UC Davis

Title

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

Permalink

Journal

ISSN

Authors

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

Publication Date

DOI

Article

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

Graphical Abstract
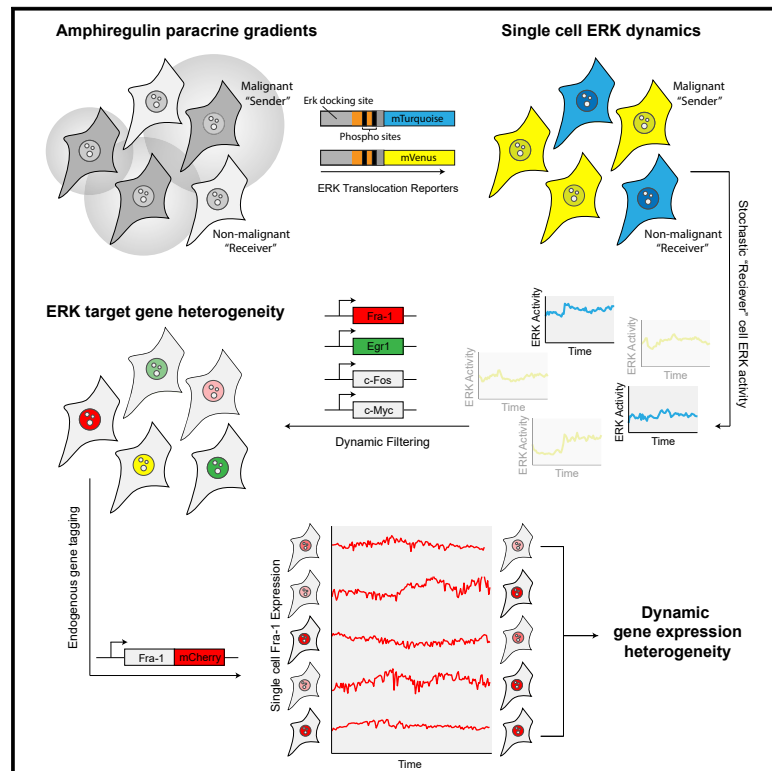
Authors

Alexander E. Davies, Michael Pargett, Stefan Siebert, ..., Gerald Quon, Mina J. Bissell, John G. Albeck

Correspondence

davies.474@osu.edu (A.E.D.), jgalbeck@ucdavis.edu (J.G.A.)

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.

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

CellPress

Article

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

Alexander E. Davies,[1,2,4,*] Michael Pargett,[1] Stefan Siebert,[1] Taryn E. Gillies,[1] Yongin Choi,[1] Savannah J. Tobin,[1,3] Abhineet R. Ram,[1] Vaibhav Murthy,[3] Celina Juliano,[1] Gerald Quon,[1] Mina J. Bissell,[2] and John G. Albeck[1,5,*]

[1]Department of Molecular and Cellular Biology, University of California, Davis, Davis, CA 95616, USA
[2]Division of Biological Systems and Engineering, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[3]Department of Veterinary Biosciences, College of Veterinary Medicine, the Ohio State University, Columbus, OH 43210, USA
[4]Present address: Department of Veterinary Biosciences, College of Veterinary Medicine, the Ohio State University, Columbus, OH 43210, USA
[5]Lead Contact
*Correspondence: davies.474@osu.edu (A.E.D.), jgalbeck@ucdavis.edu (J.G.A.)
https://doi.org/10.1016/j.cels.2020.07.004

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
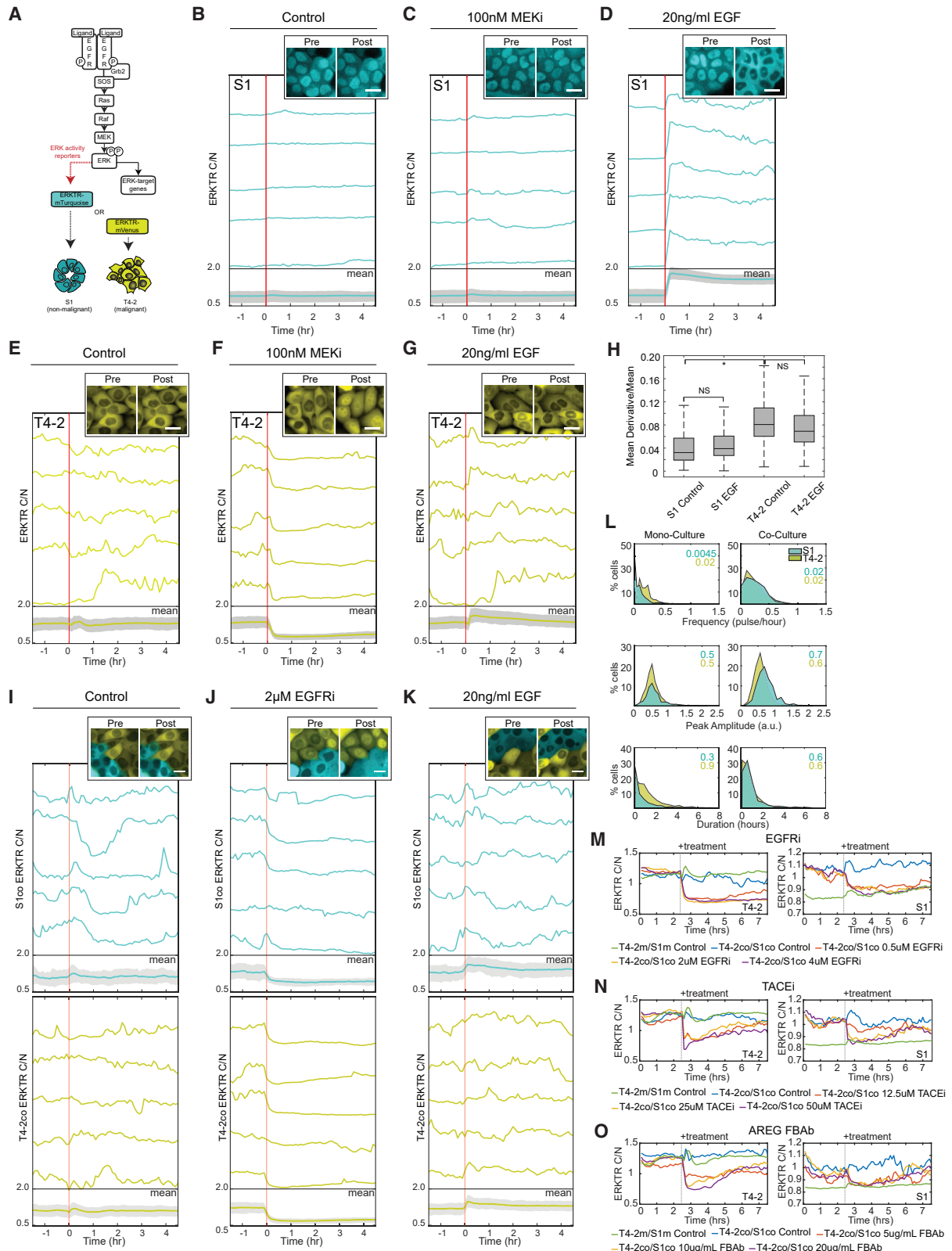
Figure 1. Progression to Malignancy Is Associated with Stochastic RAS-ERK Signaling Dynamics and Heterogeneous Target Gene Expression

(A) Schematic of the EGFR-RAS-ERK signaling pathway in mammalian cells and ERKTR.

