Title
Systems-Level Properties of EGFR-RAS-ERK Signaling Amplify Local Signals to Generate Dynamic Gene Expression Heterogeneity.

Permalink
https://escholarship.org/uc/item/70c9t9s9

Journal
Cell systems, 11(2)

ISSN
2405-4712

Authors
Davies, Alexander E
Pargett, Michael
Siebert, Stefan
et al.

Publication Date
2020-08-01

DOI
10.1016/j.cels.2020.07.004

Peer reviewed
**Highlights**

- Malignant breast cancer cells drive variable ERK activity through AREG secretion
- Non-malignant cells receiving paracrine AREG increase in gene expression variation
- Differential filtering of dynamic ERK activity diversifies gene expression states
- Paracrine signals promote cellular exploration of gene expression space

**In Brief**

This work establishes a fluorescent live-cell reporter system to monitor local communication between malignant and non-malignant breast cancer cells. This system reveals that the EGFR-RAS-ERK signaling network amplifies small ligand fluctuations to generate dynamically heterogeneous gene expression states, which may provide an adaptive advantage for tumor cells.
SUMMARY

Intratumoral heterogeneity is associated with aggressive tumor behavior, therapy resistance, and poor patient outcomes. Such heterogeneity is thought to be dynamic, shifting over periods of minutes to hours in response to signaling inputs from the tumor microenvironment. However, models of this process have been inferred from indirect or post-hoc measurements of cell state, leaving the temporal details of signaling-driven heterogeneity undefined. Here, we developed a live-cell model system in which microenvironment-driven signaling dynamics can be directly observed and linked to variation in gene expression. Our analysis reveals that paracrine signaling between two cell types is sufficient to drive continual diversification of gene expression programs. This diversification emerges from systems-level properties of the EGFR-RAS-ERK signaling cascade, including intracellular amplification of amphiregulin-mediated paracrine signals and differential kinetic filtering by target genes including Fra-1, c-Myc, and Egr1. Our data enable more precise modeling of paracrine-driven transcriptional variation as a generator of gene expression heterogeneity. A record of this paper’s transparent peer review process is included in the Supplemental Information.

INTRODUCTION

Cellular heterogeneity is a prominent feature of many tumors, including breast, colorectal, and brain (Gao et al., 2016; Stingl and Caldas, 2007). While some heterogeneity can be attributed to the genetic mosaicism of tumors, much variation arises non-genetically and involves the ability of cells to reversibly shift their gene expression profiles over time (Gupta et al., 2011). This “dynamic heterogeneity” provides an adaptive advantage for cancer cells, contributing to metastasis and drug resistance (Sharma et al., 2010). Potential drivers of dynamic heterogeneity are multifactorial (Meacham and Morrison, 2013) and include both intrinsic stochastic processes and temporal shifts in extrinsic signaling or adhesive inputs received from the tumor microenvironment (TME) (Friedl and Alexander, 2011; Tam and Weinberg, 2013).

In current models of heterogeneity, each cancer cell receives unique signaling inputs due to spatial variation in the TME; these inputs can then drive differential gene expression programs, which lead to phenotypic heterogeneity (Davies and Albeck, 2018; Cazet et al., 2018; Lu et al., 2014). However, these models are qualitative, and key quantitative and kinetic aspects of this process remain uncharacterized. First, the scope of TME heterogeneity needed to drive variable gene expression is not known. TME signals acting on tumor cells emanate from multiple cell types, extracellular matrices (ECMs), and mechanical forces, and while a paracrine signal is an obvious candidate to drive differences, it is not clear whether it alone can generate widespread transcriptional variation. Second, it is not clear whether signaling pathways act as linear transmitters of heterogeneous TME signals (Nunns and Goentoro, 2018) or instead reshape inputs through amplification or feedback (Tyson et al., 2003). Finally, it is not known how frequently gene expression states fluctuate in response to TME variation, which is essential for understanding the population dynamics of drug-resistant cells and for designing the most effective time course of drug treatments. While in vivo studies can verify the physiological importance of TME-driven heterogeneity, a defined in vitro model is still needed to address these questions quantitatively.

Heterogeneity plays a prominent role in basal-like breast cancer (BLBC), an aggressive malignancy in which cells interchange between multiple states that vary in tumor initiating capacity and...
Figure 1. Progression to Malignancy Is Associated with Stochastic RAS-ERK Signaling Dynamics and Heterogeneous Target Gene Expression

(A) Schematic of the EGFR-RAS-ERK signaling pathway in mammalian cells and ERKTR.
drug resistance (Brooks et al., 2015). Several findings suggest a role for epidermal growth factor receptor (EGFR) signaling via the proto-oncogene RAS and extracellular signal-regulated kinase (ERK) in BLBC heterogeneity. First, genes stimulated by this pathway (ERK target genes, hereafter ETGs) include transcription factors such as c-Myc and Fra-1 that have been implicated as drivers of breast cancer malignancy (Belguise et al., 2005; Berns et al., 1992; Tam et al., 2013). Second, while BLBC tumors do not frequently carry mutations in the RAS cascade, they often overexpress EGFR (Reis-Filho et al., 2006) or show protein phosphorylation profiles consistent with receptor tyrosine kinase activity (Hochgrafe et al., 2010). Third, live-cell reporters have revealed that paracrine signaling generates local, highly dynamic RAS-ERK activation (Albeck et al., 2013; Aoki et al., 2013; Hirasuuk et al., 2015) and that the dynamics of ERK influence the expression of ETGs (Bugaj et al., 2018; Wilson et al., 2017). Together, these findings suggest that paracrine signaling through EGFR-RAS-ERK may drive dynamic ETG expression, resulting in heterogeneous populations of BLBC cells, but this association has not been tested at a mechanistic level.

In this study, we investigated the link between microenvironmental heterogeneity, signaling dynamics, and gene expression heterogeneity in a cell line model of BLBC, HMT-3522 (Rizki et al., 2008). These cells were originally derived from reduction mamoplasty tissue and subjected to multiple rounds of in vitro and in vivo selection for tumorigenic behavior (Madsen et al., 1992). The non-malignant cells, termed S1, were spontaneously immortalized in culture using defined media, give rise to polarized acinar structures, exhibit growth arrest, and do not generate tumors in mouse models. Malignant T4-2 cells were derived from S1 cells through serial passage in the absence of exogenous growth factor, followed by passage through a mouse model to isolate tumor-forming invasive cells. Here, we developed a co-culture model of S1 and T4-2 cells, expressing color-coded live-cell reporters for ERK. This enabled us to track dynamic signaling profiles driven by paracrine signaling and link them to gene expression profiles of individual cells. We show that paracrine signaling between these cell types is sufficient to drive heterogeneity in gene expression similar to that found in BLBC tumors. Furthermore, we find that the pulsatile nature of ERK signaling, coupled to differentially responding target genes, expands the repertoire of transcriptional states and drives changes in ETG expression over time. Our findings validate a model in which EGFR-RAS-ERK signaling amplifies paracrine variation to generate dynamic gene expression heterogeneity.

