurements of ERK-responsive gene expression, we used CRISPR-mediated homologous recombination to insert monomeric, fast-folding fluorescent protein tags (mVenus or mCherry) as in-frame fusions at the C

terminus of the *FOSL1* locus (Fos-related antigen 1, encoding Fra-1; we refer to this genomic reporter as Fra-1::mVenus or Fra-1::mCherry). Correct integration was confirmed via western blot and genomic PCR; in cells, the nuclear Fra-1::mVenus signal was clearly distinguishable from autofluorescence (Figure S2). Integration of fluorescent protein tags into the genome is a reliable method to track gene expression from endogenous loci (Ghaemmaghami et al., 2003), and allows the net effect of transcriptional induction and stabilization by ERK to be monitored. We also introduced a destabilized, nuclear-localized fluorescent protein (NLS-mCerulean-d2, or NCd2) under a generic viral promoter/enhancer, to serve as a control for variation in the global rates of transcription and translation (Figure 1A). Constitutive degradation of NCd2 is controlled by an ornithine decarboxylase PEST sequence that is distinct from the Fra-1 PEST and is not regulated by phosphorylation.

To establish baseline ERK and Fra-1 levels, we withdrew growth factors for 48 hr. Under these conditions, baseline ERK activity was undetectable by EKAR3 (Figure S2), while Fra-1 decreased to low but detectable levels (Figures 2A and S2). We then exposed cells to different concentrations of EGF to produce initial ERK responses of varying intensity. Upon stimulation, Fra-1 began to rise within 12–24 min and continued to increase gradually over the course of 24 hr, while NCd2 began to rise after 2 hr (Figure 2A and Movie S1). To evaluate the direct relationship between ERK activity and Fra-1 expression, we focused on the initial rate of Fra-1 increase following stimulation, as this response is less likely to be affected by global cellular changes induced by active ERK. ERKTR was used to measure ERK activity, as it is most responsive to the higher amplitudes initially achieved upon stimulation by different EGF doses (see Figures 1C–1E). Collating experimental data across multiple EGF concentrations to span the full range of ERK activity levels, we found that the rate of Fra-1 increase rose linearly as a function of the amplitude of ERKTR (Figure 2B). In contrast, the rate of NCd2 accumulation was approximately constant across all ERKTR intensities (Figure 2C), indicating that the early relationship between ERK and Fra-1 induction was specific and not the result of global changes in protein production.

To comprehensively evaluate possible relationships between ERK and Fra-1 within each cell, we considered the maximum rate, peak amplitude, or mean amplitude of ERKTR during the first hour, as well as the squares of these metrics. We performed partial least-squares regression (PLSR) analysis, which reduces the complexity of the data and identifies the aspects of the input dataset most predictive of the output (Fra-1 expression). We found that mean and maximum amplitude were the most predictive of the rate of Fra-1 accumulation (Figure 2D). When comparing the predictions of the resulting PLSR linear model with the observed Fra-1 output, we find the expected linear 1:1 relationship with considerable variance, but also a discrepancy where the model fails to predict the highest levels of Fra-1 expression (Figure 2E). This discrepancy could arise because ERKTR (like EKAR) may eventually reach saturation and fail to provide an accurate measurement at very high ERK activity; alternatively, ERK-independent factors may drive high expression in some cells. Together, these analyses support a mass-action model in which, within minutes of EGF stimulation, the rate of Fra-1 accumulation is linearly regulated by the intensity of ERK activity.

Total Fra-1 Expression Responds Linearly to ERK Activity Duration

To assess the effects of the feedforward stabilization motif, we performed longer-term experiments inducing varied durations of ERK activity, directly testing the hypothesis that ETG responses are refractory to short-duration ERK activity. To generate ERK activity pulses of uniform amplitude, we used a high concentration of EGF to maximally stimulate ERK, followed by an MEK inhibitor at varying times ranging in duration from the shortest achievable pulse (approximately 24 min) to 10 hr (Figures 3A and S3). We found that even the shortest ERK activity pulses induced detectable Fra-1 expression (Figure 3B) as well as an increase in NCd2 (Figure 3C). While the response of Fra-1 depended strongly on the duration of stimulation, NCd2 displayed a more uniform increase regardless of ERK activity duration (Figures 3C and 3D), indicating that Fra-1 is specifically responsive to ERK while NCd2 reflects a general increase in expression capacity. Pretreatment with a MEK inhibitor reduced the amplitude of Fra-1 and delayed it to a similar time scale as NCd2 (Figures 3B and 3C), confirming that NCd2 reflects a non-specific component that likely results from EGF stimulation of other mitogenic pathways. Comparing the mean Fra-1 area under the curve (AUC) or peak intensity with the duration of ERK stimulation, we found a linear relationship, regardless of normalization (Figures 3E and S3). At longer durations of ERK activity (>6 hr), the expression of Fra-1 approached a saturation level, but we did not observe a minimum duration of ERK activity for Fra-1 induction (Figures 3E and S3) or a bimodal response with an increasing proportion of Fra-1-expressing cells (Figure 3F). Together, these data suggest that Fra-1 responds proportionally to cumulative ERK activity, rather than requiring a minimum duration of stimulation, across the practical range of durations of ERK activity. This analysis complements the above findings (Figure 2) that establish a linear response to ERK activity intensity after a 12- to 24-min delay, indicating that Fra-1 induction is controlled primarily by the cumulative integral of ERK activity.

Fra-1 Expression Integrates ERK Activity across Pharmacological Conditions

Pharmacological inhibitors that act at different parts of the pathway (denoted EGFRi, MEKi, and ERKi) have distinct effects on the kinetics of ERK activity (Figures 4A and S4) (Albeck et al., 2013). We therefore investigated whether Fra-1 production would still respond proportionally when ERK activity was suppressed by each of these compounds. As indicated by both EKAR3 and ERKTR reporters, ERK activity following EGFRi (gefitinib) treatment was initially suppressed, but then rebounded to a pulsatile state after a period of hours, consistent with the suppression of negative feedback loops (Duncan et al., 2012). With MEKi, a similar rebound occurred, but ERK activity reverted instead to an intermediate level with a moderate degree of pulsatility (with trametinib showing a slightly higher degree of pulsatility than PD325901). ERKi (SCH772984) induced a sharp drop in ERK activity that was maintained for 10–30 min, followed by a rapid rebound to a very stable level of ERK activity. These differences provide a means to test the response of Fra-1 to widely varying ERK kinetics.

When mean Fra-1 trajectories were examined for each drug, Fra-1 levels decreased upon the initial inhibitor treatment but

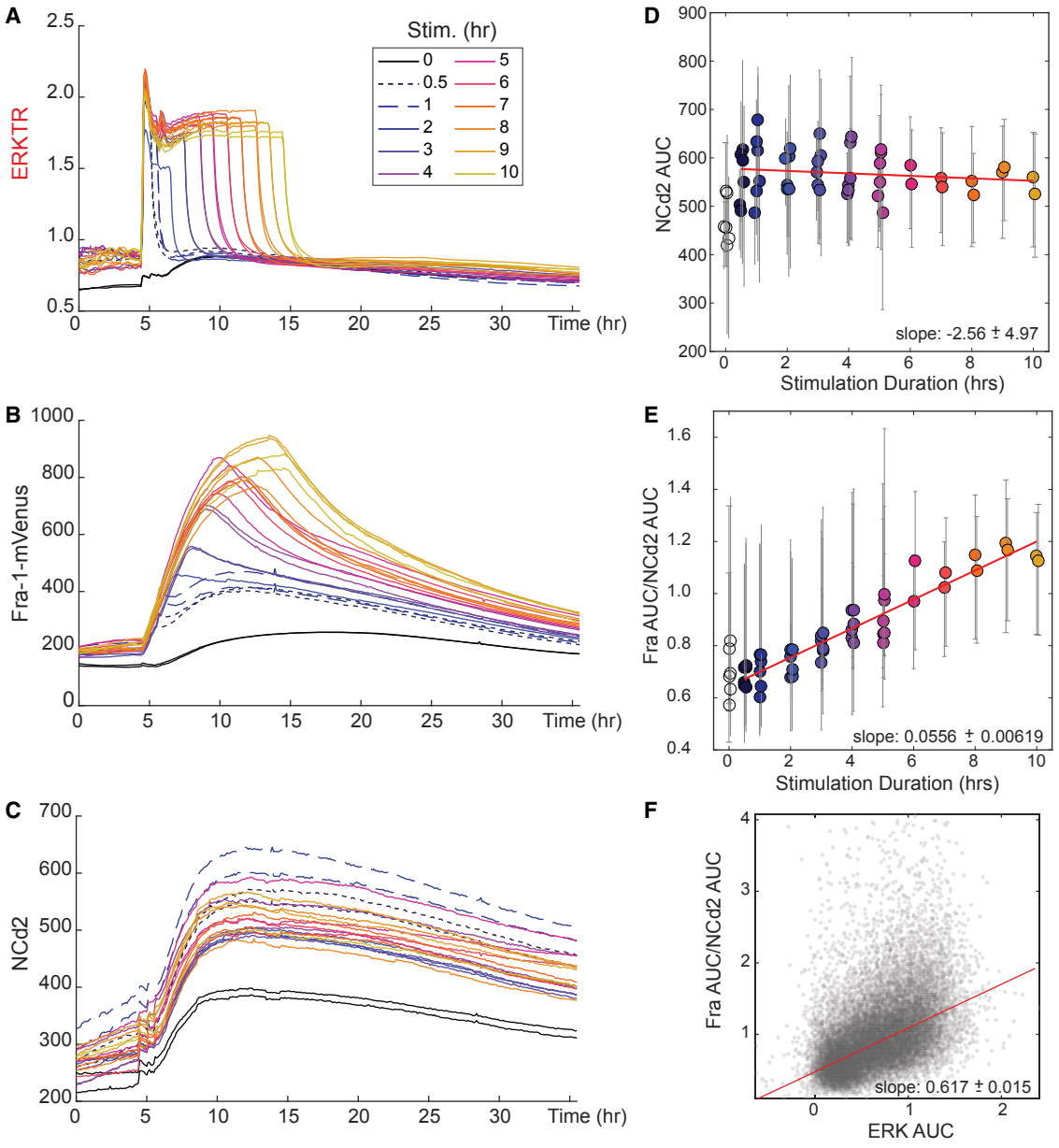


Figure 3. Fra-1 Induction as a Function of ERK Activity Duration

(A–C) Response of Fra-1 to ERK pulses of varying duration. MCF10A cells expressing Fra-1::mVenus, ERKTR, and NCd2 were cultured in the absence of growth factors for 48 hr and then stimulated by 20 ng/mL EGF. At various times following EGF treatment, ranging from 30 min to 10 hr, 100 nM PD0325901 was added to acutely inhibit ERK activity. Stimulation duration 0 received PD 15 hr prior to EGF stimulation. Plots indicate the population mean values for ERKTR cytoplasmic/nuclear ratio (A), Fra-1::mVenus intensity (B), and NCd2 intensity (C); lines are color coded according to the interval between stimulation and inhibition.

(D and E) Plot of population-averaged AUC of NCd2 (D) or Fra-1::mVenus normalized to NCd2 (E) as a function of the interval between EGF addition and PD325901 addition. Error bars show 25th–75th percentile range. Red line indicates linear fit. Data were also fit to a library of non-linear models, and the linear model was chosen as the most parsimonious by Akaike's Information Criterion. Area under the curve (AUC) was used to reflect the net Fra-1 output for the single defined ERK pulse. Regressed slopes are indicated with 95% confidence intervals. See also Figure S3.

(F) Scatterplot of single-cell measurements of Fra-1::mVenus (normalized as in E) plotted against ERKTR AUC. Red line indicates linear fit.

gradually recovered after ~24 hr, correlating approximately with the time of ERK recovery (Figure 4B). At the level of cell populations, when the integrated ERK activity was compared with final level of Fra-1 expression during a 20-hr period spanning the recovery phase, we found a correlation across all drug conditions (Figures 4C and S4), in accord with the linearity described above in the relationship between ERK activity and Fra-1 expression. The variation in ERK waveform between drugs did not appear to have a strong influence on Fra-1 recovery. At the single-cell level, we observed a high degree of variability within populations (Figure 4D). Potentially, this variability could reflect either non-ERK contributions or a failure to identify the

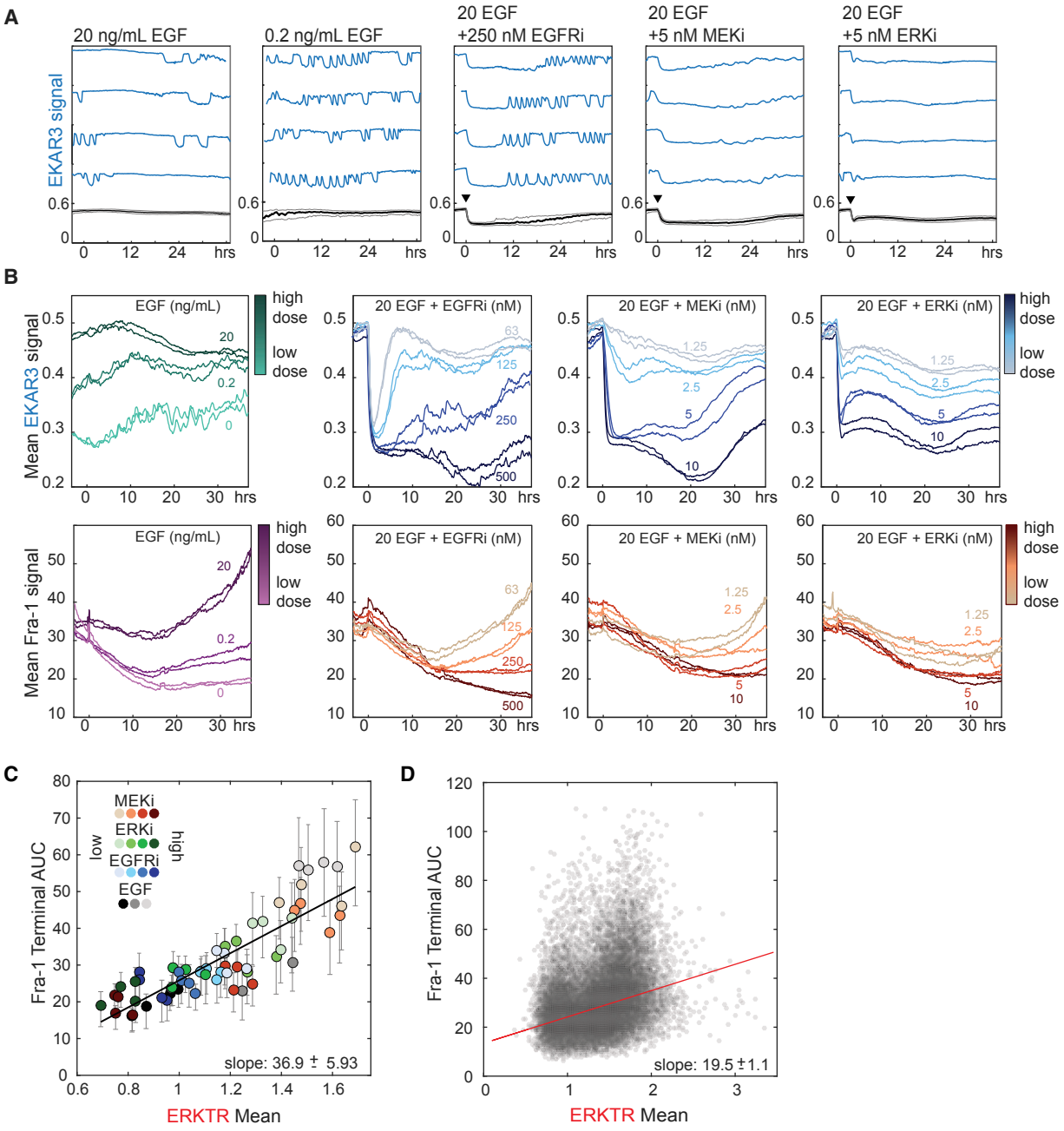


Figure 4. Fra-1 Expression in Response to Pharmacological Modulation of ERK Activity Patterns

(A) Single-cell responses of ERK activity to inhibitors. MCF10A cells expressing EKAR3 were cultured in imaging medium containing 20 or 0.2 ng/mL EGF and treated with inhibitors of EGFR (gefitinib), MEK (trametinib), or ERK (SCH772984) at time 0 (arrows). Each plot shows the mean value in black at the bottom, with lighter gray lines representing the 25th and 75th percentile. Colored lines above the mean depict four randomly selected single cells for each condition, plotted on the same scale as the mean. The higher-sensitivity EKAR3 is used here to show low ERK activity under inhibition; see Figure S4 for corresponding ERKTR data.

(B) Population responses of ERK activity and Fra-1 expression to the indicated inhibitors. Cells were cultured in the presence of EGF and treated at time 0, at concentrations indicated by line color.

(C and D) Correlation of integrated ERK activity with Fra-1 expression across pharmacological conditions. In (C), each point represents the population average and error bars for the 25th to 75th percentile range, for one drug condition, of Fra-1 intensity plotted against integrated ERKTR cytoplasmic/nuclear ratio; replicate experiments are shown individually. Measurements were made for a 20-hr window beginning 17 hr following drug treatment. Fra-1 intensity is quantified as AUC during the final hour of the 20-hr window to reflect the steady-state response to the preceding ERK activity, which is represented as AUC over the 20-hr window. In (D) each point represents a single-cell measurement of ERKTR and Fra-1, calculated over the same interval as in (C). Cells from all conditions are pooled, a total of 24,620 cells. ERKTR is used here to capture the wide dynamic range across all conditions; corresponding EKAR3 data are shown in Figure S4.

