QUANTIFYING CANCER HETEROGENEITY TO PREDICT AND IMPROVE TARGETED THERAPY OUTCOMES

By

Peter L. Frick

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Chemical and Physical Biology

December, 2014

Nashville, Tennessee

Approved:

David Piston, Ph.D.

Alissa Weaver, M.D., Ph.D.

Glenn Webb, Ph.D.

Christine Lovly, M.D., Ph.D.

Vito Quaranta, M.D.

To my family:

Thank you for all that you have done for me.

ACKNOWLEDGEMENTS

This work would not have been possible without financial support from an NCI NRSA predoctoral award. I would like to thank my mentor, Dr. Vito Quaranta, for taking a chance on me and for giving me excellent support and guidance throughout my graduate training. From his mentoring I learned how to think clearly, find interesting questions, and follow the data. At a personal level, I feel privileged to be able to have a mentor whom I could trust. I would like to thank each of my committee members for their extensive support and first-rate advice throughout my training. Your insights have truly helped in my professional development and in my scientific work.

I sincerely enjoyed the opportunities to grow as a scientist in an interdisciplinary environment and to think carefully about the impact of novel data. With that in mind, I am grateful for the lab members that brought expertise from a diverse set of disciplines. Without Shawn Garbett, the modeling and quantitative emphasis of my project would not have been possible. Darren Tyson helped by providing superb ideas and input throughout my time in the lab. Lourdes Estrada helped tremendously in my predoctoral fellowship application. Jing Hao has been very helpful throughout. The other graduate students and postdocs helped me by providing helpful suggestions and a listening ear.

I also acknowledge my family in the sacrifices they made and for taking an interest in my education from a young age. My mom has always been a source of encouragement ever since I left home. I am grateful for the friends I've made in graduate school and in Nashville. Most of all I'd like to acknowledge my wife, Kerri Grove. Thank you for loving me and providing a safe place for me, free from scrutiny. Kerri, you have supported me in everything that matters to me and given me more strength than you could know.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

LIST OF ABBREVIATIONS

ALK	anaplastic lymphoma kinase
ATCC	American type culture collection
BCR-ABL	BCR-ABL fusion protein or “Philadephia chromosome”
CCLE	cancer cell line encyclopedia
cFP	clonal fractional proliferation
CHX	shorthand for cycloheximide
DIP rate	drug-induced proliferation rate
DMSO	dimethyl sulfoxide
DS	discrete subline
DTT	dithiothreitol
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
EMT	epithelial-to-mesenchymal transition
eoe	end of experiment
Erk	extracellular signal-regulated kinase
FDA	United States food and drug administration
FPM	fractional proliferation
GI50	half maximal growth-inhibitory concentration
H2B	histone H2B protein
HER2	human epidermal growth factor receptor 2
HER3	human epidermal growth factor receptor 3
HG model	heterogeneous growth model
Hsp27	heat shock protein 27

IC50	half maximal inhibitory concentration
KS-test	Kolmogorov-Smirnov test
mAg	monomeric Azami green
MAPK	mitogen-activated protein kinase
MDR	multiple drug resistance protein
Mek	mitogen-activated protein kinase kinase
mRFP	monomeric red fluorescent protein
mRNA	messenger RNA
MWA	microwestern array
NSCLC	non-small cell lung cancer
ODE	ordinary differential equation
p38	p38 mitogen-activated protein kinase
PARP	poly (ADP-ribose) polymerase
PBS	phosphate buffered saline
PI3K	phosphatidylinositol-4,5-bisphosphate 3-kinase
PVDF	polyvinylidene fluoride
QQ plot	quantile-quantile plot
Raf	rapidly accelerated fibrosarcoma kinase
RECIST	response evaluation criteria In solid tumors
RNA	ribonucleic acid
RR	response rate
RTK	receptor tyrosine kinase
SH2	src homology 2
siRNA	small interfering RNA

STAT	signal transducer and activator or transcription
TGF-α	transforming growth factor alpha
TRAIL	TNF-related apoptosis-inducing ligand
TTR	time to rebound

Chapter I

INTRODUCTION

Background

Cancer targeted therapy

Cancer is the second leading cause of death in the United States. While outcomes to cancer treatment have improved incrementally by tailoring standard of care therapy to the cancer site of origin and histology, targeted cancer therapy has shifted the paradigm of cancer treatment (Haber et al., 2011). In the past few decades it has been recognized that, despite most cancers having an accumulation of genetic and epigenetic alterations, a subset of cancers have proven to be heavily reliant on, or 'addicted to' (Weinstein et al., 2008), oncogenic signaling for growth and survival. Therefore, oncogene addiction creates a therapeutic window to rationally target cancers based on driving molecular alterations. The earliest data supporting oncogene addiction came from conditional oncogene expression in mice, where oncogene expression was sufficient to drive tumor growth and necessary for tumor cell survival (Chin et al., 1999; Felsher and Bishop, 1999). Induction of apoptosis in cell lines by targeted inhibition also demonstrates oncogene addiction *in vitro* (Tracy, 2004). Most importantly, targeting driver oncogenes clinically has improved patient outcomes. The earliest examples include two phase III clinical trials for HER2-positive breast cancers, where trastuzumab (a monoclonal antibody targeting HER2) administration improved disease-free survival in patients with previous adjuvant chemotherapy (Piccart-Gebhart et al., 2005) and increased clinical benefit for first-line chemotherapy (Slamon et al., 2001). Subsequently, oncogene addiction has been demonstrated in other cancer types, including EGFR- and

ALK driven lung cancer (Janku et al., 2010), B-Raf V600E-driven melanoma (Chapman et al., 2011), BCR-ABL driven chronic myeloid leukemia (Druker et al., 2001), and others (Weinstein et al., 2008).

While oncogene addiction is a helpful concept to understand targeted cancer treatment, it remains ill-defined at a mechanistic level. In one proposed model, the inhibition of the driver oncogene may lead to an imbalance of signaling pathways governing apoptosis and survival leading to "oncogenic shock" and tumor cell death (Sharma et al., 2007). Alternatively, activation of the driver oncogene may induce "oncogene amnesia," where oncogenic signaling overrides inbuilt tumor-intrinsic and microenvironmental controls to allow tumor growth. Upon oncogene inactivation, therefore, the tumor cells and the microenvironment regain their respective functions to prevent DNA damage and microenvironmental conditions conducive to tumor growth (Felsher et al., 2008).

Regardless of its origins, targeting oncogene addiction is effective clinically. Patient response to therapy is refined as a thirty percent reduction in unidimensional tumor size, as defined by the RECIST criteria (Eisenhauer et al., 2009). Thus effective targeted therapies induce response rates (RR) greater than that of standard of care based on the tumor site of origin. When patients fail to respond to targeted therapeutics, by definition they have *de novo* resistance, or pre-existing resistance not selected by drug treatment (Jänne et al., 2009). One source of de novo resistance is heterogeneity in mutational status of genes associated with oncogene addiction. Yet patients with the same genetic mutations still display a range of responses (Fig. 1). Understanding patient-to-patient differences, or intertumor heterogeneity, may uncover mechanisms to make therapy more effective by overcoming *de novo* resistance. The primary vexation of

targeted therapy is the transience of clinical responses: effective responses are undermined by the inevitable onset of drug resistance. This phenomenon of tumor response followed by disease progression is defined clinically as acquired resistance (Jackman et al., 2010). Variation within a tumor, or intratumor heterogeneity, underlies the eventual acquisition of drug resistance. Accounting for intratumor heterogeneity has the potential to increase the durability of targeted therapy by extending the time-to-rebound (TTR), or the time interval for a tumor to regrow to its initial cell number.

EGFR biology

The epidermal growth factor receptor (EGFR) is a 170-kDa transmembrane receptor tyrosine kinase (RTK) of the ErbB superfamily. In the absence of stimulation, EGFR resides in a monomeric autoinhibited state. Binding of its cognate ligands – including epidermal growth factor (EGF), transforming growth factor-α (TGF-α), heregulin and others – induces conformational changes in the extracellular domain to that allow it to dimerize with itself or other ErbB family members (Burgess et al., 2003). The dimer pair is the functional unit of ErbB receptor signaling (Citri and Yarden, 2006), where receptors phosphorlylate intracellular tyrosine residues of the paired RTK. These phosphorylation events serve as binding sites for Src Homolgy 2 (SH2) domains of adaptor proteins that subsequently activate downstream intracellular signaling cascades, primarily through the PI3K-Akt, Ras-Raf-Mek-Erk mitogen activated protein kinase (MAPK), and signal transducer and activator or transcription (STAT) pathways (Fig. 1). These pathways ultimately govern gene transcriptional programs that govern cellular behavior. In the case of cancer, some of the implicated cellular behaviors are proliferation, induction of angiogenesis, motility, differentiation, and survival (da Cunha Santos et al., 2011).

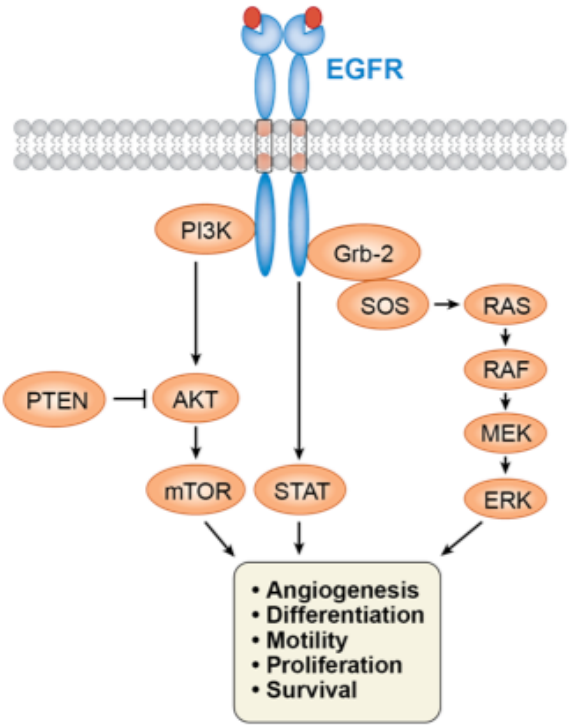

Figure 1 Aberrant EGFR signaling in cancer. Adapted with permission from (da Cunha Santos et al., 2011). Ligand stimulation activates EGFR on the cell membrane and signal transduction propagates through key intracellular signaling cascades, including PI3K-Akt, Stat, and Ras-Raf-Mek-Erk. These activated pathways initiate downstream transcriptional programs that promote cancer-related phenotypes, as listed at the bottom of the figure.

In contrast to a linear view of signaling, the EGFR pathway has several features that make it robust a robust system. These include both component redundancy and a conserved core network architecture with diverse outputs and inputs (Citri and Yarden, 2006). Several cellular control mechanisms modify EGFR signaling. For example, both positive and negative feedback loops regulate the amplitude and duration of molecular signals (Citri and Yarden, 2006). Autocrine and paracrine loops provide positive feedback, where pathway activation increases expression of EGFR cognate ligands (Singh and Harris, 2005). Alternatively, activation of EGFR signaling triggers negative feedback loops that amplifies expression of inhibitory proteins (Nicholson et al., 2005). Other competing cellular processes regulate EGFR signaling, including endocytosis and proteasomal degradation (Avraham and Yarden, 2011). Thus EGFR signaling is a robust, tightly regulated system across multiple time scales.

EGFR-addicted lung cancer

This work focuses mainly on EGFR-driven lung cancer as a model for targeted therapy because it aims to quantify heterogeneity and many of the molecular mechanisms underlying heterogeneous response to targeted therapy have been elucidated. Despite efforts to optimize treatment regimens for lung cancer, most patients present with metastatic cancer, with a one-year survival rate of less than ten percent (Sharma et al., 2007). However, a subset of patients (~10%) have tumors with mutations in the activation loop of EGFR that disrupts autoinhibitory interactions (Okabe et al., 2007). These gene mutations, primarily either exon-19 deletions (45%) or an L858R missense (40-45%) mutation (Sharma et al., 2007), encode a constitutively active EGFR-mutant receptor. Additionally, these mutations in EGFR increase its affinity for erlotinib and gefitinib (Pao and Chmielecki, 2010), both small molecule tyrosine kinase

inhibitors that specifically target EGFR by competitively binding to the ATP-pocket of EGFR. Gefitinib and erlotinib exploit this therapeutic window and induce favorable clinical responses (Lynch et al., 2004; Paez, 2004; Pao et al., 2004). Erlotinib therapy induces responses in roughly 70% patients with EGFR-addicted tumors (Chmielecki et al., 2011). It is FDA-approved for monotherapy as a second-line treatment in non-small cell lung cancer (NSCLC) (Linardou et al., 2009) as well as maintenance therapy (Gridelli et al., 2007). Erlotinib is administered at the maximally tolerated dose of 150 mg/day, corresponding to a mean plasma steady-state concentration of roughly 3 μM (Hidalgo et al., 2001). In this work, therefore, erlotinib is administered *in vitro* at 3 μM, roughly 100-fold higher than reported IC50 values of EGFR-addicted cell lines (Chin et al., 2008).

As in all targeted cancer therapies, patients responsive to erlotinib or gefitinib exhibit disease progression after about one year of continuous treatment as tumors become resistant (Chong and Jänne, 2013). While variable response to drug may derive from multiple sources, as described throughout the rest of the introduction, genetic variation is most evident in patients. Resected tumors that relapse under continuous therapy contain a set of genetic alterations, most notably (in ~50% of cases) a second-site threonine-to-methionine gatekeeper mutation (T790M) in the tyrosine kinase domain of EGFR that prevents drug binding (Sharma et al., 2007).

Intertumor heterogeneity

Patients harboring tumors with similar mutational backgrounds display a wide range of responsiveness to targeted therapy. Effective targeted therapies induce clinical responses in the majority in patients. In the case of EGFR targeted therapy, RR varies between 55-91%, depending on the study (Pao and Chmielecki, 2010). Waterfall plots

show extensive patient-to-patient variation in reduction of tumor size, even within responsive tumors (Fig. 2). Understanding the molecular factors that underlie intertumor heterogeneity may optimize responsiveness to novel therapeutic strategies. Cell lines represent an attractive model system because of widespread availability and potential for in-depth molecular characterization, though a sufficient amount of cell lines would be needed to better represent the genetic heterogeneity seen clinically (Sharma et al., 2010a). A notable example of comparing differences between grouped samples is the Cancer Cell Line Encyclopedia (CCLE), which collected genotypic and phenotypic cell line measurements across a variety of cancer types to enable predictions of drug sensitivity (Barretina et al., 2012). Alternatively, systematically varying genetic components within a cell line may reveal molecular signaling events that modulate drug sensitivity. For instance, an siRNA screen of a mutant-EGFR cell line identified targetable molecular alterations underlying *de novo* resistance (Bivona et al., 2011). In another gain-of-function genetic screening study in a melanoma model of oncogene addiction, assessing drug sensitivity while selectively inducing individual gene expression revealed a novel resistance pathway to targeted therapy (Johannessen et al., 2013). Having divided oncogenic reliance to two or more pathways is thus a method for cancer cells to attain *de novo* resistance.

Other more in-depth molecular studies clarify how complex signaling network wiring reduces drug effectiveness. Many canonical signaling pathways share common components and interact with each other (Mendoza et al., 2011). Furthermore, targeted drug treatment may itself initiate bypass signaling in a parallel pathway by relieving negative feedback of other signaling pathways (Chandarlapaty, 2012). Targeting an RTK may raise activation of a parallel pathway above baseline levels(Fan et al., 2011). This holds for targeting downstream components, as in the case with breast cancer, where

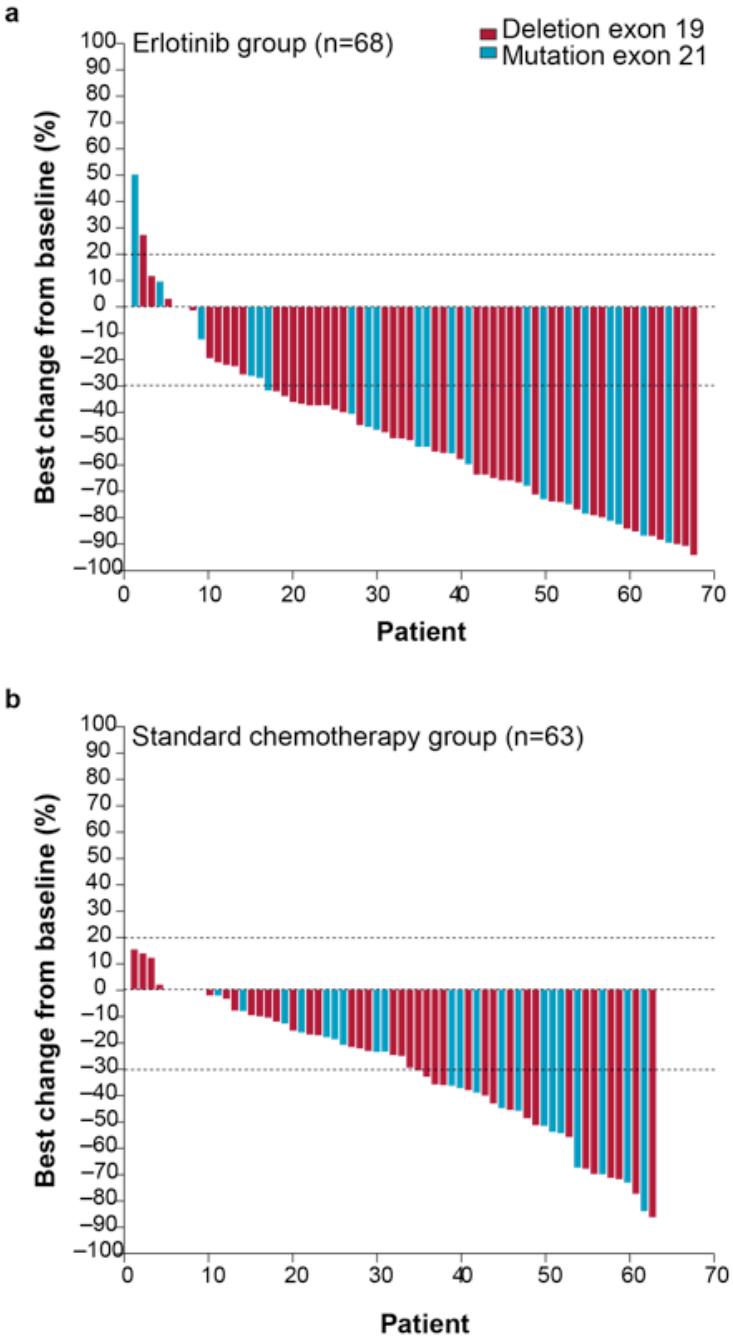

Figure 2 Erlotinib waterfall plots. Adapted with permission from (Rosell et al., 2012). (a) Waterfall plots represent the best-observed percent change in unidimensional tumor size during erlotinib treatment. Each vertical line describes the change in tumor size for an individual patient response. Colors represent which sensitizing EGFR mutation was detected in each patient. Dashed lines represent clinical cutoffs defining responsiveness according to the RECIST criteria (Eisenhauer et al., 2009). By definition, lines extending below the bottom dashed line respond to treatment. If above the top line, the patient exhibited disease progression while on treatment. If between the lines, the treatment elicited no response. (b) Same as *a*, but with standard of care chemotherapy.

inhibiting PI3K induces transcriptional upregulation of HER3 (Chakrabarty et al., 2012). Alternatively, positive feedback loops can sustain network activation in the absence of input signaling (Tyson et al., 2003). Therefore the complex feedback mechanisms that retain network robustness can provide a way for cancer cells to quickly adapt to perturbation and confer *de novo* resistance.

Intratumor heterogeneity

Despite the promise of targeted cancer therapy, patients inevitably progress after roughly a year of treatment and tumors rebound after an initial response. While microenvironmental interactions influence tumor cell drug sensitivity (Hanahan and Coussens, 2012), phenotypic heterogeneity between cancer cells, or intratumor heterogeneity, has received growing attention based on the role it plays in undermining targeted therapy (Almendro et al., 2013). Intratumor cancer heterogeneity has been recognized for decades, as exemplified by the seminal work of Fidler (Fidler and Kripke, 1977) and Heppner (Dexter et al., 1978). Single-cell technologies have shown that variability within a cancer cell population is widespread at the level of genetics (Navin et al., 2011), cell signaling (Bendall et al., 2011), and cell behavior (Gascoigne and Taylor, 2008). A challenge is to relate disparate and extensive heterogeneity to actual treatment outcomes. Genetics is widely considered the foremost culprit to therapeutic failure, because it is evident in resected clinical samples. But it may be that the fixed genetic changes arise due to early contributions by other types of heterogeneity, since different types of variation are inherited on different time scales (Rando and Verstrepen, 2007). Different sources of intratumor heterogeneity include clonal, non-genetic, and stochastic.

Clonal heterogeneity

The concept of clonal evolution was originally articulated by Nowell in 1976 (Nowell, 1976), where a tumor is a composite of phenotypically variable clones arising due to the expansion of a genetically unstable cell population (McGranahan et al., 2012). Within this framework, a tumor adapts to therapy as a function of "the frequency of mutations being produced and on the environmental pressures [such as drug selection]" (Nowell, 1976). More recent works have highlighted the importance of genomic instability (Loeb, 2011). Indeed, sequencing multiple regions of a clear-cell carcinoma displayed branched clonal evolution, leading to heterogeneity of signaling events governing cell growth (Gerlinger et al., 2012). In acute myeloid leukemia, whole-genome sequencing revealed that the major clone driving relapse to chemotherapy had acquired relapse-specific mutations (Ding et al., 2012). In the case of targeted therapy, genetically analyzing tumor tissue from patients showing disease progression has established common mechanisms underlying acquired resistance (Doebele et al., 2012; Sequist et al., 2011). Alternatively, relevant resistance mechanisms may be discovered by generating acquired resistance by culturing cell lines with increasing doses of drugs (Ohashi et al., 2012; Turke et al., 2010). Discovery of genetic mechanisms of acquired resistance has guided the development of drugs that specifically target mutated proteins underlying mechanisms of acquired resistance (Zhou et al., 2009). Additionally, real-time histological and genetic observations can guide clinical therapy based on the evolution of clonal composition (Sequist et al., 2011).

Clearly clonal genetic composition has an integral role in treatment response outcomes. Yet there are several challenges to considering only the resistant clone that emerges after about a year of therapy. Increasing evidence suggests that intratumor

heterogeneity is likely shaped by both genetic and non-genetic factors (Fig. 3). For example, multiple clones with different mechanisms of resistance may coexist within an oncogene-addicted cell population (Suda et al., 2010). Even without treatment selection, single-cell derived clones from a lung cancer cell line have variable cellular signaling states and sensitivity to chemotherapy (Singh et al., 2010). Cancer cells have remarkable ability to adapt to therapies designed to overcome acquired resistance. For example, cancer cells chronically treated with an inhibitor specifically designed to target EGFR T790M still eventually develop resistance (Ercan et al., 2012). Also, phenotypic clonal heterogeneity may derive from copy-number variations (Anaka et al., 2013) or non-genetic means (Kreso et al., 2013) as well. Another complicating factor is that measurements of clonal composition are almost always a static measurement after resistance is there already. Thus, the complex process of clonal selection is informed by clonal behavior that is inferred and not directly measured.