**RESULTS**

**A Model of Paracrine Signaling-Induced Heterogeneity in EGFR-RAS-ERK Signaling**

To investigate cell-cell signaling in a simplified microenvironment, we utilized the HMT-3522 cell line series. A defining feature of progression from S1 (non-malignant) to T4-2 (malignant) is increased production and secretion of the EGFR ligand amphiregulin (AREG), which promotes proliferation in the absence of exogenous growth factor (Kenny and Bissell, 2007). To monitor ERK activity stimulated via AREG, we generated S1 or T4-2 cells with genetically encoded fluorescent ERK translocation reporters (ERKTRs) (Figure 1A) (Regot et al., 2014; Sparta et al., 2015). These reporters contain a tandem nuclear import and export sequence that is also a substrate for ERK; phosphorylation of the reporter by ERK suppresses shuttling from nucleus to cytoplasm. Cytoplasmic shuttling is opposed by phosphorylation-mediated dephosphorylation of the reporter (Regot et al., 2014), and ERK kinase activity is, thus, measured as the ratio of cytosolic (C) to nuclear (N) fluorescence (hereafter ERKTRCN).

The different versions of the reporter are functionally equivalent, varying only in color of the fluorescent protein, facilitating identification of individual cell types in our experiments.

We measured ERK activity under: (1) baseline imaging medium lacking growth factor, (2) treatment with MEK inhibitor, PD0325901 (MEKi), or (3) stimulation with EGF. Under baseline conditions, S1 cells exhibited few fluctuations in ERKTRCN (Figure 1B; Video S1A), with infrequent low amplitude pulses. Addition of MEKi had no effect on ERKTRCN, indicating that ERK activity in S1 cells was below the detectable limit for the reporter (Figure 1C). At low EGF concentrations (0.2–2 ng/mL) we observed ERKTRCN responses ranging from undetectable to low amplitude pulses (Figure S1), while addition of 20 ng/mL EGF resulted in sustained activation of ERK signaling (Figures 1D and S1; Video S1A). These results indicate that S1 cells depend on growth factor stimulation for ERK activation, similar to other non-malignant mammary epithelial cells (Gilles et al., 2017).

In contrast to S1, T4-2-ERKTR cells exhibited elevated and variable ERKTR signals under baseline conditions (Figure 1E; Video S1B), which were reduced by MEKi (Figure 1F; Video
AEG provoking a maximal ERK response similar in magnitude to stimulated S1 cells; however, the amplitude of this response relative to baseline was much smaller than in S1 due to the pre-existing high level of ERK activity (Figures 1G and 1H; Video S1B). These results confirm that progression to malignancy in HMT-3522 is associated with elevated baseline ERK signaling, which is stochastic, self-sustaining, and influenced weakly by additional stimuli (Kenny and Bissell, 2007).

To evaluate the effects of paracrine signaling, we co-cultured S1 cells with T4-2 cells at different ratios and measured ERKTRC/N signals (Figures 1I–1K; Videos S2A–S2G). For clarity, we refer to co-cultured cells as S1co and T4-2co, and to monocultures as S1m and T4-2m. In S1co, compared with S1m, we observed increased ERKTRC/N signals that varied over time within each cell (Figure 1I; Videos S2A and S2D). A 30:70 ratio of S1 to T4-2 produced the most reliable activation in S1co, inducing ERKTRC/N signals that were quantitatively similar in pulse frequency, amplitude, and duration to the pulse characteristics of T4-2m cells (Figure 1L). Consistent with paracrine signaling as the source of ERK activity, EGFR inhibitor (EGFRi) blocked ERK signaling in both S1co and T4-2co cells with similar potency to MEKi (Figure 1J; Video S3A). The response of S1co cells to exogenous EGF was dampened relative to S1m cells and similar to T4-2m cells (Figure 1K; Video S3B), in agreement with prior reports that paracrine EGFR-mediated signals produced by T4-2m cells (Figures 2A and S2A–S2D). We assessed global gene expression profiles as a function of cell type and/or conditions (Figure 2B). Clustering analysis of S1m, S1m treated with 100 ng/mL AREG (S1m+AREG), and T4-2m cells found that cells broadly sort according to type and conditions (Figure 2B). Similarly, S1co and T4-2co cells segregated into distinct clusters, even without filtering by mRNA expression of mCherry or mVenus. We found that cells of clusters with unique but partially overlapping signatures (Figures 2C and S2E; Data S1). T4-2m and T4-2co cells were each comprised of 3 clusters enriched for genes involved in replication and cell division, as well as growth factors, reactive oxygen enzyme SOD2, and TNF-related apoptosis-inducing ligand (TNFSF10) (Figures 2C and S2E). S1m cells treated with growth factor segregated into two clusters: the bulk population of cells (cluster 2) showed broad expression changes, with a notable increase in MAPK-regulated genes compared with unstimulated controls (Figures 2C and S2E), while the second cluster (cluster 8) showed increased expression of cell-cycle-associated genes. S1co cells formed a separate cluster (cluster 9) with an expression profile that partially overlapped with S1m and T4-2m cells but also included a diverse set of genes distinct from other conditions, including inflammation-associated genes (Figure S2E; Data S1).

Since distances in tSNE plots can distort the similarity of clusters, we performed additional analysis to determine the relative gene expression similarity between S1m, S1m+AREG, S1co, and T4-2 cells. Using pairwise correlation (Figure 2D) and principal component analysis (Figure 2E) we found that S1m and T4-2m cells exhibit divergent gene expression profiles, whereas S1m+AREG and S1co cells exhibit strong correlation in gene expression profiles. These findings imply that AREG is the predominant factor driving gene expression changes under co-culture conditions and that co-culture alone is not sufficient to fully convert global S1 gene expression to a malignant-like expression profile.

We next compared transcriptional states by quantifying cell-to-cell variance in gene expression. Gene expression variance typically increases with the mean; we removed this bias by calculating, per gene, the "excess dispersion," i.e., the variance exceeding the fitted relationship between mean and variance. In both S1 and T4-2 cells, co-culture conditions promoted an increase in excess dispersion for the majority of genes (69% and 90%, respectively, Figures 2F and 2G, p = 1.7 x 10^-61). However, excess dispersion was more widespread in S1m than in T4-2m cells (Figure 2H), while treatment of S1m cells with AREG reduced excess dispersion in nearly all genes, relative to untreated S1m (Figure 2I). The ordering of variability in S1 cells (S1co > S1m > S1+AREG) is consistent with a model in which exogenous growth factor addition suppresses expression variability, because cells receive uniformly high levels of stimulation, whereas co-culture enhances variability, as cells receive variable levels of stimulation. The observation that excess dispersion is increased between S1m and S1co cells but decreased between S1m and T4-2m cells suggests that paracrine signaling can have a stronger influence on transcriptional variability than progression to a malignant phenotype, which may select for a narrowed gene expression space. The increase in excess dispersion in T4-2co relative to T4-2m cells could reflect the
effects of local regions of high S1 density, which have a suppressive effect on T4-2 signaling (Video S2), potentially by acting as a sink for secreted AREG.

To ask whether the variable ERK dynamics observed in T4-2 cells are associated with greater gene expression variance, we refined our analysis to known ETGs (c-Fos, c-Myc, Fra-1, and Egr1). We found that 60% of ETGs showed increased excess dispersion in T4-2m relative to S1m cells, with a number of genes, including Egr1, falling well outside the bulk distribution (Figure 3A; Data S1). The known classes of ETGs that are sensitive to different patterns of ERK activity (Uhlitz et al., 2017) showed similar increases in excess dispersion (Figure 3B). Thus, the variable ERK activity observed in both T4-2m and S1co correlates with increased variability in expression of ETGs, many of which are transcription factors linked to malignancy. To confirm these results, we measured Fra-1, Egr1, c-Fos, and c-Myc, protein expression by immunofluorescence (Figure 3C). Expression was minimal in S1m, but elevated in T4-2m, with coefficients of variance (CVs) ranging from 50%–100% (Figures 3C and 3D). Similarly, under co-culture conditions, ETG protein levels and heterogeneity increased in S1co cells (Figures 3E and 3F).