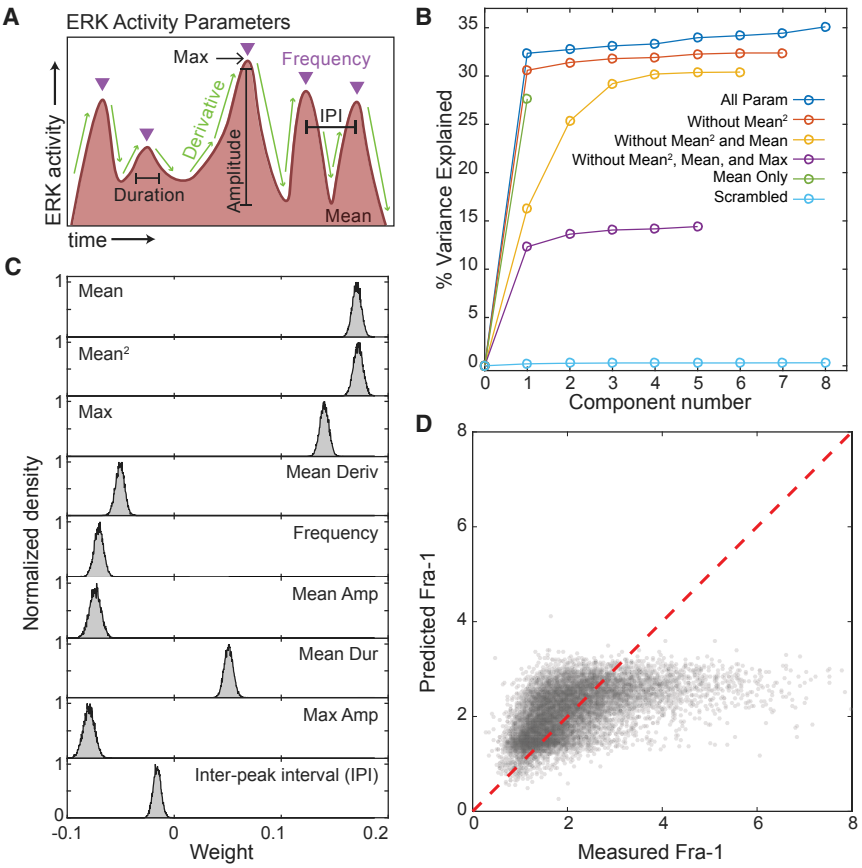


Figure 5. PLSR Modeling of the ERK-Fra-1 Relationship across All Conditions

(A) ERK activity metrics for single cells. Triangles indicate peak position; IPI, inter-pulse interval. For subsequent PLSR analysis, per-peak metrics were summarized on an individual cell basis, using the average or maximum as indicated in (C), over all measured peaks for that cell.

(B) Predictive capabilities of PLSR models for Fra-1 expression (measured by percentage of total variance explained) using different combinations of ERK activity measurements (by EKAR3). For the "scrambled" model, pairings between input and output measurements for each cell were randomly reassigned. Models generated using ERKTR data (Figure S5) showed similar trends but lower predictive capacity, possibly due to sampling bias within the dataset that included large numbers of low-ERK conditions.

(C) Contributions of ERK parameters to the PLSR model. Distributions of parameter coefficients generated by bootstrapping with replacement 10,000 times.

(D) Prediction of Fra-1 maximum value by linear modeling of ERK signal terms, plotted against measured maximum Fra-1 value. The PLSR model was constructed using the EKAR3 features in (A) over the 20-hr window, with the endpoint value of Fra-1 intensity as the output. Non-linear dependence was assessed by including mean-squared in the PLSR. Red line represents perfect prediction of the data by the model.

correct attributes of the ERK signal responsible for controlling Fra-1 expression. Therefore, to investigate the relationship of ERK to Fra-1 we applied PLSR modeling to a collated dataset containing a wide range of ERK activity patterns assembled from EGF and inhibitor concentrations over multiple independent experimental replicates, using both mVenus- and mCherry-tagged versions of Fra-1.

To create the input vector for PLSR, we reduced each single cell's ERK activity trace to a set of eight scalar values, including mean activity, pulse frequency, and mean pulse duration (Figure 5A and Table S1). The predictive capacity of these models varied between 20% and 60% depending on the conditions included in the dataset. As a control, models in which ERK and Fra-1 signals were randomly paired predicted less than 1% of Fra variance, indicating that a common trend in both ERK activity and Fra-1 induction did not artificially elevate estimates of prediction capacity. When using the broadest set of conditions and the EKAR/Fra1:mCherry configuration (as shown in Figure 4C), the model improved only negligibly beyond the first component, which explained 32% of variance in Fra-1 (Figure 5B). We examined the PLSR weights in this component to determine the relative importance of the eight dynamic features on the response variable (Fra-1 AUC over the last hour of the modeled interval). Mean and mean-squared ERK activities were weighted most heavily, with a more moderate contribution from maximum ERK activity; similar weightings were observed across all of the datasets considered (Figures 5C and S5; Table S2). These weightings again suggest that integrated (or mean) ERK activity is the primary determinant of Fra-1 expression within single cells. The weighting of the mean-squared term potentially indicates a non-linearity in the induction process, but models constructed without the mean-squared term were still able to capture 31% of the variance in Fra-1, while omitting both mean and mean-squared reduced predictive capacity to 16% (Figure 5B). Notably, the mean and mean-squared alone were capable of explaining 28% and 30% of variance, respectively. When model predictions of Fra-1 expression were compared with observed values on a cell-by-cell basis (Figure 5D), the variance not explained by the models was accounted for by both a spread around the central relationship, and by a subpopulation of 17% of cells for which the model underpredicted Fra-1 expression. Altogether, this comprehensive analysis of the dependence of Fra-1 expression on single-cell measurements of ERK indicates that (1) the relationship between ERK activity and Fra-1 expression is approximately linear across many conditions, (2) the specific dynamic form of ERK activity has little influence on the expression of Fra-1, and (3) other factors beyond ERK are likely to exert a strong influence on the Fra-1 expression level in a given cell.

Stabilization of Pre-existing Fra-1 Enables Linear Expression Responses

Given that cFFLs, such as the one connecting ERK and Fra-1, can in principle produce non-linear responses to ERK duration,

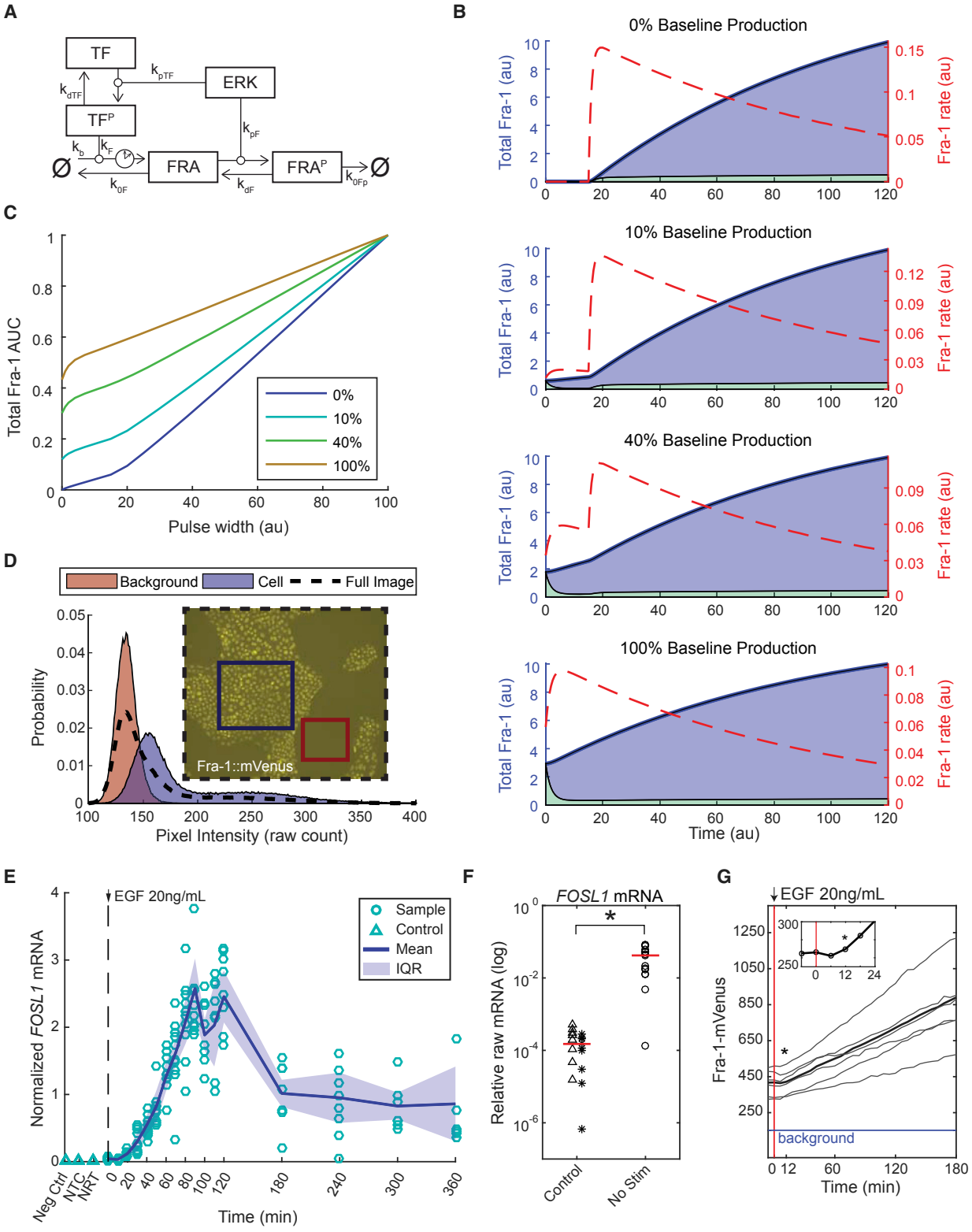


Figure 6. Model Analysis Predicts Linear Response with Baseline Fra-1

(A) Network diagram of the delay differential equation analyzed and simulated. Delay (clock symbol) is included for the stimulated production of Fra-1 by activated transcription factor (TF^P). See STAR Methods for a full description of the model.

(legend continued on next page)

we sought to understand why Fra-1 accumulation is essentially linear with respect to ERK activity duration. We constructed a minimal mathematical model comprising ERK phosphorylation of a transcription factor that promotes Fra-1 mRNA production and ERK-mediated stabilization of Fra-1 protein (Figure 6A). The minimal model retains the structure of the known system and includes an explicit delay on the production of new protein, but avoids the additional complexity of cooperativity or modeling every intermediate; it is thus situated to predict conditions giving rise to persistence detection but conservatively represent their magnitude. We analyzed model conditions necessary for persistence detection by numerically simulating the time-dependent accumulation of Fra-1 over a screen of model parameters (Figure S6 and Table S3; STAR Methods).

The model indicates that persistence detection depends critically on two parameter values in the production of Fra-1: delay and baseline. Delay arises from the synthesis time of mRNA and protein, and sets the minimal input signal duration needed to allow both arms of the cFFL to be simultaneously active. Baseline refers to mRNA (and protein) production either in the absence of stimulus (ERK) or due to pre-existing stimulus. If baseline Fra-1 mRNA production is significant, ERK stimulation immediately drives accumulation of the stabilized protein (Figure 6B). Thus, the presence of baseline expression in any cFFL motif may prime the system to respond to stimuli shorter than its processing delay. The model therefore predicts that the linear response of Fra-1 could result from a short delay, a high baseline, or both. To evaluate the feasible range of experimental outcomes in light of the theoretical analysis, we simulated ERK pulse experiments, calculating the total Fra-1 AUC as a function of pulse duration (Figure 6C). The simulated AUC varies in initial slope and curvature according to baseline levels, and is strictly linear in the long term. Comparing these curves with Figures 3E, S3C, and S3E, the long-term linearity is well corroborated and the apparent short-term linearity of the data implies an intermediate mixture of baseline and induced expression.

We further tested the applicability of the model by evaluating whether experimental measurements of Fra-1 (*FOSL1*) mRNA and protein were consistent with the prediction of significant baseline expression as well as transcriptional induction, and/or relatively short delays. Fra-1 protein expression in the prolonged absence of ERK activity, evidenced by non-zero intensity in serum-starved and inhibitor-treated cells (Figures 3B and 6D), suggests that baseline production is indeed substantial in this cell line. To evaluate transcription and processing delays, we measured *FOSL1* mRNA at the population level over a time course and observed a statistically detectable increase within 20 min of EGF stimulation ($p < 0.05$), followed by accumulation reaching a maximum at ~1.5–2 hr (Figure 6E), confirming transcriptional induction. Without stimulus, *FOSL1* mRNA is significantly higher than RT controls, confirming non-zero baseline transcription (Figure 6F, $p < 0.05$). In comparison, increases in Fra-1::mVenus or Fra-1::mCherry fluorescence occurred on an equivalent or faster time scale, detectable within 12 min following EGF stimulation (Figure 6G, $p < 0.05$), consistent with Fra-1 protein already existing at the time of initial stimulus. Furthermore, the shortest spontaneous ERK activity pulses, or those achievable by timed stimulation-inhibition, are approximately 20 min long. That these are equivalent with the time required to produce *FOSL1* mRNA and protein leaves little room for any practical ERK pulses to be effectively rejected even in the absence of any baseline *FOSL1* mRNA. Together, these measurements indicate that the baseline expression and relatively short production delay of Fra-1 bias it against selectively excluding typical ERK pulses.

To evaluate the functional impact of transcriptional versus stabilization control on Fra-1, we performed an experiment to measure the stabilization response independently. We used a reporter termed FIRE (Fra-1-based integrative reporter of ERK), in which the regulated PEST domain of Fra-1 is fused to a fluorescent protein and expressed under the control of a constitutive viral enhancer (Albeck et al., 2013), thereby reproducing only the degradation effects of Fra-1 without specific transcriptional control (Figure 7A). We measured both the full feedforward system (Fra-1) and the non-feedforward, stabilization-only system (FIRE) in individual cells following stimulation with various concentrations of EGF. Strikingly, the time series traces of Fra-1::mVenus closely resembled those of FIRE (Figure 7B), with median correlation of 0.78–0.88 versus 0.27–0.55 against total RFP as a non-specific expression control (across multiple EGF doses, Figure 7C). Directly comparing Fra-1 and FIRE fold-change responses over time (Figure 7D) the distributions of the two reporters overlap extensively, and at early times the median signals are indistinguishable. After 3 hr, the median

(B) Minimal model simulations with various relative levels of baseline Fra-1 production (percentage given by $100 \times k_b/(k_b + k_F)$). Total Fra-1 (shaded regions) is plotted against left axes; lower (green) and upper (blue) regions correspond to naive and stabilized Fra-1, respectively. Fra-1 accumulation rate (dashed line) is plotted against right axes. Models were simulated with ERK activity beginning at time 0 and a 15 time-point delay for Fra-1 production; the effect of baseline production is most apparent in this window at the start of each simulation. See also Figure S6.

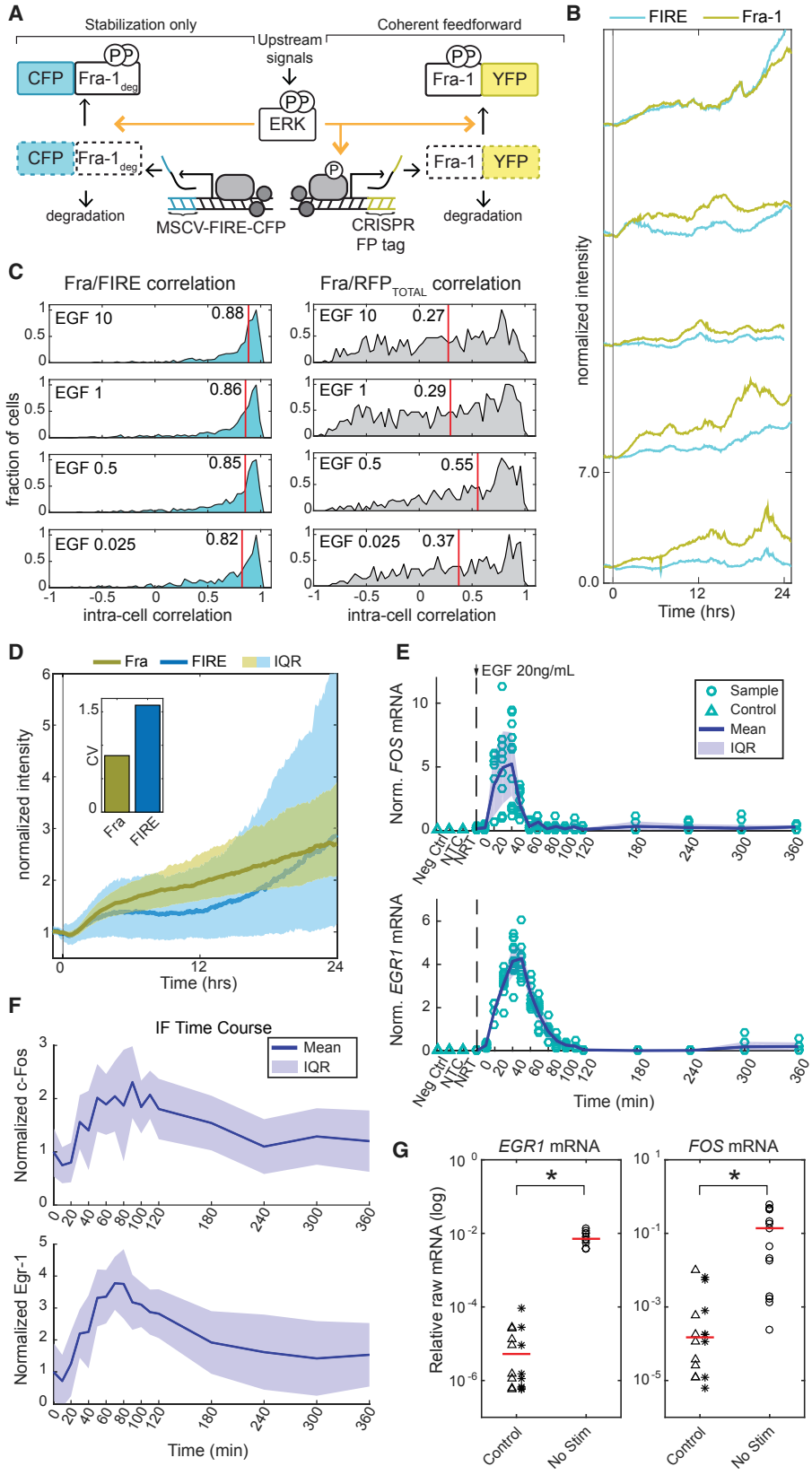
(C) Simulated experiments with varied ERK pulse width, using the same parameter values as plots in (B). The total Fra-1 AUC was calculated for square waves of ERK activity, performed with different relative baseline production levels (indicated in legend). Fra-1 AUC is plotted against the duration of ERK activity (as in Figure 3E).