(legend continued on next page)

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 (Hochgräfe 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; Hiratsuka 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 mammoplasty 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 phosphatase-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 ERKTR$_{C/N}$). 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 ERKTR$_{C/N}$ (Figure 1B; Video S1A), with infrequent low amplitude pulses. Addition of MEKi had no effect on ERKTR$_{C/N}$, 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 ERKTR$_{C/N}$ 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 (Gillies 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

---

(B–G) Measurements of ERK activity by ERKTR$_{C/N}$ in S1 or T4-2 cells following treatment with exogenous growth factors or MEK inhibitor. In each panel, mean ERKTR$_{C/N}$ is shown at bottom with the 25th–75th interquartile range (IQR) depicted by shading. Unshaded traces represent individual cells and red vertical lines indicate the addition of growth factor, inhibitor, or vehicle. n > 2,400 cells per condition.

(H) Variability of ERKTR$_{C/N}$ in control versus EGF integrated over time (n = 8,372). Boxes indicate IQR, and whiskers the range. NS, not significant, *p <0.01 by unpaired t test.

(I–K) Measurements of ERK activity by ERKTR$_{C/N}$ in co-cultured S1 (turquoise lines) and T4-2 cells (yellow lines). Annotations as in (B–G). n > 1,900 cells per condition.

(L) Distributions of single-cell pulse frequency (per hour), duration (hours), and amplitude (arbitrary units) of ERKTR$_{C/N}$ for S1 (turquoise) and T4-2 (yellow) (3,497 total cells analyzed). Numbers upper right correspond to median value color-coded to each cell type.

(M) Mean ERKTR$_{C/N}$ traces of T4-2m, S1m, or T4-2co/S1co treated with EGFR inhibitor (erlotinib) at the indicated concentrations. Numbers of cells analyzed per condition (T4-2/S1): monoculture, control 2187/692; co-culture, control 244/122; 0.5 μM erlotinib 199/89; 2 μM erlotinib 239/78, 4 μM erlotinib 281/105.

(N) Mean ERKTR$_{C/N}$ traces of T4-2m, S1m, or T4-2co/S1co treated with TACE inhibitor (TACEi). Numbers of cells analyzed per condition (T4-2/S1): monoculture, control 2374/2036; co-culture, control 249/122; 12.5 μM TACEi 233/116; T4-2/S1co 25 μM TACEi; 256/110, 50 μM TACEi 238/90.

(O) Mean ERKTR$_{C/N}$ traces of T4-2m, S1m, or T4-2co/S1co treated with AREG function-blocking antibody. Numbers of cells analyzed per condition (T4-2/S1): monoculture, control 2374/2036; co-culture, control 243/55; 5 μg FBAb 237/56; 10 μg FBAb 222/76; 20 μg FBAb 261/74.

S1C). EGF provoked 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 ERKTR$_{C/N}$ signals (Figures 1I–1K; Videos S2A–S2G). For clarity, we refer to co-cultured cells as S1co and T4-2co, and to mono-cultured cells as S1m and T4-2m. In S1co, compared with S1m, we observed increased ERKTR$_{C/N}$ 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 ERKTR$_{C/N}$ 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 that of T4-2m (Figure 1K; Video S3B), in agreement with receptor occupancy by paracrine ligands.

To confirm that ERK activity in S1co and T4-2m cells is mediated by paracrine ligands, we treated cells with TAPI-0, an inhibitor of the AREG and TGF-α shedding enzyme (TACE), compared to EGFRi (Figure 1M) or MEKi (Figures S1B–S1D). TACEi produced an initial suppression of ERK activity in both S1co and T4-2m, followed by a gradual return to dynamic signaling (Figures 1N and S1C). We found that TACEi exerts a prolonged effect if culture medium is changed immediately prior to TACEi addition (Figure S1E), suggesting that residual AREG in the microenvironment accounts for the gradual return to dynamic signaling in the absence of media washout. An AREG function-blocking antibody (FBAb) also strongly suppressed ERKTR$_{C/N}$ for ~1 h, followed by a gradual return of ERK activity (Figures 1O and S1C), which is consistent with saturation of the FBAb over time by continuous AREG secretion from T4-2 cells. High concentrations of TGF-α FBAb had no detectable effect on ERKTR$_{C/N}$ (Figure S1F). These results indicate that paracrine release of AREG by T4-2 cells is the primary driver of stochastic ERK signaling and are also consistent with the role of AREG as a low-affinity EGFR ligand that can diffuse freely to adjacent cells (DeWitt et al., 2001, 2002). Finally, we compared these observations with a co-culture lacking paracrine signaling. MCF10A-CA1D cells, which carry an activating RAS mutation but are not known to secrete AREG, showed similar ERKTR$_{C/N}$ levels to T4-2m cells but did not efficiently drive ERK signaling in S1co cells (Figure S1G). These results establish a defined system in which S1co cells serve as a receiver for paracrine EGFR-mediated signals produced by T4-2co cells.

Paracrine Signaling Drives Heterogeneous Expression of EGFR-RAS-ERK-Regulated Genes

We investigated gene expression heterogeneity using single-cell mRNA sequencing (scRNA-seq), comparing S1m cells with or

without growth factor, S1co/T4-2co mixtures, and 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 eleven clusters of cells with unique but partially overlapping signatures (Figures 2C and S2E; Data S1). T4-2m and T4-2co cells were each comprised 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 × 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
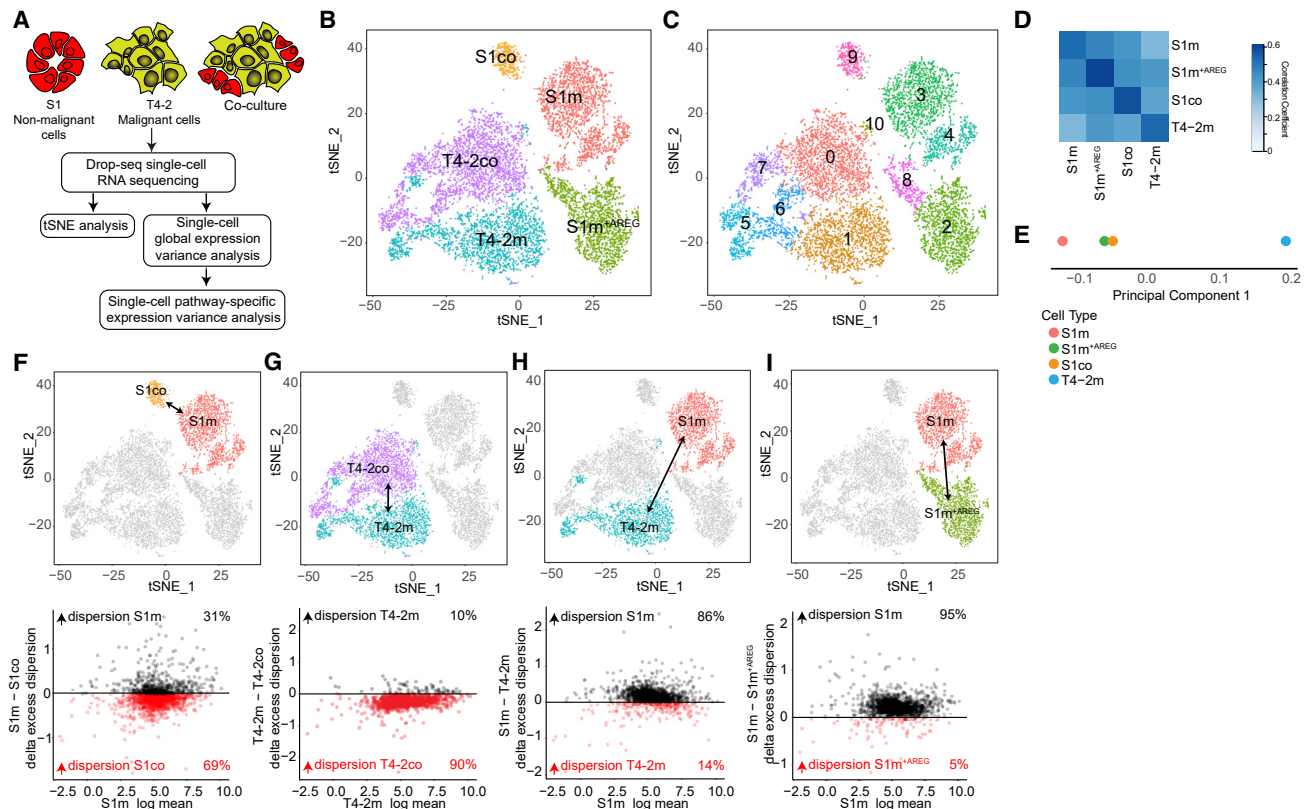
Figure 2. Induction of Global Gene Expression Variance by Cellular Crosstalk