We compared ETG heterogeneity driven by paracrine signals in T4-2m cells to models of BLBC driven by RAS mutations, MCF10A-CA1D and MB-MDA-231. ERKTRC/N displayed similar temporal variability in all cell types that was reduced by MEKi (Figure S3A and S3B). Accordingly, Fra-1 expression varied from cell to cell with CVs >45% for all cell types (Figure S3C). As expected, TACEi treatment greatly reduced Fra-1 expression in T4-2 cells but not in the RAS-driven cell types. A combination of TACEi and AREG restored Fra-1 expression in T4-2 cells, but with lower variance (CV = 29%), while the same treatment did not significantly alter Fra-1 variance in the other cell lines (Figure S3A). Thus, RAS- and paracrine-driven ERK signaling can both promote variance in ETG expression but are distinct in their dependence on extracellular ligands.

Surprisingly, despite their co-regulation by ERK, ETG mRNAs showed discordant expression in single cells, particularly between FOSL1 (Fra-1) and EGR1 (R = 0.043) and between FOSL1 and FOS (R = 0.026) (Figure 3G). Co-staining of Fra-1
Figure 3. Induction of ETG Expression Variance by Cellular Crosstalk

(A) Comparison of variance in ETG transcript levels between S1 and T4-2 cells. Red dots indicate genes that display excess dispersion in T4-2 cells compared to S1.

(B) The same graph in (A) is color coded by known ETG expression classes: immediate-early genes (IEG, light blue), immediate-late genes (ILG, purple), and delayed-early genes (DEG, orange). Insets depict each of the classes alone. Annotated as in Figures 2F–2I.

(C) Immunofluorescence of ETGs Fra-1, Egr1, c-Fos, and c-Myc (red) and nuclei (blue) in S1 and T4-2 cells in the absence of exogenous growth factors. Scale bar, 20 μm.

(D) Histograms depicting the fluorescence intensity distribution of ETG proteins under baseline conditions in T4-2 cells. n > 4,000 cells per ETG measured.

(E) S1m (turquoise) and S1co/T4-2co (turquoise and yellow, respectively) cells stained for ETG expression (red) under control or MEK inhibitor conditions with Fra-1 and Egr1 protein expression distributions plotted as histograms. Control treated (black), MEK-treated (pale red), and overlapping data (dark red) (n = 25,169). Scale bar, 20 μm.

(F) Box and whisker plots comparing mean Fra-1 and Egr1 protein levels between S1m cells co-cultured cells (n = 30,370). *p < 0.01.

(G) Pairwise single-cell gene expression correlation for the indicated ETGs. Red numbers indicate the correlation coefficient (Pearson correction).

(H) Co-staining of Fra-1 and Egr1 or Fra-1 and c-Fos in mono-cultured T4-2 cells under baseline conditions. Scale bar, 20 μm.

(I) Mono-cultured T4-2 cells stained for Fra-1 or Egr1 under baseline conditions or treatment with MEKI. Scale bar, 20 μm.

(J) Formalin-fixed resected human invasive ductal carcinoma samples stained for Fra-1 or Egr1.
Figure 4. Paracrine AREG and Gene-Specific Expression Kinetics Coordinate Induce Heterogeneous ETG Expression

(A and B) Time-dependent expression of Fra-1 (red) and Egr1 (green) in S1m or T4-2m cells following exposure to 100 ng/mL AREG or 4 μM erlotinib for the indicated times. Scale bar, 20 μm.

(C) S1m and T4-2m scatter plots of single-cell Fra-1 or Egr1 signal intensities and correlations (red lines). n > 5,900 cells per condition.

(D) T4-2 single-cell gene expression analysis of preprocessed (normalized and scaled) single-cell RNA sequencing data partitioned as a function of single-cell FOSL1 and EGR1 expression.

(legend continued on next page)
and Egr1, or Fra-1 and c-Fos proteins in T4-2m cells further corroborated their lack of correlation (Figure 3H). While this divergence could be explained by decoupling of these genes from ERK signaling in malignant cells, this possibility was ruled out by treatment of T4-2m cells with MEKi, which suppressed both Fra-1 and Egr1 expression (2.3-fold and 2.5-fold) (Figure S3). We confirmed the relevance of these findings by staining human breast cancer tissues, finding that both Fra-1 and Egr1 show similar degrees of heterogeneity in patient samples (Figure 3J). Thus, co-culture of S1 and T4-2 cells recapitulates the heterogeneity in gene expression previously observed in BLBC (Nguyen et al., 2016) and identifies new features of this heterogeneity.

Dynamic Properties of ERK-ETG Signaling Enhance Heterogeneity in Target Gene Expression
Paracrine signaling reduces cell-to-cell variability in many systems (Rand et al., 2012; Shalek et al., 2014; Handly et al., 2015) making it unclear why paracrine AREG signaling enhances ETG variability in S1co and T4m cells. One model consistent with discordant ETG expression is that, as ERK activity changes over time (Figures 1G and 1K), distinct induction and turnover parameters (Amit et al., 2007; Uhlitz et al., 2017) favor expression of each ETG at different times within individual cells. We, therefore, examined the temporal response of these genes in the absence of paracrine signaling using S1m cells, which displayed low or undetectable Fra-1 and Egr1 in baseline conditions (Figure 4A) compared with T4-2m (Figure 4B). Fra-1 and Egr1 increased in a correlated manner at 1.5 h after AREG stimulation (R^2 = 0.69, Figures 4A–4C). However, with extended stimulation, Fra-1 continued to increase while Egr1 levels peaked and then decreased. This loss of correlation at the cellular level (R^2 = 0.44, Figures 4A–4C) was similar to expression patterns observed in T4-2m cells under control and stimulated conditions (R^2 = 0.46 and 0.44, respectively; Figures 4B and 4C). Similar patterns of variance were also observed for c-Fos and c-Myc (Figure S4A), indicating that this behavior is shared among other ETGs. Thus, decoupled expression due to distinct induction and turnover kinetics can account for the low correlation of ETGs in T4-2m cells (Figures 3G and 3H), which experience asynchronous paracrine-mediated ERK activation.

To explore the functional significance of different single-cell ETG expression states, we correlated single-cell ETG expression with cell states using the MSigDB (Liberzon et al., 2015) Hallmark sets and the NCI pathway interaction database (PIDD) (Figures 4D–4F and S5; Data S2). T4-2 cells expressing high FOSL1 and EGR1 contained EMT and MYC target signatures and increased expression of ERBB, MAPK, and AP1 pathways, consistent with high ERK activity (Figure 4E) compared with FOSL1-low/EGR1-low cells. Additionally, FOSL1-high/EGR1-high cells display increased TGF-β signaling pathway expression relative to FOSL1-high/EGR1-low cells and higher expression of cell division-associated pathways relative to FOSL1-low/EGR1-high cells (Figure 4F). These findings suggest that heterogeneity in Fra-1 and Egr1 gene and protein expression within single cells is correlated with functional differences in cell behaviors.