Non-genetic heterogeneity

Non-genetic heterogeneity has also been implicated in intratumor heterogeneity. A prominent example is epigenetic maintenance of minor cell subpopulations with reduced drug sensitivity. For example, isolated leukemia cells with high expression of MDR can repopulate the full MDR expression distribution; additionally they have decreased sensitivity to chemotherapy (Pisco et al., 2013). Sharma et al. showed that sensitive cell lines maintain a minor fraction of cells that undergo epigenetic modifications to enable them to continue cycling in drug after transitioning through a cell cycle arrest of roughly nine days. Interestingly, these drug-tolerant cells were positive for CD133, a marker of stem cells (Sharma et al., 2010b). While cancer stem cells are thought to underlie relapse to chemotherapy in some contexts

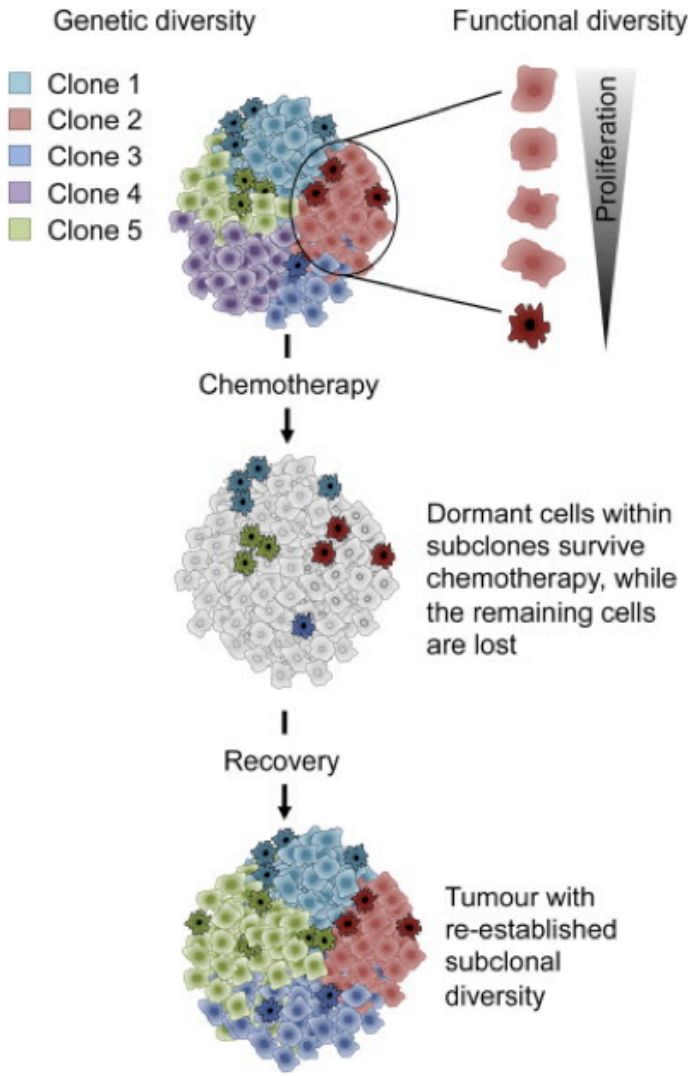

Figure 3 Genetic and epigenetic clonal dynamics during therapy. Reproduced with permission (Kreso et al., 2013). A tumor comprised of multiple genetically-distinct clones. *Top*, A multi-clonal tumor. Even genetically identical cells contain clones with unique proliferation rate. Middle, Upon chemotherapy, the majority of cells die (grey), while the remaining colored cells reside in a dormant state. *Bottom*, Upon removal of drug treatment, the tumor cells dormant during chemotherapy repopulate the tumor population with a unique clonal profile.

(Merlos-Suárez et al., 2011), their contribution to targeted therapy acquired resistance remains unclear. In melanoma, phenotype switching itself appears targetable, as co-treatment with methotrexate induces cells to transition out of a drug-tolerant transcriptional program to make more cells susceptible to targeted inhibition (Sáez-Ayala et al., 2013). Similar observations of phenotypic switching as a survival strategy are evident in bacteria (Balaban et al., 2004). One possible explanation of maintenance of multiple distinct phenotypes is that these represent local minima in the epigenetic landscape (Brock et al., 2009). Clinically, anti-EGFR therapy in non-small cell lung cancer induces a transition to small cell lung cancer phenotype in ~14% of cases (Sequist et al., 2011), though the mechanism of this transition remains unknown. Additionally, roughly five percent of acquired resistance cases in EGFR-addicted cancer display histology reminiscent of epithelial-to-mesenchymal transition (Sequist et al., 2011).

Cellular heterogeneity

Another kind of heterogeneity is cellular heterogeneity, which describes phenotypic variability between seemingly identical cells arising from stochastic biological processes (Altschuler and Wu, 2010). Cellular heterogeneity originates from gene expression noise or from altered signaling states due to random variation within a genetically homogeneous cell population (Slack et al., 2008). Because of these small stochastic cell-to-cell differences, critical decision-making processes such as response to stress (Snijder et al., 2009), differentiation (Park et al., 2012), or cell cycle progression (Yao et al., 2008), may exhibit multiple outcomes at the single-cell level, not represented by the average behavior of the cell population. Live-cell imaging studies reveal that cancer cell populations respond to drug treatment with multiple individual cell fates,

including cell death, cell division, and entry into quiescence. Cancer cells treated with chemotherapy display cell fate heterogeneity explained by competing cell cycle and cell death signaling pathways (Gascoigne and Taylor, 2008). Sibling pairs of HeLa cells treated with the TRAIL ligand, which induces extrinsic apoptosis, are more correlated in their response relative to the rest of the population, though this correlation decays over time (Spencer et al., 2009). The heritability of apoptotic response may derive in part from the robustness of the network governing the cellular behavior (Citri and Yarden, 2006). Inheritance of drug response similarity between siblings may underlie variable clinical responses, although its relevance has been questioned because the similarity decays over time (Marusyk and Polyak, 2013).

Challenges of measuring heterogeneity

As described above, intratumor heterogeneity is widespread and affects the response of a cell population. Thus an initially favorable response is undermined as therapy selects for high fitness cells within a variable population. A major challenge of relating cell-to-cell heterogeneity to treatment outcomes is the difficulty of knowing which biological differences are functionally important (Altschuler and Wu, 2010). Meeting this challenge requires assays capable of identifying relevant cell subpopulations and uncovering their phenotype in relation to the overall population.

Several experimental limitations obscure detection of cells responsible for tumor progression. First, most biological measurements are averaged, as in Western blotting or genetic sequencing. Averaging a biological measurement across a cell population masks the distribution of the data. It could be that the dominant majority of cells masks the signal from the relevant small subset of cells altogether. Or, the population may be represented by two separable subpopulations and the average may thus describe an

intermediate state that actually represents none of the cells present (Altschuler and Wu, 2010). Additionally, even if the cell population is described by a unimodal distribution, cells selected from the tails of that distribution, while statistically insignificantly different, have biologically significant functional differences (Chang et al., 2008). Finally, heterogeneous single-cell fates may follow a skewed distribution that is not appropriately described by a mean. For example, the time for single cells to pass the G1/S transition when deprived of growth factor signaling is described by an exponentially-modified Gaussian distribution (Tyson et al., 2012). Therefore a full representation of the data, rather than simply reporting the population average behavior, best describes the heterogeneous response of cells to perturbation.

An additional consideration is that many assays infer cell fate responses based on indirect measurements. This is particularly true in targeted therapy, where measuring concentration-dependent growth inhibition assesses drug sensitivity. This methodology is well-suited to high throughput applications, e.g., the Cancer Cell Line Encyclopedia (Barretina et al., 2012). But the interpretations may overextend the data. First, these assays rely on metabolic readouts as an indirect measure of the number of cells present; thus drug effects on proliferation may be convoluted with metabolic response. Also, the half-maximal concentration of growth inhibition (GI50) is often referred to as fifty-percent cell killing. However, relative to an exponentially growing control, it is possible that the GI50 effect is driven solely by quiescence and contains no cell death at all (Tyson et al., 2012). Directly observing cells using live-cell imaging shows that perturbations may induce a complex mixture of cell fates (Gascoigne and Taylor, 2008). Relating these single-cell behaviors to changes in population size helps to quantify the functional contribution of cell fate to a complex dynamic system (Tyson et al., 2012). Yet it remains

unclear how well this approach holds for long-term effects such as cancer cell rebound to therapy.

Systems biology to study targeted therapy

Tumor biology crosses many scales, including: genetics, gene expression, intracellular signaling, single-cell behavior, clonal lineages, microenvironmental interactions, and patient response (Fig. 4). At each scale, tumors display variability in respond to drug treatment. This variability underlies resistance and disease progression. Major effort continues to be invested in elucidating the spectrum of heterogeneous biological mechanisms underlying drug response. Less developed is a framework to quantitatively understand the functional contribution of heterogeneity, especially across scales. Systems biology is an approach that studies complex biological phenomena as a result of the interactions of its constituent parts. While reductionist approaches have proven remarkably successful in defining mechanisms of sensitivity and resistance to targeted therapy, systems biology may help to explain the impact of heterogeneity in undermining therapeutic response to anti-cancer therapy.

A major tool of systems biology is mathematical modeling. A model is a simplifying abstraction to understand complex phenomena. All biologists use models to formulate hypotheses and interpret data. Mathematical modeling is simply a way of translating understanding from qualitative to quantitative (Kohl et al., 2010). Thus modeling is particularly useful for phenomena with sufficient complexity to exceed intuition. In the case of cancer, heterogeneity is widespread and it is imperative to determine which differences are functionally relevant. By assigning biological traits quantitative values, mathematical modeling may help to decipher which biological parameters are most important.

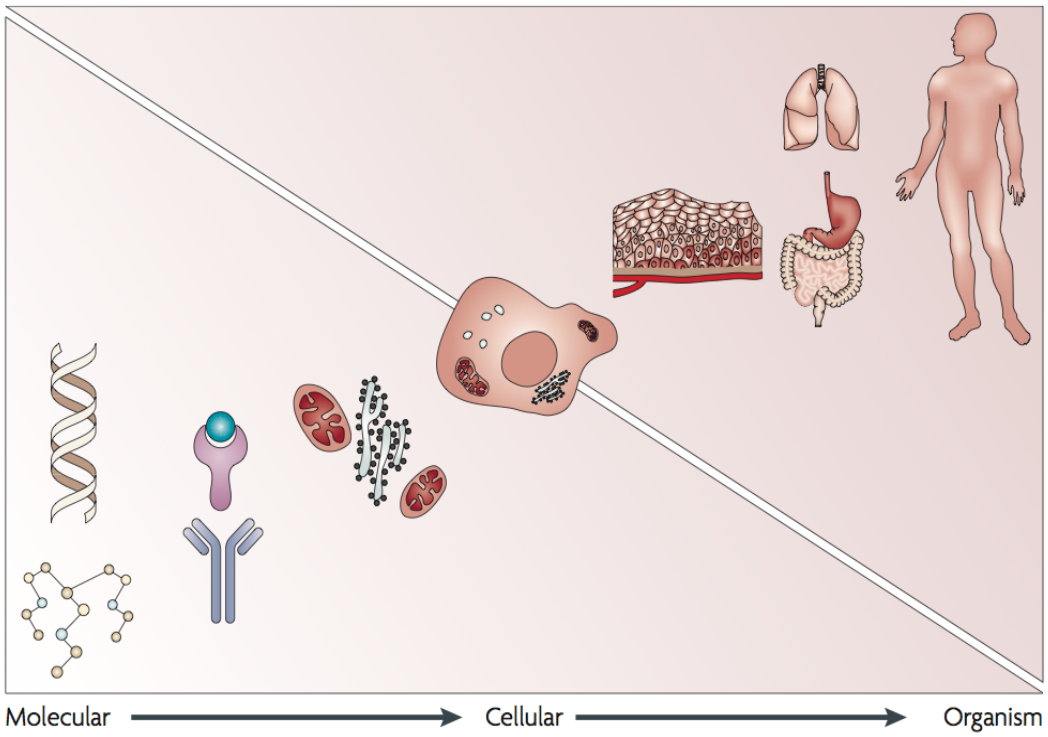

Figure 4 Multiscale biology. Reproduced with permission (Anderson and Quaranta, 2008). Biological systems cross spatial scales from genes to molecules. Systems biology approaches attempt to bridge scales in order to understand complex biological phenomena. DNA encodes molecules that govern cellular signals ultimately guiding cell behavior. Cells, in turn, interact at the level of tissues, and subsequently organisms. The cell is an excellent candidate for scale of investigation to link molecular reductionist approaches with organism-level behavior.

Beginning from theory to find interesting biology not apparent from empirical data is one strategy. For example, beginning with a set of rules governing cell growth based on microenvironment, Anderson et al. used mathematical modeling to predict that harsh microenvironmental conditions counterintuitively select for invasive tumor morphologies (Anderson et al., 2006). These model simulations then give experimentally testable predictions that can refine the model. Another successful example of predicted biological behavior is the decision of single cells to pass the restriction point, a cell cycle check point in late G1 phase, and divide. Math models predicted for cells to respond to growth stimuli with all-or-nothing bistable behavior; and the authors showed this experimentally for the first time (Yao et al., 2008).

Another modeling use is to assume knowledge of the underlying biology and determine the parameter space that fits the data. For example, Michor et al showed that a model of cellular response to chemotherapy governed by response rates of a differentiation hierarchy can explain the response and regrowth of clinical chronic myeloid leukemia data (Michor et al., 2005). Additionally, based on published mutation rate data, the model makes estimations of the number of initially present genetically resistant cells. It should be noted that the model fits represent inferences based on assumptions built into the model. So they should be viewed as a way to assess explanatory plausibility (Kohl et al., 2010) rather than empirical causation. Experimental data or literature-based knowledge can improve confidence in the model. Otherwise experimental data can constrain models to rule out alternative biological explanations. For example, Almendro et al. used modeling to conclude there was evidence of phenotype switching by ruling out clonal expansion (Almendro et al., 2014a).

Alternatively, data-driven models can be used to understand the meaning of novel biological measurements. In one case, Tyson et al. used an experimental-modeling approach to bridge single-cell and cell population responses (Tyson et al., 2012). With this framework, the contribution of single-cell behavior could be quantitatively linked to the change in population size. In another example of data-driven modeling, Lee et al. defined a sequential drug combination scheme that induces increased apoptosis in triple-negative breast cancer (Lee et al., 2012). Then, using a systems modeling approach on high-throughput molecular data, they found that the extrinsic apoptosis pathway is critical in rewiring the signaling state of cells to alter the DNA damage response. These approaches gave novel biological insights that relied upon rich cellular and molecular datasets to build and populate the model.

These examples, demonstrate how modeling can be combined with experimental data to generate and test hypotheses. Modeling can help to both interpret new biology and make new predictions based on existing knowledge. In each of these cases, modeling was used to integrate multiple interacting components to understand complex system behavior and bridge biological scales, from molecules to patients. Therefore intratumor heterogeneity, which bridges scales and contains multiple interacting components, is ideally suited to investigation by systems biology.

Purpose of this study

Cancer heterogeneity adversely affects patient outcomes to targeted therapy. Several studies have comprehensively described heterogeneity in cancer cell lines (Gascoigne and Taylor, 2008) and in patient tumors longitudinally monitored both in mice (Kreso et al., 2013) and in humans (Sequist et al., 2011). The origins of this phenotypic variability derive from sources that operate over broadly different timescales

(Rando and Verstrepen, 2007). However, few studies have sought to measure multiple sources of heterogeneity simultaneously. Additionally, these studies qualitatively interpret the contribution of heterogeneity retrospectively. Ideally, quantifying heterogeneity should give insight into how a cancer cell population will respond to perturbations. The fundamental purpose of this study is to define a quantitative link between heterogeneity and outcomes for cancer cell populations. Therefore we use a combined experimental and computational systems biology approach, suitable for understanding complex systems with emergent behavior arising from interacting components. As a model system we use live-cell imaging of EGFR-mutant lung cancer cells. These cell lines are favorable because they display all of the types of heterogeneity enumerated throughout the introduction. Furthermore, they maintain genetic and functional features similar to patient tumors. Live-cell imaging is amenable to the quantitative, data-rich experiments needed for systems biology approaches (Pepperkok and Ellenberg, 2006).

Chapter II

QUANTIFYING HETEROGENOUS POPULATION RESPONSE TO PERTURBATIONS USING CLONAL FRACTIONAL PROLIFERATION

Introduction

Proliferation is a fundamental property of living cells whose function, at it simplest level of consideration, is to produce a population-level increase in cell number. Measuring proliferation, or lack thereof, is essential to studying a wide swath of normal and pathological biological processes, including tissue homeostasis, cell differentiation, degenerative diseases and cancer progression. Such measurements are best performed *in vitro*, with assays that evaluate the proliferative response of a cell population to a perturbagen in a culture vessel. Historically, each data point in these *in vitro* assays is the average of the proliferative response of thousands of cells in a perturbagen-treated well, at single or few time points, in comparison to untreated control. However, fundamentally, the proliferative response of a cell population is the composite of the behavior of individual cells, which is often heterogeneous in terms of cell fates (Gascoigne and Taylor, 2008; Tyson et al., 2012). Furthermore, proliferation is a dynamic process that, by its very nature, would be best quantified as a rate (e.g., population doublings per unit of time), not as fold change with respect to control. Thus, current assays provide little information on two aspects of cell proliferation, heterogeneity and dynamics, which are key to consider when evaluating clonal evolution or fitness in the context of cancer progression or treatment.

We have previously shown that analyzing proliferative dynamics at the single-cell level provides a more complete characterization of the population-level proliferation as

shaped by the multiple, heterogeneous single-cell behaviors induced by perturbagens (Tyson et al., 2012). Increasing evidence indicates that both are influenced by variability amongst clonal lineages, i.e., tumor cells sharing a recent common ancestor (Huang, 2009). For example, proliferating human colorectal cancers may be comprised of clones with unique fitness properties, with no apparent genetic basis, in the presence or absence of chemotherapy (Kreso et al., 2013). Similarly, patient-derived metastatic melanoma clonal cell lines recapitulate *in vivo* tumor clonal heterogeneity (Anaka et al., 2013) since, cultured *in vitro*, they vary in both chemotherapy sensitivity and colony-formation potential. Thus, to properly evaluate population-level responses to perturbations, it is essential to examine the behavior of clones comprising that cancer cell population.

Further, it is likely that, even *within* clonal lineages, single-cell differences at the signaling level underlie the overall proliferation behavior of the clone. For example, single breast epithelial cells undergo multiple fates to form acinar structures that are driven by different transcriptional profiles (Wang et al., 2011). In the absence of microenvironmental differences, sensitivity to chemotherapy varies between breast cancer clones, but each clone is comprised of multiple signaling states (Singh et al., 2010). Nonetheless, methods to relate cell fate heterogeneity to the clonal structure of a population have not been developed. Such methods would be of great utility to understand the impact of cell fate propensity to clonal fitness.

We have previously shown (Tyson et al., 2012) that the average proliferation rate of a cell population in response to perturbations can be quantitatively deconvolved into the rates at which single cells enter specific fates (quiescence, apoptosis, changed division rate). With this technique, named Fractional Proliferation method (FPM), we

demonstrated that the effects of a perturbation on cell population size depend on the underlying perturbation-induced single cell fates, validated by high frequency single-cell tracking measurements (~10 min). We further showed that analyzing proliferative dynamics at the single-cell level provides a more complete characterization of the population-level proliferation as shaped by the multiple, heterogeneous single-cell behaviors induced by perturbagens. In this present work, we sought to relate population-averaged response and single cell fates to clonal lineages that constitute a cancer cell population. To this end we developed a methodology termed clonal Fractional Proliferation (cFP). We show that cFP assays capture both the single-cell fate heterogeneity within a clone, and clone-to-clone variability of proliferation in response to perturbations in established cancer cell lines. Furthermore, we relate clonal variability to cell morphology.

Results

High-throughput measurements of clonal behavior

The clonal Fractional Proliferation (cFP) assay is based upon high-throughput imaging methods designed to track cell numbers of many clonally-derived colonies over time (Fig. 5a). Proliferation of a colony is measured by direct, repeat counting at regular intervals of cell nuclei labeled with a live-cell fluorescent reporter (H2B-mRFP) (Tyson et al., 2012). Nuclear labeling greatly simplifies segmentation and enables efficient and accurate image processing. To scale up the throughput of biological samples, we chose to use the CellaVista High-End imager (Synentech) because of its even fluorescence illumination and an LED-based excitation source that minimizes phototoxic effects.

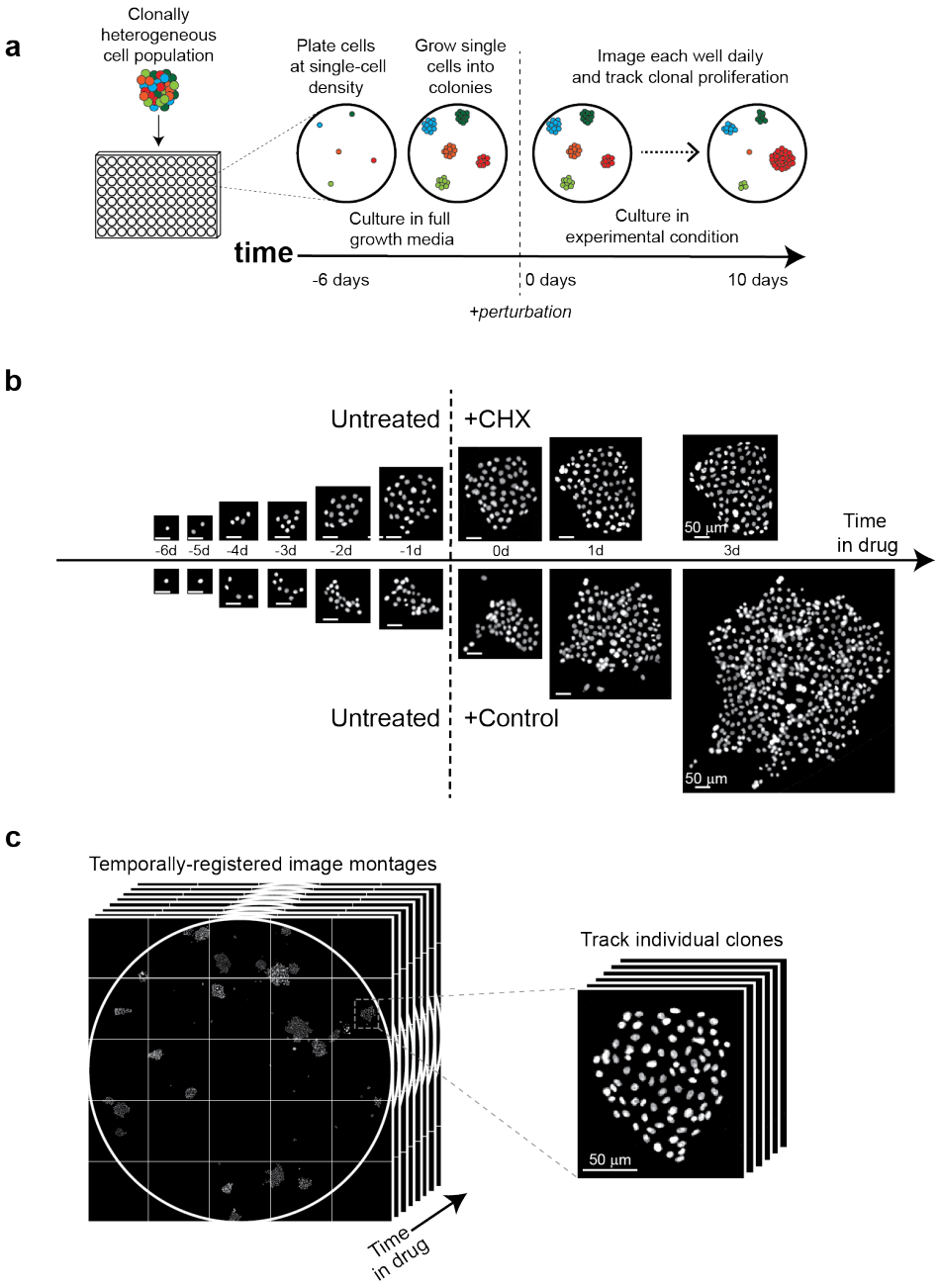

Figure 5 Schematic of clonal Fractional Proliferation experimental workflow. (a) cFP is a high-throughput, fluorescence-based assay designed to measure clonal proliferation variability within a cell population. Cells are seeded at single-cell density into microtiter imaging plates. Single-cells proliferate for 6 days in full growth media to expand into single-cell derived colonies. Once colonies reach an optimal size, cells are imaged and then the experimental perturbation is immediately added. Subsequently, each well is imaged daily until the end of the experiment. (b) Spatially and temporally registered images facilitate tracking of individual colonies. Time-ordered stacks of image montages allow sequential measurements of colony cell numbers during drug treatment. (c) Fluorescent images of cell nuclei for single-cell derived colonies. Sequential images show that single-cells form colonies when cultured in full growth media (Untreated). Tracking colonies after addition of cycloheximide (CHX) or DMSO (control), individual colonies shows drug-induced changes in cell number over time.

In setting up the cFP assay, we aimed to balance several prerequisites. We reasoned that if the cell population response to a perturbation is clonally variable, then it is imperative to quantify as many clones as feasible per experiment in order to obtain a statistically meaningful representation of the range and frequency (diversity) of clonal response. Additionally, individual colonies must contain a sufficiently high number of cells prior to treatment in order to minimize error in calculating rates due to small cell number counts. However, to ensure that colonies are in fact clonally derived, the cell population must be sparsely plated and the assay terminated prior to colony confluence. On balance, we found that using a 96-well plate format (Fig. 5a), and plating ~40 cells/well are optimal initial conditions to obtain single-cell derived colonies (Fig. 5b). The colonies are allowed to grow for six days in complete growth media, and subsequently subjected to a continuous experimental perturbation, during which the entire well is imaged daily for ten days (Fig. 5a). Plating efficiency is sufficiently high so that information on ~200 colonies per experimental condition can be obtained from 3-8 replicate wells.

In order to facilitate tracking of colonies throughout the duration of the perturbation, unordered image acquisition is spatially registered and temporally organized (Fig. 5c). To this end, we use the freely available ImageJ software (http://imagej.nih.gov/ij/, version 1.48i). First, the subset of all images belonging to a single well at a single time point is considered. Then, images are spatially ordered based on acquisition time and converted to a stitched image montage of the entire well. This step is repeated for all time points and all the montages are ordered by acquisition time, resulting in an image stack of wells over the course of the experimental treatment (Fig. 5c).