(D) Sample image and pixel intensity histograms from cells expressing the Fra-1::mVenus fusion, after growth factor withdrawal for 48 hr and treatment with MEK inhibitor for another 15 hr. Histograms shown for regions indicated in the image: background only (red), cells (blue), whole frame (dashed line).

(E) Time-course measurement of Fra-1 mRNA in MCF-10A cells, by qRT-PCR. Cells were cultured without growth factors for 48 hr, stimulated with 20 ng/mL EGF, and lysed for RNA extraction at 10-min intervals up to 2 hr, followed by 1-hr intervals up to 6 hr. Data were compiled from three replicate RNA extractions, with at least two replicate qPCR measurements per sample. NTC, no target control; NRT, no reverse transcription; IQR, inter-quartile range from 25th to 75th percentile.

(F) Comparison of qRT-PCR measurements of Fra-1 mRNA in unstimulated cells (48 hr growth factor withdrawal) to NTC (stars) and NRT (triangles), shown on a log scale. These data are not normalized to the control, HINT1. Red line indicates median. * indicates significance at the $p < 0.05$ level by t test.

(G) Early Fra-1 response following EGF stimulation (red line). Data for Fra-1::mVenus in cells cultured in the absence of growth factors for 48 hr and then stimulated with 20 ng/mL EGF are replotted from Figure 3 to clarify early dynamics. Bold line indicates the population mean, and lighter lines represent six individual cells. Asterisk indicates first time point significantly greater than baseline (*$p < 0.05$). Blue line indicates the mean intensity of background (non-cell) pixels. Inset shows further magnification of the mean signal on a shorter time scale.



(legend on next page)

Fra-1 signal continues to rise while the stabilization-only FIRE attenuates, consistent with a model in which the immediate increase following stimulation is due to stabilization of Fra-1 protein produced from pre-existing mRNA, while stimulated production of mRNA contributes to total Fra-1 levels at later times. Additionally, Fra-1 exhibits a median coefficient of variation 50% lower than that of FIRE, potentially resulting from the additional regulation acting at the endogenous promoter.

Finally, we investigated the potential for variance in mRNA production parameters governing persistence detection among ETGs. We sampled the mRNA and protein kinetics of c-Fos and Egr-1 via qRT-PCR and immunofluorescence. In comparison with Fra-1 measurements, c-Fos and Egr-1 showed an earlier peak in both mRNA and protein abundance and a return to baseline levels within 2 hr of stimulation, suggesting additional negative regulation (Figures 7E and 7F). *EGR1* mRNA appears even faster (within 10 min, versus 20 min for *FOSL1* and *FOS*), although protein levels only become detectably higher at 20 min. *EGR1* and *FOS* mRNA were also detectable even in the absence of stimulus (Figure 7G, $p < 0.05$). Thus, all three ETGs analyzed share kinetic characteristics that make them weak discriminators of characteristic ERK pulses.

DISCUSSION

Model Analysis Reconciles a Linear Fra-1 Response with Past Studies

At the center of quantitative signaling biology is the question of how the time-varying activation of an upstream regulator is converted into the time-varying outputs of downstream molecules, which ultimately alter cell fate (Kholodenko et al., 2010; Rhee et al., 2012). This process of information transfer can take many forms in principle, and remains poorly understood in many central mammalian signaling systems because of the difficulty of making adequate experimental measurements over time in the same cell. In this study, we developed a method to accurately track, in the same cell, both input and output for the ERK-ETG system, which has been heavily studied for both its relevance to human disease and its importance as a canonical model signaling system. Previous studies have highlighted the potential for this system to selectively transform the kinetics of ERK activity into different gene expression profiles. We show here that across many different conditions, the ETG Fra-1 is produced in essentially linear proportion to the integrated magnitude of ERK activity, and that the related ETGs c-Fos and Egr-1 are similar in their response characteristics.

Under some interpretations of previous data, a linear response appears to contrast with the concept of the ERK-ETG feedforward motif as a persistence detector discriminating between short and long pulses of ERK activity. This apparent conflict can be resolved by careful consideration of previous studies, in which ERK activity was manipulated using two different growth factors (e.g., EGF and nerve growth factor) (Murphy et al., 2002; Nakakuki et al., 2010), or several concentrations of a growth factor (Murphy et al., 2004) to generate either short- or long-lived average ERK activity in the cell lines examined. Because only a small range of ERK activity profiles were compared in each case, these studies were able to establish that greater duration of ERK activity increases the amount of ETG protein produced, but could not directly test the idea that longer pulses would preferentially induce the ETG. Similarly, in later studies that compared more conditions, head-to-head comparisons in which only ERK duration varied were not performed (Saito et al., 2013). Methodological differences also play a role in the apparent inconsistency, as the average responses previously measured by immunoblot have lower time resolution and more potential for non-linear readouts (Janes, 2015). Taken together, the existing body of data is neutral on the question of if (and when) ETGs respond linearly or non-linearly to ERK activity.

Our data agree with previous studies showing that the duration of an ERK activity period can be converted into differences in the peak and in the AUC (or mean) of ETG protein abundance. However, these properties are not specific to the feedforward motif of the ERK-ETG system. Intuition suggests, and modeling confirms, that creating a *selective* duration-sensitive response requires a delay between the initiation of ETG transcription and stabilization of the protein product. For the ETGs we examine here this delay is on the order of 10–20 min, and it is reasonable that we observe little evidence of selective persistence detection when growth factor-stimulated ERK responses typically last >20 min. We emphasize that our analysis does not exclude the ability of ETGs (or their active forms) to conform to a non-linear persistence detector model, depending on their particular parameter values (i.e., production and degradation rates) (Uhlitz et al., 2017). Even within an individual cell, response profiles for different genes may vary, or the same gene could be reconfigured by changes in chromatin state or general transcription factor availability. While

Figure 7. Comparative Analysis of ETG Responses between Configurations and Genes

(A) Schematic of FIRE/Fra-1 comparison experiment. FIRE consists of the Fra-1 degradation domain fused to mCerulean and expressed from a stably integrated constitutive viral promoter; relative to endogenous Fra-1 it lacks stimulated mRNA production.

(B) Simultaneous measurements of Fra-1 and FIRE signals in five individual cells treated with 100 ng/mL EGF at time 0.

(C) Distributions of correlations between Fra-1 and FIRE within individual cells. Pairwise Pearson's correlation coefficients were computed between Fra-1:mVenus and FIRE signals for 4,383 cells over a period of 60 hr following treatment with different EGF concentrations; mean values are shown as red lines and numbers. To control for the correlation of FIRE with general expression increases, we made the same calculations between FIRE and total (nuclear + cytosolic) ERKTR-RFP signal, which is not directly controlled by ERK activity (right panels).

(D) Median intensities of Fra-1 and FIRE in cells treated with 1 ng/mL EGF. Shaded region indicates inter-quartile range across all cells. Inset plot shows median (across all time points) coefficient of variation between cells.

(E) Measurements of c-Fos and Egr-1 mRNA by qRT-PCR. Cells were stimulated, analyzed, and annotated as in Figure 6E.

(F) Measurements of c-Fos and Egr-1 protein abundance by immunofluorescence following growth factor withdrawal for 48 hr and stimulation with 20 ng/mL EGF at 10-min intervals up to 2 hr, followed by 1-hr intervals up to 6 hr. c-Fos data are compiled from three replicates. Egr-1 data are compiled from two replicates.

(G) Comparison of qRT-PCR measurements of Egr-1 and c-Fos mRNA in unstimulated cells (48 hr growth factor withdrawal) to NTC (stars) and NRT (triangles), shown on a log scale. These data are not normalized to the control, HINT1. Red line indicates median. * indicates significance at the $p < 0.05$ level by t test.

the topological motif of a gene circuit is often interpreted as an indicator of its behavior and function, we demonstrate that in this case, topology can be misleading in predicting the output of a key gene regulatory event; this is consistent with theoretical studies that have emphasized the importance of parameter values over topology (Gutenkunst et al., 2007; Ingram et al., 2006). This importance is also evident at the level of EKAR3 and ERKTR, whose different parameters of response to phosphorylation yield distinct readouts to the same ERK activity pulse within the same cell. These data underscore that diversity in signaling responses is generated in part by differential tuning of the kinetic parameters specific to each effector (Pargett and Albeck, 2016).

Implications of Linear Fra-1 Response for Tissue and Cancer Biology

The lack of persistence detection in Fra-1 regulation raises the question of what function the cFFL motif plays in cellular regulation. Among ETGs, Fra-1 is prominent in the homeostasis of epithelial tissues, and expression of Fra-1 induces epithelial-to-mesenchymal transition and long-term cell motility (Vial et al., 2003). The capacity of the Fra-1 regulatory motif to integrate transient ERK signals may enable a cell to "remember" past ERK activity over hours or days, rather than reacting only to the immediate activity level. Such memory may increase the information transmission capacity of the pathway, preventing chance events from inducing inappropriate and costly cell-fate choices. Our comparison of the native cFFL configuration of Fra-1 regulation with stabilization-only control indicates that the cFFL enables both a rapid high-gain response and a more linear accumulation over time than stabilization alone, which are useful attributes for integrative memory. Our data also suggest that the cFFL motif reduces variance of Fra-1 protein abundance relative to stabilization alone, consistent with the concept that accurate information transmission may require multiple parallel pathways (Hansen and O'Shea, 2015). Notably, we also observe a high degree of variability in population-level *FOSL1* mRNA measurements (Figure 6E, >2 hr), which may suggest that the motif can perform dosage compensation. Together, the characteristics of Fra-1 regulation are consistent with the ability to reliably express this transcription factor in proportion to frequently fluctuating ERK inputs, which may be important for maintaining epithelial tissue integrity.

A precise kinetic understanding of ETG responses and their variance could allow more effective therapeutic use of ERK pathway inhibitors (Ryan et al., 2015). Overexpression of Fra-1 has been linked to metastasis in multiple cancer types (Chiappetta et al., 2007; Zhong et al., 2016), and in breast cancer drives a cancer stem cell-like gene expression program (Tam et al., 2013). Conversely, in liver pathology, loss of Fra-1 expression and its replacement by c-Fos result in fatty liver disease (Hasenfuss et al., 2014). Ideally, predictive models of multiple ETGs would identify pharmacological conditions that best manipulate the kinetic profile of ERK to restore healthy gene expression profiles. The small subset of cells (<20%) that we observe falling outside the linear trend with higher than expected Fra-1 levels may also be of particular importance under pathological conditions. Identifying parallel mechanisms of regulation, such as local histone modifications that alter the range of Fra-1 mRNA copies produced during an ERK activation event, may explain the remaining variance and provide additional strategies to manipulate Fra-1 expression.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell Culture and Media
- METHOD DETAILS
  - Reporter Cell Line Construction
  - Live Cell Microscopy and Data Processing
  - mRNA Isolation and RT-qPCR
  - Immunofluorescence Microscopy
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Image Processing
  - Linear Fitting
  - Partial Least Squares Regression
  - Normalization and Statistical Analysis
  - Dynamic Mathematical Modeling
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, three tables, and one movie and can be found with this article online at https://doi.org/10.1016/j.cels.2017.10.019.

AUTHOR CONTRIBUTIONS

T.E.G., M.P., M.M., A.E.D., and J.G.A. contributed to experimental data collection. T.E.G., M.P., and J.G.A. analyzed data and wrote the manuscript. M.P. performed mathematical modeling and analysis.

ACKNOWLEDGMENTS

Funding for this work was provided by the American Cancer Society (IRG-95-125-13), the National Institute of General Medical Sciences (1R01GM115650), the Department of Defense Neurofibromatosis Research Program (W81XWH-16-1-0085), and the American Association for Cancer Research Stand Up To Cancer (SU2C-AACR-IRG-01-16). Stand Up To Cancer is a program of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the scientific partner of SU2C. Flow-cytometry services were supported by the UC Davis Comprehensive Cancer Center Support Grant (CCSG) awarded by the National Cancer Institute (NCI P30CA093373), and we acknowledge the expert cell-sorting assistance of Dr. Bridget McLaughlin and Jonathan Van Dyke.

Received: February 27, 2017
Revised: August 29, 2017
Accepted: October 27, 2017
Published: November 29, 2017

REFERENCES

Albeck, J.G., Mills, G.B., and Brugge, J.S. (2013). Frequency-modulated pulses of ERK activity transmit quantitative proliferation signals. Mol. Cell *49*, 249–261.

Altan-Bonnet, G., and Germain, R.N. (2005). Modeling T cell antigen discrimination based on feedback control of digital ERK responses. PLoS Biol. *3*, e356.

Amit, I., Citri, A., Shay, T., Lu, Y., Katz, M., Zhang, F., Tarcic, G., Siwak, D., Lahad, J., Jacob-Hirsch, J., et al. (2007). A module of negative feedback regulators defines growth factor signaling. Nat. Genet. *39*, 503–512.

Aoki, K., Kumagai, Y., Sakurai, A., Komatsu, N., Fujita, Y., Shionyu, C., and Matsuda, M. (2013). Stochastic ERK activation induced by noise and cell-to-cell propagation regulates cell density-dependent proliferation. Mol. Cell *52*, 529–540.

Bakiri, L., Macho-Maschler, S., Custic, I., Niemiec, J., Guio-Carrion, A., Hasenfuss, S.C., Eger, A., Muller, M., Beug, H., and Wagner, E.F. (2015). Fra-1/AP-1 induces EMT in mammary epithelial cells by modulating Zeb1/2 and TGFbeta expression. Cell Death Differ. *22*, 336–350.

Birtwistle, M.R., Rauch, J., Kiyatkin, A., Aksamitiene, E., Dobrzynski, M., Hoek, J.B., Kolch, W., Ogunnaike, B.A., and Kholodenko, B.N. (2012). Emergence of bimodal cell population responses from the interplay between analog single-cell signaling and protein expression noise. BMC Syst. Biol. *6*, 109.

Chiappetta, G., Ferraro, A., Botti, G., Monaco, M., Pasquinelli, R., Vuttariello, E., Arnaldi, L., Di Bonito, M., D'Aiuto, G., Pierantoni, G.M., et al. (2007). FRA-1 protein overexpression is a feature of hyperplastic and neoplastic breast disorders. BMC Cancer *7*, 17.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Das, J., Ho, M., Zikherman, J., Govern, C., Yang, M., Weiss, A., Chakraborty, A.K., and Roose, J.P. (2009). Digital signaling and hysteresis characterize ras activation in lymphoid cells. Cell *136*, 337–351.

Duncan, J.S., Whittle, M.C., Nakamura, K., Abell, A.N., Midland, A.A., Zawistowski, J.S., Johnson, N.L., Granger, D.A., Jordan, N.V., Darr, D.B., et al. (2012). Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. Cell *149*, 307–321.

Ferrell, J.E., Jr., and Machleder, E.M. (1998). The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. Science *280*, 895–898.

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737–741.

Gille, H., Kortenjann, M., Thomae, O., Moomaw, C., Slaughter, C., Cobb, M.H., and Shaw, P.E. (1995). ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and transactivation. EMBO J. *14*, 951–962.

Gutenkunst, R.N., Waterfall, J.J., Casey, F.P., Brown, K.S., Myers, C.R., and Sethna, J.P. (2007). Universally sloppy parameter sensitivities in systems biology models. PLoS Comput. Biol. *3*, 1871–1878.

Hansen, A.S., Hao, N., and O'Shea, E.K. (2015). High-throughput microfluidics to control and measure signaling dynamics in single yeast cells. Nat. Protoc. *10*, 1181–1197.

Hansen, A.S., and O'Shea, E.K. (2015). Limits on information transduction through amplitude and frequency regulation of transcription factor activity. Elife *4*, https://doi.org/10.7554/eLife.06559.

Hao, N., and O'Shea, E.K. (2011). Signal-dependent dynamics of transcription factor translocation controls gene expression. Nat. Struct. Mol. Biol. *19*, 31–39.

Hasenfuss, S.C., Bakiri, L., Thomsen, M.K., Williams, E.G., Auwerx, J., and Wagner, E.F. (2014). Regulation of steatohepatitis and PPARgamma signaling by distinct AP-1 dimers. Cell Metab. *19*, 84–95.

Ingram, P.J., Stumpf, M.P., and Stark, J. (2006). Network motifs: structure does not determine function. BMC Genomics *7*, 108.

Janes, K.A. (2015). An analysis of critical factors for quantitative immunoblotting. Sci. Signal. *8*, rs2.

Janes, K.A., Wang, C.C., Holmberg, K.J., Cabral, K., and Brugge, J.S. (2010). Identifying single-cell molecular programs by stochastic profiling. Nat. Methods *7*, 311–317.

Jaqaman, K., Loerke, D., Mettlen, M., Kuwata, H., Grinstein, S., Schmid, S.L., and Danuser, G. (2008). Robust single-particle tracking in live-cell time-lapse sequences. Nat. Methods *5*, 695–702.