To evaluate the significance of ETG expression heterogeneity in the context of chemotherapeutic response, we examined the DNA damage response to carboplatin, a cytotoxic agent commonly used in BLBC. Under baseline conditions we found carboplatin induced minimal levels of the DNA damage marker p-γH2AX in non-malignant MCF10A cells, which have low levels of Fra-1 and Egr1 (Figure 4G), whereas T4-2 cells responded with a larger increase (Figure 4H). Both MCF10A and T4-2 cells showed a greatly increased p-γH2AX response in the presence of EGF, which was attenuated after treatment with EGFRi. At the single-cell level, p-γH2AX was heterogeneous, leading us to ask whether the ETG expression status of each cell correlated with its DNA damage response. We used partial least squares regression analysis (PLSR) to model each cell’s p-γH2AX response as a function of its ETG protein abundance (Figure 4I–4J). At the single-cell level, ETG abundance was highly predictive of p-γH2AX, more so than treatment with growth factor (GF) and carboplatin, as indicated by the relative model weights. In particular, Fra-1 staining was highly correlated with p-γH2AX, whereas Egr1 staining was only weakly correlated (Figure S4). FOSL1 expression is correlated with upregulation of the Hallmark DNA repair pathway set (Data S2), further strengthening its correlation with DNA damage response characteristics. Although a direct causal relationship between Fra-1 and DNA damage response cannot be strictly assigned based on the aggregate data, there is a strong correlation, indicating a potential link between cell-to-cell variation in ETG expression and heterogeneous responses to DNA-damaging chemotherapy.

Ligand-Specific ERK Dynamics Modulate ETG Expression Heterogeneity
Because different EGFR ligands induce distinct ERK activation dynamics (Nakakuki et al., 2010; Sparta et al., 2015), we examined how these ligands impact the heterogeneous expression of ETGs, comparing ERK activity kinetics induced by AREG or EGF in S1 cells. At saturating EGF (20 ng/mL), >80% of cells responded with sustained ERK activation (Figures 5A and 5C). At lower levels of EGF (2 ng/mL), only a fraction of S1 cells...
responded (~20%, Figures 5A and 5C), showing intermittent pulses of ERK activity. By comparison, 10 ng/mL AREG (functionally equivalent to 2 ng/mL EGF; Harris et al., 2003; Macdonald-Obermann and Pike, 2014) resulted in activation of ~80% of S1 cells (Figures 5B and 5C). However, ERK activity decreased within 2 h of AREG addition and occurred subsequently as stochastic pulses of activity (Figure 5C), but produced a maximal initial response with continual pulsatile activity (Figure 5B), unlike the sustained activity produced by 20 ng/mL EGF (Figure 5A). At an even higher level of AREG (1,000 ng/mL), essentially all cells responded with a high level of ERKTRC/N that was sustained for hours without pulsing (Figures 5B and 5C). We measured the AREG secreted into the media by T4-2 cells to be 18.7 ng/mL at equilibrium, falling within the range of tested conditions, and we found that AREG protein levels varied between individual T4-2 cells in a population (Figure 5E). Together, these observations explain the high degree of cell-to-cell ERK signaling variability in T4-2m and S1co cells, as secreted AREG is both spatially variable and prone to drive pulsatile ERK responses.

We next tested whether the different patterns of ERK activity induced by EGF and AREG impact the expression of ETGs. At an early time point (1.5 h) following addition of growth factor,
Egr1 and Fra-1 levels increased uniformly as a function of ligand concentration (Figure 5F), regardless of the stimulating ligand. By contrast, at a later time point (24 h), ETG expression heterogeneity increased (Figure 5G), and the relationships of Fra-1 and Egr1 to ligand concentration diverged. Fra-1 expression increased monotonically with the concentration of either ligand, while Egr1 expression peaked at intermediate ligand concentrations and decreased for ligand concentrations capable of stimulating a sustained ERK response. These results indicate that pulsed versus sustained ERK activity induce differential responses in Fra-1, which integrates ERK intensity over time, and Egr1, which responds transiently and is insensitive to sustained activity.

Paracrine EGFR Signaling Drives Changes in Transcriptional State on the Scale of Hours

Our results imply temporal variation in ERK-driven expression in each cell. To explore how gene expression may vary over time,
Figure 7. Live-Cell Imaging of Paracrine-Induced Fra-1 Temporal Expression Variability in Co-culture

(A) Box plots depicting the mean FRET ratio signal in MCF10A cells carrying the ERK reporter, EKAR, in mono-culture (n = 2,428) and in co-culture with T4-2 at the indicated ratios (70:30 n = 2,526, 50:50 n = 2,014, 30:70 n = 998). Boxes indicate IQR and whiskers the range. *p < 0.01 by unpaired t test.

(B) Still frames of MCF10A EKAR Fra1::mCherry cells in mono- and co-culture. Upper panels depict EKAR expression (turquoise) and tagged endogenous Fra-1 (red); T4-2 cells are not shown in co-culture images. Middle panels show Fra-1::mCherry expression only (red). Lower panels show EKAR expressing T4-2 cells (yellow). Scale bar, 20 μm.

(C) Top plot, mean Fra-1 intensity over time for each ratio of T4-2 co-culture (MCF10A [mono-cultured], 70:30, 50:50, and 30:70 MCF10A to T4-2 ratios) as labeled on the graph. Below, plots depict the same mean traces with variance (shaded, IQR) and CV (numbers upper right).

(D) Box plots depicting the mean Fra-1 expression levels under the indicated conditions. Boxes indicate the IQR and whiskers the range. MCF-10A (n = 157), 70:30 (n = 182), 50:50 (n = 160), 30:70 (n = 66). *p < 0.01 by unpaired t test.

(E) Mean Fra-1::mCherry expression in mono-cultured, AREG stimulated MCF10A cells. Plot depicts the mean traces with variance (shaded, IQR) and CV.

(F) Fra-1::mCherry expression in mono-cultured, unstimulated MCF10A cells (purple, re-plotted from C), with exogenous AREG at 20 ng/mL (dark gray), and 100 nM MEKi (light gray).

(G) Fra-1::mCherry expression in MCF10A cells co-cultured at a 30:70 ratio of MCF10A to T4-2 cells (orange, re-plotted from C), in comparison with MCF10A cells at the same co-culture ratio treated with 100 ng/mL AREG (dark gray), or 100 nM MEKi (light gray). See Video S4 for live-cell examples. For panels (A–G), mono-cultured MCF10A n > 500 cells per condition. Co-cultured MCF10A 70:30 N > 622 cells per condition.
we constructed a mathematical model extending our previous work (Gillies et al., 2017). For any given ERK activity time series, the model estimates target protein expression over time, based on mRNA and protein decay rates, as well as parameters for ERK-induced production, stabilization, and negative feedback (Figure 6A). Based on their observed kinetics, Fra-1 and Egr1 represent biological examples near the extreme behaviors expected from this model. The slow kinetics and relatively low heterogeneity of Fra-1 place it conceptually as a low-pass filter with a large time constant, which reflects average recent ERK activity.

In contrast, the rapid response of Egr1 and its apparent negative feedback inhibition under long-term stimulation place it as a high-pass filter. We constructed a model to represent these ETG archetypes, selecting example parameters based on previously published kinetic data for mRNA and protein expression of Fra-1 and Egr1 (Gillies et al., 2017; Schwahnäusser et al., 2011; Uhlitz et al., 2017). We simulated responses to artificial square wave inputs (Figures 6B and 6C) and to single-cell ERK traces measured from S1 and T4-2 cells (Figures 6D–6I). Consistent with our findings in fixed S1m and T4-2m cells, modeled expression levels of Fra-1 change slowly, reflecting average recent ERK activity regardless of rapid fluctuations. Conversely, simulated Egr1 preferentially rises when ERK activity pulses (Figures 6E, 6F, 6H, and 6I) but remains suppressed in both inactive (Figure 6D) and highly active (Figure 6G) conditions.