To facilitate quantifying proliferation response of individual clones, we developed an image processing routine using ImageJ to count the total number of cell nuclei at each time point from registered image stacks. Fluorescence intensity masks were generated for each colony by applying a user-defined threshold to the pixel intensity histogram computed from all images unique to that colony. These masks show that the fluorescence from cell nuclei is sufficiently high above background to detect cells. However, colonies remain under-segmented, compared to manual counting. To correct for this, the watershed segmentation algorithm was implemented to distinguish nearby nuclei. Then the “analyze particles” command scans the image stack for elements with optimized morphological parameters and returns the number of identified cell nuclei and the corresponding images

This technique was used to serially identify cell nuclei throughout a representative experiment (Fig. 6a). To quantify how well the automated image segmentation represented the actual cell number at each time point, we manually counted the total cell nuclei per colony at 219 total time points. The manual counts are highly correlated (adjusted $R^2 = 0.99$) with the automated cell counts (Fig. 6b) showing that automated cell counting is a faithful representation of cell number while easing the burdensome task of manually counting cells. Additionally, the residual errors of the linear model fit do not show evidence of bias (Fig. 6c) and the standard deviation of the residuals is over fifty times smaller than the smallest cell number of any colony.

Having validated the ability to quantify colony size by cFP, we set out to measure the clonal variability within a population in response to the drug cycloheximide, an inhibitor of protein translation well known to severely stunt cell proliferation. We tracked PC9 colonies for six days untreated, followed by three days in the presence or absence

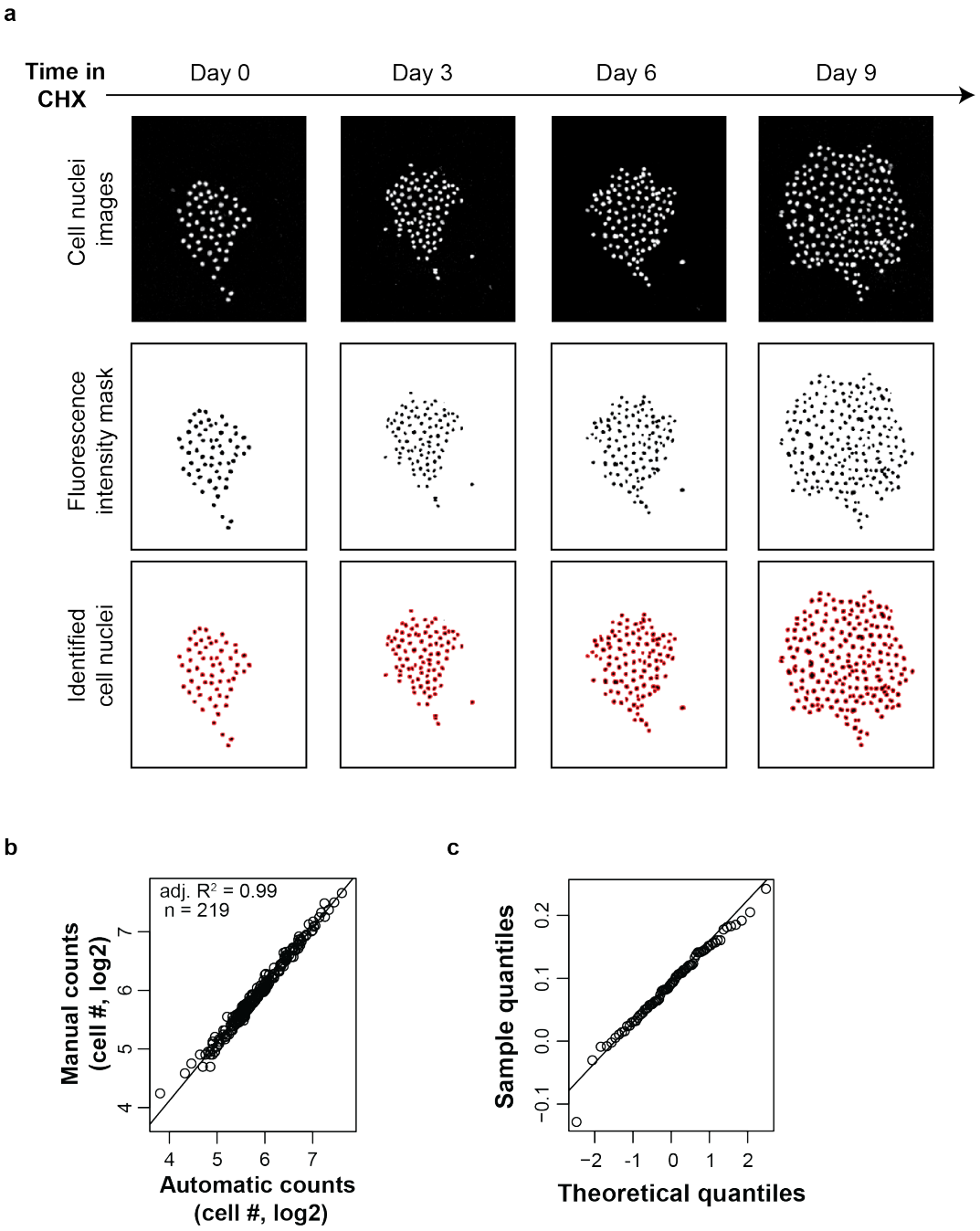

Figure 6 Validation of cFP image processing. (a) Top row: Images of a representative colony throughout cycloheximide treatment (CHX) at the indicated days. Middle row: Binary mask generated in ImageJ using the same intensity threshold at all time points. Bottom row: (b) Automated counting of cell nuclei appropriately quantifies colony cell number. Manual cell counts from colony images at various time points (n = 219) are used as a reference to validate the automatically measured colony cell numbers. The superimposed line represents the linear model fit for the data, with the corresponding adjusted R-squared value (adj. R^2) (c) Q-Q plot of the residuals of the linear model fit in used in *b*. There is insufficient bias to conclude that automated image analysis is inappropriate for colonies of certain size.

of 500 ng/ml cycloheximide. Serial images of individual colonies demonstrate that the size of colonies rapidly diverges after drug addition, and that cFP captures these proliferation differences (two representative, diverging colonies are shown in Fig. 5b).

DIP rates summarize dynamics clonal behavior

Data from large numbers of processed colonies indicate that control-treated colonies exhibit exponential growth, evidenced by an averaged log-linear plot of population doublings (Fig. 7a, dashed line). In the presence of cycloheximide, the mean proliferation of PC9 colonies is greatly reduced relative to control (Fig. 7a, solid line). A potential problem with these data is the progressive increase in the standard deviation (Fig. 7a, error bars) of colony size measurement. To investigate this more in depth, we examined the distribution of individual colony responses. Consistent with the increasing standard deviation in colony size (Fig. 7a), the interquartile range of colony sizes increases over time (Fig. 7b, boxplots). This suggests that the increasing measured error may be due to divergent colony responses. Indeed, individual colonies respond to cycloheximide uniquely (representative colony dynamics are quantified by colored lines in Fig. 7b and visualized in Fig. 7c). Notably, though colony dynamics appear complex, after three days of treatment they appear to gain a steady rate of growth. Thus we reasoned that the steady rate of proliferation for each colony would accurately simplify the dynamic colony response data. Therefore we normalized the colony response data to the colony size at 3 days (Fig. 7d), obtained linear model fits to determine the proliferation rate for each colony, and plotted the estimated colony dynamics based on proliferation rate alone (Fig. 7e). The estimated behavior of colonies (Fig. 7e, gray and colored lines) diverged over time and the range of colony responses (Fig. 7e boxplots) closely matched the range of actual colony responses

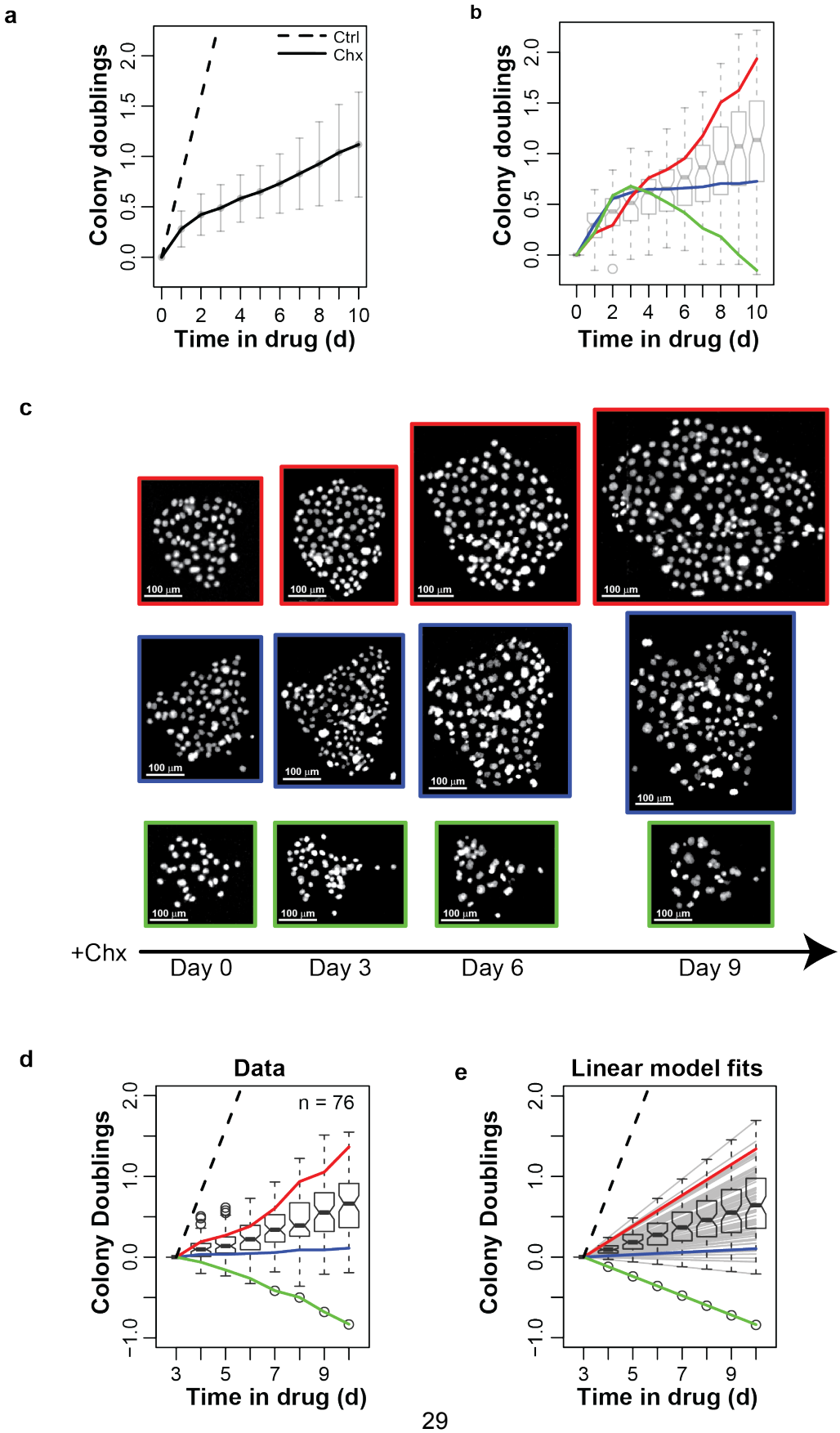
a
Colony doublings
Time in drug (d)
Ctrl
Chx
b
Colony doublings
Time in drug (d)
c
100 µm
+Chx
Day 0
Day 3
Day 6
Day 9
d
Data
n = 76
Colony Doublings
Time in drug (d)
e
Linear model fits
Colony Doublings
Time in drug (d)

Figure 7 Population response is governed by divergent drug-induced proliferation (DIP) rates unique to each clone. (a) At the cell population level, cycloheximide treatment (solid line) induces reduced proliferation relative to DMSO control (dashed line). Colony doublings represents the $\log_2$-transformed cell number per colony normalized to the initial size of that colony. Lines connect the mean of colony doublings for each unique condition and time point. Error bars represent the standard deviation of cycloheximide-treated colony doublings at each time point. (b) Increasing error in average colony size associates with divergent changes in individual colony size. Boxplots show the interquartile range of cycloheximide-treated colony doublings at each time point. Colored lines display the doublings of three representative colonies. (c) Images of the representative colonies in b. Matched colors show the correspondence of colonies in *b* and *c*. (d) Colony dynamics, same as in *b*, except the data are normalized instead to the colony size after 3 days. Lines and boxplots are the same as in *b*. (e) Linear model fits of colony dynamics simplify the colony doublings data. The linear model fit of each colony response from 3 days on is shown as a gray line. The estimated linear model fits of the representative colonies are shown as colored lines, as previously described. Boxplots represent the interquartile range of the expected values for colony doublings at each time point.

(Fig. 3d boxplots and colored lines). Thus, unique rates of colony proliferation underlie increasing variability in colony-to-colony size over time. The steady rate of proliferation for each colony suitably represents its dynamics and therefore we refer to it as the drug-induced proliferation (DIP) rate.

To validate the linear model assumption we looked at the R-squared values for the linear model fits (used in Fig. 3e) of pooled cycloheximide- or control treated colonies. Ninety-five percent of colonies displayed an R^2 value > 0.6 (Fig. 8a), generally considered an acceptable cut-off. DIP rate linearity is a powerful simplification because it is a unique property of a colony that gives a stable estimation of colony response that can predict long-term colony behavior. Indeed, the variance of colony size at 10 days correlates with and can be explained by the DIP rate variation (Fig. 8b).

Utilizing DIP rates to cross biological scales

The transformation of colony cell counts to DIP rates has the additional advantage of simplifying the relationship between colony response and whole-population response. The DIP rate can be considered as a direct measure of clonal fitness in the presence of a perturbation, i.e., how well a clone is able to cope with that perturbation and to what extent it can continue to proliferate. The DIP rate clonal profile of a population gives a sense of how the average response, and the heterogeneity between clones, evolves over time. Therefore, we represent the clonal proliferation response to a drug as waterfall plots, where each vertical line is the proliferation rate of a given colony. Both DMSO- and cycloheximide treated colony proliferation rates varied continuously (Fig. 9a). The mean clonal proliferation rate for cycloheximide (0.09 doublings per day) was much lower relative to the mean clonal proliferation rate of control (0.79 doublings per day). Furthermore, all cycloheximide-treated colonies had lower proliferation rates

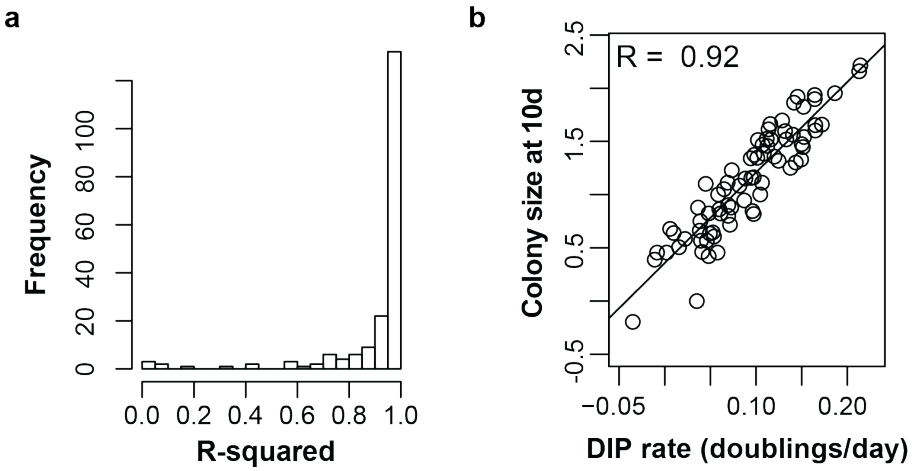

Figure 8 DIP rate correlations. (a) Colony proliferation is suitably represented by linear fits. Histogram of R-squared values from the linear model fits of all DMSO- and cycloheximide treated colonies (n=194 total). (b) DIP rates correlate with the relative colony size after 10 d cycloheximide treatment.

than DMSO-treated colonies (Fig. 9a). Waterfall plots also represent the range of colony proliferation rates in drug (-0.13 to 0.24 doublings per day) relative to control (0.44 to 1.15 doublings per day). Thus, in addition to reducing the mean clonal proliferation rate, it also collapses the range of values by over a factor of six.

Similar to FPM, cFP can also relate clonal response to the single-cell behavior that underlies it. We sought to characterize cell fates of control-treated colonies because their DIP rates displayed a larger range of values than cycloheximide. To facilitate single-cell detection of cell cycle position, we utilized the fluorescent geminin-mAg fusion marker (Sakaue-Sawano et al., 2008a), which is stably expressed after a cell has passed the G1/S transition and maintained until the cell undergoes mitosis (detectable in S/G2/M phase). In a replicate experiment, colony variation in proliferation rate was maintained. We then examined both cell number (pseudocolored red) and cell cycle position (pseudocolored green) in representative colonies with proliferation rates above or below the mean (Fig. 9b, the image outlines correspond to arrows in Fig. 9a). The faster colony had a DIP rate of 0.92 doublings per day, while the slower colony had a DIP rate of 0.51 doublings per day. Both colonies exhibited linear proliferation over three days of treatment with DMSO, both with R^2 values of 0.99. Using the same image processing methodology used to quantify all cell nuclei, we then measured the number of geminin-mAg positive cells at each timepoint relative to the total colony cell number. In both colonies, ~50% of cells are in S/G2/M phase at the beginning of the experiment (Fig. 9c). In the faster colony the percentage of geminin-mAG positive cells is maintained throughout the experiment, while it drops over time in the slower colony. This lower fraction of geminin-mAg cells may be due to an either a lengthened G1 phase or an increasing rate of transition for dividing cells to enter quiescence. Thus, examining single-cell behavior within individual cFP clones can generate hypotheses to understand

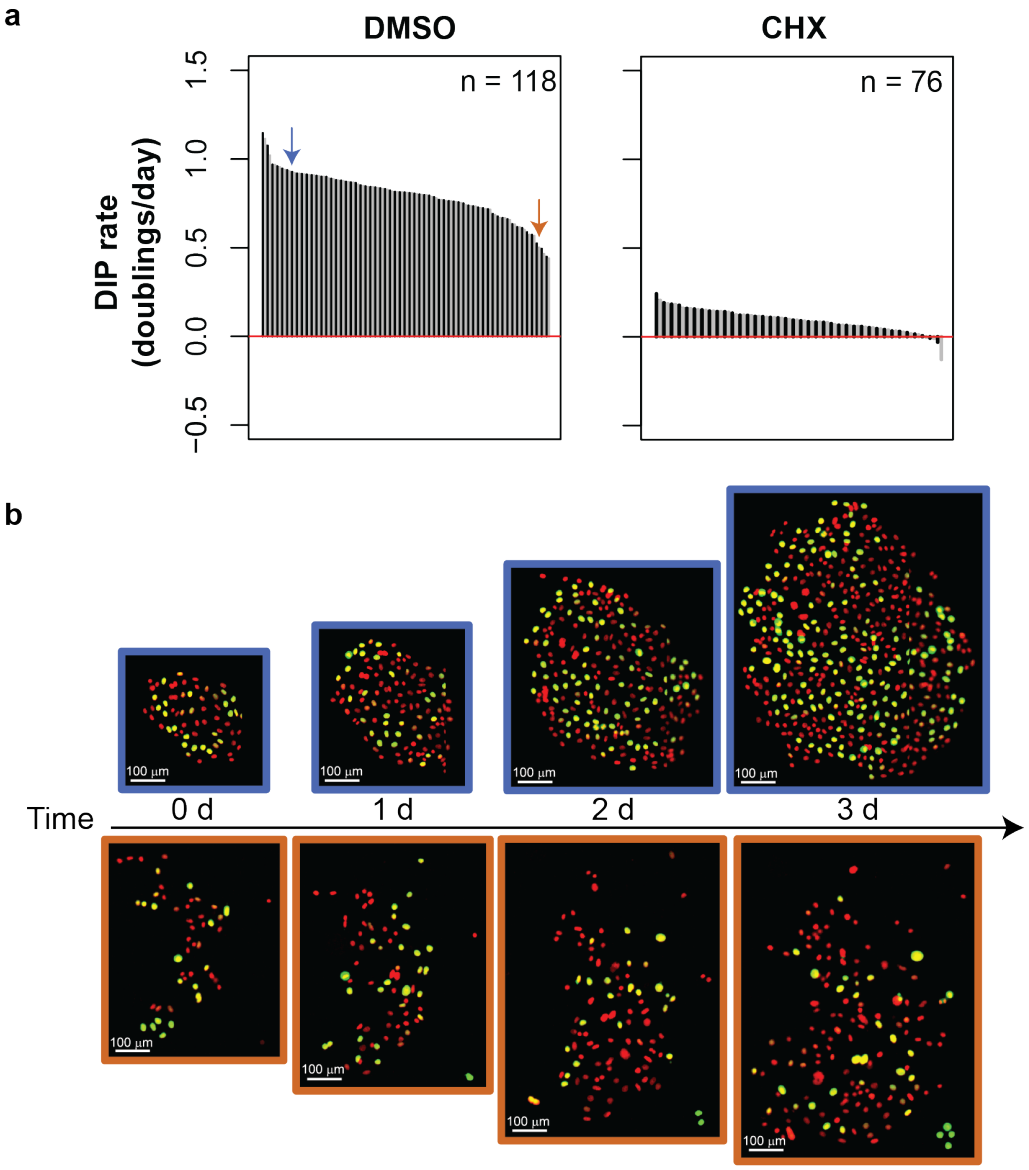
a
DMSO
CHX
n = 118
n = 76
DIP rate
(doublings/day)
1.5
1.0
0.5
0.0
−0.5
b
Time
0 d
1 d
2 d
3 d
100 µm
c
Colony doublings
3.0
2.0
1.0
0.0
0
1
2
3
days
d
Percent S/G2/M
70
50
30
10
0
0
1
2
3
days

Figure 9 Relating DIP rates to both single-cell and population responses. (a) Waterfall plots of PC9 cells treated with cycloheximide (CHX) or control (DMSO). Individual vertical lines represent the DIP rate of single colonies obtained during the subconfluent linear phase of proliferation (0-3 d for DMSO; 3-10 d for CHX). Arrows represent colonies shown in *b*. (b) Colonies with unique proliferation rates in the absence of cycloheximide at the indicated time points. All cells are fluorescently labeled with both Histone H2B-RFP (red pseudocolor) and geminin-mAg (green pseudocolor). Green cell nuclei indicate cells that have passed the G1/S transition; red cell nuclei mark cells that have not. Scale bars are 100 microns. (c) Quantitation of colony doublings for the color-matched colonies shown in *a* and *b* at the indicated time points. (d) Quantitation of the percentage of cells positive for geminin-mAg at the indicated time points.

the fundamental cellular behaviors that underlie interclonal proliferation differences. These hypotheses can then be validated for analyses more suitable to high-time resolution, single-cell tracking, such as FPM (Tyson et al., 2012).

Having established a method to rank a particular clonal DIP rate in relation to all clones, we sought to further characterize individual clonal responses. While cFP is ideal for measuring clonal behavior *en masse*, it is limited in characterizing individual clones because of experimental difficulty in matching clones across experiments. Therefore we reasoned that culturing single-cell derived sublines would recapitulate the clonal proliferation rates observed in cFP. Indeed, we isolated two PC9 clones with proliferation rates, measured in the absence of drug treatment, that were significantly different from each other (Fig. 10a–b; p=1.4e-3), yet similar to that of the colonies depicted in Figure 9b. To investigate the molecular differences associating with these unique clonal responses we quantified the protein levels of EGFR in the two PC9 clones using CellAnimation (Georgescu et al., 2012). For each clone, the individual fluorescence levels of single cells matched a skew normal distribution (Fig. 10c; KS-test=0.93 and 0.08 for clone1 and clone2 respectively; insufficient evidence to reject a skew-normal distribution). Clone1 had a higher EGFR expression level than clone2 ($p<1e-16$). We further used CellAnimation to examine single-cell speeds and determined that the average single-cell speed in the absence of drugs was faster for clone1 than for clone2 ($p<1e-16$; Fig 10d). Thus, single-cell derived clonal cell lines can expand the cFP workflow to provide in depth analysis for individual clonal behavior.