(A) Schematic depicting cell culture conditions and analyses for S1, T4-2, and S1/T4-2 co-cultures for single-cell RNA sequencing. Red: S1 cells expressing ERKTR-mCherry; yellow: T4-2 cells expressing ERKTR-mVenus.

(B and C) tSNE plots of single-cell transcript profiles, colored by identified cell type and condition (B) or by clustering of transcriptional profiles (C). n > 1,900 cells per condition.

(D) Pairwise correlation of the indicated cells and conditions using raw count data.

(E) Scatterplot of first principal component weights per the indicated cells and conditions.

(F–I) Top row: tSNE plots highlighting the groups of cells compared by excess dispersion. Bottom row: differences in excess dispersion between treatment conditions (>1,800 genes sampled); each dot represents the relative excess dispersion for one gene. Genes are color-coded according to which condition shows greater excess dispersion. Percentages indicate the proportion of genes falling on either side of the horizontal line.

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 condi-

tions, 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. ERKTR$_{C/N}$ 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
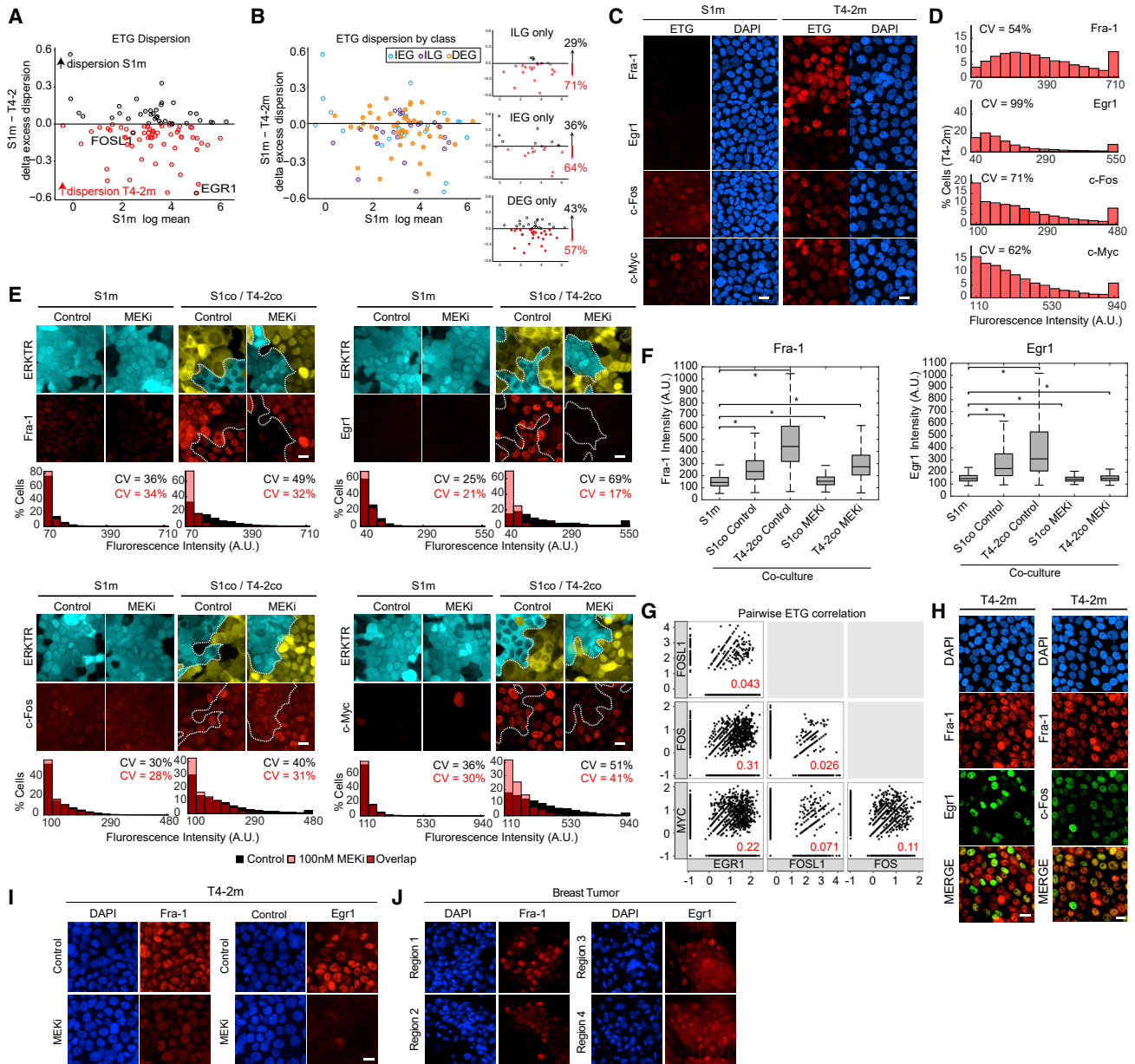
Figure 3. Induction of ETG Expression Variance by Cellular Crosstalk

(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 and 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), MEKi-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.