Kholodenko, B.N., Hancock, J.F., and Kolch, W. (2010). Signalling ballet in space and time. Nat. Rev. Mol. Cell Biol. *11*, 414–426.

Kwong, L.N., Costello, J.C., Liu, H., Jiang, S., Helms, T.L., Langsdorf, A.E., Jakubosky, D., Genovese, G., Muller, F.L., Jeong, J.H., et al. (2012). Oncogenic NRAS signaling differentially regulates survival and proliferation in melanoma. Nat. Med. *18*, 1503–1510.

Lee, R.E., Walker, S.R., Savery, K., Frank, D.A., and Gaudet, S. (2014). Fold change of nuclear NF-kappaB determines TNF-induced transcription in single cells. Mol. Cell *53*, 867–879.

Legewie, S., Bluthgen, N., and Herzel, H. (2005). Quantitative analysis of ultra-sensitive responses. FEBS J. *272*, 4071–4079.

Little, S.C., Tikhonov, M., and Gregor, T. (2013). Precise developmental gene expression arises from globally stochastic transcriptional activity. Cell *154*, 789–800.

Mackeigan, J.P., Murphy, L.O., Dimitri, C.A., and Blenis, J. (2005). Graded mitogen-activated protein kinase activity precedes switch-like c-Fos induction in mammalian cells. Mol. Cell. Biol. *25*, 4676–4682.

Mangan, S., Zaslaver, A., and Alon, U. (2003). The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. J. Mol. Biol. *334*, 197–204.

Murphy, L.O., MacKeigan, J.P., and Blenis, J. (2004). A network of immediate early gene products propagates subtle differences in mitogen-activated protein kinase signal amplitude and duration. Mol. Cell. Biol. *24*, 144–153.

Murphy, L.O., Smith, S., Chen, R.H., Fingar, D.C., and Blenis, J. (2002). Molecular interpretation of ERK signal duration by immediate early gene products. Nat. Cell Biol. *4*, 556–564.

Nakakuki, T., Birtwistle, M.R., Saeki, Y., Yumoto, N., Ide, K., Nagashima, T., Brusch, L., Ogunnaike, B.A., Okada-Hatakeyama, M., and Kholodenko, B.N. (2010). Ligand-specific c-Fos expression emerges from the spatiotemporal control of ErbB network dynamics. Cell *141*, 884–896.

Niepel, M., Hafner, M., Pace, E.A., Chung, M., Chai, D.H., Zhou, L., Muhlich, J.L., Schoeberl, B., and Sorger, P.K. (2014). Analysis of growth factor signaling in genetically diverse breast cancer lines. BMC Biol. *12*, 20.

O'Shaughnessy, E.C., Palani, S., Collins, J.J., and Sarkar, C.A. (2011). Tunable signal processing in synthetic MAP kinase cascades. Cell *144*, 119–131.

Okazaki, K., and Sagata, N. (1995). The Mos/MAP kinase pathway stabilizes c-Fos by phosphorylation and augments its transforming activity in NIH 3T3 cells. EMBO J. *14*, 5048–5059.

Pargett, M., and Albeck, J.G. (2016). Mapping the spectrum of gene expression responses. Cell Syst. *2*, 221–222.

Pargett, M., Gillies, T.E., Teragawa, C.K., Sparta, B., and Albeck, J.G. (2017). Single-cell imaging of ERK signaling using fluorescent biosensors. Methods Mol. Biol. *1636*, 35–59.

Purvis, J.E., and Lahav, G. (2013). Encoding and decoding cellular information through signaling dynamics. Cell *152*, 945–956.

Regot, S., Hughey, J.J., Bajar, B.T., Carrasco, S., and Covert, M.W. (2014). High-sensitivity measurements of multiple kinase activities in live single cells. Cell *157*, 1724–1734.

Rhee, A., Cheong, R., and Levchenko, A. (2012). The application of information theory to biochemical signaling systems. Phys. Biol. *9*, 045011.

Ryan, M.B., Der, C.J., Wang-Gillam, A., and Cox, A.D. (2015). Targeting RAS-mutant cancers: is ERK the key? Trends Cancer *1*, 183–198.

Saito, T.H., Uda, S., Tsuchiya, T., Ozaki, Y., and Kuroda, S. (2013). Temporal decoding of MAP kinase and CREB phosphorylation by selective immediate early gene expression. PLoS One *8*, e57037.

Shah, M., Smolko, C.M., Kinicki, S., Chapman, Z.D., Brautigan, D.L., and Janes, K.A. (2017). Profiling subcellular protein phosphatase responses to coxsackievirus B3 infection of cardiomyocytes. Mol. Cell. Proteomics *16*, S244–S262.

Sparta, B., Pargett, M., Minguet, M., Distor, K., Bell, G., and Albeck, J.G. (2015). Receptor level mechanisms are required for epidermal growth factor (EGF)-stimulated extracellular signal-regulated kinase (ERK) activity pulses. J. Biol. Chem. *290*, 24784–24792.

Sturm, O.E., Orton, R., Grindlay, J., Birtwistle, M., Vyshemirsky, V., Gilbert, D., Calder, M., Pitt, A., Kholodenko, B., and Kolch, W. (2010). The mammalian

MAPK/ERK pathway exhibits properties of a negative feedback amplifier. Sci. Signal. *3*, ra90.

Tam, W.L., Lu, H., Buikhuisen, J., Soh, B.S., Lim, E., Reinhardt, F., Wu, Z.J., Krall, J.A., Bierie, B., Guo, W., et al. (2013). Protein kinase C alpha is a central signaling node and therapeutic target for breast cancer stem cells. Cancer Cell *24*, 347–364.

Tay, S., Hughey, J.J., Lee, T.K., Lipniacki, T., Quake, S.R., and Covert, M.W. (2010). Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. Nature *466*, 267–271.

Treisman, R. (1985). Transient accumulation of c-fos RNA following serum stimulation requires a conserved 5′ element and c-fos 3′ sequences. Cell *42*, 889–902.

Uhlitz, F., Sieber, A., Wyler, E., Fritsche-Guenther, R., Meisig, J., Landthaler, M., Klinger, B., and Bluthgen, N. (2017). An immediate-late gene expression module decodes ERK signal duration. Mol. Syst. Biol. *13*, 928.

Vial, E., Sahai, E., and Marshall, C.J. (2003). ERK-MAPK signaling coordinately regulates activity of Rac1 and RhoA for tumor cell motility. Cancer Cell *4*, 67–79.

Wang, L., Brugge, J.S., and Janes, K.A. (2011). Intersection of FOXO- and RUNX1-mediated gene expression programs in single breast epithelial cells during morphogenesis and tumor progression. Proc. Natl. Acad. Sci. USA *108*, E803–E812.

Woods, D., Parry, D., Cherwinski, H., Bosch, E., Lees, E., and McMahon, M. (1997). Raf-induced proliferation or cell cycle arrest is determined by the level of Raf activity with arrest mediated by p21Cip1. Mol. Cell. Biol. *17*, 5598–5611.

Zhong, G., Chen, X., Fang, X., Wang, D., Xie, M., and Chen, Q. (2016). Fra-1 is upregulated in lung cancer tissues and inhibits the apoptosis of lung cancer cells by the P53 signaling pathway. Oncol. Rep. *35*, 447–453.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Fra-1, clone C-12	Santa Cruz Biotechnology	Cat#sc28310; RRID: AB_627632
Anti-Egr-1, clone 15F7	Cell signaling	Cat#4153; RRID: AB_2097038
Anti-cFos, clone 9F6	Cell signaling	Cat#2250; RRID: AB_2247211
Chemicals, Peptides, and Recombinant Proteins		
Epidermal growth factor	Peprotech	Cat#AF-100-15
Amphiregulin	Peprotech	Cat# 100-55B
Gefitinib (ZD1839)	Selleck Biochemicals	Cat#S1025
PD0325901	Selleck Biochemicals	Cat#S1036
SCH772984	Selleck Biochemicals	Cat#S7101
Trametinib (GSK1120212)	Selleck Biochemicals	Cat#S2673
Cholera Toxin	Sigma-Aldrich	Cat#C8052
Hydrocortisone	Sigma-Aldrich	Cat#H0888
Insulin	Sigma-Aldrich	Cat#I9278
Bovine Serum Albumin	Sigma-Aldrich	Cat#A7906
Heat Inactivated Horse Serum	Life Technologies	Cat#26050
Puromycin	Life Technologies	Cat#A113803
Fugene HD	Promega	Cat#E2311
DMEM/F-12 1:1	Life Technologies	Cat#11320
Critical Commercial Assays		
RNAaqueous Phenol-free total RNA Isolation Kit	Invitrogen	Cat#AM1912
Nucleospin RNA	Machery-Nagel	Cat#740955.50
Experimental Models: Cell Lines		
Human: MCF-10A, clone 5E	Joan Brugge, Harvard Medical School (Janes et al., 2010)	RRID:CVCL_0598
Human: 5E/Fra1::mVenus/ERKTR-mCherry/NCd2	This paper	Available on request
Human: 5E/Fra1::mCherry/EKAR3	This paper	Available on request
Human: 5E/EKAR3/ERKTR-mCherry	This paper	Available on request
Human: 5E/Fra1::mVenus/FIRE-mCerulean/ERKTR-mCherry	This paper	Available on request
Recombinant DNA		
Plasmid: pPBJ-EKAR3nes-puro	(Sparta et al., 2015)	Addgene # forthcoming
Plasmid: pMSCV-ERKTR-mCherry-puro	(Sparta et al., 2015)	Addgene # forthcoming
Plasmid: pAAV-Fra1-mVenus	This paper	Addgene # forthcoming
Plasmid: pAAV-Fra1-mCherry	This paper	Addgene # forthcoming
pX330-U6-Chimeric_BB-CBh-hSpCas9	(Cong et al., 2013)	Addgene # 42230
Plasmid: pX330-FOSL1	This paper	Addgene # forthcoming
Plasmid: pMSCV-NCd2-puro	This paper	Addgene # forthcoming
Plasmid: pMSCV-FIRE-mCerulean-puro	This paper	Addgene # forthcoming
Oligonucleotides		
Fra1-forward sequencing primer: TCAGCTCATCGCAAGAGTAGC	This paper	N/A
Fra-1 reverse sequencing primer: AGCTCAAGCAGCTCAACCG	This paper	N/A
FOSL1 forward qPCR primer: TCCACCCAACCCTATCTCCTG	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FOSL1 reverse qPCR primer: TCTCAAAGCTGAGATCCGCAG	This paper	N/A
EGR1 forward qPCR primer: TCAGGAGTTGGAATGTTGTAGTTAC	This paper	N/A
EGR1 reverse qPCR primer: TGGGCAATAAAGCGCATTCAATG	This paper	N/A
FOS forward qPCR primer: GGGGCAAGGTGGAACAGTTA	This paper	N/A
FOS reverse qPCR primer: AGGTTGGCAATCTCGGTCTG	This paper	N/A
HINT1 forward qPCR primer: GCCTTGCTTTCCATGACATT	(Wang et al., 2011)	N/A
HINT1 reverse qPCR primer: GCCTTGCTTTCCATGACATT	(Wang et al., 2011)	N/A
Software and Algorithms		
NIS-Elements AR ver. 4.20	Nikon	RRID:SCR_014329
Bio-Formats ver. 5.1.1 (May 2015)	OME	RRID:SCR_000450
uTrack 2.0	(Jaqaman et al., 2008)	http://www.utsouthwestern.edu/labs/danuser/software/
MATLAB	Mathworks	SCR_001622
Other		
Glass Bottom Plates, #1.5 cover glass	In Vitro Scientific	Cat#P24-1.5H-N, P96-1.5H-N

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, John Albeck (jgalbeck@ucdavis.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Culture and Media

MCF-10A 5E cells (Janes et al., 2010) were used for all experiments. Cells were cultured in DMEM/F12 supplemented with 5% horse serum, 10 ng/mL insulin, 500 ng/mL hydrocortisone, 100 ng/mL Cholera Toxin, 20 ng/mL EGF, and 50 μg/mL Penicillin/Streptomycin. Cells were maintained in 10 cm plates and passaged approximately every 4 days. For imaging, cells were cultured in DMEM/F12 lacking phenol red, folate, and riboflavin, supplemented with 0.1% BSA, hydrocortisone, and cholera toxin; EGF and inhibitors were added at the indicated concentrations.

METHOD DETAILS

Reporter Cell Line Construction

MCF10A cells with FP tags inserted at the endogenous *FOSL1* locus were generated by transfecting cells with a pX330 plasmid (Cong et al., 2013) carrying the Cas9 gene and an sgRNA guide specific for *FOSL1* together with a plasmid homology repair template carrying either mCherry or mVenus, flanked on either side by 1.4 kB of *FOSL1* genomic sequence. Positive integrants were selected by fluorescence-based cell sorting, and clonal populations were isolated. Correct integration was confirmed by sequencing of PCR products amplified from genomic DNA and by immunoblotting to confirm a 26 kDa shift in the molecular weight of Fra-1 corresponding to the FP tag. NCd2, ERKTR, and FIRE reporters were introduced using retroviral gene transfer, and the EKAR3 reporter was stably integrated using the piggyBAC transposase system (Pargett et al., 2017). For each reporter configuration shown in the paper, clonal cell lines were derived and used for analysis to eliminate heterogeneity arising from different integration sites.

Live Cell Microscopy and Data Processing

Multi-well plates with #1.5 glass bottoms were coated with collagen and seeded with reporter cell lines. Prior to imaging, growth medium was replaced with a custom imaging medium, iGM-base (DMEM/F12 lacking phenol red, riboflavin, and folic acid and supplemented with 0.1% BSA, hydrocortisone, and cholera toxin). Prepared culture plates were imaged on a Nikon Ti-E inverted microscope with a stage-top incubator to maintain the culture at 37°C and 5% CO_2 throughout the experiment. Microscopy and

image processing were performed as described previously (Pargett et al., 2017). Briefly, imaging sites within each well were selected and imaged sequentially at each acquisition time, automated via the NIS-Elements AR software. Images were captured using a 20x/0.75 NA objective and an Andor Zyla 5.5 scMOS camera.

mRNA Isolation and RT-qPCR

MCF-10a cells were cultured in 10 cm dishes to approximately 50% confluence, and growth factors were withdrawn for 48 prior to the start of a time course. Cells were stimulated with 20 ng/ml EGF, followed by lysis for RNA isolation at 17 time points: 0 (no stimulation), 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 180, 240, 300, and 360 minutes. Lysis and isolation employed the Macherey-Nagel NucleoSpin RNA kit or the Invtrogen RNAqueous total RNA Isolation kit. RNA isolates were reverse transcribed using a poly-T primer to select for mRNA, and amplified under conditions preserving linearity (Janes et al., 2010), accompanied by No-Target and No-Reverse-Transcription controls. qPCR of the resulting cDNA was performed with Taq polymerase and the SYBR Green dye on a Bio-Rad CFX-96 Touch, using the primers listed under Key Resources. Three replicate time courses and RNA isolations were performed. For each, two replicate qPCR plates were prepared per target gene, with some replicate wells per plate (typically for early time points). All qPCR measurements were fit to a standard curve made by serially diluting purified amplicon on each plate. All qPCR measurements were normalized to HINT1, which was, on average, invariant across the time course. Control data (negative, NRT and NTC) were normalized by the mean of the HINT1 samples. Comparisons against control values were made without normalization to HINT1. Data were excluded from the analysis only for experimental error (e.g. missed wells or contamination indicated by grossly outlier values).

Immunofluorescence Microscopy

After growth and treatment as indicated on glass-bottom 96-well plates, cells were fixed for 20 min at room temperature with a freshly prepared solution of 4% paraformaldehyde in PBS and permeabilized with 1% Triton X-100. Samples were then stained with primary and secondary antibodies in PBS+0.1% Triton X-100+2% bovine serum albumin, and images were captured on a Nikon Ti-E inverted microscope with a 20x/0.75 NA objective with an Andor Zyla 5.5 scMOS camera.

QUANTIFICATION AND STATISTICAL ANALYSIS

Image Processing

Imaging data were processed to segment and average pixels within each identified cell's nucleus and cytoplasm, using a custom procedure written for MATLAB (Pargett et al., 2017). The procedure accessed image data from ND2 files generated by NIS Elements, using the Bio-Formats MATLAB toolbox, and tracked single cell positions over time using uTrack 2.0 (Jaqaman et al., 2008). The resulting single cell time series traces were filtered for quality (minimum length of trace, maximum number of contiguous missing or corrupt data points), and ratiometric reporter levels calculated. EKAR3 level was calculated as $1 - \frac{CFP/YFP}{R_P}$, where *CFP* and *YFP* are the pixel intensities of the cyan and yellow channels, respectively, and R_P is the ratio of total power collected in cyan over that of yellow (each computed as the spectral products of relative excitation intensity, exposure time, molar extinction coefficient, quantum yield, light source spectrum, filter transmissivities, and fluorophore absorption and emission spectra). ERKTR level was calculated as I_{Cyto}/I_{Nuc}., where I_{Cyto} and I_{Nuc} are the average fluorescent intensities in the cytoplasm and nucleus, respectively.

Linear Fitting

Linear regression analysis was preformed using a MATLAB implementation of the Theil-Sen estimator, which is insensitive to outliers and robust to heteroskedastic data.

Partial Least Squares Regression

To parameterize dynamic data, each single cell ERK time trace was processed to calculate features (see Table S1). Distinct pulses in dynamic data were identified by a customized procedure implemented in MATLAB. To be robust to noise and baseline drift, pulses were defined by where the intensity rises/falls across both a minimum peak threshold and maximum valley threshold, defined as 33% and 60% of the local signal (35 time points). Mean derivative was defined as the average absolute value point-to-point difference over the entire time window. Mean duration, mean amplitude, and max amplitude calculated from all peaks within the window. Inter-peak-interval (IPI) calculated as the distance between the centers of sequential peaks. The MATLAB implementation of the SIMPLS partial least squares regression algorithm was employed to fit paired data sets to a linear model. Data were first centered by subtracting the mean from each parameter and input parameters were scaled to unit variance.