Since PC9 cells are hypersensitive to EGFR inhibition (Faber et al., 2011), and EGFR varies between clones, we measured clonal PC9 response to erlotinib. We find

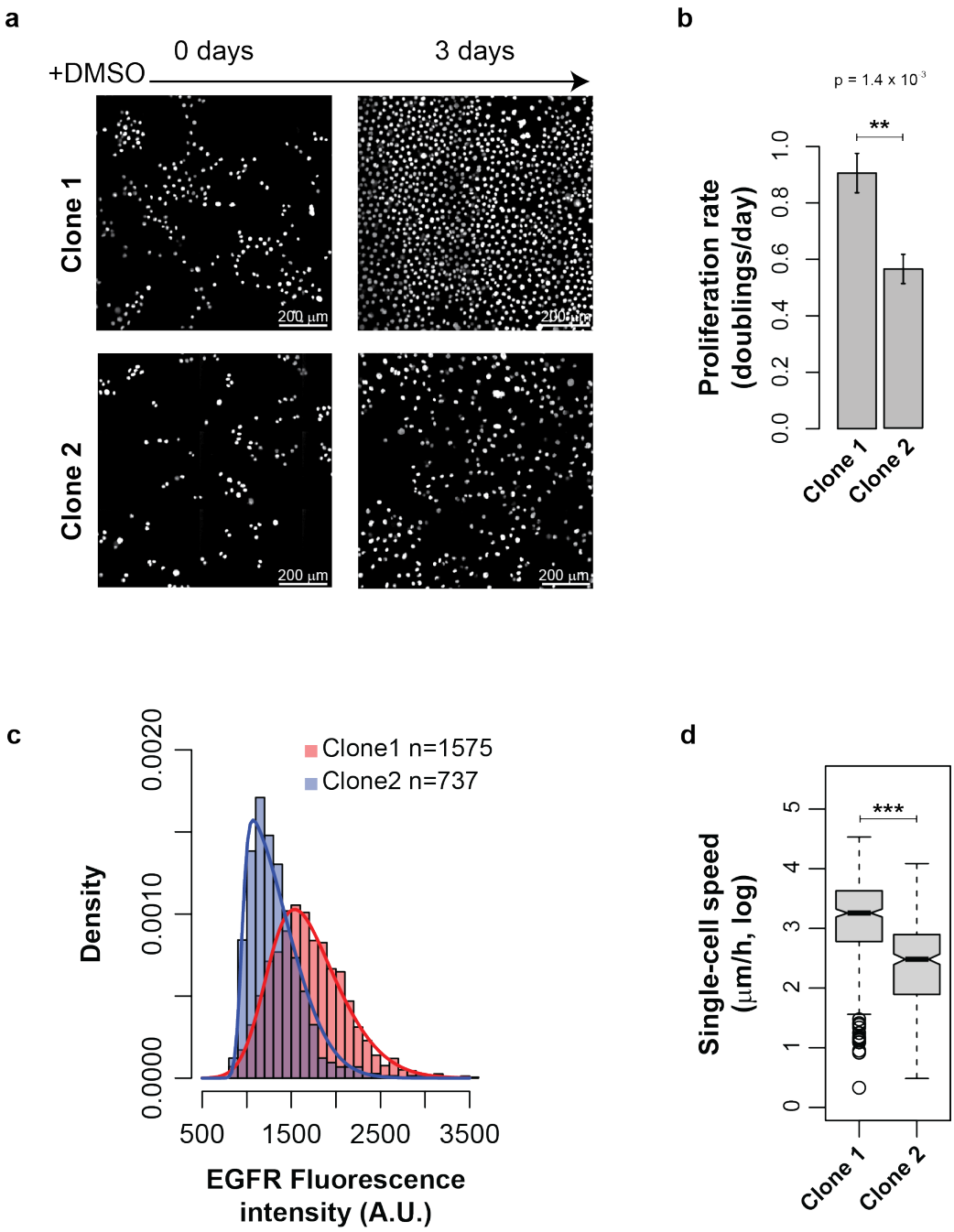

Figure 10 Cultured single-cell derived sublines enable comprehensive clonal analysis. (a), Representative serial images of cell nuclei from clone1 and clone2 at the indicated times cultured in the absence of drugs. Error bars=200 μm. (b) Calculated untreated clonal proliferation rates based on serial hourly cell number measurements from a three day experiment (triplicate wells, n=3 experiment). (c), Histograms of single-cell EGFR intensity derived from a representative immunofluorescence experiment. Images were segmented and quantified using CellAnimation (Georgescu et al., 2012). Colored lines represent the best fit to a skew normal distribution. (d), Single-cell speed measurements of control-treated cells quantified by CellAnimation (Georgescu et al., 2012). *** indicates p-value<0.001.

that several representative clones display negative, variable individual DIP rates to erlotinib (Fig. 11a). Interestingly, these clones also have morphological differences.

Are clonal differences actually relevant in a population that has not been artificially manipulated by either fluorescence labeling or culturing all cells at single-cell density conditions? To test this, we looked at unlabeled PC9 cells treated with erlotinib by brightfield live-cell imaging. Consistent with clonal variation, nearby cells within the same microtiter well respond similarly, in both morphology and cell fate. For example, some groups of cells respond to erlotinib by increasing their cell size and displaying morphological traits evidencing either increased cell-cell contacts or cytoplasmic protrusions (Fig. 11b, dashed circle). These cells display relatively higher cell survival and the morphological features may give clues to molecular pathway activation underlying cell survival. Other cells treated with erlotinib do not increase in cell size or undergo morphological adaptations, but rather display membrane ruffling and subsequent cell death (Fig. 11b, outside dashed circle). Thus we conclude that clonal variation appears to underlie population responses in the absence of sparse clonal seeding conditions.

Methods

Cell culture and reagents. PC9 cells were obtained as a gift from William Pao (Vanderbilt University School of Medicine). PC9 cells were fluorescently labeled as previously described(Tyson et al., 2012). Cells were cultured in RPMI 1640 (obtained from ATCC) supplemented with 10% fetal bovine serum and kept at CO2- and temperature-controlled humidified incubators. Cells were confirmed negative for mycoplasma before used. Cycloheximide was obtained from Sigma and used at a final

concentration of 500 ng/ml. Erlotinib was obtained from LC laboratories and stored as single-use aliquots. Both erlotinib and cycloheximide were stored at -20C.

Immunocytochemistry. Immunofluorescence detection of EGFR utilized an EGFR antibody obtained from Cell Signaling Technologies and Alexa Fluor 647-conjugated secondary antibody from Life Technologies. Cells were grown in a BD 96-well imaging plate and at the appropriate time fixed using 4% paraformaldehyde-PBS for ten minutes, washed in PBS, and stored in PBS at 4C. Cells were permeabilized with a blocking buffer containing 0.3% Triton-X and 5% normal goat serum. Primary antibody (1:100, Calbiochem) incubation went overnight at 4C. Cells were washed three times with PBS, and then secondary antibody (1:1,000) was added for 1 hour in blocking buffer. Cells were counterstained with Hoechst 33342 (Invitrogen, 1:10,000 in PBS, 15 min) and imaged. Single-cell quantitation of EGFR intensity was performed using CellAnimation (Georgescu et al., 2012).

clonal Fractional Proliferation (cFP) assay. As described in the legend to Figure 5. Briefly, subconfluent cells are split and seeded at optimized single-cell density into Benton Dickson 96-well imaging plates. For PC9, the ideal conditions are 40 cells seeded per well in a final volume of 100 mL growth media. Imaging plates are subsequently stored in tissue culture incubators for six days to allow single cells to expand into single-cell derived colonies. After this period, all wells imaged and then the experimental perturbation is immediately added. Subsequently, each well is imaged daily until the end of the experiment.

Image registration and processing. To prepare raw images for analysis, images were sequentially organized into spatially- and temporally registered image montages (Fig. 5c) using the freely-available ImageJ software (http://imagej.nih.gov/ij/).

Subsequent image processing scripts were applied on a per-colony basis. Raw and processed images were both stored for reference.

Generation of discrete sublines. PC9 cells were isolated as single cells and expanded until frozen stocks could be obtained. Sublines were kept under ten passages to ensure consistency across experiments.

Statistical analysis. All statistical analysis was performed using the R statistical software (R-project.org). Linear model fits utilized the *lm* function and incorporated the datapoints indicated in the text. Adjusted R^2 values were calculated from the *lm* function. Pearson correlation coefficients were calculated using *cor.test*. And skew normal distribution fits for the immunofluorescence data were generated using the selm function contained within the SN package (http://azzalini.stat.unipd.it/SN, version 1.0).

Discussion

We present clonal Fractional Proliferation (cFP), a high-throughput imaging methodology to relate the heterogeneous drug response of a cell population to its clonal structure (Fig. 12). The cFP assay reports response to perturbation as a distribution of clonal proliferation responses. The average population response can also be obtained from cFP data as the sum of all clonal responses. cFP feasibility relies upon reliable image processing routines and the novel observation that the colony proliferation rate becomes stable (linear) a few days after exposure to perturbation, obviating the need to

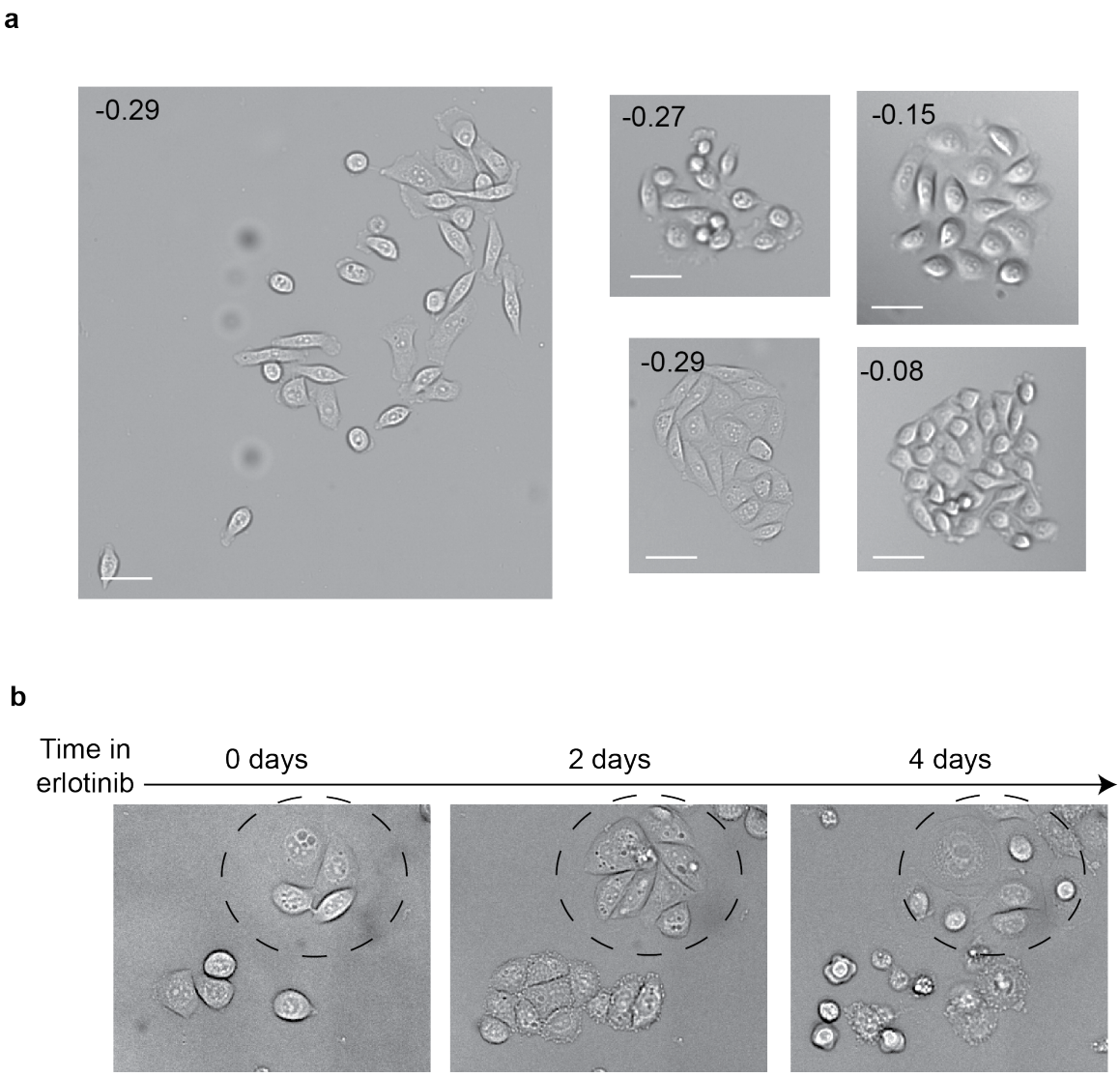

Figure 11 Morphological variation and DIP rate amongst clones. (a), Colonies display characteristic cell morphologies prior to treatment. Shown are several brightfield images of PC9 colonies acquired immediately before drug treatment with 3 μM erlotinib. Scale bars represent 50 microns. Numbers in the upper left corner indicate the calculated erlotinib DIP rate for each colony calculated during 10 d treatment. (b), Unsorted, high density PC9 cells display clonal response by morphology and by cell fate. Timecourse erlotinib response images of unlabeled PC9 parental cells. Cells within the dashed circle increase cell-cell contacts and size and survive until the end of the experiment. Cells outside the dashed circle undergo apoptosis.

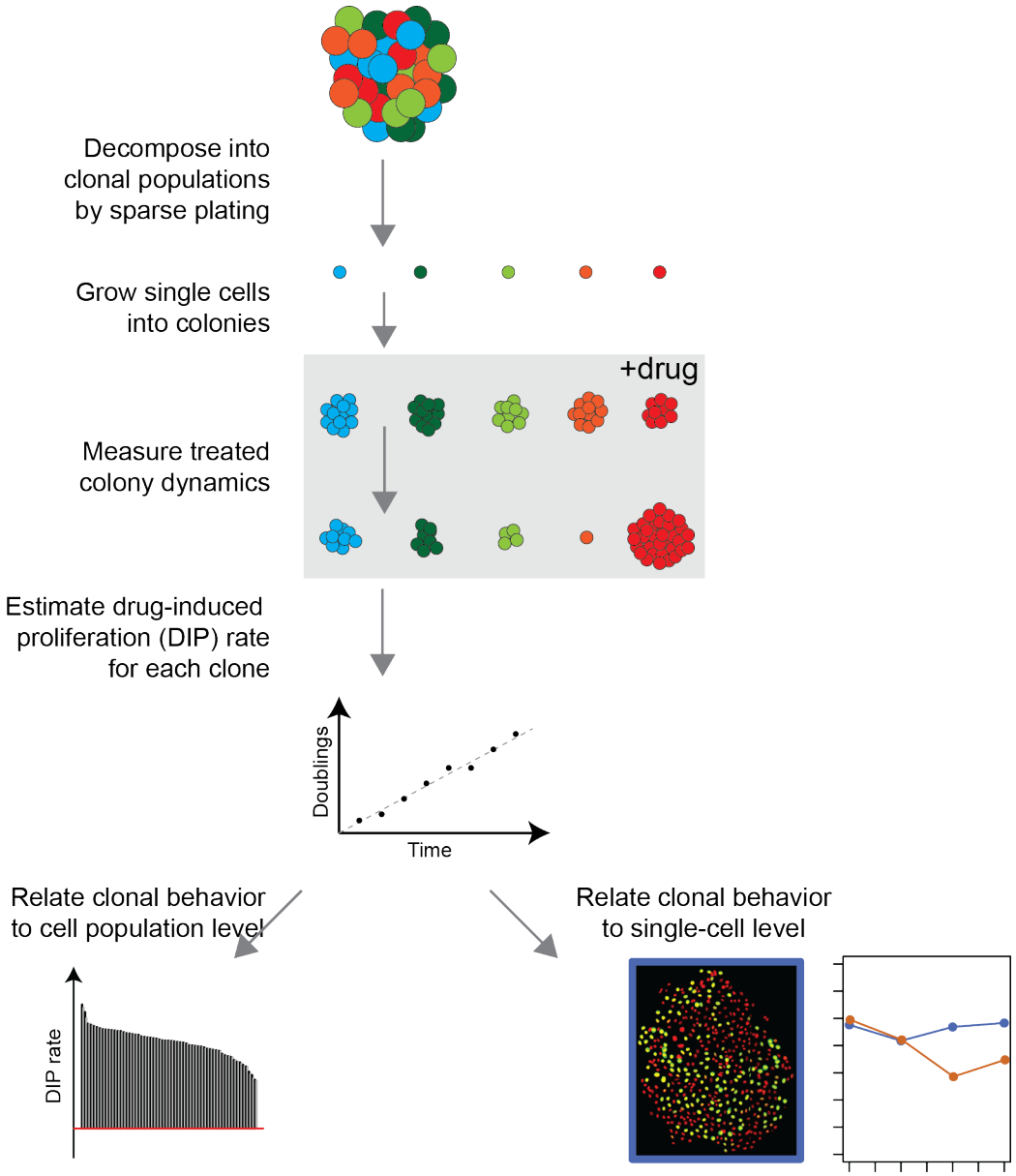

Figure 12 Schematic of entire cFP approach. Figure described in descending order. A clonally heterogeneous cell population is seeded at single cell density as in Figure 5. Individual cells are grown into clonally-derived colonies in full growth media. After the colonies have a sufficiently high cell number drug treatment is initiated and changes in cell sizes are quantified. Drug-induced proliferation rates are estimated for each colony by a linear model best fit. Clonal DIP rates can then be related to either the population-level response by waterfall plots or to the derivative cellular behavior within that colony by single-cell analysis.

directly track individual cell fates. Clonal behaviors can be further examined in larger cell numbers by using expanded clonal cell lines. We envision that cFP will be broadly applicable to study how perturbations induce changes in the clonal fitness structure of a cell population, especially in biological processes that display single-cell heterogeneity, such as cancer progression and differentiation.

Cell proliferation is a commonly reported metric to quantify the functional effects of genetic or chemical perturbations. It is most commonly measured by an indirect estimation of cell number relative to control at the population level. These assays are attractive due to their high sample throughput (Barretina et al., 2012), but can lead to misinterpretation of drug effects because they report relative, rather than direct, proliferation. For example, it is common to infer “fifty percent cell killing” if the number of cells remaining is half of control after 72 h. But measuring cell populations over time can reveal that reduction relatively fewer cells can involve minimal apoptosis (Tyson et al., 2012). Therefore population metrics, naïve to cell fates, can provide an incomplete picture of perturbation effects on proliferation. cFP, in contrast, measures drug effects directly for clonal populations constituting a cell population (Fig. 7c).

It is well established that variability of biological traits is widespread at the single-cell level(Altschuler and Wu, 2010). In depth investigation of single-cell behavior in response to perturbations is an active field of investigation and molecular mechanisms of cellular heterogeneity are progressively being described (Gascoigne and Taylor, 2008; Spencer et al., 2013). However, the functional consequences of single-cell heterogeneity on the whole population becomes apparent when some kind of organizing structure is discovered. This is crucial to gaining a fundamental understanding of the biological system, as in bacterial resistance, organ development, or tumor evolution (Johnston and

Desplan, 2010). Previously, in FPM (Tyson et al., 2012), we developed a computational framework to link comprehensive single-cell tracking data with population dynamics for limited samples. cFP parallelizes FPM across hundreds of clones within a sample to uncover perturbation-induced changes to the entire clonal profile. Thus cFP, by relatively few measurements, reveals the governing clonal structure that underlies extensive single-cell fate heterogeneity (Fig. 9a).

The key feature that enables cFP to quickly summarize the clonal fitness profile of a cell population is the DIP rate. While a clonal response may involve multiple cell fates, it achieves a steady rate of proliferation after a few days (Fig. 7d), which can be summarized as its slope (Fig. 7e), or DIP rate. Thus DIP rates provide sufficient information to infer the predominance of single-cell fates: If DIP rate is negative, then death must prevail; if positive, then division prevails. Since DIP rates are linear, entry of cells into quiescence plays by necessity a minor role, as quiescence functions primarily to shape non-linear proliferation responses (Tyson et al., 2012).

What does it mean that clones proliferate at different rates in the presence of a perturbation? The cell lines that we test are “programmed” to continuously proliferate; therefore the rate of proliferation can be taken as a direct measure of fitness of a particular clone to a particular microenvironment. DIP rate is expressed as a continuous variable, in contrast to the colony formation assay, which reports the relative percentage of single cells capable of adhering and forming a colony (Franken et al., 2006). Because DIP rates are stable, they can be used to explain a majority of the varying long-term outcomes between clones (Fig. 8b). This may relate to, for example, the general ability of cancer cells to adapt to harsh environments through selection (Nowell, 1976), and should be the focus of further work.

Measurement of DIP rates relies upon accurate quantitation of clonal proliferation. By sparsely plating a cell line at single-cell density, cFP separates a population into its clonal constituents (Fig. 5a). Automated image processing of fluorescent cell nuclei yields accurate cell counts (Fig. 6a). Other methods of cell counting are equally amenable, provided that they closely match manually validated data (Fig. 6b–c). cFP was designed to measure clonal proliferation originating from single cells (Fig. 5b). Future work should test if it matters if colonies originate from multiple cells. Furthermore, addition of microenvironmental factors may aid in more closely mimicking cell growth *in vivo*.

Because all colonies are measured, the waterfall plots can be taken as a representation of the global clonal fitness of a cell population in a particular microenvironment (Fig. 9a). In this manuscript, we used cycloheximide as a proof-of-concept example. Cycloheximide treatment induced clonal DIP rates that were continuously distributed, were reduced relative to control, and had a decreased the range of values (Fig. 9a). These need not be the case. In fact, we envision that the clonal profile of a cell population would be most interesting for cases where the crucial clonal proliferation is substantially different from the bulk population. For example, it may be that rare clonal cells with positive DIP rate that drive drug resistance. Or potentially slow dividing cells necessary to maintain tumor proliferation (Roesch et al., 2010). Alternatively, cFP captures a range of cell morphologies (Fig 11); this may be interesting to relate drug response to morphology, since epithelial-to-mesenchymal transition has been implicated as a mechanism of drug resistance(Chong and Jänne, 2013). The establishment of expanded clonal populations should enable more in depth clonal analyses (Fig. 10).

In summary, our approach represents a way to capture clonal fitness variability within a cell population (Fig. 12). DIP rates act as a bridge to connect heterogeneous cell fates within a cell population to the dynamic population response by assigning them into clonal lineages that proliferate at different rates.

Chapter III

CLONAL DRUG-INDUCED PROLIFERATION RATES REVEAL STRUCTURE TO THERAPY-INDUCED CELL FATE HETEROGENEITY

Introduction

Targeting specific mutated gene products in cancer has fast become the most promising approach to defeat this incurable disease (Haber et al., 2011). However, even in patients carefully selected for well-defined targetable genetic mutations, as in lung cancer and melanoma, depth and duration of response is variable and relapse inevitable. This variability is a major obstacle to better, more permanent outcomes. Causes of variability are multiple and still ill defined, but increasing evidence implicates intratumor heterogeneity (Almendro et al., 2013; Bedard et al., 2013). Recognized for decades (Dexter et al., 1978; Fidler and Kripke, 1977), tumor heterogeneity has been unequivocally demonstrated by several molecular approaches (Gerlinger et al., 2012). Lately, its impact on anticancer drug treatment is being studied roughly along two lines.

The first considers response to treatment in terms of clonal variation (Greaves and Maley, 2012), long known to adversely affect cancer therapy due to selection of clones with advantageous mutations (Nowell, 1976). Historically, genetics alone (Nowell, 1976) was considered as a source for variation of clonal fitness (intended as the variation allowing a population to accommodate environmental change, including drug-treated tumors). Recent in vivo studies (Kreso et al., 2013) elegantly show that non-genetic sources may additionally contribute to clonal fitness in response to chemotherapy. Data-driven modeling of cell population dynamics also shows promise for understanding the effect of clonal genetic and non-genetic fitness variation on treatment

outcomes (Anderson et al., 2006; Michor et al., 2005). A limitation of these studies is that clonal fitness in the presence of drug is inferred, rather than directly measured with a quantitative metric.

A second line of investigation focuses on the heterogeneity of cell-to-cell fates in response to drugs (Niepel et al., 2009), postulated to contribute to treatment failure. Molecular and functional single-cell analyses have revealed many sources of drug-induced cell fate heterogeneity in a cell population, including epigenetic (Fan et al., 2011; Kreso et al., 2013; Sharma et al., 2010b), stochastic (Altschuler and Wu, 2010; Gascoigne and Taylor, 2008) and non-genetic (Pisco et al., 2013; Spencer and Sorger, 2011). However, these studies did not consider alternative cell fates in the context of the clonal structure of a cell population.

Clearly, clonal fitness and cell fate heterogeneity during drug treatment are likely intertwined, yet a framework integrating the two has yet to be attempted. We reasoned that a synthesis of clonal fitness variation with cell fate heterogeneity might provide a platform to realistically model global cell population dynamics in the presence of drug, and predict treatment time-course and outcomes. Such predictive power would advance tumor heterogeneity studies closer to preclinical and clinical applications. We attempt to integrate clonal fitness variation with cell fate heterogeneity using a well-established experimental model for targeted therapy—human lung cancer cultured cell lines harboring the clinically relevant exon19 epidermal growth factor receptor (EGFR) deletion. These cells are oncogene-addicted (Weinstein et al., 2008) to mutated EGFR, and sensitive to inhibition of proliferation by EGFR-targeted drugs such as erlotinib. In these cells, both clonal resistance (Chmielecki et al., 2011) and cell fate heterogeneity (Sharma et al., 2010b; Tyson et al., 2012) in response to drug have been shown.