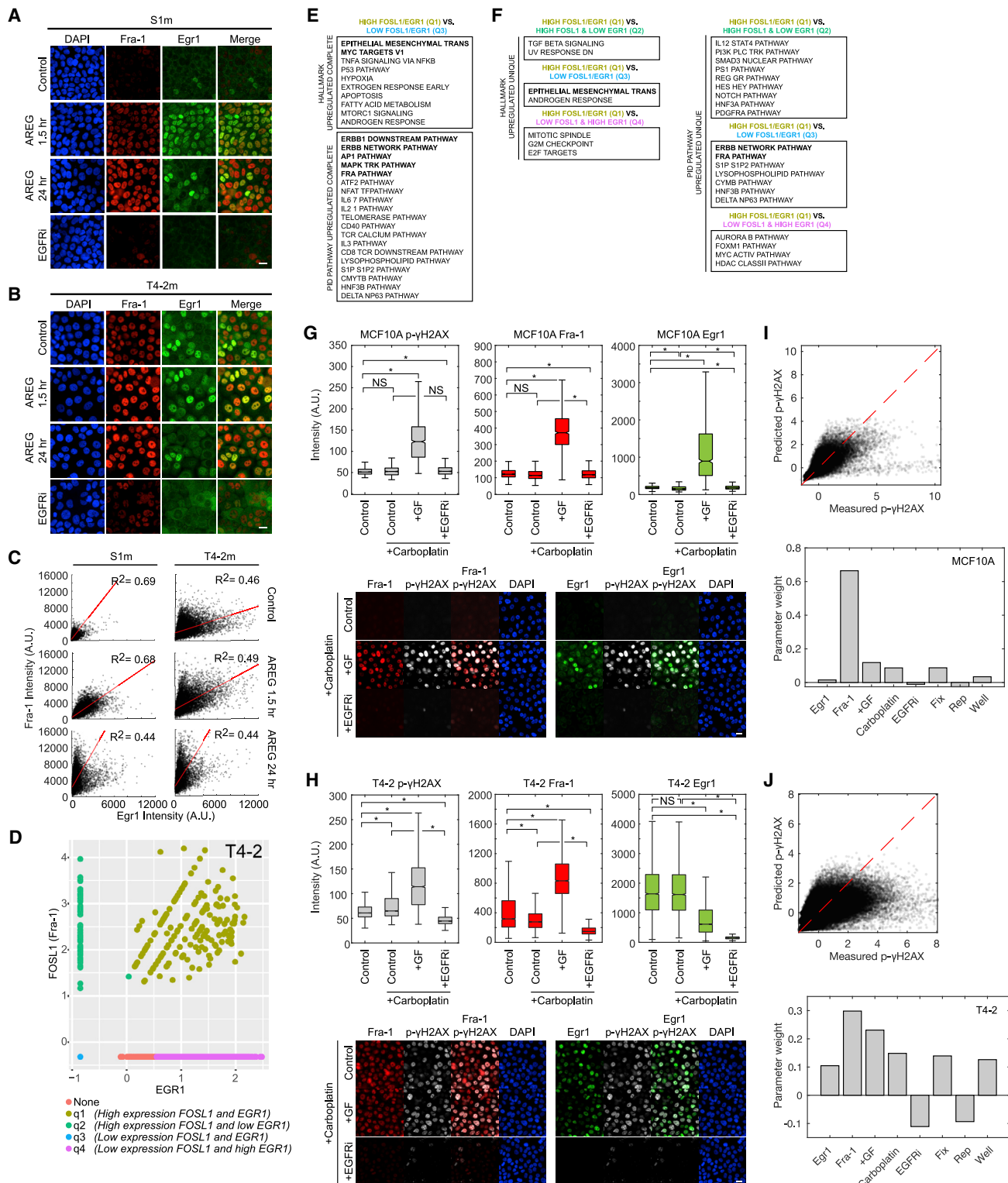(J) Formalin-fixed resected human invasive ductal carcinoma samples stained for Fra-1 or Egr1.

Figure 4. Paracrine AREG and Gene-Specific Expression Kinetics Coordinately 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.

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 3I). 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 (PID) (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

(E) Differential gene expression analysis based on FOSL1 and EGR1 expression partitioning. A complete list of upregulated molecular signature database (MSigDB) hallmark and NCI pathway database (PID) pathways in FOSL1 or EGR1 high- versus low-expressing cells.

(F) As in (E), differential gene expression analysis based on ETG expression was analyzed with MSigDB and PID. The list represents uniquely upregulated pathways for the indicated comparisons.

(G) Analysis of p-γH2AX, Fra-1, and Egr1 staining levels in MCF10A cells under the indicated conditions: control, +GF (growth factor, 100 ng/mL AREG), or 5 μM erlotinib, and 50 μM carboplatin. n > 1,000 cells per condition. NS, not significant; *p < 0.01 unpaired t test. Scale bar, 20 μm.

(H) Analysis of pH2AX, Fra-1, and Egr1 staining levels in T4-2 cells under the indicated conditions: control, +GF (growth factor, 20 ng/mL EGF), or 5 μM erlotinib and 50 μM carboplatin. n > 5,000 cells per condition. NS, not significant; *p < 0.01 by unpaired t test. Scale bar, 20 μm.

(I) PLSR analysis on MCF10A cells indicating model fit to p-γH2AX signals (predicted versus measured) and parameter weighting by Egr1, Fra-1, erlotinib (EGFRi), carboplatin, time of fixation (Fix), replicates (Rep), or intraexperiment replicates (Well). n = 56,300.

(J) PLSR analysis on T4-2 cells indicating model fit to p-γH2AX (predicted versus measured) and parameter weighting by Egr1, Fra-1, erlotinib (EGFRi), carboplatin, time of fixation (Fix), replicates (Rep), or intraexperiment replicates (Well). n = 209,940.
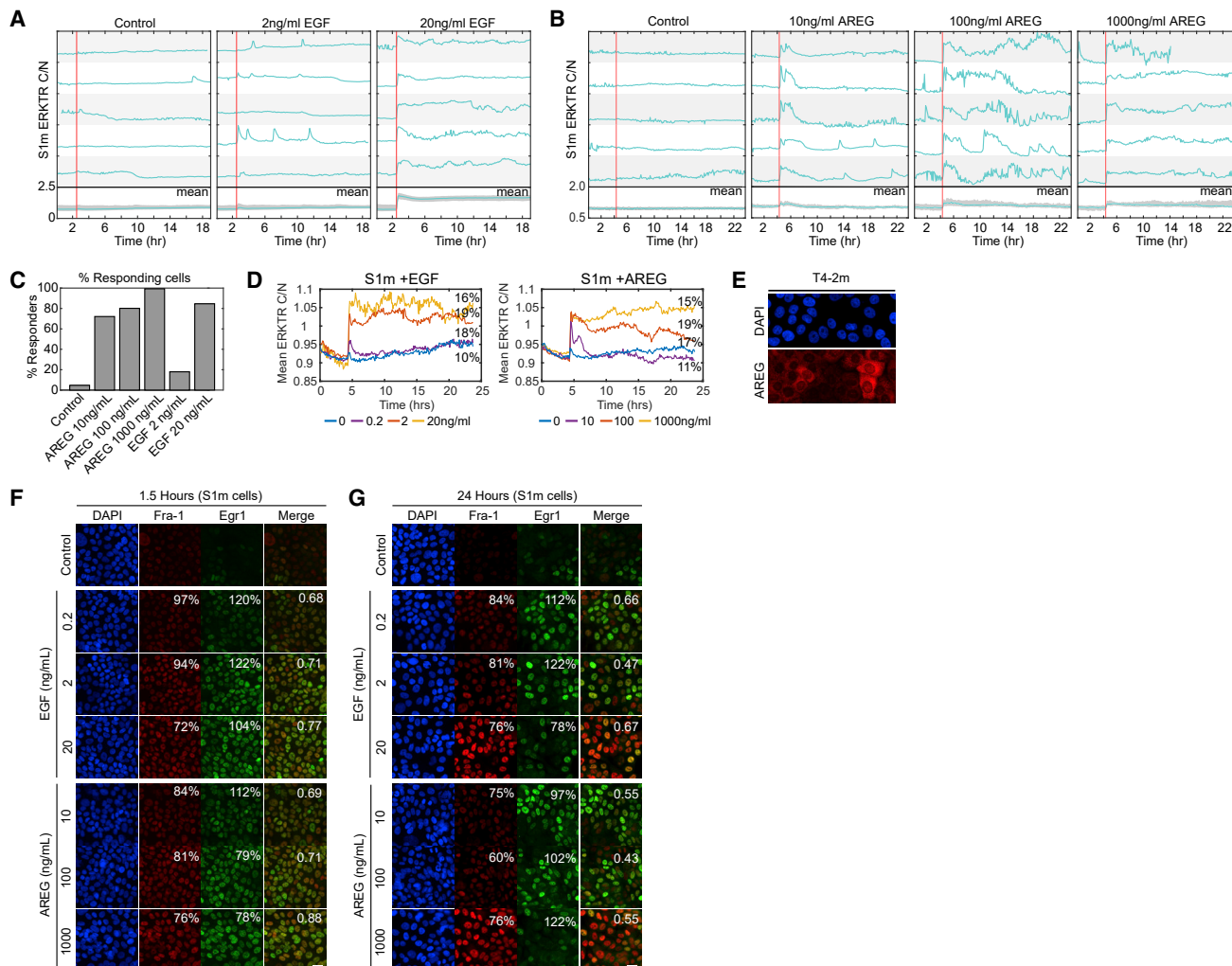
Figure 5. AREG Drives Stochastic ERK Signaling to Induce Heterogeneous ETG Expression