Normalization and Statistical Analysis

For all experiments shown, a minimum of 100 cells were imaged and tracked for each condition. Single-cell data points were excluded as outliers if greater than six standard deviations from the dataset mean. For all analyses, at least two independent experimental replicates were performed. Where feasible, replicate datasets were merged by normalizing for experimental differences by minimizing the mean variance over a scaling factor attributed per experiment (not per replicate well in a multi-well plate). For visual clarity, some figures show individual replicate experiments, and we verified that similar trends were observable in every replicate.

A p value of 0.05 was considered significant for all hypothesis testing. Multiple comparison statistics were controlled for false discovery rate (at 5%) via the step-up procedure of Benjamini and Hochberg. All statistical and computational tasks were performed using MATLAB. Data dispersion and precision measures are provided in the figure legends.

Dynamic Mathematical Modeling

Mathematical modeling of ERK-to-Fra-1 regulation was designed as a system of delay differential equations (DDE). Simulation of the model and calculation of the associated metrics were performed using MATLAB, and its built-in DDE solver dde23. The purpose of this theoretical treatment was to evaluate the conditions required for a "persistence detector" in this production-stabilization context. More formally, a persistence detector requires a super-linear response to increasing input duration, meaning that the longer the input lasts, the faster the stable product is produced. We begin from the putative theoretical formulation around ERK and Fra-1. This basic model of Fra-1 regulation by ERK is analyzed to identify conditions (i.e. parameter ranges) corresponding with duration sensitivity that could give rise to a persistence detector system. From these parameter ranges, we infer qualitative conditions necessary for a persistence detector.

A basic model of the Fra-1 response to ERK activity is constructed (Figure 6A), considering ERK acting to upregulate Fra-1 via two mechanisms: activation of a transcription factor, and inhibition of Fra-1 protein degradation by phosphorylation. To facilitate analysis of the basic regulatory effects, we restrict the model to summary mass action kinetics without detailed accounting for the processes of transcription and translation. Although for some ETGs, such as c-Fos, multiple phosphorylation events appear to be involved (Murphy et al., 2002), we implicitly make the simplifying assumption that a single phosphorylation of Fra-1 protein is sufficient to confer stabilization, or that subsequent modifications are relatively rapid. We include the delay of protein production via the term τ (Equation 2) to account for latency in processes such as transcriptional initiation and mRNA splicing and transport. The resulting 3 differential equations are linear in their states, and describe the dynamics of an active transcription factor TF^P, Fra-1 protein FRA, and stabilized Fra-1 protein FRA^P (parameters are described in Table S3).

$$\frac{dTF^P}{dt} = k_{pTF} \cdot ERK \cdot \left(TF^T - TF^P\right) - k_{dTF} \cdot TF^P \quad \text{(Equation 1)}$$

$$\frac{dFRA}{dt} = k_b + k_F \cdot TF^P(t-\tau) + k_{dF} \cdot FRA^P - k_{\varnothing F} \cdot FRA - k_{pF} \cdot ERK \cdot FRA \quad \text{(Equation 2)}$$

$$\frac{dFRA^P}{dt} = k_{pF} \cdot ERK \cdot FRA - k_{dF} FRA^P - k_{\varnothing FP} FRA^P \quad \text{(Equation 3)}$$

To evaluate super-linearity with respect to duration, we first establish a suitable measure of FRA accumulation, given that some portion of FRA may degrade rapidly (recall that the two forms of FRA may decay at differing rates, which depend on the parameter values for any particular model run). We seek a metric suitable to describe the overall FRA concentration remaining in the future (after ERK is deactivated) based only on current levels of FRA and FRA^P. This is possible if we assume a rapid deactivation of ERK and TF^P, as the subsequent decay of FRA species is completely defined by the model parameters. Without ERK and TF^P, modeled FRA and FRA^P levels decay according to Equations 4 and 5, i.e. exponentially depending only on their respective degradation and de-phosphorylation rates.

$$\frac{dFRA}{dt} = k_b + k_{dF} \cdot FRA^P - k_{\varnothing F} \cdot FRA \quad \text{(Equation 4)}$$

$$\frac{dFRA^P}{dt} = -\left(k_{dF} + k_{\varnothing FP}\right) \cdot FRA^P \quad \text{(Equation 5)}$$

We define and numerically evaluate this "future Fra" metric, derived from the explicit solution of the system after ERK deactivation (Equations 4 and 5). Solving first the homogenous equation for FRA^P, then the non-homogenous for FRA (using the FRA^P solution), results in the following (Equations 6 and 7). The sum of the two represents the total FRA, as a function of time after ERK deactivation, i.e. $FRA^{TOTAL}(t) = FRA(t) + FRA^P(t)$.

$$FRA = \left(FRA_0 - \frac{k_b}{k_{\varnothing F}}\right) e^{-k_{\varnothing F} t} + \frac{k_{dF} \cdot FRA_0^P}{k_{\varnothing F} - k_{dF} - k_{\varnothing FP}} \cdot \left(e^{-(k_{dF} + k_{\varnothing FP})t} - e^{-k_{\varnothing F} t}\right) + \frac{k_b}{k_{\varnothing FP}} \quad \text{(Equation 6)}$$

$$FRA^P = FRA_0^P e^{-(k_{dF} + k_{\varnothing FP})t} \quad \text{(Equation 7)}$$

To arrive at a scalar metric that avoids the dependence on time since ERK deactivation (while retaining the differences in degradation rate), we evaluate (Equations 6 and 7) at a non-zero value of t. We choose t = 1 for convenience, as any value of t > 0 will retain the qualitative relationships between different parameter values or starting FRA concentrations. The resulting metric, termed *fF* for

"future FRA", depends on the model parameters and the concentrations of FRA and FRA^P when ERK is deactivated. The question of persistence detection (super-linear accumulation) regards the rate of *fF* accumulation, i.e. it's time derivative (*dfF/dt*). A rise in this derivative (Equation 8, Figure S6A, red line) indicates super-linearity, while a constant value is perfectly linear and a falling value sub-linear.

$$\frac{dfF}{dt} = \left(k_b + k_F \cdot TF^P + k_{dF} \cdot FRA^P - \left(k_{\varnothing F} - k_{pF} \cdot ERK\right) \cdot FRA\right) \cdot e^{-k_{\varnothing F}} + \left(k_{pF} \cdot ERK \cdot FRA - \left(k_{dF} + k_{\varnothing FP}\right) \cdot FRA^P\right) \cdot \left(e^{-\left(k_{dF} + k_{\varnothing FP}\right)} + \frac{k_{dF}}{k_{\varnothing F} - k_{dF} - k_{\varnothing FP}} \cdot \left(e^{-\left(k_{dF} + k_{\varnothing FP}\right)} - e^{-k_{\varnothing F}}\right)\right) \quad \text{(Equation 8)}$$

In a super-linear response, this value will rise to a peak significantly after the stimulation time, while an approximately linear response will start high. In both cases, peak rates cannot feasibly be sustained indefinitely and will decrease as the reactions saturate. We use the time value at which *dfF/dt* peaks as a scalar measure of this super-linear duration dependence. To evaluate a single parameter set, we simulate the system (Equations 1, 2, and 3), evaluate *dfF/dt* at each time point and identify the peak time. We numerically sample over the parameter space, evaluating the peak *dfF/dt* time, to map the requirements for duration sensitivity. All sampling is log-spaced and relative, with nominal values being 1 (10^0), and ranging 2 orders of magnitude in each direction (10^{-2} to 10^2). Due to system stiffness under certain parameters with the delay term, numerical solution is intractable with the readily available dde23 algorithm in MATLAB (a Runge-Kutta order 2,3 procedure). For these parameter screens, we neglect the explicit delay term for the convenience of employing MATLAB's ode15s, a variable step solver designed for stiff systems. In this case, feasible persistence detection requires relatively slow phosphorylation of the transcription factor ($k_{ptf} = 10^{-1.5}$, $k_{dtf} = 10^{-1.5}$); this approximates a delay between TF phosphorylation and FRA production, an expected necessary condition.

Parameter regions where duration sensitivity is feasible are shown in Figure S6B (top), by plotting level sets (i.e. planes along which the value of the peak time metric is uniform). The region where k_{0F} is low (i.e. FRA is not rapidly degraded) corresponds to stabilization-independent super-linearity, due to the slow rise of the TF increasing the rate of FRA production; this is illustrated by producing the same plot with the phosphorylation rate for FRA (k_{pF}) set to zero (Figure S6B, bottom). The remaining region of super-linearity in Figure S6B (top) corresponds to the current conceptual model: FRA degradation (k_{0F}) is fast, while FRA^P degradation and dephosphorylation (k_{dF}, k_{0Fp}) are slow. Selecting a point in this region ($k_{0F} = 10^1$, $k_{dF} = 10^{-1}$, $k_{0Fp} = 10^{-1}$, red dot in Figure S6B, top), we explore the effects of varying remaining model parameters, especially as relates to the production of FRA (Figure S6C). Variation of FRA baseline and stimulated production clearly shows a tradeoff, that duration sensitivity can only exist when baseline production is small compared to stimulated production. From this model-based hypothesis, we designed experiments and analyses to evaluate FRA baseline levels and the contribution of ERK-dependent transcription, via phosphor-specific western blot time series and live cell imaging of the FIRE reporter that mimics the FRA degron on a naive promoter.

Simulations of ERK pulse experiments were performed using the DDE model. Starting from steady state conditions with no ERK, we simulated ERK pulses as square waves of various durations; simulations were continued for twice as long as the longest pulse (200 arbitrary time units). For each simulation, the total Fra-1 AUC was calculated by the trapezoidal method. We validated the AUC of Fra-1 abundance as a metric to assess persistence detection and confirmed that this metric is strictly linear in the long term with significant early curvature differences dependent on baseline and delay. From the biological perspective, the AUC is the metric most likely to reflect its net functional effect as a transcription factor, as the number of transcripts produced will scale with both the rate and duration of transcript production (Little et al., 2013), and we therefore focused our subsequent analysis on this output parameter.

DATA AND SOFTWARE AVAILABILITY

Datasets are available at https://www.mcb.ucdavis.edu/faculty-labs/albeck/data.htm.



Current Protocols

Live-cell imaging and analysis with multiple genetically encoded reporters

Journal:	*Current Protocols*
Manuscript ID	Draft
Wiley - Manuscript type:	Protocol
Date Submitted by the Author:	n/a
Complete List of Authors:	Pargett, Michael; University of California, Davis, Molecular & Cellular Biology Albeck, John; University of California, Davis, Molecular & Cellular Biology
Keywords:	Live-cell microscopy, signal transduction, fluorescent protein, biosensor, FRET, image informatics
Abstract:	Genetically encoded live-cell reporters measure signaling pathway activity at the cellular level with high temporal resolution, often revealing a high degree of cell-to-cell heterogeneity. By using multiple spectrally distinct reporters within the same cell, signal transmission from one node to another within a signaling pathway can be analyzed to quantify factors such as signaling efficiency and delay. With other reporter configurations, correlation between different signaling pathways can be quantified. Such analyses are useful in establishing the mechanisms and consequences of cell-to-cell heterogeneity and can inform new models of the functional properties of signaling pathways. In this unit, we describe an approach for designing and executing live-cell multiplexed reporter experiments. We also describe approaches for analyzing the resulting time-course data to quantify correlations and trends between the measured parameters at the single-cell level.

SCHOLARONE™
Manuscripts

Live-cell imaging and analysis with multiple genetically encoded reporters

Michael Pargett and John G. Albeck

Department of Molecular and Cellular Biology, University of California, Davis

Significance statement

Live-cell reporters have become an increasingly important tool for the study of intracellular signaling pathways. A reporter can provide data on detailed kinetics and cell-to-cell heterogeneity in signaling that are needed to update and refine models of signal transduction. However, multiple reporters are often essential to link these new views of signaling to other pathways and cellular fate choices. The simultaneous analysis of two or more reporters within the same cell presents a new set of challenges in experimental design, implementation, and data processing. The protocols described here enable multiple signals to be tracked over time and related to each other, providing the data needed to quantify links between signaling pathways in the same cell.

Abstract

Genetically encoded live-cell reporters measure signaling pathway activity at the cellular level with high temporal resolution, often revealing a high degree of cell-to-cell heterogeneity. By using multiple spectrally distinct reporters within the same cell, signal transmission from one node to another within a signaling pathway can be analyzed to quantify factors such as signaling efficiency and delay. With other reporter configurations, correlation between different signaling pathways can be quantified. Such analyses are useful in establishing the mechanisms and consequences of cell-to-cell heterogeneity and can inform new models of the functional properties of signaling pathways. In this unit, we describe an approach for designing and executing live-cell multiplexed reporter experiments. We also describe approaches for analyzing the resulting time-course data to quantify correlations and trends between the measured parameters at the single-cell level.

Keywords

Live-cell microscopy; signal transduction; fluorescent protein; reporter; biosensor; image informatics

Introduction

Live-cell imaging has become a very useful tool for signal transduction research, and can be essential for making accurate measurements of dynamic signaling activity, particularly in heterogeneous cell populations (Locke and Elowitz, 2009; Purvis and Lahav, 2013). However, measuring any single signaling event in isolation leaves many question unanswered, due to the highly interconnected nature of intracellular signaling networks. Each measurement made by a signaling reporter includes the influence of many different effects, including the activity of other pathways, metabolism, and the cell cycle, which vary from cell to cell at any time. Recording multiple signaling components within the same cell allows the dynamics of related pathways to be directly correlated without the interference of cell-to-cell variance. The techniques in this unit are focused on reliably measuring and determining the relationships among intracellular signaling components.

This unit describes techniques and design principles for the simultaneous dynamic measurement of multiple gene products or enzyme activities in single cells. Applied to multiple parts of a signaling

pathway, this approach facilitates the identification and modeling of intracellular signaling and regulatory processes. An increasing array of signaling molecules can now be observed in live cells via fluorescence microscopy. Molecular processes including kinase activity (Komatsu et al., 2011; Regot et al., 2014), transcription (Chubb et al., 2006), translation (Han et al., 2014), and intracellular translocation (Nelson et al., 2002) can be reliably imaged using existing live-cell reporters. Moreover, CRISPR-based knock-in of fluorescent proteins to genomic loci is expected to add to the number and diversity of regulatory events that can be imaged in living cells. In each case, live-cell imaging allows the observation of signal dynamics in single cells (Albeck et al., 2013) and the evaluation of intra- vs. extra-cellular noise (Elowitz et al., 2002).

This unit is designed to provide a conceptual overview of the experiment design, and deliver critical details for managing practical implementations. Four basic topics are covered: 1) Construction of suitable cell lines, 2) Long term time-lapse imaging of live cells, 3) Image processing to automate cell segmentation and extract fluorescent reporter values, and 4) Signal analysis of relationships among the reporters.

Strategic Planning

General

Efficiently executing a multiple reporter live-cell imaging experiment requires coordination of several key operations. Because the construction of cell lines bearing stable fluorescent reporter genes can be quite lengthy, careful planning is critical. In addition to the principles of the experiment, the targets to measure, fluorophores to use, and final analyses should be planned prior to embarking. Illustrated in Figure 1, the complete experiment consists of: 1) Acquiring/producing DNA constructs for reporters, 2) Inserting all constructs into desired cell line(s), 3) Imaging cells under desired treatments, and 4) Processing image data and analyzing time-series data. Review the *Time Considerations* section following the protocols for suggestions regarding time-consuming steps.

Selection of reporters and fluorophores

Reporter selection is a balance of choosing targets suitable for the study that are feasible to measure and minimally challenging to implement. Establishing the fundamental reporter function can be very application-specific (Komatsu et al., 2011; Slattery and Hahn, 2014) and is beyond the scope here. Instead, this protocol addresses the choice of fluorophores, especially fluorescent proteins, and methods to drive their expression. Five aspects of the fluorescent proteins themselves are emphasized: spectral overlap, chromophore maturation rate, aggregation, photobleaching, and brightness. Each candidate fluorophore should be evaluated against alternatives based on these features.

Basic protocol 1: Paired reporter cell line construction

This protocol addresses the preparation of cell lines expressing multiple fluorescent reporters, including the selection of reporters, transfection and transduction into the desired cell line, selection, and validation of the line. Specific DNA constructs are very application-specific, and the details of their construction are varied, but are well established using a variety of methods and are not addressed here.

Materials

Cell line(s) to receive reporters

Culture medium appropriate for selected cell line(s)

Sterile culture materials: dishes, pipettes, conical tubes

DNA construct(s) for reporters

Transfection reagents, depending on the technique chosen, e.g. Optimem, Fugene, Lipofectamine, PEI (polyethylenimene), Electroporation apparatus, etc.

Selection antibiotics (optional)

Epifluorescence microscope (to view cells and evaluate reporter expression)

Select fluorophores for each reporter

Table 1 enumerates critical fluorescent protein properties and indicates the concerns when considering each. Comparative lists of properties for different fluorescent proteins are available from a variety of sources (such as www.fpvis.org), though none are exhaustive.

1. Evaluate fluorophore options for spectral overlap. Review the excitation and emission spectra for candidates and compare with available filters to ensure that each chosen filter will allow less than 1% of the on-target intensity from off-target fluorophores. Minimizing cross-talk becomes increasingly important if reporters are not expressed at the same levels, as higher concentration reporters will affect other filtered images more. Contact filter manufacturers for details and recommendations. See Figure 2 for example spectral comparisons.

   Ensure that a nuclear marker is available for image processing, either by a reporter confined to the nucleus or a dedicated nuclear marker. If a nuclear dye is to be added at the time of imaging, ensure that its spectra are evaluated alongside those of the other reporters. Typically, three colors can be imaged with little difficulty, e.g. with cyan, yellow, and red fluorescent proteins (a UV dye, such as DAPI, may be used as well).