Results

Clonal Drug Induced Proliferation (DIP) rates encapsulate cell fate heterogeneity.

We readily obtained evidence that the response of PC9 cells to erlotinib is a composite of heterogeneous cell fates. Erlotinib potently inhibited PC9 proliferation (IC50=44nM, Fig. 13a), abrogated EGFR phosphorylation and downstream signaling (Fig. 1b), and induced markers of apoptosis (cleaved PARP1, Fig. 13b). Yet, 17% of cells continued to divide on day 3 in erlotinib, and 28% on day 10 (Fig. 13c, d), suggesting the coincidence of differential cell fates (cell death, quiescence and division).

To quantify this erlotinib response in terms of rates of division (intermitotic time), death and quiescence, we expressed geminin degron (mAG-gem) (Sakaue-Sawano et al., 2008b), a live-cell reporter of cell cycle progression, in PC9 cells and tracked single cells over 6 days in the presence of drug (Quaranta et al., 2009). In PC9 cells treated with erlotinib, single-cell tracking at high time resolution (~120 observations per day) over 6 days revealed that 16% of cells died, 52% continued to divide with elongated cell cycle time, and 33% had not committed to any fate by end of experiment (eoe) (Fig. 14a–b). From these data, we derived a Fractional Proliferation Method (FPM) graph (Tyson et al., 2012) that estimates the contribution of individual cell fates as fractions of the size and overall proliferation rate of the population (Fig. 14c). The proliferation rate (which integrates rates of division and death) reaches a steady state after an initial period of nonlinear growth (~72 hours, vertical dashed gray line in Figure 14c). The dividing cell fraction remains detectable throughout the experiment (over ~150 h) suggesting the possibility that these cells represent a separate subpopulation that is insensitive or resistant to erlotinib.

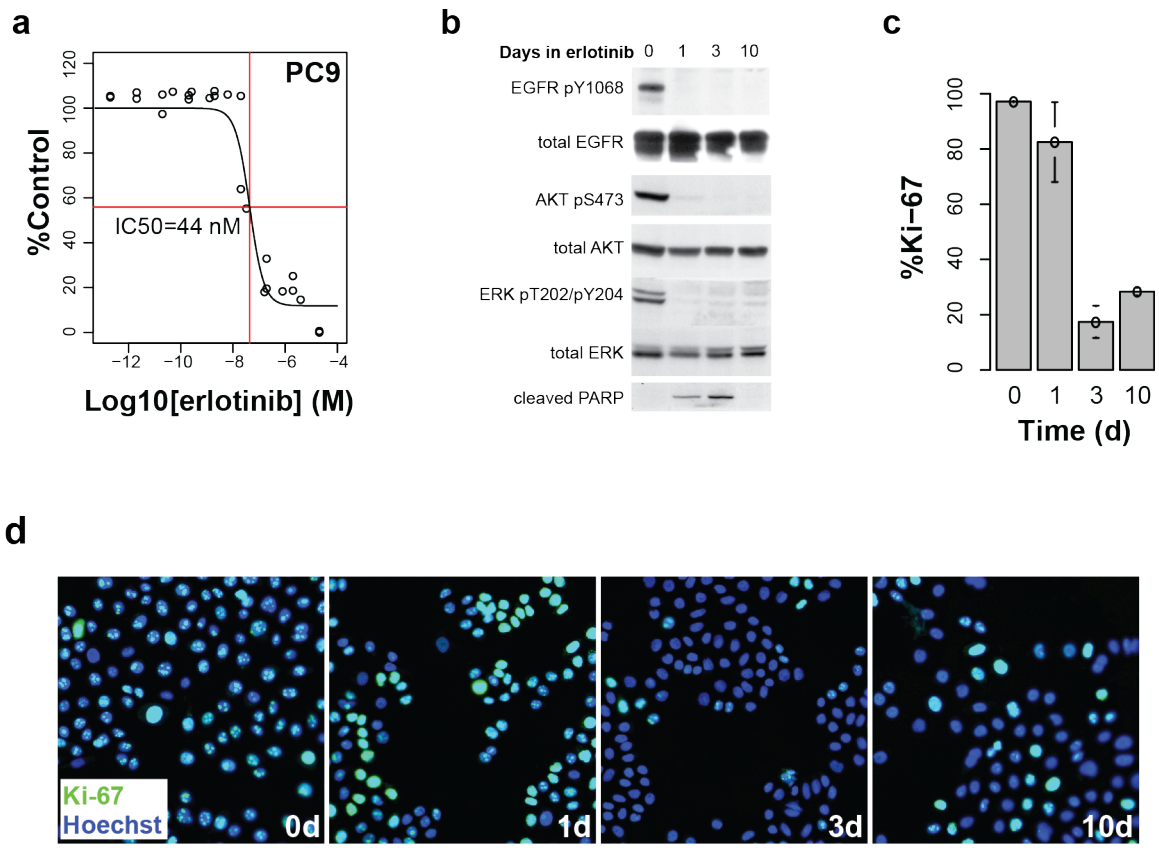

Figure 13 PC9 cells respond to erlotinib with multiple cell fates. (a) PC9 cell proliferation is inhibited at very low concentrations of erlotinib (IC50 = 44 nM). (b) Phosphorylation of epidermal growth factor receptor (EGFR), AKT and ERK detected by immunoblotting PC9 lysates is inhibited by 3 μM erlotinib, indicating drug activity. Cleaved PARP presence indicates apoptosis activation. (c) Erlotinib decreased the abundance of Ki-67 nuclear antigen-expressing cells at 3d indicating decreased proliferation (d), Quantitation of %Ki-67-positive cells in *c*.

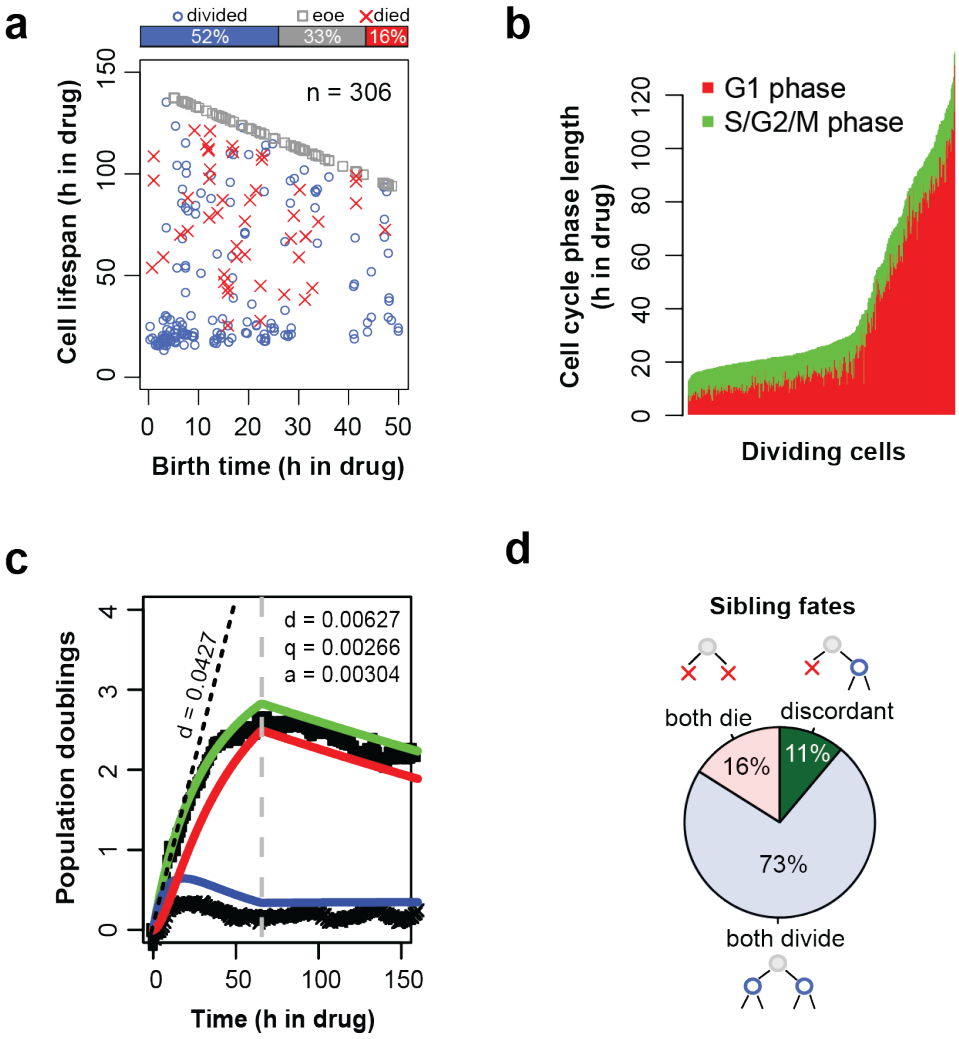

Figure 14 Erlotinib-treated PC9 cells captured at high time resolution exhibit heterogeneous cell fates. (a) Single-cell tracking of erlotinib-treated PC9 cells (Tyson et al., 2012) demonstrates multiple cell fates: Cells born in the presence of drug within the first 50 hours (x-axis) undergo division (blue circles), death (red X) or neither by the end of experiment (eoe) (grey squares). (b) Cells that divide in erlotinib display extended cell cycle time due to extended G1 phase. Individual stacked vertical lines represent the measured cell cycle length of single cell tracks of all dividing cells from *A*. G1 phase is defined as the time from cell division to the detectable expression of geminin degron. S/G2/M is defined as the remainder of the cell cycle until the subsequent cell division. (c) A Fractional Proliferation Graph (Tyson et al., 2012) depicting the dynamic PC9 cell response to erlotinib. Nonlinearity of erlotinib-treated PC9 cell population doublings (circles) is predicted (green line) by the single-cell behavior shown in *a* and deconvolved into dividing (blue) and non-dividing (red) fractions. Rates of division (*d*), death (*a*) and entry into quiescence (*q*) are estimated from the single-cell behavior. Dotted line depicts division rate of vehicle-treated cells. (d), The pie chart depicts the percent of concordant (both die, light blue; both divide, light red) and discordant (one sibling dies and the other divides again, green) cell fates in sibling pairs with observable cell fates before eoe (n = 81), tracked in the same experiment as *a*.

This heterogeneity of fates in response to drug is consistent with earlier studies (Gascoigne and Taylor, 2008; Tyson et al., 2012) but had not been previously placed in the context of clonal lineages. To begin relating cell fate heterogeneity to clonal lineages, we compared the fate of cell sibling pairs originating from a single mitotic event occurring after drug addition. From the tracked cells in Figure 14a, there were 81 informative sibling pairs, i.e., for which the fate (death or division) of both siblings was observed within the time frame of the experiment. The sibling pairs that exhibited concordant cell fates (both either died or divided, Fig. 14d) was 89%, supporting the conventional view of inheritance within a clonal lineage. However, 11% of sibling pairs had discordant fates (Fig. 14d), supporting a stochastic source of cell fate decision (Altschuler and Wu, 2010; Gascoigne and Taylor, 2008; Spencer et al., 2009). To reconcile both inherited and stochastic fates with the dynamics of a drug-treated population, the logical next step was to examine cell fate heterogeneity directly within the context of clonal lineages.

In previous work (Tyson et al., 2012), we linked perturbation-induced heterogeneous cell fates to the proliferation dynamics of an entire cell population with the FP method. To enable bookkeeping of heterogeneous cell fates within PC9 clonal lineages we utilize the novel *clonal* Fractional Proliferation (cFP) assay in which hundreds to thousands of single-cell-derived colonies (clonal lineages) are tracked in a cell population over extended time periods (~10 days; Fig. 1c). We observed a mixture of cell fates (death, division, quiescence) *within* clonal colonies of treated PC9 (Fig. 15a), indicating that cells that continue to divide in drug do not reside in separate subpopulations insensitive to erlotinib. Rather, it appears that the multiple cell fates detected at the population level (Fig. 13, Fig. 14a) in fact occur *within* a clonal lineage. This mixture of cell fates within each colony results in clones with distinct rates of

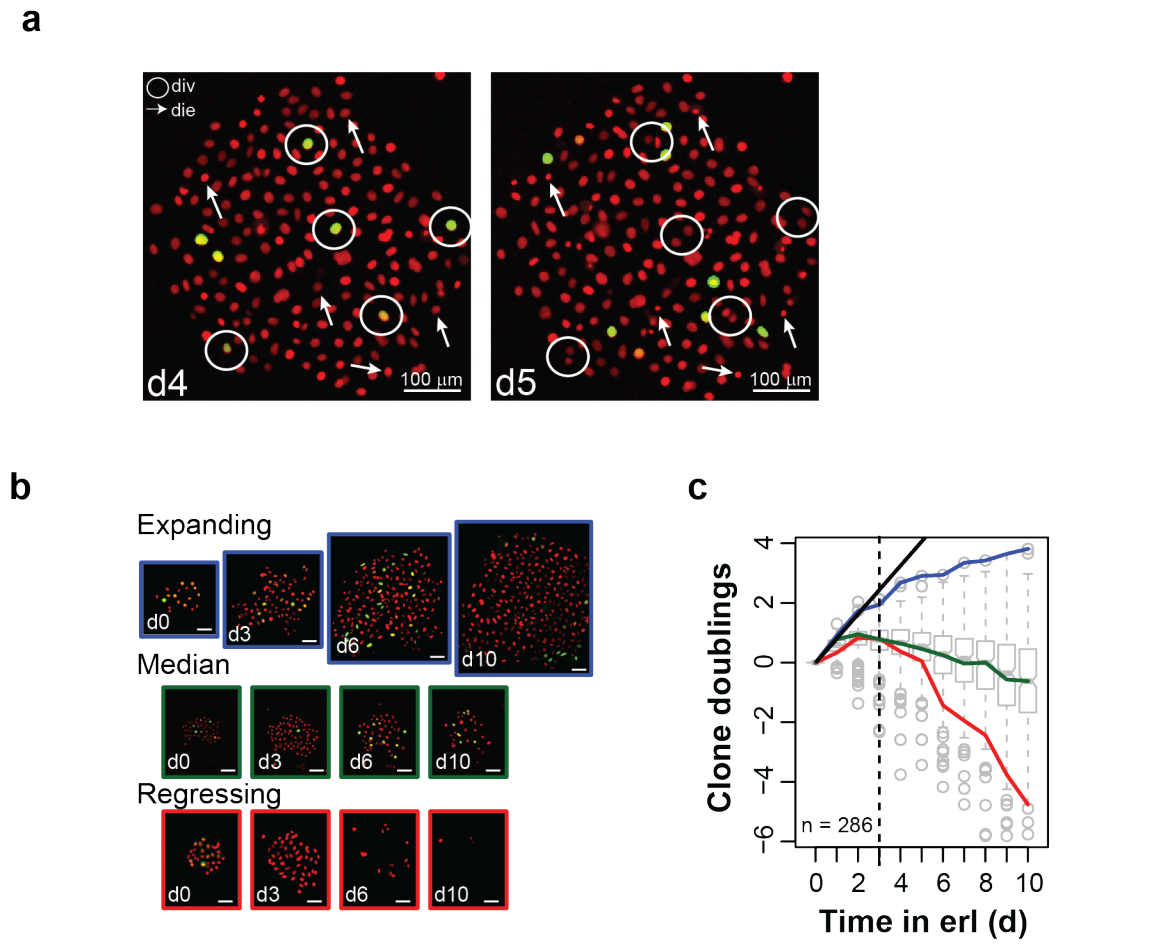

Figure 15 Clonal response of PC9 cells to erlotinib. (a) Multiple cell fates in a single PC9 clone from a cFP microtiter well, tracked in the presence of erlotinib and imaged after 4 and 5 days of treatment. Circled green nuclei are cells that have past the G1/S transition(Sakaue-Sawano et al., 2008a) (d4) and subsequently divide (d5). Arrows represent interphase cells (d4) that undergo nuclear shrinkage and dysmorphism (d5) characteristic of cell death(Tyson et al., 2012). Scale bars = 100 μm. (b) Growth variability among erlotinib-treated clones in the cFP assay. Three clones were imaged on indicated days, showing distinct representative growth outcomes, as indicated. Scale bars = 100 μm. (c) Quantification of PC9 clonal dynamics in the cFP assay. The proliferation rate of each clone (286 total) is calculated from the slope of the line fit to data ≥ 72h (dashed line); average DMSO-treated proliferation rate is represented by the solid black line. The representative clones from *b* are depicted by color-matched lines. Boxplots represent the range of clone sizes at each time point. erl=erlotinib.

proliferation, e.g., they expand or regress depending on whether division or death prevails as a fate (Fig. 15b).

In vehicle control (DMSO), proliferation of all colonies was exponential, exhibiting a characteristic linear behavior in $\log_2$ plots (Fig. 15c, solid black line). In contrast, colony proliferation in erlotinib was nonlinear over the 10-day period, both for individual colonies (Fig. 15c, colored traces) and on average (box plots in Fig. 15c). However, proliferation becomes linear after approximately 72h (Fig. 15c) indicating that a steady state rate is reached, similar to the results in the FPM graph of Figure 14c.

We then investigated whether the post-72h proliferation rate could be a useful erlotinib response metric at the level of clonal lineages. In the pre-72h phase, most colonies (87%) exhibit a positive rate, as best visualized in waterfall plots (Fig. 16a). In contrast, in the post-72h phase the majority of colonies (85%) have negative rate (Fig. 16a). In individual colonies, pre- and post-72h rates were poorly correlated (R=0.25), as depicted by color-coded rank reordering (Fig. 16a). The calculated pre-72h colony rates are poorly correlated with the total change in colony cell number at the end of the 10-day experiment (Fig. 16b). In contrast, the steady state post-72h rate, defined by the slope of the line best fitting the data between days 3 and 10, is highly correlated with the 10-day outcome (Fig. 16b), indicating that it is an effective metric of long-term cell colony (and population) response to erlotinib.

Therefore, after an initial period (~72 hours, vertical dashed line in Fig. 15c), each clone reaches a steady state proliferation rate that is suitable to represent long-term clonal drug response. To our knowledge, this is a novel observation enabled by the cFP assay, which tracks cell populations over an extended time frame. This Drug-Induced Proliferation (DIP) rate at steady state can be considered a stable trait of a

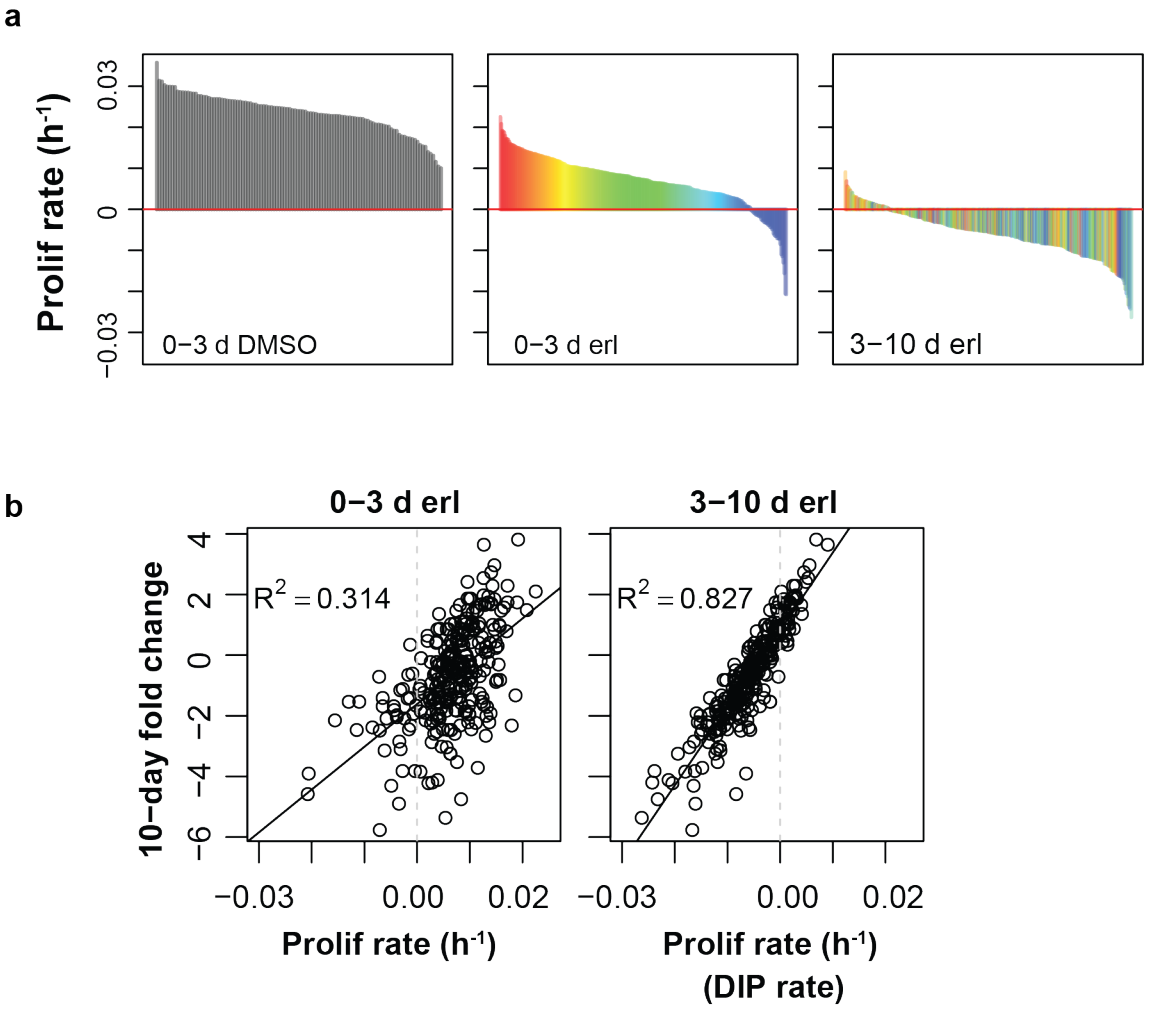

Figure 16 DIP rates effectively capture long-term colony response. (a) No colonies have negative proliferation rates in the absence of drug (DMSO). Erlotinib induces some negative proliferation rates during 0–3 d but most (>85%) of colony proliferation rates calculated over days 3–10 are negative. Color-coded ranks of colony proliferation rates from 0–3 d were redistributed across the graph when the colonies were reordered based on 3–10 d proliferation rates. (b) Overall change in colony cell number (10 d fold change) correlates with 3–10 d colony proliferation rate but not 0–3 d rates. The proliferation rates obtained from 3–10 d colony cell counts in erlotinib are henceforth referred to as drug-induced proliferation (DIP) rates.

clonal lineage, and therefore can be taken as a measure of clonal fitness in the presence of that drug concentration. Furthermore, since the underlying mixture of cell fates determines the clonal proliferation rate, as shown above (Fig. 15), the DIP rate effectively synthesizes clonal fitness with intra-clonal cell-to-cell heterogeneity of cell fates (Fig. 14d).

Positive DIP rate clones are part of a normally distributed continuum

In erlotinib-treated PC9 cells, DIP rates varied from clone to clone over a broad range (-0.038 to +0.013 doublings/hour), with approximately 15% exhibiting a positive proliferation rate (Fig. 17a). To determine whether this DIP rate variability is a general phenomenon, we tested other human lung cancer cell lines harboring mutant EGFR. In every cell line, we observed clonal DIP rate variation (Fig. 17a). Note the proliferation rate of a clone prior to treatment did not correlate with its DIP rate in erlotinib (Fig. 17b, p=0.13). Overall, the global profile of clonal DIP rates is a continuum, but since the distribution in every case crosses the zero DIP rate line (Fig. 17a, red lines), clones with qualitatively divergent behavior can be observed: some drug-treated clones continue to exponentially increase in size (expanding) while others decrease (regressing). The percent of clones with positive DIP rates varied from cell line to cell line (~2–15%). It should be noted that the positive DIP rates (Fig. 17a) were in every case much less than the proliferation rates of DMSO-treated cells (Fig. 17a, left panels) or erlotinib-resistant $EGFR^{T790M}$ PC9-BR1 cells(Chmielecki et al., 2011) (Fig. 17a, bottom panels), indicating again that they do not represent drug-insensitive clones.