(A and B) Single cell and mean traces of ERKTR$_{C/N}$ in S1m cells exposed to the indicated EGFR ligand and concentrations for a duration of >18 h. Vertical red lines indicate the time of ligand or vehicle addition, the bottom plot shows the mean ERKTR$_{C/N}$ in bold, with IQR shaded. Above the mean trace, 5 representative single-cell measurements of ERKTR$_{C/N}$ are shown. >500 cells were for each condition.

(C) Bar graphs depict the percent of cells responding to the indicated dose of EGFR ligand within 30 min of treatment. n > 500 cells per condition

(D) Mean ERKTR$_{C/N}$ traces for S1 cells receiving the indicated EGF or AREG treatments. Percentages represent the temporal variability score for each condition. n > 200 cells per condition.

(E) Immunofluorescence imaging of AREG expression in T4-2 cells, AREG (red) and nuclei (blue).

(F and G) Co-staining of Fra-1 and Egr1 in S1m cells at the indicated timepoints and conditions. Percentages indicate the coefficient of variation for Fra-1 or Egr1 respectively. Numbers inset on "merge" images indicate the R^2 value for Fra-1 and Egr1. n > 1,000 cells per condition. Scale bar, 20 μm.

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 (Figures 5B and 5D). Treatment with 100 ng/mL AREG resulted in a similar fraction of responding cells (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 ERKTR$_{C/N}$ 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,
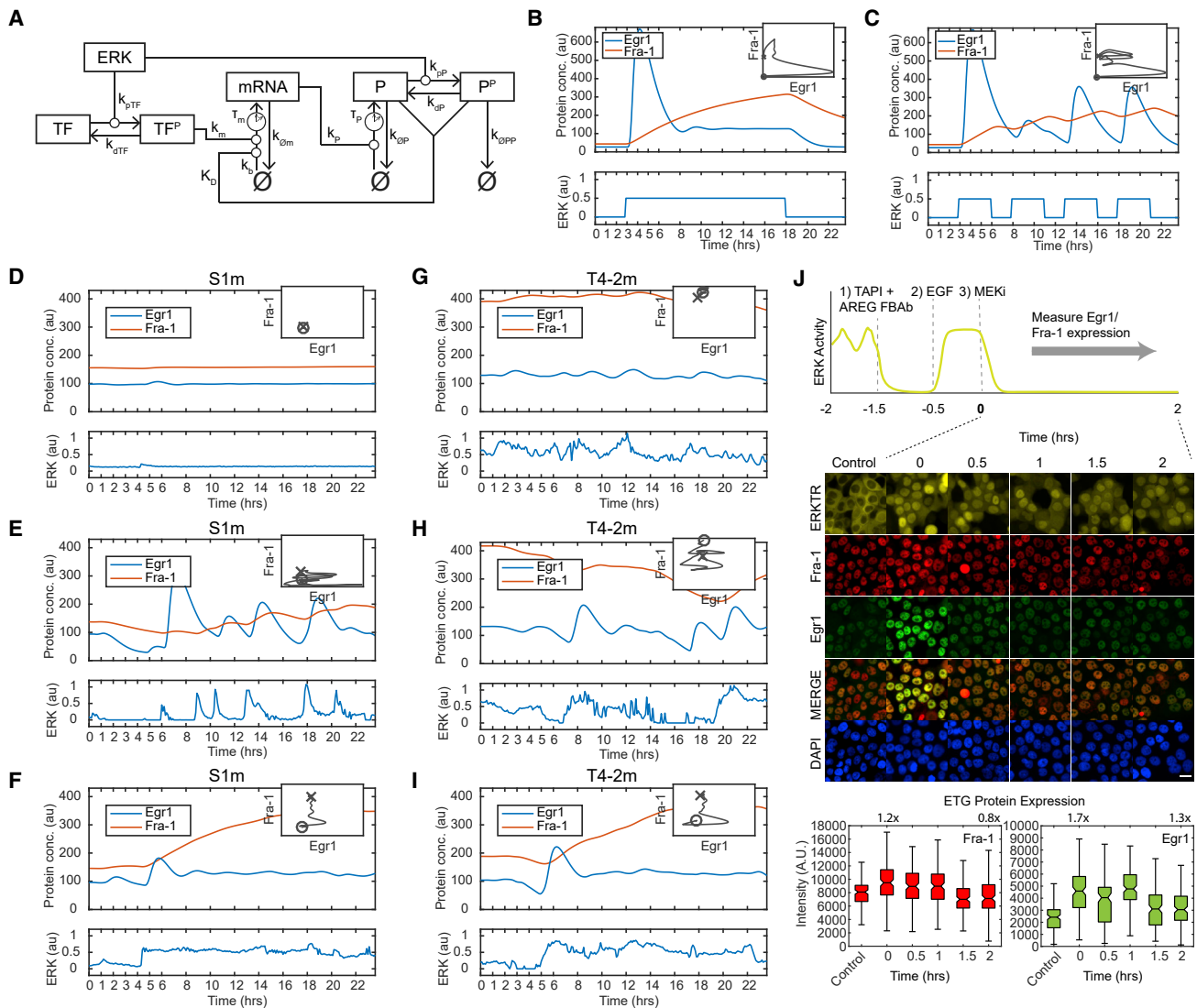
Figure 6. Mathematical Simulation of ETG Expression Heterogeneity

(A) Model diagram describing ERK-dependent modification of a transcription factor (TF), expression of mRNA, and a protein (P) product. Superscript P indicates phosphorylation, clock icons indicate a time delay τ, all lowercase k's indicate rate parameters (see Table S1 for definitions and values used), and uppercase K indicates a dissociation constant for feedback effects.

(B and C) Simulated responses to hypothetical square wave ERK signals and corresponding ETG expression levels (presented in arbitrary units).

(B–I) Insets show the trajectory over time in expression space, starting at the circle and ending at the X.

(D–I) Simulated responses to measured single-cell ERK activity traces from S1m (D and F) or T4-2m (G–I) cells in control conditions (D, E, G, and H) or treated with 1,000 ng/mL AREG (F and I). Insets show the modeled trajectory over time in expression space, starting at the circle and ending at the X.

(J) Temporal ETG responses to a single pulse of ERK activity in T4-2 cells. Fixed images were taken at the indicated time points, in hours, and compared relative to control. Images presented were equally scaled to time point 0 to prevent over-scaling of peak intensity. Graphs correspond with mean ETG intensity and numbers represent the fold change in ETG intensity relative to control. Scale bar, 20 μM.