2. For protein expression level reporters, rule out any fluorescent proteins with long maturation times (greater than 1 hour). Notably fast maturing fluorescent proteins include mNeonGreen, mVenus, mCherry, and mKate2.

   Availability of maturation time measurements is not consistent. Refer to www.fpvis.org and primary literature regarding the development of particular fluorescent proteins.

3. Evaluate the tendency of remaining candidates to aggregate, a common issue for fluorescent proteins. While tolerable in some cases, aggregation is generally detrimental and leads to punctate imagery and increased risk of perturbing the function of anything fused to the protein.

   Direct data on aggregation are often difficult to find, and decisions rely on more qualitative commentary. It is often necessary to evaluate the trade-off between maturation rate, aggregation and brightness (see step 4). Selections are generally qualitative based on which characteristics are most important for a particular application.

4. Consider the quantum yield, extinction coefficient and relative photobleaching time of remaining candidates and determine an acceptable compromise between brightness (estimated as quantum yield times extinction coefficient) and bleaching.

   Compiled properties may be found at www.fpvis.org, and primary literature typically includes quantification of these properties.

Select reporter gene delivery and co-expression strategy

5. Compare co-expression options to determine a suitable compromise between the expected expression conditions and the complexity of the reporter insertion process. Multiple reporters may be included on one plasmid using an internal ribosome entry site (IRES), or a "self-cleaving" 2A peptide. Multiple plasmids or viruses may also be used to introduce reporters.

   It is useful to consult experimental comparisons of bicistronic expression systems (Goedhart et al., 2011) and promoter expression output (Qin et al., 2010) when choosing a co-expression strategy. Due to possible interference between promoters, it is not recommended to use multiple promoters on any one plasmid.

6. Determine the reporter insertion technique to use. See Table 2 for a comparison of insertion techniques, including references for appropriate protocols.

   Transient expression can be achieved by simple transfection with the appropriate plasmid DNA (e.g. using polyethylenimine, FuGene, Lipofectimine, electroporation, etc.). However, expression levels will be variable over time and typically fade 5-8 days after transfection as cell division dilutes the plasmid DNA. The useful length of the experiment is thus limited and frequently yields noisy data that are difficult to interpret; transient transfection is therefore not recommended.

Insert reporter constructs into target cells

7. Prepare purified plasmid DNA for each reporter, and for each associated genomic insertion method as needed. Modify any reporters via molecular cloning techniques of choice. Similarly, prepare any additional constructs necessary (helper plasmids for transposon expression, genomic targeting, or retroviral packaging). See step 9 for commentary on selection antibiotic resistance genes that may be included.

   DNA for transfection into a cell line should be highly purified, for optimal performance. While DNA from a standard alkaline lysis "mini-prep" can be used, it is advantageous to purify by alcohol precipitation or, preferably, by paramagnetic beads.

8. Introduce reporters to the target cells by transfection, transduction, etc., depending on the technique chosen in steps 5 and 6.

9. Select and/or sort cells for adequate stable expression of reporter constructs. Selection can be carried out by inclusion of an appropriate antibiotic resistance gene in the same construct as the reporter. To select for multiple reporters, each construct must bear a different resistance gene, such as puromycin, neomycin, blasticidin, or hygromycin. Alternatively, cells may be sorted by flow cytometry, avoiding the need for selection antibiotics. As multiple antibiotic resistance

genes can become cumbersome, sorting is recommended for more than 2 insertions. If the insertion efficiency for each reporter is high enough, it is feasible to transfect/transduce in tandem and sort (or select) for multiple insertions rather than sorting after each.

Validate reporter insertion and function

10. Evaluate the insertion of fluorescent reporters initially by viewing cells on an epifluorescence microscope. In particular, look for side-by-side pairs of positive cells suggestive of a successful cell division.

    If expression is weak and the insertion efficiency low, it may be quite difficult to visually identify the presence of positive cells even when there are enough to successfully sort. Analysis by flow cytometry provides the best assessment of brightness and frequency.

11. After selection or sorting, evaluate reporter function via a preliminary experiment (following Basic Protocol 2). Identify two conditions that will induce minimal and maximal reporter response. Plate cells for a live-cell experiment (Basic Protocol 2 and Support Protocol 1), and apply both conditions in the same experiment, preferably in the same well via spike-in of a drug or stimulant. Evaluate the resulting movie for appropriate responses.

    For example, with an ERK reporter in a cell line responsive to EGF stimulation, withdraw growth factors for >2 hours before imaging. After beginning imaging, spike in EGF to a saturating concentration, image for enough time to observe the response (1-2 hours). Then spike in an ERK, MEK or EGFR inhibitor to saturating conditions and continue imaging to observe the inhibition (1-2 hours).

12. If any insertions are targeted to specific sites in the genome, validate the insertion site by sequencing the local region of the genome and/or the appropriate transcript.

13. Store frozen aliquots of cells lines with successful reporter integration for future use.

    It is often useful to freeze cells after each successful insertion to retain stocks that can be used for future combinations of reporters.

    If a genetically uniform clonal line is desired, it can be produced at this point (e.g. by limiting dilution).

Basic protocol 2: Time-lapse imaging of living cells

This protocol describes the culture and imaging of cells for a long-term live cell imaging experiment. The process with multiple reporters is nearly identical to that with a single reporter. See (Pargett et al.) for a specific implementation focused on a single reporter in MCF-10A epithelial cells. The goal is to keep the cells viable in an environment accessible to the microscope with high throughput and the ability to add treatments. To achieve this, cell monolayers are produced on a glass-bottom multi-well plate by coating with collagen, laminin, or another appropriate substrate. To provide adequate media for longer experiments (those longer than 24 hours), use larger wells (e.g. a 24-well plate) and spot only a small portion with the substrate. To facilitate robust image processing, it is

advantageous to plan for one well to be left without cells; background images will be taken from this well.

Materials

Glass-bottom culture plates for imaging. (24-well: Cellvis P24-1.5H-N, 96-well: Cellvis P96-1.5H-N, square well 96-well: Brooks Life Science Systems MGB096-1-2-LG-L)

Substrate coating solution, e.g. collagen or laminin

Pipetting supplies, typ. 20 µl, 1 µl pipettors, 5-300 µl 8- or 12- channel multipipettor (or similar), multichannel aspirating tool, sterile reagent reservoirs.

Imaging medium for cells, preferably without phenol red, riboflavin and folic acid, to reduce background fluorescence (e.g. Fluorobrite from Invitrogen, or custom media formulations that are available from a variety of vendors)

Wide-field epifluorescence microscope equipped with a digital camera, automated stage and capture controls, and an environmental control chamber for samples. E.g. Nikon Eclipse Ti with NIS Elements control software, Andor Zyla 5.5 camera, Lumencor SOLA light source, and OKO Labs stage-top incubator.

Prepare imaging culture plate

1. Coat each well of the glass-bottom plate with the substrate solution, using enough solution to cover the desired region. For long-term experiments (>24 hours), it is preferable to increase the ratio of media volume to cells by using a square well 96-well plate or a 24-well plate and applying only a 3 µl "spot" of the substrate solution in the center of each well. Cover and incubate at least 30 minutes.

2. Aspirate any remaining substrate solution using a Pasteur pipette or aspirating tool. Wash gently with sterile PBS, and aspirate. Avoid touching the plate with pipette tips, as it will mar the coated substrate.

3. Obtain a suspension of the desired cells by treating with trypsin and passaging as usual. Count cell concentration by hemocytometer or automated cell counter. Cells will settle; remix suspension prior to use.

4. Add the desired number of cells to each well. Determining a suitable cell count may require experimentation. For medium density of MCF-10A cells, 15,000 cells per well is sufficient for 96-well plate. If only a "spot" was coated in step 1, add 5,000 cells in 10 ul media pipetted directly onto the substrate spot to yield near confluence. It is recommended to leave a well empty of cells to provide a convenient background measurement during imaging.

5. Add PBS or sterile water to spaces between wells to provide humidity and prevent drying of the cell suspension. Cover the plate and incubate for 2 to 2.5 hours.

6. Observe with an inverted microscope to determine if cells have adhered. Elongated cells and small processes indicate adhesion and that more media may be added. If cells still appear spherical, incubate for an additional hour.

7. When cells are adhered, gently add 0.5 ml media. Pipet slowly and avoid contacting the adhered cells to minimize disturbance. Incubate overnight to allow cells to fully attach and acclimate.

8. Check cells with an inverted microscope to verify attachment and appearance. When cells appear fully attached and healthy, experimental treatments may be started. Switch to an imaging media, if available, well before imaging (4 hours or more) to avoid observing transient perturbations from the media change.

Prepare microscope and imaging procedure

9. Prior to desired start of imaging, power on the microscope and environmental controller. Set environment temperature and gas mixture and allow time for the system to stabilize (time required will vary over different hardware).

10. When the environmental controller is stable at the set point, transfer the plate to the microscope stage and mount it securely. Any flexibility in the mounting often results in frame shifts during imaging, especially when any treatments are added.

11. Start the microscope control software, focus on cells, and set desired acquisition parameters in the software (i.e. filters to use, plate positions to image, acquisition interval, total number of time points, etc.). See Support Protocol 1 for instructions and recommendations specific to NIS Elements control software.

> *All mechanical movements require time, so it is beneficial to minimize movement distances by choosing plate positions in an order that always moves to an adjacent well and minimizes the total travel. Similarly, selecting filters in the order they are physically positioned limits the movement from one to the next.*
>
> *When choosing excitation intensity and exposure time, there are several factors to consider: 1) maximum image intensities should be well below the saturation point of the camera; 2) the intensities of different channels should be kept similar for later image processing steps; 3) cellular phototoxicity can often be reduced by using lower intensity excitation for longer exposure times.*

Perform imaging experiment

12. When all imaging parameters are set, begin imaging. It is recommended to review a trial run of all filters at all plate positions to verify that the views are suitable.

13. If timed treatments are required, pause imaging after completion of all plate positions, prior to the start of the next time point. Apply treatments (e.g. spike-in of a drug, or washing and replacement with a new media) as quickly as possible while being careful to avoid disturbing or contaminating the plate.

Contacting the plate directly often results in minor image frame shifts even when it is firmly secured. Shifts can be troublesome for image processing, but software implementations of image registration can correct minor cases.

A spike solution containing the treatment at 20-fold the final concentration diluted in imaging medium allows for easy pipetting with a minimal (5%) change in the culture volume. Direct addition of solutions with different viscosity than the imaging medium (such as DMSO or glycerol) should be avoided due to the potential for poor mixing. The disturbance of the imaging medium upon treatment can also stimulate responses in signaling pathways due to redistribution of localized signaling molecules or changes in fluorescence readings due to local photobleaching of imaging medium; vehicle-only treatment is an important control

14. After all imaging is complete, copy the image file(s) to a suitable location for processing, and retain a backup copy in a separate location.

Basic protocol 3: Processing for multi-channel time-series images

This protocol describes the procedure and algorithms for automated processing of time-lapse images to extract single cell data. As computation procedures often vary widely among implementations, we describe here the concepts of the procedure and provide notes regarding specific software implementations using MATLAB. Image processing procedures are complex with multiple layers of subroutines, so this protocol should be reviewed carefully to understand the full process prior to developing or implementing software. A detailed implementation for MATLAB is available in (Pargett et al.). Alternative software solutions that perform the same functions are equally appropriate; for example, CellProfiler provides a publicly available solution with many of the features described here in a graphical user interface (Carpenter et al., 2006).

Materials

Computer with at least 4 GB RAM

Scripting software (e.g. MATLAB)

(Optional) Bio-Formats software (http://www.openmicroscopy.org/site/products/bio-formats)

Particle tracking software, e.g. uTrack (Jaqaman et al., 2008)

Preparation of an image processing pipeline

1. Install all necessary software per distributor instructions (e.g. MATLAB, Python, etc.). The Bio-Formats library is useful for reading images from proprietary microscopy files. Particle tracking software (such as uTrack) is used to track cells over time.

2. Make images available to the processing computer by mounting on a network drive or copying the image data file(s) locally.

As accessing network storage is often relatively slow, it is preferable to copy the data to the processing computer ahead of time.

3. Set up access to microscopy files, or extract data to image files (e.g. tiff). If using Bio-Formats, review the documentation (http://www.openmicroscopy.org/site/support/bio-formats5.4/) for methods to access imagery data from different microscopy file types using the chosen scripting software. Microscopy data or image files will be read one frame at a time, each followed by the processing below.

    It is convenient to load all color channels for each frame. If these are exported to different tiff files, all channels will need to be loaded at once (e.g. into a Xpixels-by-Ypixels-by-Channels matrix).

Correct images for known biases and background intensity

4. (Optional) Correct non-uniform intensity via a recorded image of a uniform field. Scale each pixel intensity by dividing by the relative "intensity" it receives from a uniform field, with the highest intensity pixel valued at 1. With a recorded image of a uniform field, simply divide by the maximum value to give the pixel scalings. It is prudent to filter (i.e. smooth) noise or remove outlier pixels first.

5. Identify background intensity in each color channel and subtract from images. To get intensity values, define the well and region to sample for background, load the relevant image(s) if necessary and take the mean value of the selected pixels for each color channel.

    Ideally, a background value should be determined for each time point. Retaining and imaging a well with media and treatment compounds, but no cells, provides a convenient reference for background. In this case, it is more efficient to evaluate and store background values prior to processing other frames. Without a designated well, a cell-free region from the current or other wells may be used.

Segment images to estimate nuclear regions

6. Reduce noise for segmentation by filtering (i.e. smoothing). A Gaussian filter sized to cover approximate $1/10^{th}$ of a cell nucleus is typically sufficient.

7. (Optional) If the only reliable marker in the cells is a nuclear-excluded cytoplasmic marker, use a gradient magnitude filter to highlight "edges", i.e. where the image intensity changes quickly, targeting the border between the bright cytoplasm and dark nucleus. Invert the resulting image intensities by subtracting it from its maximum value (this makes the cytoplasm dark and both the nuclei and the background bright, which aligns with the following procedures).

    Gradient magnitude filtering may be carried out by applying a Sobel filter in both the X and Y directions and taking the L^2 norm of the two filtered images (square root of sum of squares).

8. Segment nuclei by selecting regions that exceed a threshold value and filtering for shape and size. To be robust to varied expression levels, consider multiple thresholds spanning the range of intensities in the image. Using 20 thresholds linearly spaced between the 5^{th} and 95^{th} percentile intensity values has been successful and robust over a variety of imaging conditions and cell types.

Binary morphological operations and analyses are commonly included in packages focused on image analysis. For example, MATLAB includes the function `regionprops` *that calculates metrics for each region of a binary image (e.g. area, perimeter, etc.)*

9. Remove "salt and pepper" noise in the binary image by morphological dilation and erosion. Both dark and light spots can be removed by first dilating by 1 pixel, eroding by 2 pixels, and re-dilating by 1 pixel. Dilation and erosion sizes may be scaled up to suit the application, e.g. when desired regions (nuclei) are larger.

10. Identify each distinct binary feature and select only those matching the feasible size and shape of a cell nucleus. For size, evaluate the total area, and for shape, use the squared ratio of the ideal perimeter of a circle (with the measured area) to the actual perimeter. Store passing features by setting corresponding pixels in a logical mask to TRUE.

A feature found again for the next threshold will be the same size or larger, and will simply overwrite the logical mask.

Create representative masks over the nucleus and cytoplasm of each segmented cell

11. Create nuclear masks. The logical mask from segmentation may be used directly for nuclear masks. If appropriate, regions may be eroded to confine masks further within nuclei to ensure robust sampling. Such approaches may be adjusted based on performance for a particular cell type.

12. Create cytoplasmic masks by dilating the nuclear mask to extend over a portion of the (presumed) cytoplasm region. From that expanded mask, remove a smaller dilation of the nuclear mask, making a hole over the nuclear region. The size of dilations may be scaled based on the expected nucleus size.

13. (Optional) Remove pixels in the cytoplasmic mask that could be associated with more than one nucleus. If using label matrices as masks (each nucleus or cytoplasm is represented by regions of a particular integer value), perform cytoplasm mask construction again with inverted (negative) label values. In grayscale morphological operations, this will give dilations that overlap in the opposite manner. Invert the resulting mask, then set to zero (i.e. background) all pixels with values that differ between the two masks.

14. Extract channel data by taking the average over each mask for each color channel. Optionally, to guard against outliers, exclude a maximum and minimum fraction of the intensity values within each mask. Robust performance can be achieved using only the values between the 20^{th} and 80^{th} percentiles, provided masks are large enough (hundreds of pixels). Store values in a matrix alongside X- and Y-coordinate values for the centroids of each nucleus, one row for each cell.

Track cells over time

After all images have been processed via steps 4-14, the cells may be tracked across frames using the X and Y coordinate data stored in step 14.

15. Track cells over time using the centroid coordinates and a particle tracking algorithm (or similar software). uTrack is a suitable software for the MATLAB environment, with a usage example available in (Pargett et al.).

16. Align track coordinates with segmentation centroid coordinates. The values extracted in step 14 need to be rearranged (within each time point) to provide coherent cell tracks. Find the nearest centroid coordinate to each tracked coordinate, and rearrange the order of cells (e.g. the row entries in a cells-by-time-by-channel data matrix) for each time point to create a dataset where each track is single row.

Basic protocol 4: Paired signal analysis

This protocol provides a framework and describes a variety of analyses that may be valuable for paired time-series data. Analyses comprise quality control, correlation metrics, and modeling. These procedures are independent and may be included as desired, though initial quality control is highly recommended. All detailed computational procedures are presented for use with MATLAB, but any such scripting software is suitable with some variation in the availability and features of packages. See Anticipated Results for discussion of the outputs from these analyses.

Materials

Computer with at least 4 GB RAM

Scripting software, e.g. MATLAB

Quality control

Preliminary processing to ensure the quality of tracks is critical for the accuracy of all further analyses. Quality control corrects for the errors that arise in automated image processing, as well as the occasional outlier behavior from the biological system.