Using these simulations, we plotted the ETG expression space that a single cell occupies, driven by the stochastic ERK conditions observed in our experiments (Figures 6B–6I, insets). In both T4-2 and S1 cells, our model demonstrates that stochastic fluctuations in ERK induce larger excursions in ETG expression state over time (Figures 6E and 6H), compared with sustained ERK activity (Figures 6F and 6G). This model demonstration is consistent with our experimental findings that pulses in ERK activity lead to rapid changes in ETG expression (<30 min) followed by decay in protein levels (Figure 6J). Thus indicating that paracrine-induced expression is not persistent and instead highly sensitive to temporal changes in signaling. These findings illustrate how pulses in a single microenvironmental signal can enable cells to take on a range of expression states, as represented by two differently responding genes.

To directly observe the ETG fluctuations predicted by the model, we used MCF10A cells carrying an mCherry fusion at the endogenous FOSL1 locus (Fra-1::mCherry) and the ERK sensor EKAR3 (Gillies et al., 2017; Harvey et al., 2008). Thus, ERK activity and the expression of Fra-1 protein can be visualized in a cell line that acts as a “receiver,” similar to the “sender,” for paracrine signals from T4-2 cells (Figure 7A; Videos S4A and S4B). The mean and variance of Fra-1::mCherry expression under each condition were consistent with the level of EKAR activity (Figures 7A–7D). While 20 ng/mL AREG treatment increased the mean Fra-1 intensity in MCF10A cells by 24% (Figure 7E), relative variation remained essentially constant (CV = 27% untreated, 28% AREG-treated, Figures 7E and 7C), which is consistent with the sustained ERK response to AREG in MCF10A (Gillies et al., 2017). However, in co-cultured cells, mean Fra-1::mCherry intensity was similar to the maximum induced by AREG (Figures 7F and 7G) but displayed a higher degree of variation (Figures 7C and 7D, CV = 35% at 30:70 ratio). At the single-cell level, temporal fluctuations in Fra-1::mCherry expression were evident under co-culture conditions, in comparison to mono-cultured cells (Figure 7H). To exclude high frequency noise, we filtered our data to focus on changes consistent with expression regulation (>30 min). We then calculated the mean derivative, normalized by the mean Fra-1::mCherry intensity on a single-cell basis, to measure variance of Fra-1::mCherry expression over time (Fra-1 temporal variability index) in mono- versus co-culture. By this measure, time-dependent variability in Fra-1::mCherry increased by 20% in co-culture relative to mono-culture (Figure 7I, p = 1.06e−21). Thus, consistent with model, Fra-1 expression varies dynamically in individual cells over the scale of hours, resulting in the observed heterogeneity in Fra-1 expression.

**DISCUSSION**

By exploring diverse cellular phenotypes, tumor cells gain a selective survival advantage in adverse physiological environments (Michor and Polyak, 2010). Genetic mutations are important drivers of tumor cell diversification but are constrained to operate on timescales of weeks or longer. Immediate cell survival in response to the time-varying stress conditions within a tumor may require a more rapid and flexible means of diversification.

Our analysis demonstrates how such variation can be generated by the EGFR-RAS-ERK signaling pathway at levels equivalent to those observed in human breast cancer tissues. This expression variance results from a specific mode of operation in which heterogeneity in the TME is amplified by intracellular signal processing in the EGFR-RAS-ERK pathway to drive fluctuation of cancer related genes on the timescale of hours (Figure 7J). While this system is simplified relative to actual tumors, these data establish a potentially important source of variation in the expression of key tumor genes, driven by spatial variation within the TME.

The mixture of malignant (T4-2) and non-malignant (S1 or MCF10A) cells within our co-culture model represents an approximation of the cellular and genetic diversity within a tumor, where clones containing different mutational burdens would interact. Although the arrangement of the cells in vitro does not specifically replicate the cellular organization of a tumor, we expect AREG gradients to occur on similar length scales in vivo because tumor cells are likely to express AREG heterogeneously, as we observe for T4-2 cells. The proficiency of T4-2 cells to self-stimulate through AREG secretion is consistent with previous work (Fischer et al., 2003; Madsen et al., 1992), and we find that this mode of RAS-ERK signaling induces both a high magnitude of ERK activity as well as time-dependent variation that propagates to the level of gene expression.
Co-cultured non-malignant cells respond with similar characteristics of RAS-ERK activity and ETG expression, even in the absence of activating mutations in the RAS-ERK pathway. Within our experimental system, the low level of constitutive ERK signaling and high capacity for stimulation of these cells makes them ideal receivers to quantify the paracrine stimuli from the nearby malignant cells. An intriguing but unresolved question is whether, within the context of a real tumor, such receiver cells represent a population that is important for tumor progression. In principle, such cells, supported by paracrine signals from more highly mutated cells nearby, could represent a reservoir of uncommitted possibilities with higher adaptive potential than heavily mutated cells that cannot tolerate further genetic alterations. Such flexibility could play a role in tumor cell survival when new stresses, such as cytotoxic chemotherapeutics, are encountered.

The general process we demonstrate, by which spatial fluctuations in paracrine ligands are amplified and converted to temporal variability in gene expression, may also apply in other signaling pathways, such as IL-6 (Hartman et al., 2013; He et al., 2013) or TGF-β (Oft et al., 1996; Wang et al., 2014). However, the kinetic parameters of the AREG-EGFR-RAS-ERK pathway appear especially well suited for generating global diversity in gene expression. The amplification characteristics of the EGFR pathway are capable of driving high-intensity ERK activity bursts in response to even sub-saturating concentrations of ligand, and the attenuation of these pulses by negative feedback is rapid, resulting in ERK output that magnifies external changes in AREG concentration. Subsequently, the genes responding to ERK activity do so with a wide range of kinetic properties (Uhlitz et al., 2017), continuously translating the temporal dynamics of the pathway into different gene expression profiles. It is intriguing to speculate that the prevalence and potency of the RAS-ERK pathway in tumorigenesis may be a consequence not only of its target functions (as other, less prevalent pathways, also stimulate cell proliferation and migration) but also of its ability to act as a generator of cellular diversity. It is also possible that the capacity to diversify gene expression profiles may be an important aspect of the function of RAS-ERK signaling within the developmental processes in which it acts. In vivo data indicate that transient gradients of EGFR ligand signaling on the scale of 50 μm drive pulsatile ERK activity in C. elegans vulval patterning (de la Cova et al., 2017), or in the mouse epidermis (Hiratsuka et al., 2015). These dynamic patterns could create variation in ERK-controlled gene expression similar to that observed here, which could play a role in determining cell fate (Hamilton et al., 2019). A key question yet to be resolved is how the ligand gradients within a tumor differ quantitatively from those in normal developing tissues. Our data suggest that ETGs, such as EGR1, FOSL1, MYC, and FOS, could help to address this question by creating an expression signature that reflects the dynamic status of ERK signaling, which may be useful in identifying EGFR signaling gradients in vivo using transcriptional profiling.