Plotting these data as histograms revealed that clonal DIP rates within a cell line are distributed normally (Fig. 17c). Since a normal distribution is continuous, it reinforces the notion that expanding clonal lineages are not a distinct subpopulation from

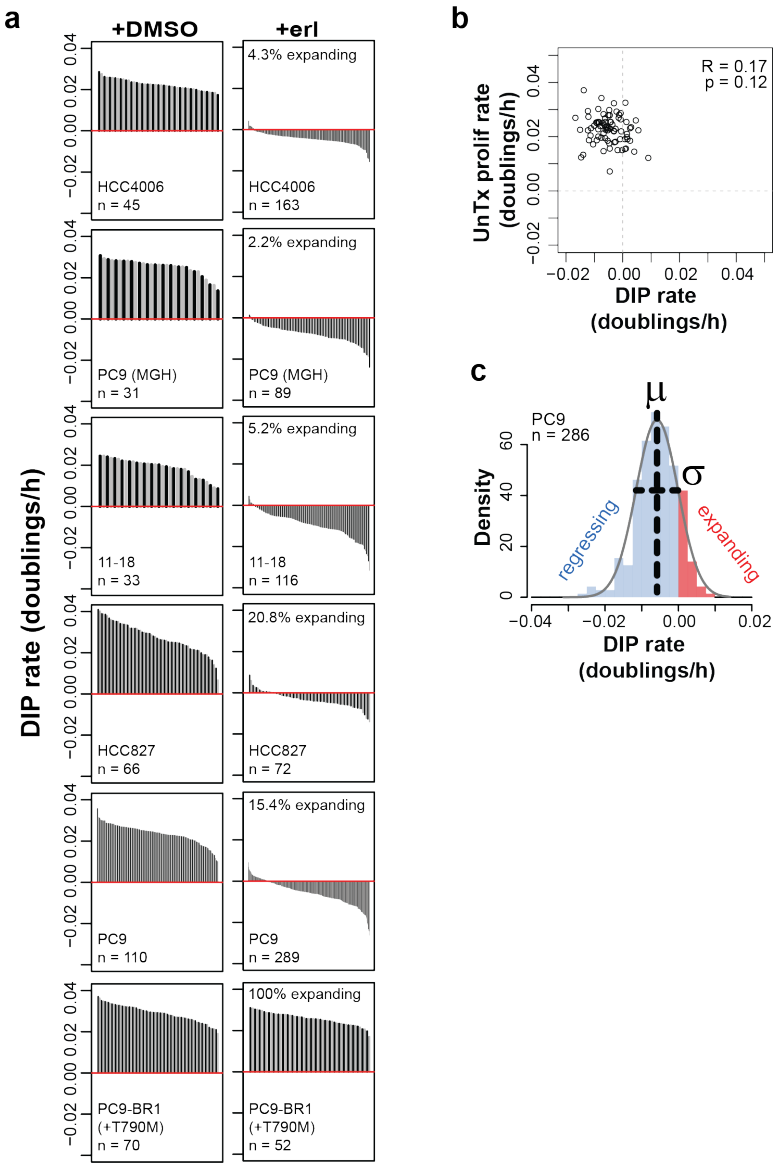

Figure 17 DIP rates of EGFR-addicted lung cancer cell lines exist along a continuum. (a) Waterfall plots of clonal DIP rates (from cFP assays) in cell lines with EGFR-activating mutations treated with erlotinib (+erl) or control (+DMSO). In DMSO-treated cell lines and in the erlotinib-resistant PC9-BR1(Chmielecki et al., 2011) cell line all clonal DIP rates (left panel) are positive. In sensitive, treated cell lines, a small, variable percentage of clones continue expanding in the presence of drug, i.e., exhibiting positive DIP rate. (b) Lack of correlation between individual PC9 clonal proliferation rates measured before (UnTx) and after (DIP rates) erlotinib treatment (R=0.17, p=0.12). (c) PC9 clonal DIP rates are binned, plotted as a histogram, and colorized according to positive (*red*; "expanding") or negative (*blue*; "regressing"). Clonal DIP rates are normally distributed (Kolmogorov-Smirnov test p = 0.35; insufficient evidence to reject the null hypothesis of a normal distribution). The probability density function of the normal distribution fit to the data (gray curve) overlays the data. Dashed lines represent the two fit parameters that define the distribution, μ (vertical, indicates position) and σ (horizontal, indicates magnitude).

regressing clones; rather, they represent the extreme of a continuum. Note that though the data is suitably fit by a normal distribution, it is likely that different perturbations may induce a variety of distributions.

Clones with a positive DIP rate are definitely sensitive to drug because they proliferate at a much-reduced rate compared to control or drug-resistant PC9-BR1 (Fig. 17a). However, since both negative and positive DIP rates are a part of a Gaussian distribution, the positive DIP rate clones are not a separate subpopulation. Furthermore, unlike antibiotic resistant bacterial persisters (Balaban et al., 2004), they cannot be identified in advance based on their untreated proliferation rate, since there is no correlation between a colony proliferation rate prior to drug addition and its DIP rate (Fig. 17b).

Based on their continued proliferation in erlotinib, positive DIP rate clones may contribute to relapse, warranting more in-depth analyses. To this end, we isolated single-cell derived “sublines” from PC9 parental in order to further investigate the nature of clones with positive DIP rate.

PC9 isogenic sublines recapitulate the parental PC9 clonal DIP rate distribution

Ninety-five discrete isogenic single-cell derived clonal sublines from the PC9 parental line (PC9-DS sublines) were prepared independently of cFP, in the absence of any drug selection. DIP rates in the presence of erlotinib were then measured. As in the cFP clones, the post-72h drug response (Fig. 18a) observed in the DS sublines is linear over 3–7 days, and predicts long-term outcome (Fig. 18b), consistent with the DIP rate metric. The DIP rate distribution of the DS sublines is indistinguishable from that of parental clones assessed by cFP (Fig. 19a). These data support the validity of DS

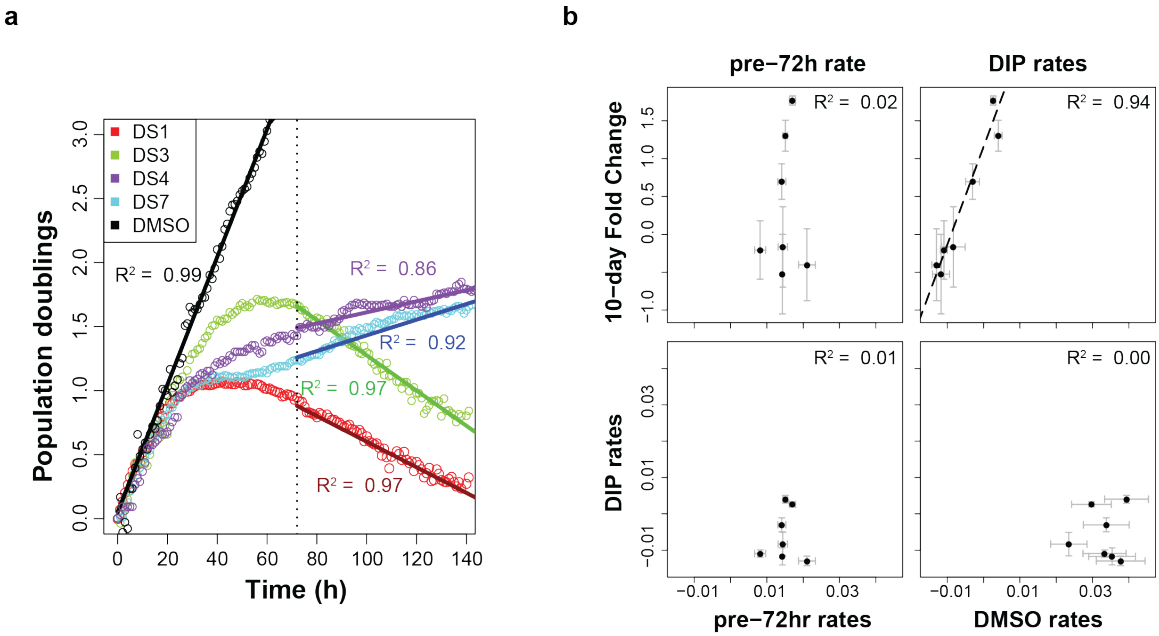

Figure 18 Drug-induced subline dynamics match colony behavior. A panel of seven PC9 discrete sublines, DS1-DS7, was imaged during a continuous 6-day erlotinib treatment. (a) Three representative subline growth response curves indicate the proliferation rate during erlotinib treatment linearizes after 72 hours and assumes a characteristic DIP rate (linear model fit to estimate DIP rate is overlaid in solid colored lines). (b) When treated with erlotinib, DIP rates, but not pre-72h proliferation rates (in erlotinib), correlate with the 10-day fold change (cell number at day 10 relative to the cell number of cells at day 0) in PC9 sublines. Data shown are from independent erlotinib-treated wells (n=3, in triplicate). Error bars represent 95% confidence intervals. DIP rates of PC9 sublines do not correlate with the proliferation rate measured during the first 72 hours of either erlotinib or DMSO.

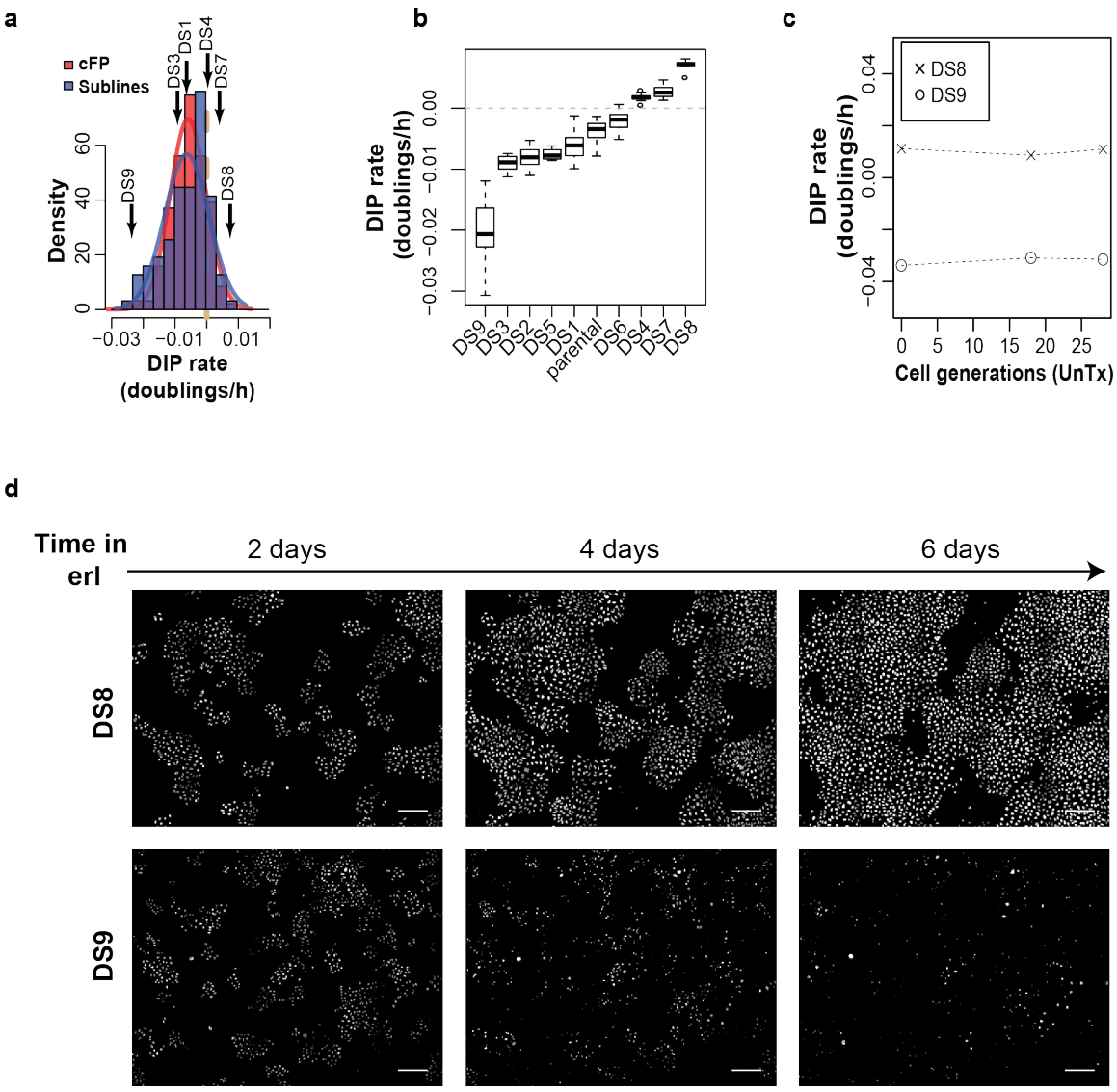

Figure 19 Discrete sublines act as a surrogate experimental tool for colonies. (a) DIP rate histograms of colonies (red) and sublines (blue) are superimposed. Arrows indicate mean DIP rates of sublines sampled from the distribution. Lines indicate best fits of Gaussian distributions to the data. The KS-test had insufficient evidence to reject the null hypothesis that the distributions were significantly different ($p = 0.35$). (b) Measured DIP rates are stable within each subline. Boxplots indicate mean and interquartile range (n = 3 experiments, in triplicate). (c) DS8 and DS9 were kept in culture in the absence of drugs and tested for DIP rate at several intervals. The left-most data points represent the data used to generate the PC9 Subline DIP rate distribution (n=1). Other data points represent the average DIP rate of triplicate wells. (d) Erlotinib treatment timecourse for DS8 and DS9. DS8 and DS9 represent the two PC9 sublines with the highest and lowest DIP rates, respectively. Representative serial image montages of cell nuclei are shown during a six-day erlotinib treatment. Scale bars = 200 μm

sublines as a surrogate for cFP clones, with the advantage that they can be repeatedly subjected to a battery of tests.

An important question to address is whether a DIP rate is an intrinsic property of a subline. A panel of nine representative DS sublines (Fig. 19b) was cultured in the absence of any drug selection over ~25 generations. From these cultures, DIP rates in the presence of erlotinib were measured at regular intervals, and found to remain constant (representative data in Fig. 19c). Thus, a DIP rate is a stable trait realized in the presence of drug but inherited over multiple generations in its absence.

The heritability of DIP rates indicate that positive DIP-rate clones may undermine the outcome of targeted therapy, prompting further investigations. Traditionally, the ability to proliferate in the presence of drug has been associated primarily with genetically acquired resistance resistance (Takezawa et al., 2012). To test this possibility, we subjected positive DIP-rate sublines to SNaPshot analysis (Su et al., 2011b), which detects drug-resistance mutations. Five out of five positive DIP rate DS sublines maintained EGFR exon 19 deletions and lacked each of 38 common clinically relevant resistance mutations (Table 1). Another traditional measure of drug insensitivity is the IC50. Each of the DS sublines had IC50s in the same low nanomolar range, similar to the drug-sensitive PC9 parental (Fig. 20).

It is not unexpected that DS sublines, whether DIP-rate negative or positive, cannot be distinguished by traditional measures of drug resistant/insensitive cells, since they are embedded in a normal DIP rate distribution that suggests an underlying stochastic process. Indeed, in the parental PC9, alternative cell fates in response to erlotinib appeared to be stochastically assigned (Fig. 14d). To determine the relationship between multiple cell fates and DIP rates in the DS sublines, we examined

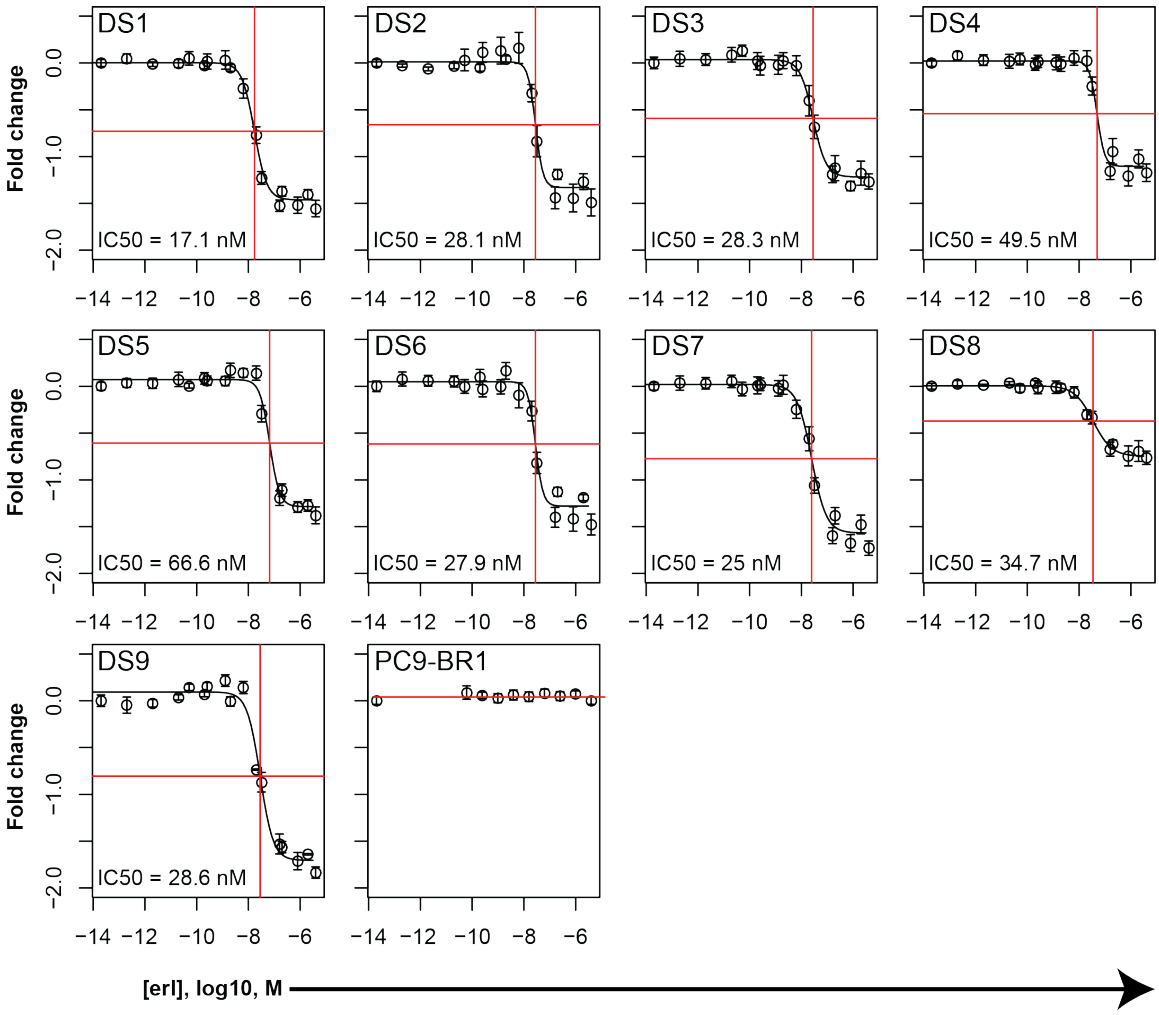

Figure 20 PC9 DS sublines remain sensitive to erlotinib as assessed by IC50 values. PC9 sublines were plated at 4,000 cells per well and allowed to adhere overnight before treatment with erlotinib at various concentrations (in triplicate, n=3 experiments). The cell number for each well was obtained by directly counting cell nuclei using automated image processing software. The data are expressed as $\log_2$(cell number) normalized to vehicle control at 72 hours. Data points represent the average value of all wells at a given concentration. Error bars represent the standard error of all samples for each unique erlotinib (erl) concentration. Curve fits and IC50 values were obtained by using R statistical software. The red lines indicate the concentration of the calculated IC50 value (vertical red line) and the relative change in cell number at that value (horizontal red line).

	PC9-BR1	DS8	H12	H07	DS7	DS4	Parental	DS9
DIP rate	0.0370	0.0102	0.0084	0.0087	0.0040	0.0026	-0.0052	-0.0265
EGFR exon19del	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
EGFR T790M	Yes	No	No	No	No	No	No	No
All other mutations by SNaPshot[36]	No	No	No	No	No	No	No	No

Table 1 Mutational analysis of positive DIP rate PC9 clonal sublines. The NSCLC SNaPshot platform (Su et al., 2011a) was used to assess the mutational status of PC9 parental, PC9-BR1, DS9 (negative DIP rate) and five other PC9 clonal sublines with positive DIP rate. All of the samples retained the activating EGFR exon 19 deletion mutation, while only PC9-BR1 contained EGFR T790M. No additional mutations were detected in any of the samples.

the erlotinib response at high time resolution (every 10 min over 6-days) in two positive and two negative DIP rate sublines.

In agreement with their negative DIP rates, DS1 and DS3 are enriched for dying cells, but quiescent and dividing cells are still observed out to 150h (Fig. 21a). By contrast, essentially no dying cells are observed in sublines with positive DIP rate. However, FP analysis (Tyson et al., 2012) indicates that DS4 and DS7 DIP rates, although similar, are rooted in different combinations of cell fates: DS4 has more cells that enter quiescence and fewer cells with a reduced division rate; vice versa, DS7 has fewer quiescent cells and more with a reduced division rate (Fig. 21b–c). Thus, the DIP rates of DS sublines, including positive ones, are determined by a composite of multiple cell fates. By contrast, cells with acquired resistance mutations in drug (PC9-BR1 (Chmielecki et al., 2011)) and control- treated cells exhibit uniform fate survival and division (Fig. 21b–c).

This raised the question: If a distribution of clones explains the single-cell fate heterogeneity of the parental population, then will discrete sublines, also having cell fate heterogeneity, also display clonal variance of DIP rates? To test this, we performed cFP on three select DS sublines (Fig. 22). Each of the distributions is suitably fit with a normal distribution. We find that the mean of clonal DIP rate distribution matches with the previously measured subline data (Fig. 19b). The DIP rate variance of DS8 and DS9 closely matches that of the parental, whereas DS6 displayed a narrower variance. Nevertheless, DIP rate variance is clearly evident, even in low-passage clonal cell populations.

Together, these results define unique features of positive DIP rate clones: 1) they are detectable in a cell line population at relatively high frequency (~2–15%; Fig. 17a);

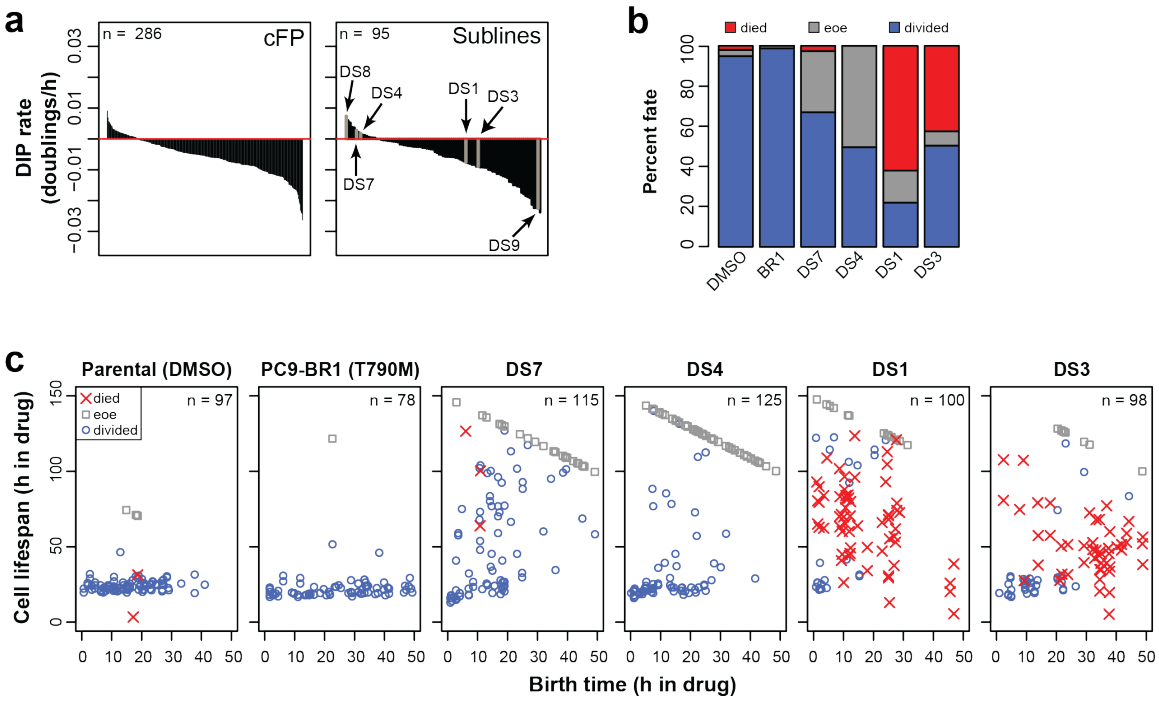

Figure 21 Clonal PC9 DS sublines act as a surrogate for cFP clones. (a) Waterfall plots of clonal DIP rates from cFP assay (*left*) or a collection of 95 PC9 clonal DS sublines (*right*), with the position of six representative sublines (used in *c*) indicated. (b, c) Single-cell fates fate tracking of DMSO-treated PC9 parental cells, four DS sublines continuously treated with erlotinib, and erlotinib treated PC9-BR1 (with a second-site T790M EGFR mutation), a previously-derived model (Chmielecki et al., 2011) of acquired resistance. (b) Summary of cell fates for cell populations tracked in *c*. (c) Single-cell tracking indicates different mixtures of cell fates can result in similar DIP rates. Single cells born during the first 50 h of erlotinib exposure (x-axis) exhibit variable cell fates: Blue circle = division; red X = death; grey square = neither, end of experiment (eoe).