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,
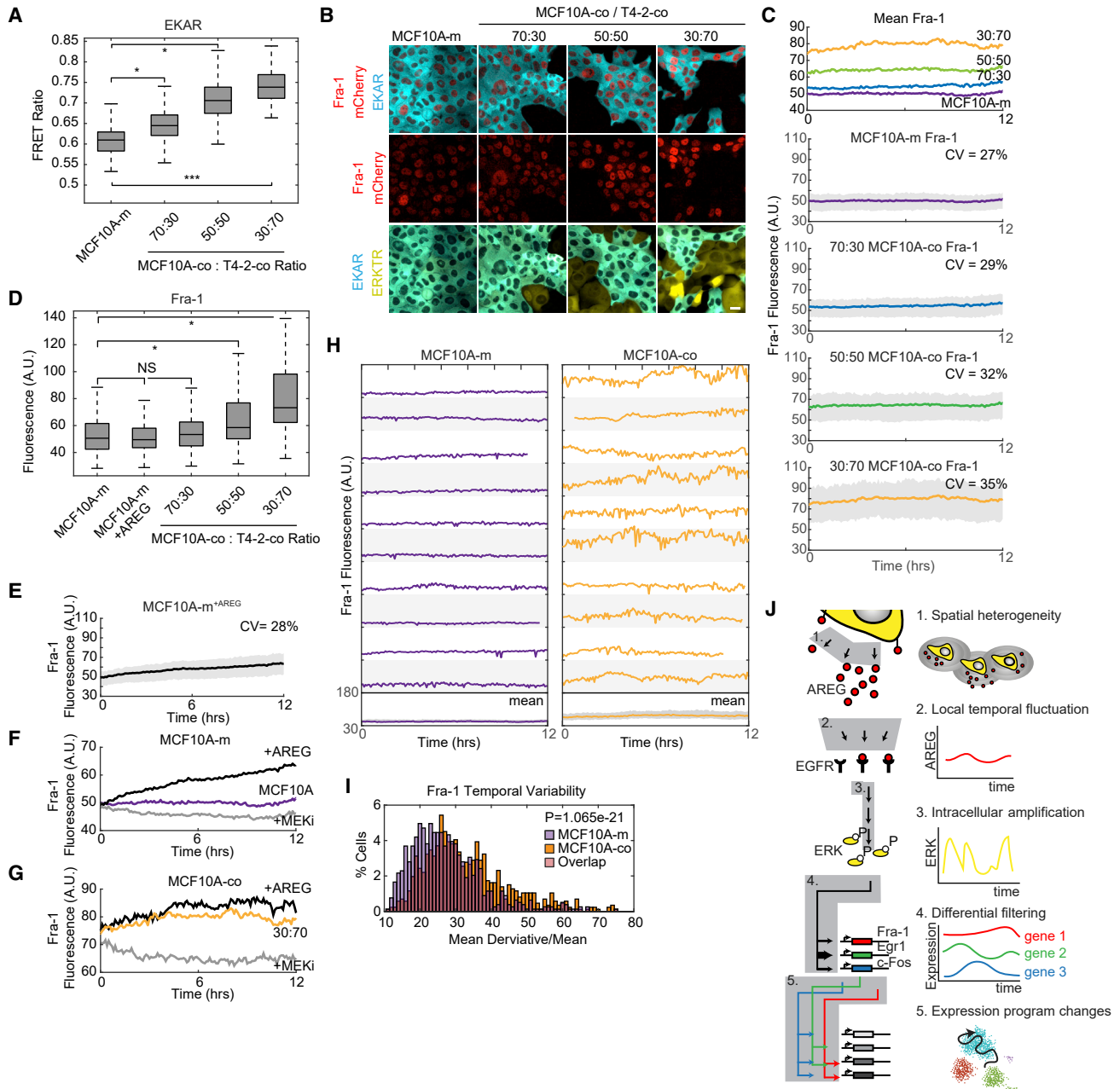
Figure 7. Live-Cell Imaging of Paracrine-Induced Fra-1 Temporal Expression Variability in Co-culture

(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 10A cells (turquoise) and ERKTR 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.

(legend continued on next page)

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; Schwanhäusser 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 6I). 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 S1 cells, 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 MCF10Am 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 = 1.065e−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.

(H) Fra-1::mCherry fluorescence over time (12 h) in individual MCF10A-m (purple) or MCF10A-co cells (orange). The bottom trace represents mean Fra-1 signal with IQR shaded; the traces above depict representative single-cell traces of Fra-1 expression over time. N > 700 cells per condition.

(I) Distribution of temporal variability in individual cells, defined as the absolute value of the derivative Fra-1 signal divided by the mean Fra-1 signal for each cell.

(J) A model of heterogeneity generation by paracrine-driven EGFR-RAS-ERK signaling. Molecular processes are shown at left, and the information processing function performed by each biochemical step is shown at right (numbered 1–5).

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. Futhermore, 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 reworded 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 STATSTICAL ANALYSIS
  - Imaging, Data Processing, and Statistics and Normalization
  - Dynamic Modeling
  - Single Cell RNA Sequencing and Analysis
  - Sequencing Strategy

- ○ Reference for Read Alignment
- ○ Drop-seq Pipeline and Generation of the Gene Expression Matrix
- ○ Species Mixing Experiment
- ○ Cell QC and Clustering
- ○ Excess Variance Calculation and Delta-Excess Variance

SUPPLEMENTAL INFORMATION

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

AUTHOR CONTRIBUTIONS

A.E.D., J.G.A., and M.J.B. conceptualized the study, analyzed data, and wrote the manuscript. A.E.D., S.J.T., 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. G.Q. 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

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.

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.

Belguise, K., Kersual, N., Galtier, F., and Chalbos, D. (2005). FRA-1 expression level regulates proliferation and invasiveness of breast cancer cells. Oncogene *24*, 1434–1444.

Berns, E.M., Klijn, J.G., van Putten, W.L., van Staveren, I.L., Portengen, H., and Foekens, J.A. (1992). c-myc amplification is a better prognostic factor than HER2/neu amplification in primary breast cancer. Cancer Res. *52*, 1107–1113.

Briand, P., Petersen, O.W., and Van Deurs, B. (1987). A new diploid nontumorigenic human breast epithelial cell line isolated and propagated in chemically defined medium. In Vitro Cell Dev. Biol. *23*, 181–188.

Brooks, M.D., Burness, M.L., and Wicha, M.S. (2015). Therapeutic implications of cellular heterogeneity and plasticity in breast cancer. Cell Stem Cell *17*, 260–271.

Bugaj, L.J., Sabnis, A.J., Mitchell, A., Garbarino, J.E., Toettcher, J.E., Bivona, T.G., and Lim, W.A. (2018). Cancer mutations and targeted drugs can disrupt dynamic signal encoding by the RAS-Erk pathway. Science *361*, eaao3048.

Cazet, A.S., Hui, M.N., Elsworth, B.L., Wu, S.Z., Roden, D., Chan, C.L., Skhinas, J.N., Collot, R., Yang, J., Harvey, K., et al. (2018). Targeting stromal remodeling and cancer stem cell plasticity overcomes chemoresistance in triple negative breast cancer. Nat. Commun. *9*, 2897.

Cong, Le, Ran, F. Ann, Lin, Shuailiang, Barretto, Robert, Naomi, Habib, Hsu, Patrick D., Wu, Xuebing, Jiang, Wenyan, Marraffini, Luciano A., and Zhang, Feng (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Davies, A.E., and Albeck, J.G. (2018). Microenvironmental signals and biochemical information processing: cooperative determinants of intratumoral plasticity and heterogeneity. Front. Cell Dev. Biol. *6*, 44.