1. Correct for gaps in tracking. Tracking software will typically allow for some gaps in the data (e.g. linking points that appear in frames 3 and 5, but not 4), but many signal analyses require the track to be complete. Choose a maximum gap size that is acceptable, and fill any gaps up to that size by interpolating between the data on either side. If multiple track segments remain that are separated by oversized gaps, keep the longest of these tracks.

> *It is recommended to choose a maximum gap size according to the expected rate of change for the signals at hand, relative to the imaging interval. If the signal changes rapidly, and especially if it may be pulsatile or oscillatory, a shorter gap size will be necessary to ensure that relevant activity is not missed. Gaps are expected even with effective image processing, as cells may transiently change shape such that they no longer fit the filters, and passing debris in the media may occlude the view of some cells for a short time.*

2. Filter tracks for desired length by completely removing any tracks that are shorter than a desired threshold.

> *Choose a minimum track length that is suitable for the desired analyses (e.g. based on the duration of the expected activity). When a timed stimulation or drug treatment has*

been applied early in the experiment, it may be convenient to remove tracks that do not start at the beginning of imaging (the first time point).

3. Calculate any derived signals, such as normalized values of translocation reporters (ratio of cytoplasmic to nuclear signal) or bi-color reporters like FRET where the values of two color channels are divided. This calculation may be performed at any point up to here, but it can be important to have complete for the following filtering step.

4. Filter data for consistency by removing tracks that contain grossly outlier behavior. Typically, examples of outlier behavior include values outside of a reasonable range, and changes that occur too quickly (i.e. derivative magnitudes beyond a bound). Bounds need to be determined based on the system being measured and imaging interval.

Reasonable bounds should be chosen based on the combination of the biochemical process and underlying signal being recorded. For example, both FRET and translocation reporters are automatically normalized with a district range of signal value for any given reporter.

Direct signal comparisons

Because each cell contains multiple reporters, direct comparisons may be made between the reporters in each cell. These comparisons are essentially forms of correlation, some of which accommodate non-linearity or dynamic lag.

5. Compute correlation coefficients for each track. The most common correlation coefficient, Pearson's rho, specifically evaluates linear relationships, while Spearman's tau does not diminish for non-linear relationships. Both are typically implemented in basic and/or statistics packages for scripting software.

6. Compute mutual information for each track. Mutual information is the most general form of "correlation", and evaluates how much knowing one signal's value restricts possible values of the other. It is implemented in some statistical packages for scripting software, though less widely than basic correlation coefficients.

Computation of mutual information relies on binning data to make a histogram, or on summing up nearest neighbors. As such, it is particularly important to have many data points for a robust estimate of mutual information.

7. Compute cross-correlation for each track. Cross-correlation is equivalent to correlation, but allowing for a dynamic lag between signals. Properly normalized, it represents the linear correlation coefficient evaluated for each possible time lag (i.e. shifting one signal backward in time and calculating correlation, neglecting points that no longer overlap at the ends). Typically, the critical values to record are the peak (maximum) correlation value and the lag at which it occurs. Cross-correlation is typically included in signal processing packages.

Signal response modeling

Signal response models may be constructed to estimate the specific relationship between signals across a population of cells. Each pair of input and output signals from individual cells becomes

one observation, and the mean relationship is estimated over all observations. The time-series signals are broken down into a smaller number of "characteristics" to comprise an observation. Partial least squares modeling (PLS) is particularly useful as it extracts the most informative features from the covariance structure, allowing many input and output characteristics to be used simultaneously.

8. Determine a suitable set of characteristics to describe the single cell tracks. Basic statistics, such as mean value, variance, maximum and minimum values, etc. are likely to be informative. Signals that fluctuate significantly may be informed by mean (and/or max and min) rates of change. Signals with pulsatile or oscillatory behavior may require evaluation of mean (and/or max, min) frequency and amplitude. It is often useful to apply some degree of smoothing when calculating maximum and minimum values, or rates of change, as these are highly sensitive to noise.

> *There are no hard limitations on what decompositions are applicable. If convenient for the data and further desired analyses, rigorous mathematical expansions may be applied, such as the Fourier Transform or an orthogonal basis decomposition. Such expansions offer a relatively unbiased breakdown of the signal, but are often more difficult to interpret conceptually.*

9. Calculate the characteristic values for every track. Each track will result in an observation vector containing the values for each of the characteristics, for each of the signals recorded.

10. Infer models using one or more characteristics of two selected signals as input and output. Any model fitting routine is applicable, and the selection and application of such a scheme is beyond this protocol. However, it is highly recommended to use PLS regression as it can simultaneously consider all of the characteristics in the input and/or the output.

> *PLS regression and similar procedures are fundamentally linear problems with each of the terms considered being one of the signal characteristics. However, any non-linear combinations of characteristics may be added as additional terms when desired, e.g. x^2 or e^(x*y).*

Support Protocol 1: Imaging with NIS Elements

1. In NIS Elements, open the ND Acquisition panel.

2. Set path and filename for saving data. Data are automatically saved when experiment is complete.

3. In the Time tab, select the imaging interval and total imaging duration.

4. In the Wavelength tab, select the desired fluorescent channels for the experiment.

5. Navigate to the first well in the imaging sequence and find cells. Turn on "Perfect Focus" and set the appropriate focal plane using the focus wheel.

6. Open the Zyla Settings panel and set the exposure for each channel to be used in imaging experiment.

7. Using the joystick, navigate to a suitable area for imaging; try to avoid areas where cells are overcrowded.

8. In the ND Acquisition panel select the XY tab. Navigate to the first desired ROI using the joystick and select the XY point by checking the box in the XY selection tab.

9. Continue selecting ROIs in each desired well by navigating to each ROI and selecting the coordinate using the checkbox. Adjust focal plane at each ROI as necessary using Perfect Focus.

10. To select a background ROI (optional), navigate to an XY position that contains no cells in the field of view. This will be used later to calculate background during image processing.

11. When all ROIs have been selected return to first ROI and check that focus and exposure are set correctly and that the plate has not shifted.

12. Start imaging experiment by hitting the "Run Now" button.

13. To add treatments during imaging, pause imaging using the "Pause" button on the Acquisition Window. Then, remove plate lid and environmental control box, add treatments to each well, minimize contact with plate. Return lid and environmental control box to plate, and resume imaging using the "Resume" button on the Acquisition Window.

14. Once the time course has completed, the file will be saved in the location specified during imaging set up. Alternatively, terminate the image acquisition using the "Finish" button.

15. When finished, close NIS Elements and shut down the microscope.

Reagents and Solutions

Collagen coating solution: 20 μg/ml rat tail type I collagen (BD Biosciences) in 0.1 N acetic acid. Sterilize acetic acid and water by filtration through 0.2 μm filter before adding irradiated collagen; do not filter after adding collagen. Store at 4° C.

Laminin coating solution: 50 mg/mL laminin 111 (Invitrogen) in 20 mM sodium acetate (pH 4), containing 1 mM $CaCl_2$. Prepare fresh for each experiment.

Commentary

Background

Live-cell studies of signal transduction have revealed kinetic aspects of signaling that are invisible to both population-based measurements (such as immunoblots) and static single-cell methods (such as immunofluorescence). For example, live-cell reporters of the p53 tumor suppressor protein have revealed that this key regulator of DNA repair is activated in discrete hour-long bursts following certain forms of DNA damage, with cells varying in both the spacing and number of bursts (Batchelor et al., 2011; Lahav et al., 2004). Kinetics observed in single cells are often incompatible with existing models for signaling pathways, prompting significant revisions to long-standing models (Batchelor et al., 2008). However, to move beyond simple description, the relationship of the single-cell kinetics for a given molecule to its upstream and downstream processes must be established, and this task typically

requires that the variables of interest be imaged simultaneously in the same cell. In the case of p53, simultaneous imaging of p53 activity and the resolution of DNA damage (upstream) or activation of the target gene p21 (downstream) was necessary to understand the importance of p53 kinetics in encoding responses to different types of DNA damage (Loewer et al., 2013; Purvis et al., 2012). Similar examples can be found in multiplexed single-cell studies of death receptor-induced apoptosis (Albeck et al., 2008), EGFR/Ras/ERK-induced proliferation (Albeck et al., 2013; Shankaran et al., 2009), and NF-kB-mediated gene induction (Nelson et al., 2004).

Single-cell experiments have revealed that “noise” plays a significant role in determining the signaling response of a given pathway to a stimulus, resulting in a distribution of responses for a population of cells responding to a uniform stimulus (Cheong et al., 2011; Raser and O'Shea, 2005). While this variation may result from stochastic effects in the biochemical process of signaling, the observed variation in any one pathway can also result from deterministic but unmeasured cell-to-cell variation in other pathways and cellular processes (Selimkhanov et al., 2014). Multiparameter analysis as presented here can be useful in identifying the sources of such variation by making it possible to quantify the amount of variance in one pathway that can be explained by measured variance in a second (or further additional) pathways. When choosing reporters for multiparameter analysis, it is important to perform a careful literature search to take into account known dependencies (or “crosstalk”) between the pathways of interest and consider how these will influence the quality of the data collected.

Critical Parameters

Throughout the processes of establishing, imaging and analyzing a multiplexed reporter cell line, it is important to double check the status and quality at every opportunity. In cell line construction, this comprises evaluating the cells after each reporter insertion, sorting/selection, and/or cloning step. Verify that cells retain the phenotype and any characteristic behavior of the parental line and that reporter expression levels are suitable for measurement, but not too high (excess expression can both perturb the cell itself and interfere with the measurement of other reporters). Additionally, be certain to perform the experiments described (Basic Protocol 1, step 11) to validate the function of the reporters.

When setting up the microscopy experiment, take care to choose frame positions with minimal obstructions (floating debris, distortions from the substrate, cells overlaying each over, etc.), and carefully double check the focus. Ensure that a background region is available (see Basic Protocol 3, step 5). After starting imaging, it is recommended to observe at least one frame of each position to ensure that it meets expectations. When adding a treatment during imaging (Basic Protocol 2, step 13), be especially cautious not to disturb the plate, and observe the first frames taken after restarting imaging to verify that they align with the previous frames. If the plate was disturbed and a large shift occurred, it is possible to realign it (to some degree) with minimal loss of data by carefully pressing the direction needed and checking the view (select a notable feature in the desired view as a point of reference).

Prior to processing images (Basic Protocol 3), ensure that the parameters are appropriate for the current experiment. Especially after multiple experiments have been processed, previously used parameters can be overlooked and give sub-optimal results. In particular, check that the limits on nucleus size (see step 8) are suitable for the cell type and the magnification that was used. Similarly, an inappropriate background can degrade processing. If the background is to be taken from frames containing cells, carefully ensure that neither cells nor fluorescent materials are present in the

background regions. Accurate background removal is particularly important when using ratiometric measures, such as FRET or translocation reporters.

While a well-constructed image processing and quality control pipeline enable high-throughput analysis, it is critical to review the data at different points. Review the raw imagery for microscopy issues, such as out-of-focus frames, misaligned filters, or obstructions in some frames. After processing, visualize tracks in bulk to ensure both that a reasonable number were generated and that overall quality is acceptable. Having many jagged or extremely noisy tracks is suggestive of a processing error. Review tracks again after quality control (Basic Protocol 4, steps 1-4) to ensure that final data are high quality. In such review, it is best to view all individual tracks, as well as averages, to ensure all outliers have been addressed. Individual tracks may be viewed simultaneously by plotting them overlaid, as grossly aberrant tracks will remain apparent. If the data is used for modeling (steps 8-10), it is important to carefully consider how the models are interpreted, as with any modeling endeavor. In particular, it is essential to evaluate for a parsimonious model, i.e. the simplest model that satisfies the data available (Burnham and Anderson, 2003).

Anticipated results

Each imaging experiment in which multiple reporters are measured (Basic Protocol 2) will result in a series of digital images corresponding to each time point, fluorescence channel, and microscope stage position. Software such as ImageJ, MATLAB, or packages from the microscope manufacturer can be used to view each recorded image series. Processing of the raw images using Basic Protocol 3 will result in a dataset of 10s to 1,000s of individual cell "tracks" per stage position, with each track consisting of time-course measurements of all measured reporters for an individual cell. Heatmap visualizations are often helpful in quickly obtaining an overview of the full dataset and identifying major trends within each reporter (Fig. 1). It is also helpful to display a line plot of the mean or median for each of the measured reporters, along with lines corresponding to 25^{th} and 75^{th} percentile values for the population; overlaying line plots of individual cells on top of the mean can also be of use in determining the extent to which individual cells depart from the average population behavior. Visualizing the relationship between reporter measurements within each cell is also critical but can be more tedious, as it is difficult to meaningfully display more than one cell's multiparameter data on a single 2-dimensional axis, and it is often necessary to adjust the axes limits or normalization to obtain the best visualization for each cell. An appropriate strategy is to combine visual inspection of at least 20 individual cells with the multiparameter analyses described in Basic Protocol 4 for each condition.

The paired signal analyses described in Basic Protocol 4 will result in scalar values that allow for quantitation of the relationship between the reporters analyzed. Correlation analysis, for example, will produce a single value between 0 and 1 for each cell analyzed, with higher numbers indicating increasing similarity between the reporter signals. Cross-correlation analysis will produce two values for each cell, one indicating the lag time between signals and one indicating correlation value (between 0 and 1). Mutual information will also produce a single positive value for each cell, with higher values indicating an increased level of dependence between the reporter signals analyzed. In the case of signal response modeling, each dataset modeled will produce a series of coefficients that indicate the weight with which each considered parameter contributes to the overall response, with negative values indicating an opposing influence and positive values a supporting influence. Interpretation of the biological significance of these values will depend on the particular experimental context. It may be

appropriate to compare, for example, the average cross correlation for all cells in one condition in which two reporters are expected to behave similarly (such as when both pathways are induced by the same stimulus) to the same measurement for all cells in a condition in which the reporters are expected to behave differently (such as when one pathway is stimulated and the other is blocked by an inhibitor). Altenatively, it may be useful to compare the calculated values for one pair of reporters to similar measures in a different pair of reporters, such as a control fluorescent proteins that are not expected to vary specifically in response to similar stimuli. In the absence of either different conditions or different reporters that can serve as a point of reference, reporter signals from each cell in the dataset can be "scrambled" such that reporter signals are mismatched; this calculation can be used as a point of comparison to determine whether correlations in the original dataset are meaningful.

Independent replicates of experiments are also critical for establishing reproducibility and statistical significance for comparisons between conditions. It is not recommended to perform a standard t-test between conditions using each cell as an independent measurement, as the large number of cellular measurements collected will result in very low p values and statistically "significant" differences even between replicates of the same condition. It is more appropriate to perform a t-test on measurements of the average correlation (or other metrics) value for each conditions performed as experimental replicates on different days.

Time Considerations

Several aspects of establishing successful live-cell experiments require long-term time investment. The most time-consuming operations are in producing the necessary DNA constructs and cell lines, which can require weeks to months. Naturally, establishing the computational pipeline for processing image data may also require significant time and troubleshooting, and should be started early. Once in place, image processing should consume only hours to days, depending on processor speeds and availability of parallel computing (via desktop machines or clusters).

Some of the experimental operations can be performed in parallel to avoid unnecessary delays, especially the preparation of different DNA constructs, and the (co-)transfection of cells for reporter insertion, when feasible. However, for cell lines or techniques with low insertion rates (such as targeted genomic insertions via CRISPR), transfection is more efficiently done in series, selecting for positive cells between insertions. Establishing a clonal population of reporter-bearing cells can require an additional month; evaluate the trade-off between the extra time invested and having uniform genotype and reporter expression. In practice, it is productive to begin performing and refining imaging experiments as soon as cells bearing a reporter are available, rather than waiting for the perfect cell line to be complete.

Acknowledgments

Funding for this work was provided by the American Cancer Society (IRG-95-125-13), the National Institute of General Medical Sciences (1R01GM115650), the Department of Defense Neurofibromatosis Research Program (W81XWH-16-1-0085) and the American Association for Cancer Research Stand Up To Cancer (SU2C-AACR-IRG-01-16). Stand Up To Cancer is a program of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the scientific partner of SU2C. Flow cytometry services were supported by the UC Davis Comprehensive

Cancer Center Support Grant (CCSG) awarded by the National Cancer Institute (NCI P30CA093373), and we acknowledge the expert cell sorting assistance of Dr. Bridget McLaughlin and Jonathan Van Dyke.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Literature Cited

Albeck, J.G., Burke, J.M., Spencer, S.L., Lauffenburger, D.A., and Sorger, P.K. (2008). Modeling a snap-action, variable-delay switch controlling extrinsic cell death. PLoS Biol *6*, 2831-2852.
Albeck, J.G., Mills, G.B., and Brugge, J.S. (2013). Frequency-modulated pulses of ERK activity transmit quantitative proliferation signals. Mol Cell *49*, 249-261.
Batchelor, E., Loewer, A., Mock, C., and Lahav, G. (2011). Stimulus-dependent dynamics of p53 in single cells. Mol Syst Biol *7*, 488.
Batchelor, E., Mock, C.S., Bhan, I., Loewer, A., and Lahav, G. (2008). Recurrent initiation: a mechanism for triggering p53 pulses in response to DNA damage. Mol Cell *30*, 277-289.
Burnham, K.P., and Anderson, D.R. (2003). Model selection and multimodel inference: a practical information-theoretic approach (Springer Science & Business Media).
Cadinanos, J., and Bradley, A. (2007). Generation of an inducible and optimized piggyBac transposon system. Nucleic Acids Res *35*, e87.
Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J.*, et al.* (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol *7*, R100.
Cepko, C., and Pear, W. (2001). Overview of the retrovirus transduction system. Curr Protoc Mol Biol *Chapter 9*, Unit9 9.
Cheong, R., Rhee, A., Wang, C.J., Nemenman, I., and Levchenko, A. (2011). Information transduction capacity of noisy biochemical signaling networks. Science *334*, 354-358.
Chubb, J.R., Trcek, T., Shenoy, S.M., and Singer, R.H. (2006). Transcriptional pulsing of a developmental gene. Curr Biol *16*, 1018-1025.
Dickinson, D.J., Ward, J.D., Reiner, D.J., and Goldstein, B. (2013). Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. Nat Methods *10*, 1028-1034.
Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science *297*, 1183-1186.
Geurts, A.M., Yang, Y., Clark, K.J., Liu, G., Cui, Z., Dupuy, A.J., Bell, J.B., Largaespada, D.A., and Hackett, P.B. (2003). Gene transfer into genomes of human cells by the sleeping beauty transposon system. Mol Ther *8*, 108-117.
Goedhart, J., van Weeren, L., Adjobo-Hermans, M.J., Elzenaar, I., Hink, M.A., and Gadella, T.W., Jr. (2011). Quantitative co-expression of proteins at the single cell level--application to a multimeric FRET sensor. PLoS One *6*, e27321.
Han, K., Jaimovich, A., Dey, G., Ruggero, D., Meyuhas, O., Sonenberg, N., and Meyer, T. (2014). Parallel measurement of dynamic changes in translation rates in single cells. Nat Methods *11*, 86-93.
Jaqaman, K., Loerke, D., Mettlen, M., Kuwata, H., Grinstein, S., Schmid, S.L., and Danuser, G. (2008). Robust single-particle tracking in live-cell time-lapse sequences. Nat Methods *5*, 695-702.
Komatsu, N., Aoki, K., Yamada, M., Yukinaga, H., Fujita, Y., Kamioka, Y., and Matsuda, M. (2011). Development of an optimized backbone of FRET biosensors for kinases and GTPases. Mol Biol Cell *22*, 4647-4656.
Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A.J., Elowitz, M.B., and Alon, U. (2004). Dynamics of the p53-Mdm2 feedback loop in individual cells. Nat Genet *36*, 147-150.