2) they enter multiple fates in response to drug (Fig. 20c); 3) they continue to proliferate in the presence of erlotinib at a reduced, drug-induced proliferation rate that is stable after 72 hours (Fig. 18a); 4) they maintain IC50s in the nanomolar range similar to parental (Fig. 20); 5) they harbor no known acquired resistance mutations (1); 6) due to DIP rate variance, they arise even with clonal populations (Fig. 22).

Methods

Cell culture. All cells were cultured in RPMI 1640 (ATCC) growth media supplemented with 10% fetal bovine serum. Cells were maintained in a humidified cell culture incubator kept at 37C with 5% CO2. PC9, PC9-BR1, 11-18, HCC827, and H1975 cells were kindly provided by William Pao (Vanderbilt University Medical Center). HCC4006 and PC9 (MGH) cells were kindly provided by Jeff Engelman (Massachusetts General Hospital). Cells were checked regularly for mycoplasma contamination and no mycoplasma positive cell lines were used in this study.

Antibodies and inhibitors. Erlotinib (LC labs) was solubilized in dimethyl sulfoxide (DMSO, 10mM) and stored at -20C in single-usage aliquots and used at 3 μM unless otherwise noted. Negative controls for erlotinib received an equal volume DMSO treatment. Primary antibodies for EGFR, pAkt, Akt, pErk, Erk, and cleaved PARP were obtained from Cell Signaling Technologies; p-EGFR (Y1068) obtained from Abcam and Ki-67 obtained from Calbiochem. Secondary HRP-conjugated antibodies were obtained from GE Healthcare. Alexa Fluor fluorescence- conjugated secondary antibodies (Invitrogen) were used for immunofluorescence detection.

Western blotting. Cells were grown to 70% confluence in 100mm tissue culture dishes. Then the individual plates were treated with erlotinib or vehicle, replaced every 3

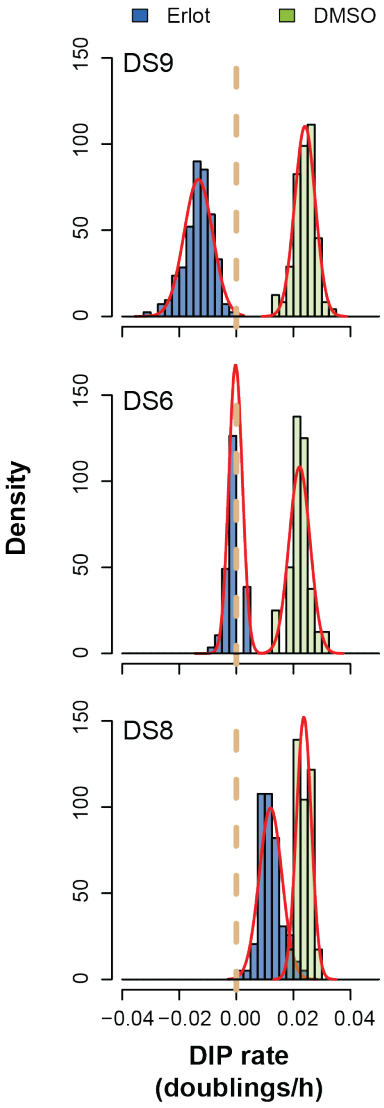

Figure 22 Clonal DIP rate distributions of discrete sublines. Clonal DIP rate distributions of three PC9 DS sublines in the presence (blue) or absence (green) of 3μM erlotinib (*left*). Red curves represent the normal distribution best fit to the data.

days for longer timepoints. To collect cell lysates, the plates were washed once with 10 mL ice-cold PBS and pelleted by centrifugation at (15,000g) in a 4C centrifuge. Cell pellets were resuspended in 100 µL lysis buffer (0.2 M Tris-HCl, 1 M NaCl, sodium deoxycholate, 100 mM NaF, 100 mM NaVO4, Tergitol, Sigma protease inhibitor (cat# P8340) for 10 minutes. Cell lysates were cleared by centrifugation. Lysates were snap frozen and stored at -80°C until analysis. Protein content was determined using a Pierce protein assay. Lysates were thawed on ice and combined with sample buffer (NuPage) and dithiothreitol (DTT). 25 µg of lysates were separated using 4–12% Bis-Tris gels (Novex) and MOPS running buffer (Invitrogen). The proteins were then transferred at 80 mV for 2 hrs (4C) to a PVDF membrane through using 20% MeOH and NuPage transfer buffer (Invitrogen). Membranes blocked for 1 hour using Tris-buffered saline containing 0.1% Tween-20 (TBST) and 5% (w/v) milk. Primary antibodies added at recommended dilution and incubated overnight at 4°C with gentle rocking. The membranes were then washed 5X with TBST and horseradish peroxidase-conjugated secondary antibodies were added in 1% milk-TBST. Secondary antibodies were detected using enhanced chemilluminescence (PerkinElmer) and exposure to autoradiography film. Exposed and processed film was scanned using a Umax Powerlook 1000 scanner.

IC50 assay. Cells were seeded at 4,000 cells per well and allowed to adhere overnight. Then erlotinib was added to the cells at various concentrations in triplicate. After 72 hours, cells were fixed in 4% paraformaldehyde-PBS for 15 minutes, then stained with Hoechst 33342 (Invitrogen, 1:10,000 in PBS) for 15 minutes. The cells were imaged using the Synentec CellaVista imager (10X objective, 3X3 montage) and the total cell number per well was quantified using automatic segmentation software. Percent control was calculated as the average cell number of three wells per condition (n=3 experiments per cell line) normalized to the cell number of DMSO treated cells. The

fold change represents the normalized log_2-difference in final cell number, relative to untreated, induced by the drug. For each subline, the pooled data from three separate experiments were used to compute the IC50 value. The IC50 value were fit using the 'nls' function in R (http://r-project.org) using a 4-parameter logistic model formula (Sebaugh, 2011).

Immunofluorescence. Cells were grown in a BD 96-well imaging plate and at the appropriate time fixed using 4% paraformaldehyde- PBS for ten minutes, washed in PBS, and stored in PBS at 4C. Cells were permeabilized with a blocking buffer containing 0.3% Triton-X and 5% normal goat serum. Primary Ki-67 antibody (1:100, Calbiochem) incubation went overnight at 4C. Cells were washed three times with PBS, and then secondary antibody (1:1,000) was added for 1 hour in blocking buffer. Cells were counterstained with Hoechst 33342 (Invitrogen, 1:10,000 in PBS, 15 min) and imaged using the Synentec Cellavista high-end imager using a 10X objective and 3X3 montage. The percentage of Ki-67 positive cells was quantified as the number of cells exceeding a fixed intensity threshold normalized to the total number of cells quantified by Hoechst 33342 (4 replicates per condition). Samples from all experiments (n=3) were labeled and imaged at the same time to reduce technical variability.

Live-cell imaging of PC9 sublines. For single cell tracking and dynamic cell population measurements, cells were labeled, imaged, and analyzed as previously described (Tyson et al., 2012). Rate determination was made using the R programming software (R-project.org) by fitting linear models to normalized data sub-sampled every hour.

Clonal Fractional Proliferation (cFP) Assay. Cells were seeded in full growth media at single-cell density (50-200 cells per well, depending on cell line growth

characteristics) and allowed expand (~1 week) until the cells per colony was sufficiently high enough to ensure stable typical colony growth characteristics for over a ten day response. Colony sizes were limited so that most colonies would grow independently. When colonies reached the appropriate size, 3 μM erlotinib was added immediately before the day zero data point and replaced every 3 days. Subsequent data were obtained by fluorescence imaging using the CellaVista imager (SynenTec) with a 10x objective across whole wells of a 96-well plate. Between timepoints, cells were maintained in cell culture incubators. After 10-days of experimentation, fluorescence images were converted into image stacks such that serial time points were spatially registered to facilitate single-colony tracking. Data quantitation was performed as previously described (Tyson et al., 2012).

High-throughput derivation of PC9 sublines. In order to isolate drug-naïve clonal sublines in a high throughput manner, we sorted single fluorescently labeled PC9 cells by flow cytometry into wells of multiple 96-well plates. These single-cell derived sublines were expanded in culture for 21 days, and then split into two cultures: one culture for DIP rate measurement and one frozen down.

Discussion

Overall, our results show that a synthesis of cell fate heterogeneity with clonal fitness variation can advance our understanding of targeted therapy dynamics at the cell population level. Essential to our approach is the Drug-Induced Proliferation (DIP) rate, a novel metric that encapsulates multiple cell fates into a steady rate of clonal proliferation. In a nutshell, the DIP rate is a metric of long-term clonal fitness in a population. The temporal quality of the DIP rate is essential to predictive power because, by tracking speed of proliferation of a clone, it can project forward in time its functional contribution

to the overall population dynamics. In contrast, previous metrics of cell fate heterogeneity are at the single-cell level without consideration of proliferation rate and/or relationship to clonal expansion, or at a single (or few) time point that imposes severe constraints on predictive ability. More broadly, a time component is also largely absent from current industry-wide standard assays for anticancer drug screening. The DIP rate removes these limitations and provides predictive power. Since the cFP assay can be deployed in a high-throughput fashion and requires only a few thousand cells to perform, it is suitable to both preclinical applications, such as single or combination drug screening with cell lines, and clinically relevant samples from primary tumor specimens or patient-derived xenografts.

That clonal DIP rates vary along a normal distribution in an oncogene-addicted cell line is a surprising finding (Fig. 17c), because clonal lineages from cell lines harboring oncogene-activating mutations have previously been classified as either sensitive or resistant to a targeted drug. In contrast, our findings suggest a graded clonal response along a continuum, even within cell populations with minimal genetic background differences (as in the PC9-DS sublines that were strictly derived from single cells and maintained at low passage number). Thus, it appears that profiling clonal response by means of DIP rates provides a more realistic assessment of oncogene addiction, revealing a previously unsuspected structure of clonal response heterogeneity. Furthermore, the clonal DIP rate normal distribution structure immediately identifies which clones which may be responsible for the inevitable rebound to treatment. These expanding clones can be studied as sublines, which may reveal the molecular underpinnings of their phenotype.

Individual single-cell fates are present within clones at different proportions, giving rise to steady DIP rates (Fig. 21c). Our data suggest that the propensity to a particular mixture of cell fates is partly inherited, partly stochastic within a clonal lineage. The finding that that inheritance of cell fate can be distributed amongst clonal lineages intrinsic to the population expands upon previous studies that demonstrate heritability of heterogeneous cell fate (Spencer et al., 2009) in the entire population. Accounting for cell fate heterogeneity by single-cell tracking is a laborious task (Tyson et al., 2012), thus substantially limiting both sample throughput and discovery of an underlying structure that may explain the total distribution of cell fates. In contrast, measuring clonal DIP rates effectively accounts for the functional contribution of multiple cell fates without having to measure them directly. This allows for conclusions to be drawn from much smaller datasets. Therefore measuring proliferation rates, rather than single cells, greatly enhances the biological sample throughput to allow highly parallelized proliferation assays.

Expanding clones emerge at the right tail of the normal distribution (Fig. 17c). Thus, though treated oncogene-addicted cell populations may display massive apoptosis initially, our results suggest that DIP rate variance ensures the existence of positive DIP rate clones at some initial frequency. Future studies are warranted to establish mechanisms that may control the proliferation rate of this trait in cancer vs. normal cell populations or cell lines. Clonal selection is already a well-established means of tumor progression and drug resistance, yet it typically refers to a rare, random genetic mutation conferring higher fitness that is apparent under heavy and prolonged selective pressure. In contrast, we see a relative high frequency of clones with positive DIP rate (Fig. 17c) prior to any drug selection. Further studies should determine the capacity of these clones in contributing to relapse to therapy.

Even though single-cell derived clones that proliferate at a positive DIP rate appear to be intrinsic to a cell line (inasmuch as the right tail of the DIP rate distribution extends into positive values) their contribution to evading targeted therapy has remained obscured in traditional assays because they rely on either averaged or static measurements. Instead of a functional heterogeneity continuum, conventional assays have emphasized discrete clonal subpopulations, generally rare ones, that are altogether unresponsive to the targeted drug, and that behave as either resistant or persister cells. Our experimental data indicate that expanding clones, intrinsic to the cell line, are unequivocally distinct from clones exhibiting genetically acquired resistance (Bozic et al., 2013) because of unchanged erlotinib IC50 sensitivity (Fig. 20), lack of known acquired-resistance mutations (1), and reduced proliferation in erlotinib (Fig. 17). Non-genetic mechanisms in rare clonal subpopulations may also explain rebound to cancer drugs. In one case, Sharma et al. showed that a discrete persister subpopulation of non-cycling PC9 cells survives under continuous drug selection via a chromatin-mediated state (Sharma et al., 2010b). The relationship, if any, between positive DIP rate clones and previously described resistant or persister clones remains to be established. It is however tempting to speculate that positive DIP rate clones may represent a reservoir for cells that eventually acquire genetic (or non-genetic) resistance, as in the assimilation phenomenon (Waddington, 1953).

The DIP rate normal distributions of parental cell lines can be recapitulated, but not discretized, by clonal sublines. That is, a clonal subline itself responds to drug with a normal distribution of DIP rates (Fig. 22). The mean of this distribution differs from parental, but the variance is roughly maintained. This finding suggests that variance is not constrained by the mutation rate required to diversify a phenotypic trait and thus may not require a genetic basis. Thus, we speculate that it may have its origins in stochastic

processes, as is the case for other phenotypic traits that exhibit a normal distribution in a population. Along these lines, Kreso et al. found no detectable genetic differences that correlated with differential behavior or treatment response among colorectal cancer clones serially co-passaged as tumor xenografts (Kreso et al., 2013). In that report, clonal lineages were presumably phenotypically discrete, possibly due to discrete epigenetic events (Kreso et al., 2013; Marusyk and Polyak, 2013). In contrast, we observed that clonal DIP rate variation is continuous and normally distributed, strongly indicating it arises from stochasticity in biological processes. This finding is reminiscent of biological noise-based survival strategies in unicellular organisms, such as bet-hedging strategies in bacteria (Fraser and Kaern, 2009). Nevertheless our results, consistent with Kreso et al. (Kreso et al., 2013), suggest that functional clonal profiles of relatively homogeneous cancer cell populations are more complex than random genetic diversification alone.

In summary, DIP rates effectively act as a bridge to join heterogeneous cell fates in response to perturbation to the underlying clonal structure of a cell population. A main motivating question behind improving cancer therapeutics is: why do some cells respond to therapy while others don't? In light of widespread biological heterogeneity (Altschuler and Wu, 2010), finding mechanistic explanations of cellular behavior at the single-cell level is complex and can quickly become unwieldy. In contrast to understanding the mechanistic basis of single-cell behavior, dynamic measurements of clonal proliferation integrate these multiple single-cell fates and cast them as an ordered distribution of rates. Interestingly, positive DIP rate clones appear to be an intrinsic component of a mainly negative distribution. The mechanisms that govern clonal DIP rate remain to be determined. Nevertheless positive DIP rate clones appear unique from previously described genetic and epigenetic mechanisms of cell survival during therapy. Thus these

clones may provide a new therapeutic window to eliminate more tumor cells before the population develops resistance thus prolonging treatment durability.

Chapter IV

INTEGRATING CLONAL FITNESS HETEROGENEITY PREDICTS THERAPY RESPONSE DYNAMICS

Introduction

Intratumor cancer heterogeneity has been recognized for decades, as exemplified by the seminal work of Fidler (Fidler and Kripke, 1977) and Heppner (Dexter et al., 1978). A resurgence of interest in heterogeneity is driven by its essential role in causing short-lived remissions observed with the latest targeted therapies (Bozic et al., 2013). Both genetic and non-genetic mechanisms contribute to intratumor heterogeneity (Marusyk et al., 2012). Recent studies have elegantly shown that a tumor is a composite of clones (Ding et al., 2012; Gerlinger et al., 2012; McGranahan et al., 2012; Welch et al., 2012) with common origin but genetically diversified by mutations and genomic instability, which may become substrates for selection towards increased fitness against microenvironmental stresses such as therapy (Bozic et al., 2013; Dexter et al., 1978; Fidler and Kripke, 1977; Nowell, 1976).

While intratumor heterogeneity underlies poor clinical outcomes, it remains difficult to quantitatively infer its functional role in therapy failure, especially in a predictive manner. Mathematical modeling can provide a theoretical framework to link intratumor heterogeneity to outcomes (Michor et al., 2004). A classic example is drug-induced clonal selection: treatment applies a selective pressure on a mixture of cells defined as sensitive and resistant. Over time the population initially shrinks due to depletion of treatment-sensitive cells. However, continuous treatment selects for the emergence of a rare resistant clone that is either pre-existing or stochastically arises and

the population eventually regrows despite continued therapy. By translating biological assumptions of cellular behavior into steady rates of change from a set of initial conditions, models can simulate the changes in complex processes such as cancer treatment response and disease progression (Mumenthaler et al., 2011). These models, in turn, can be examined to quantify the relative contribution biological features have on driving the phenotype of interest.

One example of a tool for modeling evolving population dynamics is an ordinary differential equation (ODE) compartment model. ODE compartment models are well established in fields such as disease epidemics (Webb et al., 2005) and pharmacology (Sheiner et al., 1979). For this type of model, a population is defined as the sum of multiple subpopulations denoted as compartments. The dynamics of each compartment are defined by steady rates of entry or exit from that compartment, which depend on the current size of the compartment from which individuals are exiting. Thus, given knowledge of the initial composition of a population and the rates that individuals transition to and from subpopulations, a compartment model can predict the population dynamics over time. In the case of cancer, clinical data match model expectations, especially in the case of targeted therapy. The assumption is that the onset of resistance can be explained by either selection of rare resistant clones or transition into a treatment-resistant compartment (Bozic et al., 2013). In the model, this transition may be defined by estimated rates of DNA mutation or differentiation.

The power of these models is illustrated when interfacing model outputs with clinical data to infer the cellular dynamics that underlie tumor rebound in patients. A notable example of modeling clinical data with an ODE compartment model is the response of patients with chronic myeloid leukemia to targeted therapy (Michor et al.,

2005). The authors show that a combined differentiation-genetic resistance model is sufficient to explain the clinical phenomena of biphasic reduction in cell number and the onset of resistance. By constraining the model output with clinical data, then estimations can be made for experimentally challenging measurements, such as the pretreatment number of resistant cells present. Thus, the output (clinical outcome) is known, but the parameters that drive the model are allowed to float to fit the response data. Because this approach begins with clinical data and builds a model that can accommodate that can fit that data, this strategy can be considered a top-down modeling approach. The biologically-informed model can then be used to explore the phenotypic space that the model occupies to inform novel therapeutic regimes with optimized outcomes. For example, models predict that the fitness disadvantage of resistant cells in the absence of drug treatment can be exploited by pulsing drug administration at optimal intervals (Chmielecki et al., 2011). Additionally, modeling approaches can predict the impact of treatment combinations on reducing the likelihood of genetic resistance (Bozic et al., 2013).

One limitation of modeling from a top-down perspective is that it must assume knowledge of the underlying structure of the population. While models may provide insight into the relevance of biological parameters, it is possible that the model can fit the data phenomenologically while omitting the relevant disease biology entirely. In the examples above, the model structure is constructed with the help of literature-based knowledge to fit the clinical data. If the model prediction can fit the data, then it is plausible that the model explains the data (Kohl et al., 2010). In contrast to top-down approaches that work backward to construct the model and fit the model parameters to match output data, model assumptions can be defined explicitly by data-driven bottom-

up modeling. This can help to constrain the model to better reflect the actual biology and further enable potential detection of novel biology.

Experiments define the biology precisely, models account for their contribution in a complex system. Thus math modeling and experimentation can therefore be used iteratively to simplify complex biological systems to their most relevant constituent traits (Anderson and Quaranta, 2008). Experimental measurements are especially important in cancer targeted therapy models because, as described throughout Chapter 1, there are many sources of heterogeneity that can account for treatment failure. Therefore assuming that tumor rebound is driven by only one source of heterogeneity may be a major biological oversight. Furthermore, compartment models assume that the population can be subdivided into uniquely defined subpopulations with certain behaviors. In reality, drug response may be a continuous variable (Brock et al., 2009) unsuitable for compartment models. Experiments are thus a powerful tool to directly test the biological assumptions that the model relies upon.

We have demonstrated this previously with single-cell behavior and population dynamics (Tyson et al., 2012). There are several advantages that can be exploited once the model is constructed and thoroughly validated: 1) With output measurements, a model can then estimate the input data; 2) With input measurements, the model can extrapolate population dynamics; 3) Varying input parameters can give predictions of which parameters are most critical in modulating the population outcome; 4) Fixing output parameters can show the range of model inputs that can generate a particular population behavior. Thus examining model predictions can then quickly test a combinatorial matrix of conditions and highlight potentially high-impact biological hypotheses to be tested experimentally.

Chapter 2 described novel experimental methods for quantifying heterogeneous cellular response to perturbation within a cell population; Chapter 3 characterized the clonally heterogeneous response of EGFR-mutant cell lines to erlotinib; this chapter focuses on predicting the impact of heterogeneous cellular and clonal drug responses to the overall population outcome using bottom-up, experimentation-driven modeling. Beginning with assumptions based on the biology defined in Chapters 2 & 3, we construct a model, the Heterogeneous Growth (HG) model, which relates clonal growth to population rebound under continuous therapy. We then test both the model assumptions and the model predictions experimentally. Then we subsequently use the model to find that, counterintuitively, that ideal drug combinations should focus on reducing positive DIP rate clones by constraining the proliferation heterogeneity.

Results

The variance of a clonal DIP rate distribution drives time-to-rebound (TTR)

In Chapter 3 we described that clonal response to erlotinib is suitably described by a normal distribution of drug-induced proliferation (DIP) rates. A normal distribution is entirely defined by two numerical parameters, mean (μ) and variance (σ^2), enabling translation of the erlotinib-induced heterogeneity of cell fates, encapsulated within clonal DIP rates, into population dynamics by mathematical modeling. All measured cell lines can be suitably represented with a normal distribution (Fig. 23). Since traditional exponential growth models generally only consider the mean (μ), we derived a Heterogeneous Growth (HG) model of exponential growth that incorporates both mean and variance of DIP rate distributions (Methods).

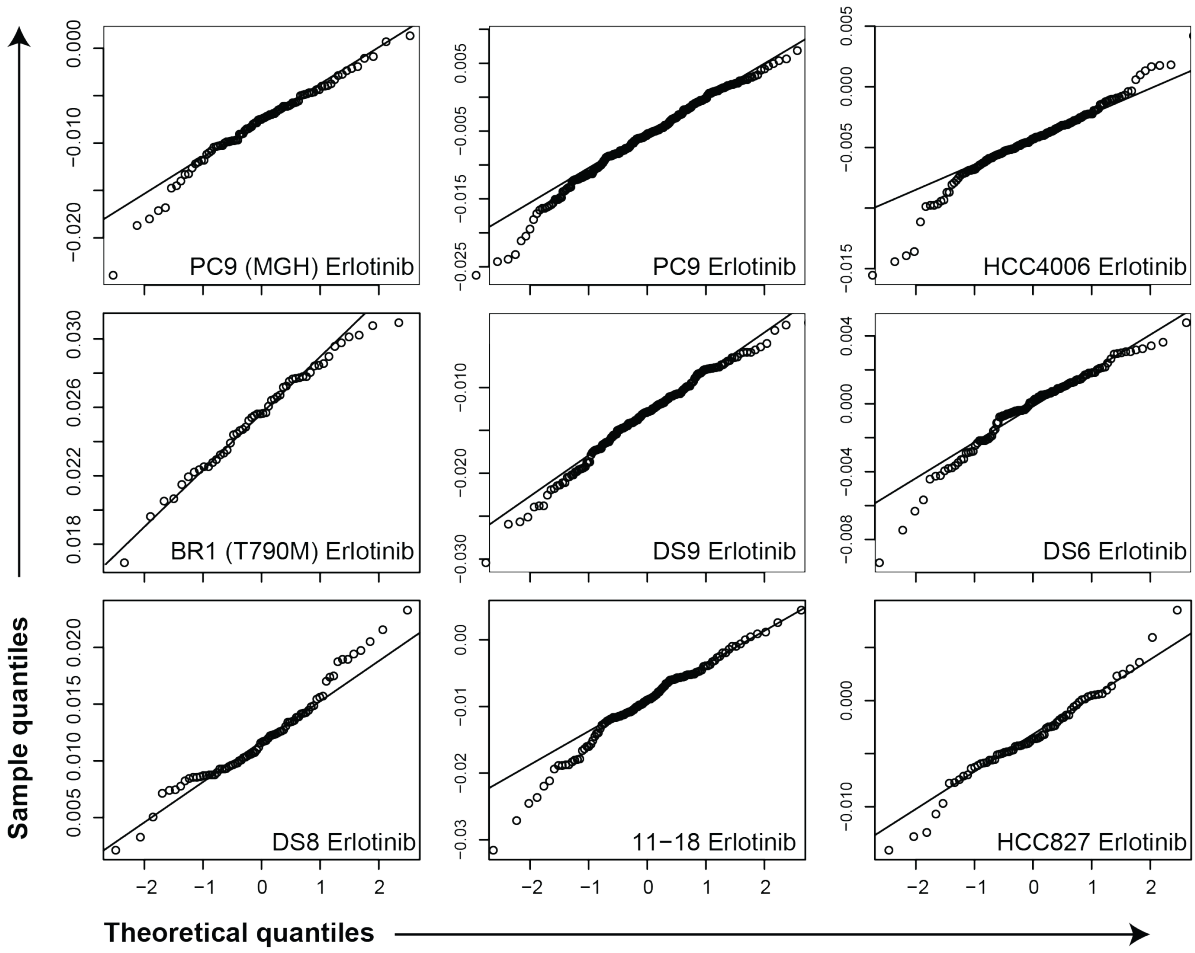

Figure 23 QQ-plots validate assumption of normality. QQ plots were generated using the R (http://r-project.org) function qqnorm on the colony DIP rates. These plots show the DIP rate versus the normal quantiles of the sample. Data was sufficiently close to normal in most cases, with deviation from normality on the left (low DIP rates, i.e. rapidly dying colonies). This validates the assumptions of normality when integrating to develop the HG model.