DeWitt, A., Iida, T., Lam, H.Y., Hill, V., Wiley, H.S., and Lauffenburger, D.A. (2002). Affinity regulates spatial range of EGF receptor autocrine ligand binding. Dev. Biol. *250*, 305–316.

de la Cova, Claire, Townley, Robert, Regot, Sergi, and Greenwald, Iva (2017). A Real-Time Biosensor for ERK Activity Reveals Signaling Dynamics during C. elegans Cell Fate Specification. Dev. Cell *42*, 542–553.

DeWitt, A.E., Dong, J.Y., Wiley, H.S., and Lauffenburger, D.A. (2001). Quantitative analysis of the EGF receptor autocrine system reveals cryptic regulation of cell response by ligand capture. J. Cell Sci. *114*, 2301–2313.

Fischer, O.M., Hart, S., Gschwind, A., and Ullrich, A. (2003). EGFR signal transactivation in cancer cells. Biochem. Soc. Trans. *31*, 1203–1208.

Friedl, P., and Alexander, S. (2011). Cancer invasion and the microenvironment: plasticity and reciprocity. Cell *147*, 992–1009.

Gao, R., Davis, A., McDonald, T.O., Sei, E., Shi, X., Wang, Y., Tsai, P.C., Casasent, A., Waters, J., Zhang, H., et al. (2016). Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. Nat. Genet. *48*, 1119–1130.

Gillies, T.E., Pargett, M., Minguet, M., Davies, A.E., and Albeck, J.G. (2017). Linear integration of ERK activity predominates over persistence detection in Fra-1 regulation. Cell Syst. *5*, 549–563.e5.

Gupta, P.B., Fillmore, C.M., Jiang, G., Shapira, S.D., Tao, K., Kuperwasser, C., and Lander, E.S. (2011). Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell *146*, 633–644.

Hamilton, William B., Mosesson, Yaron, Monteiro, Rita S., Emdal, Kristina B., Knudsen, Teresa E., Francavilla, Chiara, Barkai, Naama, Olsen, Jesper, and Brickman, Joshua M. (2019). Dynamic lineage priming is driven via direct enhancer regulation by ERK. Nature *575*, 355–360.

Handly, L.N., Pilko, A., and Wollman, R. (2015). Paracrine communication maximizes cellular response fidelity in wound signaling. eLife *4*, e09652.

Harris, R.C., Chung, E., and Coffey, R.J. (2003). EGF receptor ligands. Exp. Cell Res. *284*, 2–13.

Hartman, Z.C., Poage, G.M., den Hollander, P., Tsimelzon, A., Hill, J., Panupinthu, N., Zhang, Y., Mazumdar, A., Hilsenbeck, S.G., Mills, G.B., and Brown, P.H. (2013). Growth of triple-negative breast cancer cells relies upon coordinate autocrine expression of the proinflammatory cytokines IL-6 and IL-8. Cancer Res. *73*, 3470–3480.

Harvey, Christopher D., Ehrhardt, Anka G., Cellurale, Cristina, Zhong, Haining, Yasuda, Ryohei, Davis, Roger J., and Svoboda, Karel (2008). A genetically encoded fluorescent sensor of ERK activity. PNAS *105*, 19264–19269.

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.

Hiratsuka, T., Fujita, Y., Naoki, H., Aoki, K., Kamioka, Y., and Matsuda, M. (2015). Intercellular propagation of extracellular signal-regulated kinase activation revealed by in vivo imaging of mouse skin. eLife 4, e05178.

Hochgräfe, F., Zhang, L., O'Toole, S.A., Browne, B.C., Pinese, M., Porta Cubas, A., Lehrbach, G.M., Croucher, D.R., Rickwood, D., Boulghourjian, A., et al. (2010). Tyrosine phosphorylation profiling reveals the signaling network characteristics of Basal breast cancer cells. Cancer Res. 70, 9391–9401.

Hochman-Mendez, C., Cantini, M., Moratal, D., Salmeron-Sanchez, M., and Coelho-Sampaio, T. (2014). A fractal nature for polymerized laminin. PLoS One 9, e109388.

Janes, Kevin A., Wang, Chun-Chao, Holmberg, Karin J., Cabral, Kristin, and Brugge, Joan S. (2010). Identifying single-cell molecular programs by stochastic profiling. Nature Methods 7, 311–317.

Jaqaman, Khuloud, Loerke, Dinah, Mettlen, Marcel, Kuwata, Hirotaka, Grinstein, Sergio, Schmid, Sandra L., and Danuser, Gaudenz (2008). Robust single-particle tracking in live-cell time-lapse sequences. Nat. Methods 5, 695–702.

Kenny, P.A., and Bissell, M.J. (2007). Targeting TACE-dependent EGFR ligand shedding in breast cancer. J. Clin. Invest. 117, 337–345.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25.

Liberzon, Arthur, Birger, Chet, Thorvaldsdóttir, Helga, Ghandi, Mahmoud, Mesirov, Jill P., and Tamayo, Pablo (2015). The Molecular Signatures Database Hallmark Gene Set Collection. Cell Systems 1, 417–425.

Lu, H., Clauser, K.R., Tam, W.L., Fröse, J., Ye, X., Eaton, E.N., Reinhardt, F., Donnenberg, V.S., Bhargava, R., Carr, S.A., and Weinberg, R.A. (2014). A breast cancer stem cell niche supported by juxtacrine signalling from monocytes and macrophages. Nat. Cell Biol. 16, 1105–1117.

Macdonald-Obermann, J.L., and Pike, L.J. (2014). Different epidermal growth factor (EGF) receptor ligands show distinct kinetics and biased or partial agonism for homodimer and heterodimer formation. J. Biol. Chem. 289, 26178–26188.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 161, 1202–1214.

Madsen, M.W., Lykkesfeldt, A.E., Laursen, I., Nielsen, K.V., and Briand, P. (1992). Altered gene expression of c-myc, epidermal growth factor receptor, transforming growth factor-alpha, and c-erb-B2 in an immortalized human breast epithelial cell line, HMT-3522, is associated with decreased growth factor requirements. Cancer Res. 52, 1210–1217.

Meacham, C.E., and Morrison, S.J. (2013). Tumour heterogeneity and cancer cell plasticity. Nature 501, 328–337.

Michor, F., and Polyak, K. (2010). The origins and implications of intratumor heterogeneity. Cancer Prev. Res. (Phila) 3, 1361–1364.

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.

Nguyen, A., Yoshida, M., Goodarzi, H., and Tavazoie, S.F. (2016). Highly variable cancer subpopulations that exhibit enhanced transcriptome variability and metastatic fitness. Nat. Commun. 7, 11246.

Nunns, H., and Goentoro, L. (2018). Signaling pathways as linear transmitters. eLife 7, e33617.

Oft, M., Peli, J., Rudaz, C., Schwarz, H., Beug, H., and Reichmann, E. (1996). TGF-beta1 and Ha-RAS collaborate in modulating the phenotypic plasticity and invasiveness of epithelial tumor cells. Genes Dev. 10, 2462–2477.

Pargett, M., and Albeck, J.G. (2018). Live-cell imaging and analysis with multiple genetically encoded reporters. Curr. Protoc. Cell Biol. 78, 4.36.1–4.36.19.

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.

Rand, U., Rinas, M., Schwerk, J., Nöhren, G., Linnes, M., Kröger, A., Flossdorf, M., Kály-Kullai, K., Hauser, H., Höfer, T., and Köster, M. (2012). Multi-layered stochasticity and paracrine signal propagation shape the type-I interferon response. Mol. Syst. Biol. 8, 584.

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.

Reis-Filho, J.S., Pinheiro, C., Lambros, M.B., Milanezi, F., Carvalho, S., Savage, K., Simpson, P.T., Jones, C., Swift, S., Mackay, A., et al. (2006). EGFR amplification and lack of activating mutations in metaplastic breast carcinomas. J. Pathol. 209, 445–453.