Our system makes it possible to model and quantify EGFR-mediated generation of gene expression variability in vitro. However, it by no means recapitulates the full complexity of tumor physiology. Many additional sources of variability are present in the TME, including other cytokines, metabolites, ions, and ECMs. Furthermore, many genes other than Egr1 and Fra-1 contribute to tumor cell survival and drug response. Nonetheless, this experimental system can be used to further investigate particular mechanisms behind spatial heterogeneity and to evaluate applicable candidate inhibitors for their ability to reduce tumor cell heterogeneity, a property that is likely closely related to their efficacy (Shea et al., 2018), but for which no direct measurement has yet been available. It remains important to determine which forms of gene expression heterogeneity play a functional role in cancer cell spread or survival and which forms are simply an epiphenomenon of deregulated tumor cell states. Our simulations characterize the behavior of cells within a two-gene expression landscape, but overall transcriptional variability occurs in a highly multidimensional space. An essential, though challenging, next step relies on the effort to identify, at the single-cell level, the relationship between variation in these many dimensions and relevant cellular phenotypes. We expect that the data and model presented here will help to guide experiments in which heterogeneity is manipulated in vivo on the appropriate time and length scales.

**Key Changes Prompted by Reviewer Comments**

In response to the reviewer comments, we added additional analysis of single-cell gene expression data and performed analysis of DNA damage responses in relation to ETG expression. We reorganized the presentation of the computational model and reordered our interpretation regarding the correlation between Fra-1 and DNA damage response to indicate that they covary and a direct relationship cannot be established in the current context. We also added a section to the STAR Methods denoting the cell density and conditions used to determine the concentration of free AREG in our experiments by ELISA. For context, the complete transparent peer review record is included within the Supplemental Information.

**STAR METHODS**

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

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Cell Culture and Media
  - Human Tissue Samples
- **METHOD DETAILS**
  - Reporter Line Construction
  - Live-Cell Microscopy and Co-culture Conditions
  - Immunofluorescence Microscopy
  - DNA Damage Response Assay
  - Amphiregulin ELISA
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Imaging, Data Processing, and Statistics and Normalization
  - Dynamic Modeling
  - Single Cell RNA Sequencing and Analysis
  - Sequencing Strategy
SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cels.2020.07.004.

ACKNOWLEDGMENTS

Funding for this work was provided by an American Association for Cancer Research Stand Up To Cancer Innovative Research Grant (SU2C-AACR-IROC-01-16) and by the National Institute for General Medical Sciences (RO1-GM115650) to J.G.A. Stand up to Cancer (SU2C) is a program of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the scientific partner of SU2C. A.E.D. was funded by the California Institute of Regenerative Medicine/Children’s Hospital of Oakland Clinical Research Fellowship and the Ohio State University Comprehensive Cancer Center and the National Institutes of Health under grant number (P30 CA016508). Funding was also provided by NSF CAREER award 1846559 to G.Q. Y.C. was supported by an NIH T32 training grant in Molecular and Cellular Biology grant number T32 GM007377 from NIH NIGMS. A.E.D. and M.J.B. were also funded by the Breast Cancer Research Foundation and the Woodland Fund. We would like to thank the members of the Bisell, Juliano, Quon, and Albeck Labs for their helpful suggestions on the manuscript.

AUTHOR CONTRIBUTIONS

A.E.D., J.G.A., and M.J.B. conceptualized the study, analyzed data, and wrote the manuscript. A.E.D., J.G.A., and A.R.R. performed imaging experiments. S.S. performed single-cell RNA sequencing and analysis with assistance from A.E.D. A.E.D. and T.E.G. performed data processing and statistical analysis. M.P. performed statistical analysis and model generation. Q.G. and Y.C. performed excess variance calculations and statistical analysis on single-cell sequencing data. M.J.B., V.M., and C.J. analyzed data and contributed to writing the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES


He, G., Dhar, D., Nakagawa, H., Font-Burgada, J., Ogata, H., Jiang, Y., Shalapour, S., Seki, E., Yost, S.E., Jepsen, K., et al. (2013). Identification of
liver cancer progenitors whose malignant progression depends on autocrine IL-6 signaling. Cell 155, 384–396.


# STAR★METHODS

## KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-Fra-1, clone C-12</td>
<td>Santa Cruz Biotechnology</td>
<td>Cat#sc28310; RRID: AB_627632</td>
</tr>
<tr>
<td>Anti-Egr-1, clone 15F7</td>
<td>Cell signaling</td>
<td>Cat#4153; RRID: AB_2097038</td>
</tr>
<tr>
<td>Anti-c-Fos, clone 9F6</td>
<td>Cell signaling</td>
<td>Cat#2250; RRID: AB_2247211</td>
</tr>
<tr>
<td>Anti-c-Myc</td>
<td>Cell Signaling</td>
<td>Cat#9402; RRID: AB_2151827</td>
</tr>
<tr>
<td>Anti-pH2AX</td>
<td>Millipore Sigma</td>
<td>Cat#05-636; RRID: AB_2755003</td>
</tr>
<tr>
<td>TGFα function blocking antibody</td>
<td>R&amp;D Systems</td>
<td>Cat#AF-239; RRID: AB_2201779</td>
</tr>
<tr>
<td>Amphiregulin function blocking antibody</td>
<td>R&amp;D Systems</td>
<td>Cat#MAB262; RRID: AB_2060676</td>
</tr>
<tr>
<td><strong>Biological Samples</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human Invasive Ductal Carcinoma</td>
<td>Comparative Human Tissue Network, NCI</td>
<td>De-identified</td>
</tr>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epidermal growth factor</td>
<td>Peprotech</td>
<td>Cat#AF-100-15</td>
</tr>
<tr>
<td>Amphiregulin</td>
<td>Peprotech</td>
<td>Cat# 100-55B</td>
</tr>
<tr>
<td>TAPI-0</td>
<td>Tocris</td>
<td>Cat#5523</td>
</tr>
<tr>
<td>Erlotinib</td>
<td>Selleck Biochemicals</td>
<td>Cat#S1023</td>
</tr>
<tr>
<td>Gefitinib (ZD1839)</td>
<td>Selleck Biochemicals</td>
<td>Cat#S1025</td>
</tr>
<tr>
<td>PD0325901</td>
<td>Selleck Biochemicals</td>
<td>Cat#S1036</td>
</tr>
<tr>
<td>Sodium Selenite</td>
<td>Sigma-Aldrich</td>
<td>Cat#S5261</td>
</tr>
<tr>
<td>Beta-Estradiol</td>
<td>Sigma-Aldrich</td>
<td>Cat#E2758</td>
</tr>
<tr>
<td>Transferrin</td>
<td>Sigma-Aldrich</td>
<td>Cat#T8158</td>
</tr>
<tr>
<td>Prolactin</td>
<td>Sigma-Aldrich</td>
<td>Cat#L6520</td>
</tr>
<tr>
<td>Cholera Toxin</td>
<td>Sigma-Aldrich</td>
<td>Cat#C8052</td>
</tr>
<tr>
<td>Hydrocortisone</td>
<td>Sigma-Aldrich</td>
<td>Cat#H0888</td>
</tr>
<tr>
<td>Insulin</td>
<td>Sigma-Aldrich</td>
<td>Cat#I9278</td>
</tr>
<tr>
<td>Bovine Serum Albumin</td>
<td>Sigma-Aldrich</td>
<td>Cat#A9706</td>
</tr>
<tr>
<td>Heat Inactivated Horse Serum</td>
<td>Life Technologies</td>
<td>Cat#26050</td>
</tr>
<tr>
<td>10μM Dragon Green</td>
<td>Bangs Laboratory</td>
<td>Cat# FC06F-10163</td>
</tr>
<tr>
<td>Puromycin</td>
<td>Life Technologies</td>
<td>Cat#A113803</td>
</tr>
<tr>
<td>Laminin-111</td>
<td>Life Technologies</td>
<td>Cat#23017015</td>
</tr>
<tr>
<td>Fugene HD</td>
<td>Promega</td>
<td>Cat#E2311</td>
</tr>
<tr>
<td><strong>Cell Culture Media</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMEM/F-12 1:1</td>
<td>Life Technologies</td>
<td>Cat#11320</td>
</tr>
<tr>
<td><strong>Critical Commercial Assays and Sequencing Regents</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amphiregulin ELISA</td>
<td>R&amp;D Systems</td>
<td>Cat#DY262</td>
</tr>
<tr>
<td>DSQ 3x9 array microfluidic device</td>
<td>Nanoshift</td>
<td>N/A</td>
</tr>
<tr>
<td>Barcoded Beads SeqB</td>
<td>ChemGenes Corp.</td>
<td>MACOSKO-2011-10B</td>
</tr>
<tr>
<td>Agencourt AMPure XP beads</td>
<td>Beckman Coulter</td>
<td>Cat#A63881</td>
</tr>
<tr>
<td>Nextera XT (Illumina) sample preparation kit</td>
<td>Illumina</td>
<td>FC-131-1024</td>
</tr>
<tr>
<td><strong>Deposited Data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw sequencing data</td>
<td>GEO</td>
<td>GSE118312</td>
</tr>
<tr>
<td><strong>Experimental Models: Cell Lines</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human: MCF-10A, clone 5E</td>
<td>Joan Brugge, Harvard Medical School (Janes et al., 2010)</td>
<td>RRID:CVCL_0598</td>
</tr>
<tr>
<td>Human: 5E/Fra1::mCherry/EKAR3</td>
<td>(Gillies et al., 2017)</td>
<td>Available on request</td>
</tr>
</tbody>
</table>