Locke, J.C., and Elowitz, M.B. (2009). Using movies to analyse gene circuit dynamics in single cells. Nat Rev Microbiol *7*, 383-392.
Loewer, A., Karanam, K., Mock, C., and Lahav, G. (2013). The p53 response in single cells is linearly correlated to the number of DNA breaks without a distinct threshold. BMC Biol *11*, 114.
Nelson, D.E., Ihekwaba, A.E., Elliott, M., Johnson, J.R., Gibney, C.A., Foreman, B.E., Nelson, G., See, V., Horton, C.A., Spiller, D.G.*, et al.* (2004). Oscillations in NF-kappaB signaling control the dynamics of gene expression. Science *306*, 704-708.
Nelson, G., Paraoan, L., Spiller, D.G., Wilde, G.J., Browne, M.A., Djali, P.K., Unitt, J.F., Sullivan, E., Floettmann, E., and White, M.R. (2002). Multi-parameter analysis of the kinetics of NF-kappaB signalling and transcription in single living cells. J Cell Sci *115*, 1137-1148.
Pargett, M., Gillies, T.E., Teragawa, C.K., Sparta, B., and Albeck, J.G. Single-cell imaging of ERK signaling using fluorescent biosensors. Methods in Molecular Biology *in press*.
Purvis, J.E., Karhohs, K.W., Mock, C., Batchelor, E., Loewer, A., and Lahav, G. (2012). p53 dynamics control cell fate. Science *336*, 1440-1444.
Purvis, J.E., and Lahav, G. (2013). Encoding and decoding cellular information through signaling dynamics. Cell *152*, 945-956.
Qin, J.Y., Zhang, L., Clift, K.L., Hulur, I., Xiang, A.P., Ren, B.Z., and Lahn, B.T. (2010). Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. PLoS One *5*, e10611.
Raser, J.M., and O'Shea, E.K. (2005). Noise in gene expression: origins, consequences, and control. Science *309*, 2010-2013.
Regot, S., Hughey, J.J., Bajar, B.T., Carrasco, S., and Covert, M.W. (2014). High-sensitivity measurements of multiple kinase activities in live single cells. Cell *157*, 1724-1734.
Selimkhanov, J., Taylor, B., Yao, J., Pilko, A., Albeck, J., Hoffmann, A., Tsimring, L., and Wollman, R. (2014). Systems biology. Accurate information transmission through dynamic biochemical signaling networks. Science *346*, 1370-1373.
Shankaran, H., Ippolito, D.L., Chrisler, W.B., Resat, H., Bollinger, N., Opresko, L.K., and Wiley, H.S. (2009). Rapid and sustained nuclear-cytoplasmic ERK oscillations induced by epidermal growth factor. Mol Syst Biol *5*, 332.
Slattery, S.D., and Hahn, K.M. (2014). A High-Content Assay for Biosensor Validation and for Examining Stimuli that Affect Biosensor Activity. Curr Protoc Cell Biol *65*, 14 15 11-31.

Internet Resources (annotated)

www.openmicroscopy.org/site/products/bio-formats

Website for Bio-Formats

www.utsouthwestern.edu/labs/danuser/software

Website hosting u-Track software

www.fpvis.org

Website hosting aggregated fluorescent protein properties

Figure Legends:

Figure 1. Schematic example figure showing the construction of a cell line with multiple reporters, live-cell imaging, and data processing.

Figure 2. A. Absorption (upper) and emission (lower) spectra of an example fluorescent protein combination suitable for multiplexed imaging, comprising a CFP (mTurquoise2, in cyan), a YFP (mVenus, in yellow) and a RFP (mCherry, in orange). Appropriate filters are indicated by dashed lines, color-coded according the protein they are intended to measure. The relative contribution of each protein through each filter is shown as the shaded area beneath the filter band. The contributions of RFP to YFP excitation and CFP to YFP emission are allowable because neither contributes significantly to both filters. B. Example filter selection for YFP, by plotting the interference ratio of the other fluorophores (RFP in orange, CFP in cyan). The ratio is defined as the interfering spectrum divided by the intended spectrum, bounded on 0 to 1 for convenient visualization. Suitable filter regions are shown by dashed black lines, with shaded regions indicating the interference contribution. Filters should be placed where the ratio is minimal, and contribution of any interfering fluorophore through both filters should be strictly avoided.

Figure 3. Examples of paired signal analyses. A. Correlation, computed for each cell, yielding a vector of coefficients. B. Cross-correlation, computed from the correlation of shifted (lagged) traces. For each cell, the peak correlation and corresponding lag are stored, yielding a vector for each. C. Signal response modeling example, using a feature decomposition (left) extracting the: mean value, max value, frequency, amplitudes, derivatives, pulse durations and inter-pulse intervals (IPI). Mean and/or maximum values of the latter four features would be stored (center). Stored features are used to fit model weights (right, e.g. via PLS), so that output values may be estimated by these weights and the input values.

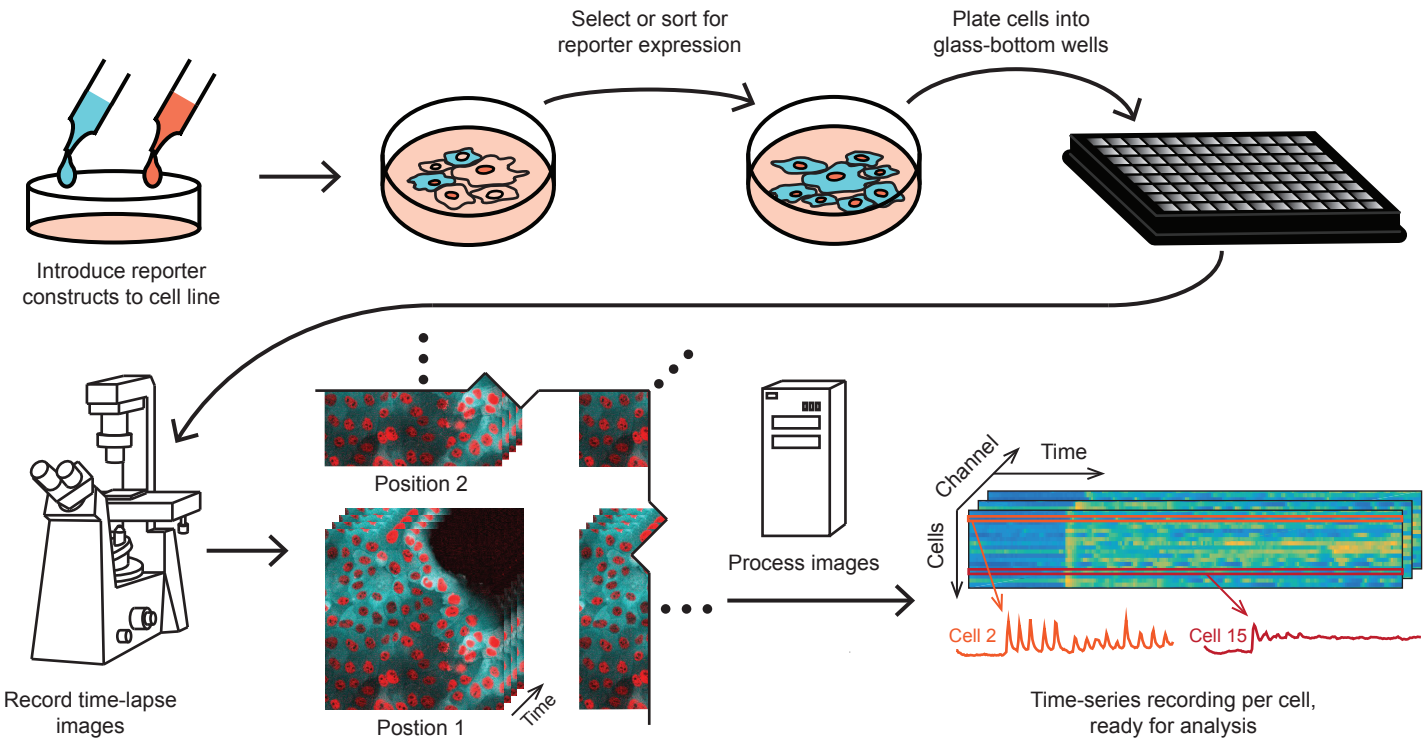
Table 1: Criteria for choosing FPs and reporters.

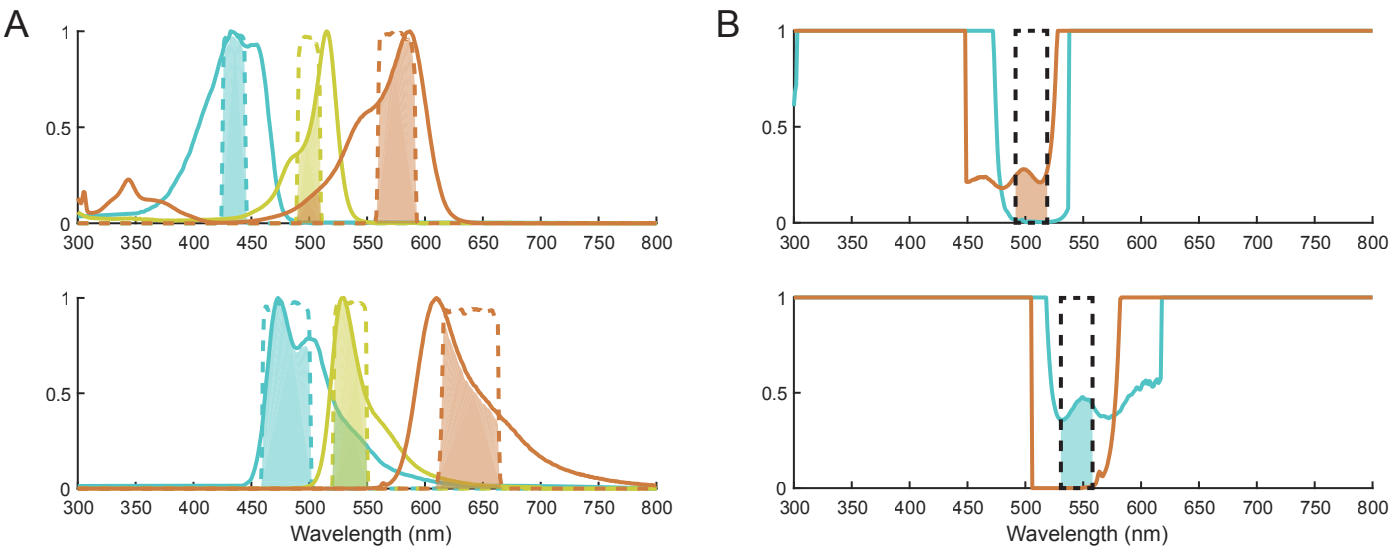
Property	Design preference	Complications
Spectra	Narrow, and distinct, for separability by filters (see Figure 2)	Spectra (excitation or emission) that are particularly broad, have multiple distant peaks, or overlap greatly (for two FPs) can lead to bleed-through noise
Quantum yield (QY)	High, for brightness	When low, FP is inefficient and dim, requiring more intense excitation
Extinction coefficient	High, for brightness	When low, FP absorbs light poorly, requiring more intense excitation
Maturation time	Low, for expression sensors	When high, protein is slow to develop its chromophore, and not suitable for reporting expression levels dynamically
Aggregation	Low, for image quality	When high, imagery is corrupted by puncta and activity may be altered by dimer/multimer formation
Photobleaching time	Long (slow bleaching), for dynamics	When short (rapid bleaching), dynamic measurements may be corrupted
pKa	Low, for selectivity	When near physiological pH (6-7.4), fluorescence can be sensitive to pH changes, esp. in lysosomes

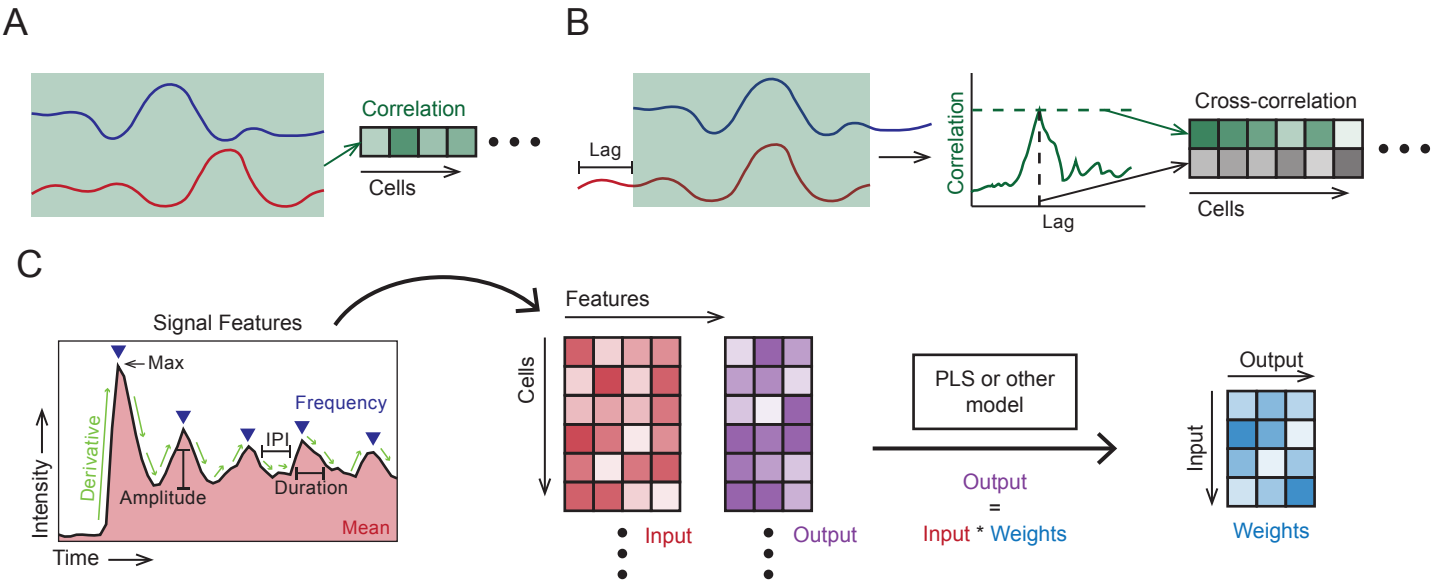
Table 2. Comparison of reporter gene delivery and expression strategies.

Technique	Description	Add'l Requirements	Pros/Cons
Transfection	Insertion of plasmid DNA only; genomic integration occurs at a low frequency	None	↓ Transient (X-Y days)
Transposon	Use a co-transfected transposase to insert the reporter into genome (Cadinanos and Bradley, 2007; Geurts et al., 2003)	Transposase plasmid	↑ Stable integration with high efficiency in transfected cells ↓ Limited by transfection frequency of target cell
Retrovirus	Use retroviral particles to insert the reporter into the genome (Cepko and Pear, 2001)	Packaging cell line Viral packaging plasmids Filtration devices Polybrene	↑ Stable integration ↑ High efficiency in a broad range of target cells ↓ Virus handling
CRISPR	Introduce a fluorescent protein under the control of a genomic promoter of interest. (Dickinson et al., 2013)	Cas9 Homology plasmid	↑ Stable integration ↑ Targeted insertion ↓ Low success rate

Figure 1


Introduce reporter constructs to cell line
Select or sort for reporter expression
Plate cells into glass-bottom wells
Record time-lapse images
Position 2
Postion 1
Time
Process images
Channel
Time
Cells
Cell 2
Cell 15
Time-series recording per cell, ready for analysis


A
B
1
0.5
0
300 350 400 450 500 550 600 650 700 750 800
Wavelength (nm)


A
Correlation
Cells
B
Lag
Correlation
Lag
Cross-correlation
Cells
C
Signal Features
Max
Frequency
IPI
Derivative
Intensity
Amplitude
Duration
Mean
Time
Features
Cells
Input
Output
PLS or other model
Output
=
Input * Weights
Output
Input
Weights