Rizki, A., Weaver, V.M., Lee, S.Y., Rozenberg, G.I., Chin, K., Myers, C.A., Bascom, J.L., Mott, J.D., Semeiks, J.R., Grate, L.R., et al. (2008). A human breast cell model of preinvasive to invasive transition. Cancer Res. 68, 1378–1387.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. 33, 495–502.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature 473, 337–342.

Shalek, A.K., Satija, R., Shuga, J., Trombetta, J.J., Gennert, D., Lu, D., Chen, P., Gertner, R.S., Gaublomme, J.T., Yosef, N., et al. (2014). Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature 510, 363–369.

Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A., et al. (2010). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell 141, 69–80.

Shea, M.P., O'Leary, K.A., Fakhraldeen, S.A., Goffin, V., Friedl, A., Wisinski, K.B., Alexander, C.M., and Schuler, L.A. (2018). Antiestrogen therapy increases plasticity and cancer stemness of prolactin-induced ERα+ mammary carcinomas. Cancer Res. 78, 1672–1684.

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.

Stingl, J., and Caldas, C. (2007). Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. Nat. Rev. Cancer 7, 791–799.

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.

Tam, W.L., and Weinberg, R.A. (2013). The epigenetics of epithelial-mesenchymal plasticity in cancer. Nat. Med. 19, 1438–1449.

Tyson, J.J., Chen, K.C., and Novak, B. (2003). Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr. Opin. Cell Biol. 15, 221–231.

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

Wang, C.C., Bajikar, S.S., Jamal, L., Atkins, K.A., and Janes, K.A. (2014). A time- and matrix-dependent TGFBR3-JUND-KRT5 regulatory circuit in single breast epithelial cells and basal-like premalignancies. Nat. Cell Biol. 16, 345–356.

Wilson, M.Z., Ravindran, P.T., Lim, W.A., and Toettcher, J.E. (2017). Tracing information flow from erk to target gene induction reveals mechanisms of dynamic and combinatorial control. Mol. Cell 67, 757–769.e5.

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

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HMT-3522	Mina Bissell, Lawrence Berkeley National Laboratory (Briand et al., 1987; Rizki et al., 2008)	Available on request
Recombinant DNA		
Plasmid: pPBJ-EKAR3nes-puro	(Sparta et al., 2015)	Addgene # forthcoming
Plasmid: pLJM1-ERKTR-mCherry-puro	This paper	Addgene # forthcoming
Plasmid: pLJM1-ERKTR-mTurquoise2-puro	This paper	Addgene # forthcoming
Plasmid: pLJM1-ERKTR-mVenus-puro	This paper	Addgene # forthcoming
Plasmid: pAAV-Fra1-mCherry	(Gillies et al., 2017)	Addgene # forthcoming
pX330-U6-Chimeric_BB-CBh-hSpCas9	(Cong et al., 2013)	Addgene # 42230
Plasmid: pX330-FOSL1	(Gillies et al., 2017)	Addgene # forthcoming
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
Seurat	Satija Laboratory	https://satijalab.org/seurat/
R Studio	R Studio	https://rstudio.com/
Other		
Glass Bottom Plates, #1.5 cover glass	In Vitro Scientific	Cat#P24-1.5H-N, P96-1.5H-N

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 are 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-estrogen, 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% CO_2 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-EKAR-Fra-1::mCherry cell lines were generated previously using the EKAR3 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). 50ug/ml Laminin-111 was deposited onto glass wells overnight in 20mM sodium acetate buffer containing 1mM $CaCl_2$ 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 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 STATSTICAL ANALYSIS

Imaging, Data Processing, and Statistics and Normalization

For all experiments >500 cells were analyzed per condition unless otherwise indicated. Each imaging dataset presented 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$_F$*) over time divided by the mean.

$$TVI = \left| \frac{d}{dt} \left(Fra1_F \right) \right| \Big/ \overline{Fra1_F}$$

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 (*TF^P*), (2) transcription of target mRNA (*mRNA*), (3) translation of target

protein (P), and (4) potential stabilization of target protein by ERK-dependent phosphorylation (P^P). 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.

$$\frac{d}{dt}TF^P(t) = k_{pTF} * ERK(t) * (TF^T - TF^P(t)) - k_{dTF} * TF^P(t)$$ (Equation 1)

$$\frac{d}{dt}mRNA(t) = \frac{k_b + k_m * TF^P(t - \tau_m)}{\left(\frac{P(t-\tau_m) + P^P(t-\tau_m)}{K_D}\right)^v + 1} - k_{\varnothing m} * mRNA(t)$$ (Equation 2)

$$\frac{d}{dt}P(t) = k_P * mRNA(t - \tau_P) + k_{dP} * P^P(t) - (k_{\varnothing P} + k_{pP}) * P(t)$$ (Equation 3)

$$\frac{d}{dt}P^P(t) = k_{pP} * ERK(t) * P(t) - (k_{\varnothing P^P} + k_{dP}) * P^P(t)$$ (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 100ng/ml amphiregulin. Species mixing experiments were performed using mouse embryonic fibroblasts (MEFs) (Figure S2A).

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://mccarrolllab.com/dropseq/). Droplet size was determined using fluorescent beads (P=S/2%) as described (Measuring-Droplet-Volume-in-Home-Made-Devices, http://mccarrolllab.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). For each library, three test PCRs (50μl) each containing a bead equivalent of 100 STAMPs were conducted to determine the optimal cycle number for library amplification. 35μl of each test PCR were purified using Agencourt AMPure XP beads (21μl beads (0.6X) and 7μl of H_2O for elution) and the DNA concentrations were determined using a Qubit 4 Fluorometer. A concentration in between 400-1000pg/μl was taken as optimal. A variable number of PCR reactions was conducted to amplify the available 1st strand cDNA also with 100 STAMP bead equivalents per reaction with the optimal cycle number (50 μl volume; 4 + 8-11 cycles, Figure S2C). 12-μl fractions of each PCR reaction were pooled, then double-purified with 0.6X volumes of Agencourt AMPure XP beads and eluted in H_2O using 1/3rd of the bead volume. 1μl of the amplified cDNA libraries were quantified Qubit 4 Fluorometer and library size distribution verified on a BioAnalyzer High Sensitivity Chip (Agilent). 600 pg cDNA of the library was fragmented and amplified (12 cycles) using the Nextera XT (Illumina) sample preparation kit. The libraries were double-purified with 0.6× volumes of AMPure XP Beads and quantified.

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://mccarrolllab.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 (50cells/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 $\mu_{i,g}$ and variance $v_{i,g}$, then calculated gene-wise dispersion parameters $\alpha_{i,g}$ from the expression $v_{i,g} = \mu_{i,g} + \alpha_{i,g}\mu_{i,g}^2$. We observed a linear trend between log $\mu_{i,g}$ and log $\alpha_{i,g}$ 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 $E_{i,g} = (\mu_{i,g} - \hat{\mu}_{i,g})$, where the fitted overdispersion estimate $\hat{\mu}_{i,g}$ is generated by performing Local Polynomial Regression Fitting (loess) using R version 3.4.4 with span=1, simultaneously regressing log $\alpha_{i,g}$ on log $\mu_{i,g}$ and log $\alpha_{j,g}$ on log $\mu_{j,g}$ over all genes g. Statistical significance was calculated using a paired, two-sided Wilcoxon signed rank test between all $E_{i,g}$ and $E_{j,g}$. 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).