(Continued on next page)
RESOURCE AVAILABILITY

Lead Contact
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).

Materials Availability
Plasmids generated in this study are forthcoming to Addgene. Plasmids will be available upon request from the Lead Contacts.

Data and Code Availability
MATLAB code and R markdown files are available at https://www.mcb.ucdavis.edu/faculty-labs/albeck/data.htm. Imaging data requests will be fulfilled by the Lead Contacts. Sequencing data have been uploaded to the GEO repository: GEO accession GSE118312.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Culture and Media
HMT-3522 cell lines were maintained in DMEM/F12 supplemented with prolactin, insulin, sodium selenite, hydrocortisone, β-estradiol, transferrin, and EGF (S1 only), as previously described (Kenny and Bissell, 2007). MCF10A-5E cell lines were maintained in DMEM/F12 media supplemented with 5% horse serum, insulin, cholera toxin, hydrocortisone, and EGF. All cell lines were maintained in 5% CO2 at 37°C.

Human Tissue Samples
De-identified human breast cancer tissue samples were provided by the Cooperative Human Tissue Network, as described in STAR Methods, in accordance with IRB protocols. Other investigators may have received samples from these same tissue specimens.

METHOD DETAILS

Reporter Line Construction
Reporter cell lines were created using lentiviral transduction from plasmid DNA containing the ERK Translocation Reporter (ERKTR) (Regot et al., 2014) fused to fluorescent proteins as indicated below. T4-2 reporter lines were constructed with ERKTR-mVenus. S1
reporter lines were constructed using ERKTR-mTurquoise2 or ERKTR-mCherry. Reporter expressing cells were selected with puromycin followed by sorting using flow cytometry. Flow sorting was conducted using a wide gating strategy to maximally preserve the inherent heterogeneity of S1 and T4-2 cell populations. MCF10A-ERKAR-Fra-1::mCherry cell lines were generated previously using the ERKAR3 reporter inserted into MCF10A cells carrying mCherry fused to the C terminus of endogenous FOSL1 (Fra-1) locus (Gillies et al., 2017).

**Live-Cell Microscopy and Co-culture Conditions**

Live-cell microscopy was conducted on multi-well plates with #1.5 glass bottoms that were coated with laminin-111 (Invitrogen, Carlsbad, CA). 50μg/ml Laminin-111 was deposited onto glass wells overnight in 20mM sodium acetate buffer containing 1mM CaCl₂ to generate a fractal laminin-coated surface (Hochman-Mendez et al., 2014). Immediately prior to plating, the wells were washed once with PBS to remove excess buffer and laminin. For mono-culture experiments, cells were plated at a density of 9000 cells/ well for both S1 and T4-2 cells. For co-culture experiments, total cell density was kept constant and the ratio of S1:T4-2 cells adjusted for a particular experiment (e.g. 30:70 S1 to T4-2 ratio corresponds to a combination of 2700 and 6300 cells per well, respectively). Plates were then incubated for 24 h in complete media. After 24 h, plates were washed twice in media containing no additives then placed in custom imaging media (DMEM/ F12 without phenol red, riboflavin, and folic acid) containing hydrocortisone, β-estrogen, transferrin, and Hoechst stain (1:100,000), then incubated overnight prior to live-cell microscopy. Following preparation, plates were imaged on a Nikon Ti-E inverted microscope fitted with an environmental chamber. A single stage position was chosen within each well of the plate and time lapse images were captured every 6 min under the indicated conditions with 20X 0.75 NA objective and Andor Zyla 5.5 scMOS camera. Automated imaging was performed using NIS-Elements AR software.

**Immunofluorescence Microscopy**

For fixed staining experiments cells were plated exactly as described for live cell imaging experiments. Following treatment and incubation under the indicated conditions the cells were fixed in 4% paraformaldehyde in phosphate buffered saline for 20 min. Wells were then blocked with buffer containing 0.1% Triton and 4% bovine serum albumin. Primary antibodies were used at 1:400 dilution. Secondary antibodies were used at a 1:200 dilution. Nuclei were stained with Hoechst-33342 and imaged using a DAPI filter set; nuclear stain images are labeled ‘DAPI’ for brevity.

**DNA Damage Response Assay**

MCF10A or T4-2 cells were treated with vehicle control, amphiregulin or EGF, respectively, and with erlotinib or carboplatin. Media were then washed out at 12 h and cells fixed at time scales of 0, 1, 2, and 3 h post-washout using paraformaldehyde. Cells were then co-stained with pH2AX-Alexa 488, Fra-1, and Egr1 antibodies for imaging and analysis.

**Amphiregulin ELISA**

To ascertain the concentration of free AREG present in our imaging experiments, media was removed from multiple wells of an imaging plate at a time point corresponding with the start live cell imaging experiments and subjected to ELISA. Seeding density was 9000 cells per well. Measurement of media amphiregulin levels were made using the Quantikine human amphiregulin ELISA kit per the manufacturer instructions.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

**Imaging, Data Processing, and Statistics and Normalization**

For all experiments >500 cells were analyzed per condition unless otherwise indicated. Each imaging dataset presented in this report is representative of at least two independent replicate experiments. Image, data processing, and statistics were performed using custom MATLAB software as previously described (Pargett and Albeck, 2018; Pargett et al., 2017). Linear regression was performed using a Theil-Sen estimator as previously described (Gillies et al., 2017). Temporal Variability Index (TVI) was calculated by taking the absolute value of the mean differential of filtered Fra-1 intensity (Fra₁) over time divided by the mean.

\[
TVI = \frac{\frac{d}{dt} (\text{Fra}_1)}{\text{Fra}_1}
\]

Fra-1 filtering was performed in MATLAB employing an infinite input response and Butterworth filter design set to lowpass filtering to remove high frequency noise with a periodicity < 30 min. All t-tests were calculated using an unpaired approach, significance was considered \( P < 0.05 \).

Partial least squares regression (PLSR) analysis was performed as previously described (Gillies et al., 2017) by incorporating all replicates and fixed time point data into the analysis.

**Dynamic Modeling**

The model of ERK dependent gene expression (Equations 1, 2, 3, and 4) was constructed from a mass action approximation of four steps: (1) phosphorylation of a transcription factor by ERK (TF⁵), (2) transcription of target mRNA (mRNA), (3) translation of target
protein ($P$), and (4) potential stabilization of target protein by ERK-dependent phosphorylation ($P^\text{F}$). Additionally, a regulatory term is included in the transcription process allowing negative feedback from target protein onto its own production. Stabilization and negative feedback are included as at least several ERK target genes are known to be phosphorylated by ERK, inhibiting their degradation, and feedback is a common feature in gene expression, evidenced for ERK target genes by decreasing mRNA expression after long term stimulus (Gillies et al., 2017; Uhlitz et al., 2017). The model is formulated as a delay differential equation to account for the effective lag times of transcription and translation without explicitly addressing the complex processes involved. Parameters, and their values for Fra-1 and Egr1, are listed in Table S1.

$$
\frac{d}{dt} TF^\text{F}(t) = k_{p,TF} * \text{ERK}(t) * (TF^\text{F}(t) - TF^\text{F}(t)) - k_{dT,TF} * TF^\text{F}(t) \quad \text{(Equation 1)}
$$

$$
\frac{d}{dt} \text{mRNA}(t) = \frac{k_b + k_m * TF^\text{F}(t - \tau_m)}{(1 + \frac{TF^\text{F}(t - \tau_m)}{k_p})} - k_{d,\text{mRNA}} + \text{mRNA}(t) \quad \text{(Equation 2)}
$$

$$
\frac{d}{dt} P(t) = k_p * \text{mRNA}(t - \tau_p) + k_{d,PP} + P^\text{F}(t) - (k_{d,PP} + k_{d,PP}) + P(t) \quad \text{(Equation 3)}
$$

$$
\frac{d}{dt} P^\text{F}(t) = k_{d,PP} * \text{ERK}(t) * P(t) - (k_{d,PP} + k_{d,PP}) + P^\text{F}(t) \quad \text{(Equation 4)}
$$

**Single Cell RNA Sequencing and Analysis**

**Cell Lines Used for Sequencing and Treatments**

To conduct single cell RNA sequencing experiments the following cells were grown under standard culture conditions for 4 days then dissociated with 0.25% trypsin/EDTA and rinsed into phosphate buffered saline immediately before Drop-seq. S1 and T4-2 monoculture cells were grown as described previously with the media changed daily. S1/T4-2 cells were grown under the same conditions except that no exogenous growth factors were added. S1a cells received new media daily containing 100 ng/ml amphiregulin.

**Drop-seq**

Droplets were generated using a DSQ 3x9 array microfluidic part (Nanoshift, Emeryville, CA, USA) using a Drop-seq setup according to Macosko et al. (2015) (Online-Dropseq-Protocol-v.3.1-Dec-2015, http://mccarrollab.com/dropseq). Barcoded beads and cells were loaded at concentrations specified in (Figure S2C). Prior to the collection, cell syringes and tubing were blocked using PBS + 0.1% BSA. A magnetic mixing disc was inserted into the cell syringe to allow for manual cell mixing during the run and the cell pump was used in a vertical position. Droplets were collected in 50mL Falcon tubes and the target volume of aqueous flow varied in between 1-1.2mL of the cell suspension. Droplets were broken immediately after collection and reverse transcription, exonuclease-treatment and further processing was conducted as described previously (Macosko et al., 2015).

**Sequencing Strategy**

Nextera libraries were sequenced on Illumina Nextseq 500 sequencers using the Nextseq High Output v2 kit (75 cycles), using a custom primer and a custom paired end sequencing strategy, R1 20bp, index 8bp, R2 remaining bp (Macosko et al., 2015). The five Nextera libraries were pooled and a total of 9.9k anticipated STAMPs were loaded on the flow cell with the following library portions; S1/MEF: 5%, T4: 17.1%, T4/S1: 30.3%, S1: 17.2%, S1 amphiregulin: 13.1%. The pool was sequenced on two Nextseq 500 runs. Raw sequencing data have been submitted to the GEO repository and are available under GEO accession GSE118312.
Reference for Read Alignment
For the species mixing experiment we used a mixed reference (human+mouse) available at GEO accession GSE63269 (Macosko et al., 2015). For mapping of T4 and S1 reads we used the human genome primary assembly (GRCh38) available at https://www.ensembl.org/ and release 92 gene models. Sequences for transgenic markers mCherry and mVenus were added to the reference.

Drop-seq Pipeline and Generation of the Gene Expression Matrix
All Drop-seq data were processed using Drop-seq Tools v1.12 as described previously (http://mccarrollab.com/wp-content/uploads/2016/03/DropseqAlignmentCookbookv1.2Jan2016.pdf) (Macosko et al., 2015). Bowtie2 (v2.2.6) (Langmead et al., 2009) was used for read alignment with the settings –phred33 –very-sensitive -N 1.

Species Mixing Experiment
To determine whether single cells were generated, we mixed S1 cells and MEF (mouse) cells, then mapped the resultant data against a hybrid reference containing both human and mouse genes. Cells were loaded at approximately 100 cells/uL (50 cells/μl final), which would be expected to give ~2.5% doublet rate, similar to the 1.6% doublet rate observed in the experiment (Figure S2A). This library was sequenced very shallow since its sole purpose was doublet assessment.

Cell QC and Clustering
Cells with more than 500/2k and less than 6k/30k genes/UMIs were considered in the analyses (Figure S2B DropseqStats). Cells with more than 10% of mitochondrial expression were filtered out and excluded from downstream analyses. 10,212 cells remained after filtering with medians of 2,653 genes and 6,283 UMIs. Analyses were conducted using R package Seurat (v2.2.1) (Satija et al., 2015).

Excess Variance Calculation and Delta-Excess Variance
UMI counts for each cell were converted to transcripts per million (TPM) by scaling all cell-wise counts to 1,000,000. For each cell population i and each gene g, under the assumption of a negative binomial distribution, we computed its mean expression μ_ig and variance v_ig, then calculated gene-wise dispersion parameters α_ig from the expression v_ig = μ_ig + α_igμ_ig². We observed a linear trend between log μ_ig and log α_ig for all cell populations, indicating as the mean expression of a gene increased, the overdispersion tends to decrease. Therefore, to measure change in dispersion for each gene between two populations (when the mean may also change), we instead compare “excess variance”. For every pair of populations i and j being compared, we fit an “excess variance” term E_ig = (μ_ig − μ_jg), where the fitted overdispersion estimate μ_jg is generated by performing Local Polynomial Regression Fitting (loess) using R version 3.4.4 with span=1, simultaneously regressing log α_ig on log μ_ig and log α_jg on log μ_jg over all genes g. Statistical significance was calculated using a paired, two-sided Wilcoxon signed rank test between all E_ig and E_jg. See Data S1 for a complete list of genes and relative excess variance.

Parameters were selected to match previously data collected on average mRNA and protein expression kinetics (Gillies et al., 2017; Schwanhäusser et al., 2011; Uhlitz et al., 2017).