We plotted the HG model prediction of erlotinib response over a 20-day time course (Fig. 24a–b) using parameter values (μ, σ^2) from the clonal DIP rate distribution of parental PC9. Initially, the size of the population declines in the HG model simulation, as expected from the fact that the majority of the clones exhibit negative DIP rate and the μ of the DIP rate distribution is negative. However, it is important to recognize that even during this initial phase the small fraction of expanding clones are proliferating, though with minimal impact on the overall size of the population. Due to exponential expansion, the contribution of this minority of clones to the population dynamics increases over time and overtakes the regressing fraction at the inflection point (~5 days, Fig. 24b). After this minimum (depth of response = -0.60 doublings) the population size begins to increase steadily, reaching initial size at 10.3 days (Fig. 24b). We refer to the time required for the population to regain its initial size in continuous drug treatment as 'time-to-rebound' (TTR). Note that the same model, with variance excluded, gives rise to a very different prediction in which there is no rebound (red line in Fig. 24b).

An intriguing insight from the HG model is that TTR is disproportionately affected by the σ parameter relative to the μ parameter, as indicated by the analytical solution ($\hat{t} = -2\frac{\mu}{\sigma^2}$). More precisely, decreasing the μ results in a deeper initial response but only incremental gains in TTR (Fig. 25). In contrast, small decrements in σ substantially lengthen TTR (Fig. 25).

To assess the general value of these findings, we then used the HG model to predict responses across a panel of mutant EGFR-addicted cell lines. Each cell line has a characteristic clonal DIP rate distribution and the HG model predicted both a characteristic depth of response and TTR for each of the cell lines, as a function of its DIP rate distribution parameters (Fig. 26, Table 2). The predicted time course highlights

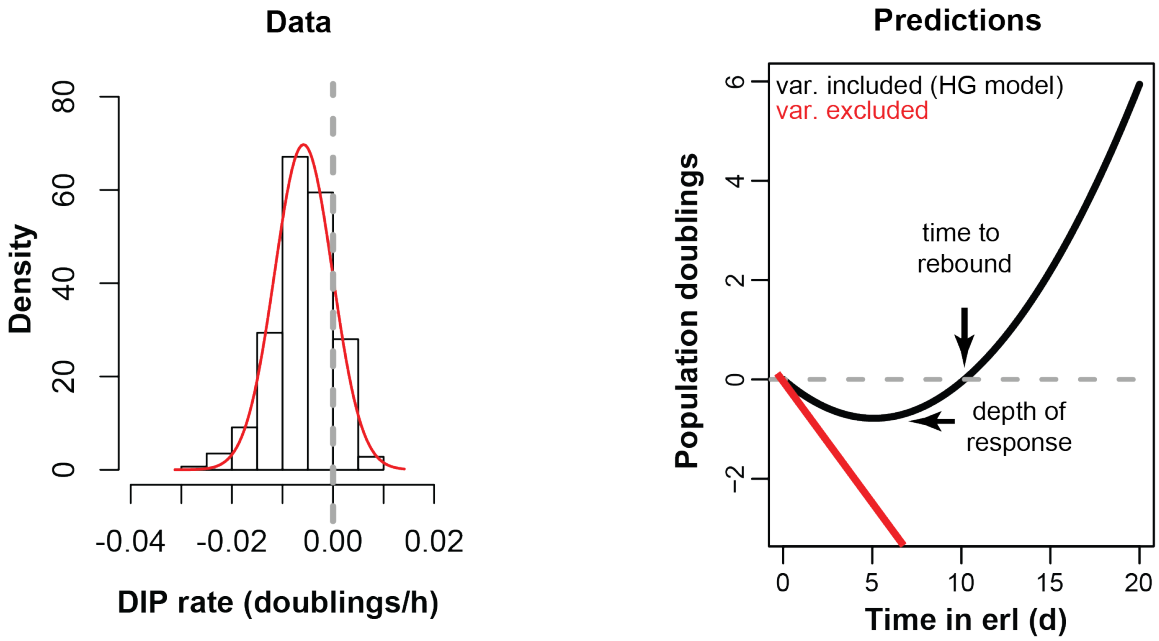

Figure 24 HG model integrates of clonal DIP rates to predict rebound. Incorporation of clonal DIP rate variance into a model of exponential growth predicts time-to-rebound (TTR). (a) Histogram of erlotinib-treated PC9 clonal DIP rates. Red curve represents a normal distribution best-fit line. There is insufficient evidence to reject the normal distribution fit (K-S test, p=0.35). Gray dashed line represents the divide between positive and negative DIP rate clones. (b) The heterogeneous growth (HG) model replaces the single proliferation rate parameter of the basic exponential growth model with the Gaussian probability density function that describes the full DIP rate distribution with the μ and σ parameters. The black line shows the HG model output of continuous erlotinib treatment ($y = y_0 e^{\mu t + \frac{\sigma^2 t^2}{2}}$). This model predicts the depth of response ($y_{min} = \frac{-\mu^2}{2\sigma^2 \log(2)}$) and time-to-rebound ($\hat{t} = -2\frac{\mu}{\sigma^2}$) as a function of the data. Red line shows the output of an exponential model ($y = y_0 e^{\mu t}$) of erlotinib response that incorporates only the mean DIP rate of PC9 (variance excluded).

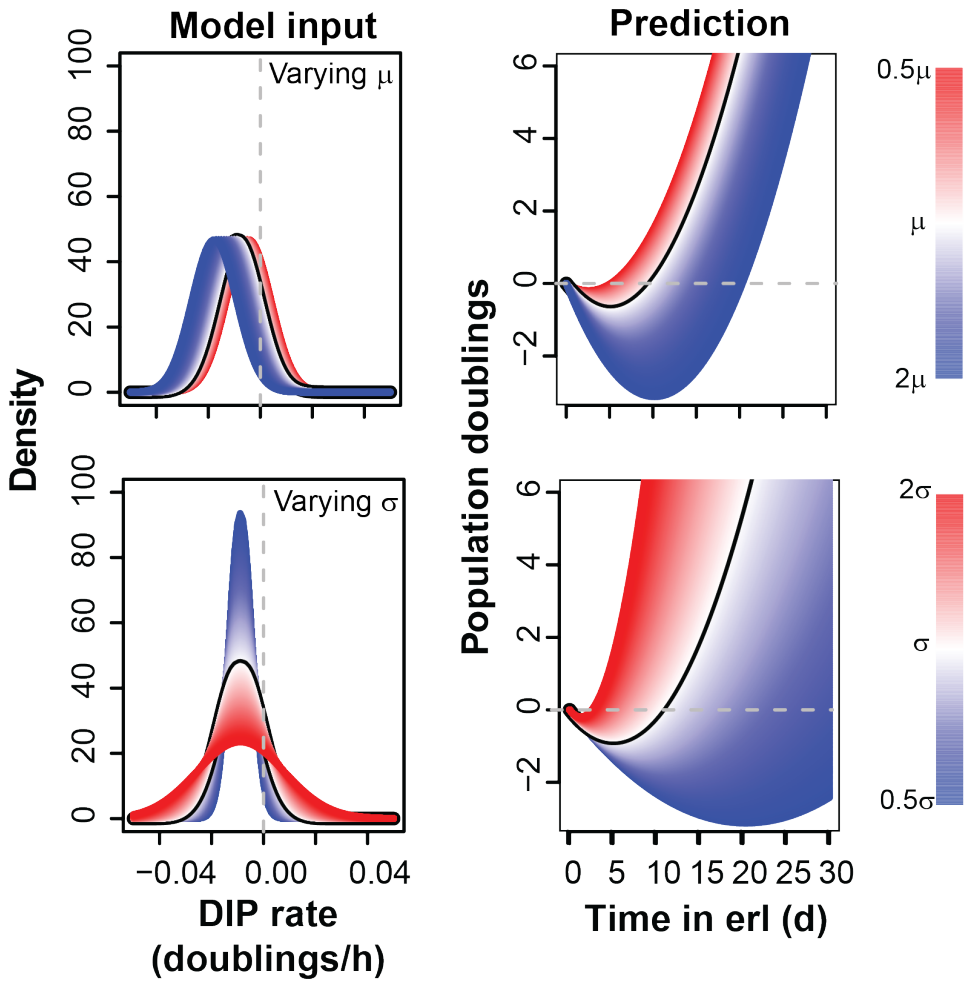

Figure 25 Analysis of HG model parameters. Varying the μ parameter value (*top*) affects depth of response but TTR is modestly influenced, whereas varying the s parameter value (*bottom*) dramatically affects TTR. Color scale: blue = 0.5σ, white = 1σ, red = 2σ. Black line represents the prediction using the parameter values obtained experimentally from parental PC9 cells.

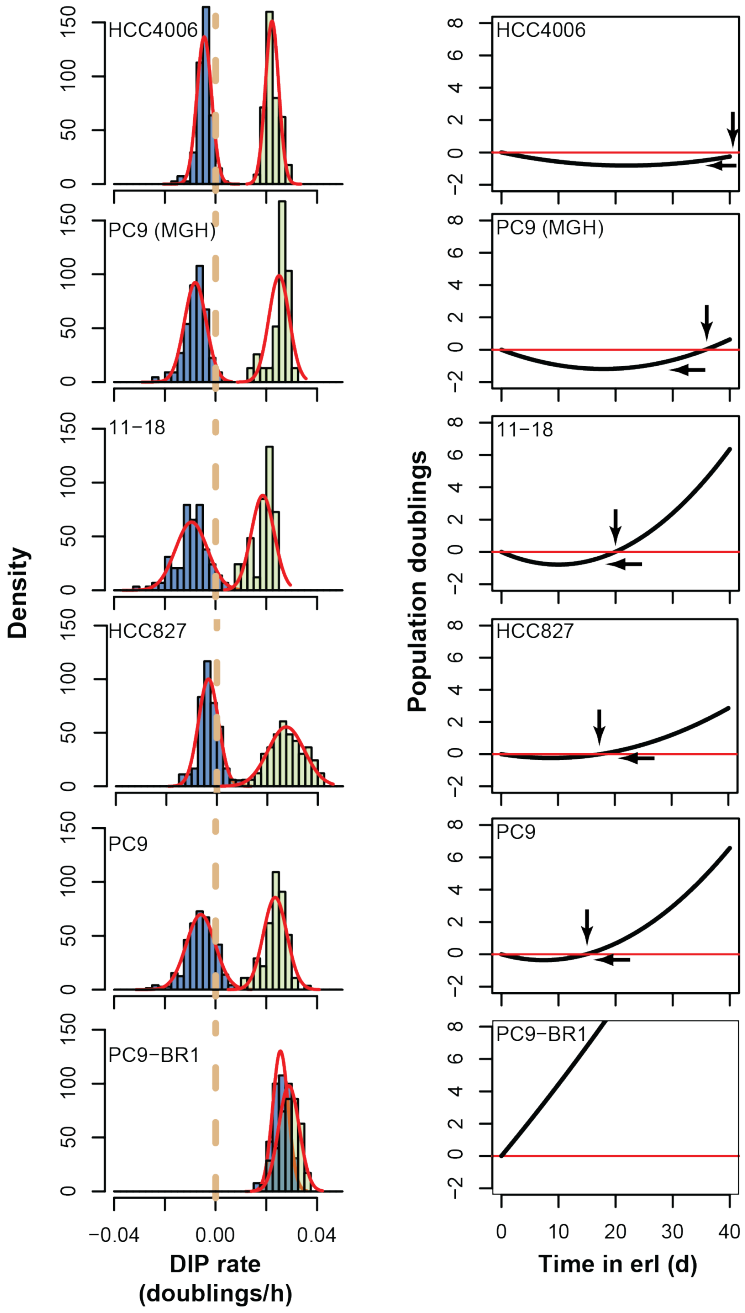

Figure 26 Clonal DIP rates and HG model predictions for EGFR-addicted cell lines. Incorporation of clonal DIP rate variance into a model of exponential growth predicts time-to-rebound (TTR) in erlotinib-treated EGFR mutant cell lines. Wide variation of clonal DIP rate distributions across cell lines in the presence (blue) or absence (green) of 3μM erlotinib. Note that in all instances cFP data were fit by a Gaussian distribution (*left*). Population-level dynamics specific for each erlotinib-treated cell line are predicted by incorporating the Gaussian parameter values of each DIP rate distribution into the HG model (*right*). Horizontal arrows = depth of response; vertical arrows = TTR.

	p-value	D	n	μ	σ	TTR (days)
DS9	0.4617	0.0656	169	-0.0192	0.00724	30.52365
11-18	0.3054	0.0899	116	-0.0123	0.00803	15.92171
PC9 (MGH)	0.4922	0.0864	89	-0.0115	0.00623	24.767
PC9	0.3183	0.0566	286	-0.0084	0.00825	10.28498
HCC4006	0.1369	0.0907	163	-0.0065	0.00421	30.35713
HCC827	0.7182	0.0992	46	-0.0039	0.00533	11.47047
DS6	0.04019	0.1309	114	-0.0004	0.00343	2.528069
H1650	0.2042	0.1211	75	0.01514	0.00872	NA
DS8	0.5909	0.0853	60	0.01732	0.00578	NA
BR1	0.9663	0.066	52	0.03687	0.00442	NA

Table 2 DIP rate distribution parameter values and prediction of relapse. The p-value for the Kolmogorov-Smirnov test is shown, where the normal distribution is the null hypothesis; p-value < 0.05 indicates sufficient evidence that data are not sampled from a normal distribution. The D statistic in the KS-test is the largest observed difference between the model's cumulative density function and the data's cumulative density. n represents the number of colonies used in the DIP rate distribution. μ and σ represent the mean and standard deviation, respectively, of the DIP rate distribution. TTR represents the time-to-rebound, as predicted by the HG model. NA, not applicable because the cell population never decreases in size.

the importance of the DIP rate distribution shape for each cell line. For example, the HCC4006 distribution shifted leftward only half as far as 11-18 (HCC4006 μ=-6.5e-3 doublings/h; 11-18 μ=-1.23e-2 doublings/h), yet results in double the TTR (HCC4006 TTR 30 days; 11-18 TTR 16 days) due to a narrower variance.

We experimentally tested the HG model predictions of erlotinib-treated PC9 parental cells by live-cell imaging 14 separate wells of high-density PC9 cells cultured in erlotinib over 20 days. In each culture well there was an initial decrease in the population followed by a slow, steady rebound, as predicted by the HG model (Fig. 27a). For comparison, an alternative model assuming the presence of rare drug-resistant cells in the PC9 parental population predicted "V-shaped" growth curves and did not agree with the rebound experimental data (Fig. 27b). In contrast, the HG model predictions, incorporating parameter values from the parental population DIP rate distribution, closely tracked curvilinear growth of the majority of the wells, validating the expectation. Furthermore, the observation of rebound corresponded to the exponential outgrowth of positive DIP rate clones (Fig. 27c).

These data and model simulations established the relevance of the clonal DIP rate normal distribution and the power of HG model to predict outcomes of erlotinib therapy in terms of depth of response and TTR. Furthermore, the HG model indicated that the variance is predominantly responsible for the TTR duration. That is, a smaller variance would lead to a desirably longer TTR.

Clonal DIP rate variance compression by combination therapy lengthens TTR

To interrogate the molecular underpinnings of clonal DIP rate variation within a cell population, we utilized a set of PC9 DS sublines spanning the continuum of clonal

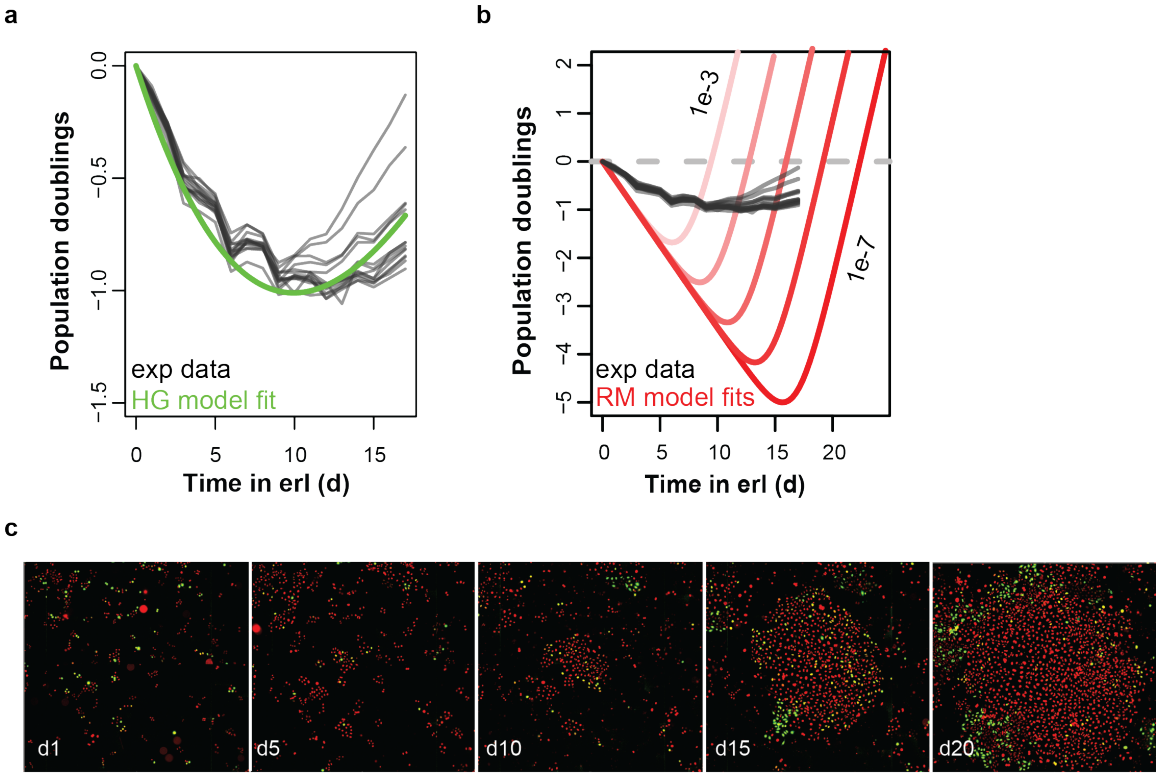

Figure 27 PC9 Rebound validation. (a) The experimental population-level dynamics of erlotinib-treated PC9 cells (gray lines; n=14) were accurately predicted (green line) by the HG model incorporating μ and σ parameter values from PC9 experimental clonal DIP rate distribution. PC9 cells were seeded at high density in 14 microtiter wells, continuously treated with erlotinib over 20 days, and imaged daily. The well-to-well variability is likely due to uneven clonal sampling. (b) A rare mutant (RM) model assuming a population of cells with a single proliferation rate in the presence of erlotinib (the μ parameter value of PC9 cells) and a rare subpopulation of acquired-resistance cells (the μ parameter value of PC9-BR1 cells (Chmielecki et al., 2011)) at indicated ratios cannot explain the experimental data (same as in *a*). (c) Rebound to erlotinib is accompanied by expanding clones. Shown is the imaging timecourse of the well displaying the fastest rebound to erlotinib (gray lines in *a*).

DIP rates derived from cFP. DIP rates act as a continuous variable that can be directly linked to molecular changes across the DS subline panel. Thus, we selected 10 DS sublines that cover the extremes and middle portion of the DIP rate continuum (Fig. 19b), treated them for 5d with erlotinib and then applied microwestern array (MWA) analysis (Ciaccio et al., 2010) (Fig. 28). In total, the MWA covered 48 analytes for each subline (Table 3). To assess significance, we considered analytes significant if they had R^2>0.6. By this criterion, seven analytes correlated with DIP rate by linear regression (Fig. 29a and Table 3). These correlations were validated and extended by traditional Western blots (Fig. 29b). Within the scope of this work, we focused on the EGFR, apoptosis and stress response pathways. In particular we attempted to find analytes that were either correlated or anticorrelated with DIP rate that could be chemically modulated to test their contribution to DIP rate.

To better accommodate the drug combination data, we reasoned that different perturbations may not exert their influence evenly amongst clones and a normal distribution may not be the optimal fit for the data. Therefore we chose to fit the combination data with a skew-normal fit, which is simply a generalization of the normal distribution with an additional parameter, α, in addition to μ and σ. In the case where α is equal to zero, the skew normal distribution yields the original normal distribution (Azzalini, 2005). The HG model, with an assumed skew-normal distribution of DIP rates, still gives numerical predictions of the depth of response and TTR, although these solutions cannot be solved analytically (Methods). Indeed, a strength of the HG model is that it can accommodate a wide variety of distributions. In the EGFR pathway the levels of phospho- and total EGFR, MEK and ERK did not vary sufficiently across the sublines to draw meaningful conclusions regarding their contribution to DIP rates, in spite of the prominent role of the RAS/MAPK pathway in mediating mutant EGFR cellular effects.

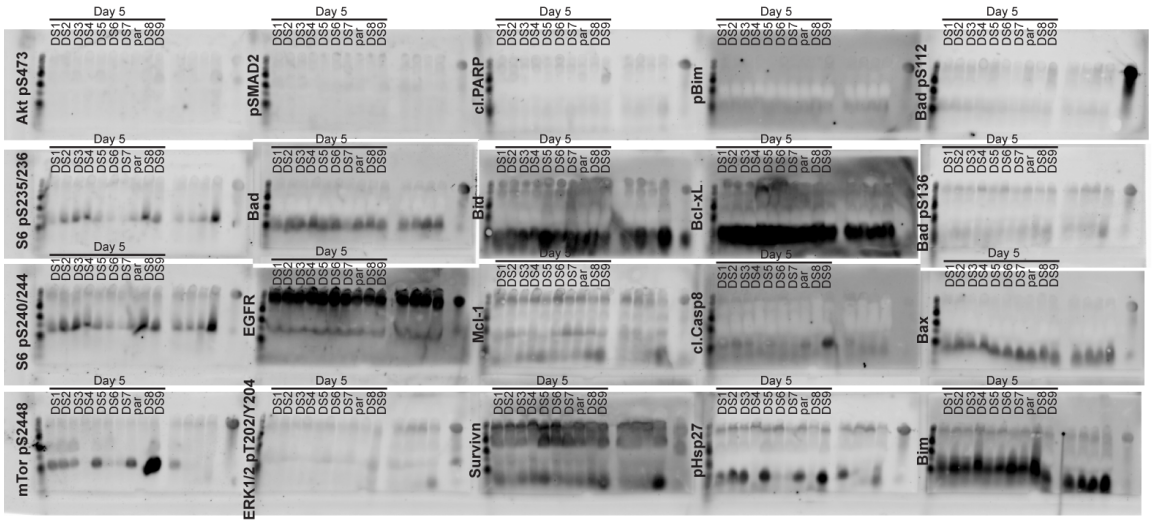

Figure 28 Molecular correlates of clonal DIP rate variation. Representative raw image data of microwestern array. Cell lysates of DS1–9 and parental PC9 were obtained from cells after 5 days of erlotinib treatment. Samples were used at concentrations above 5 μg/μl probed using 48 separate antibodies by microwestern array.

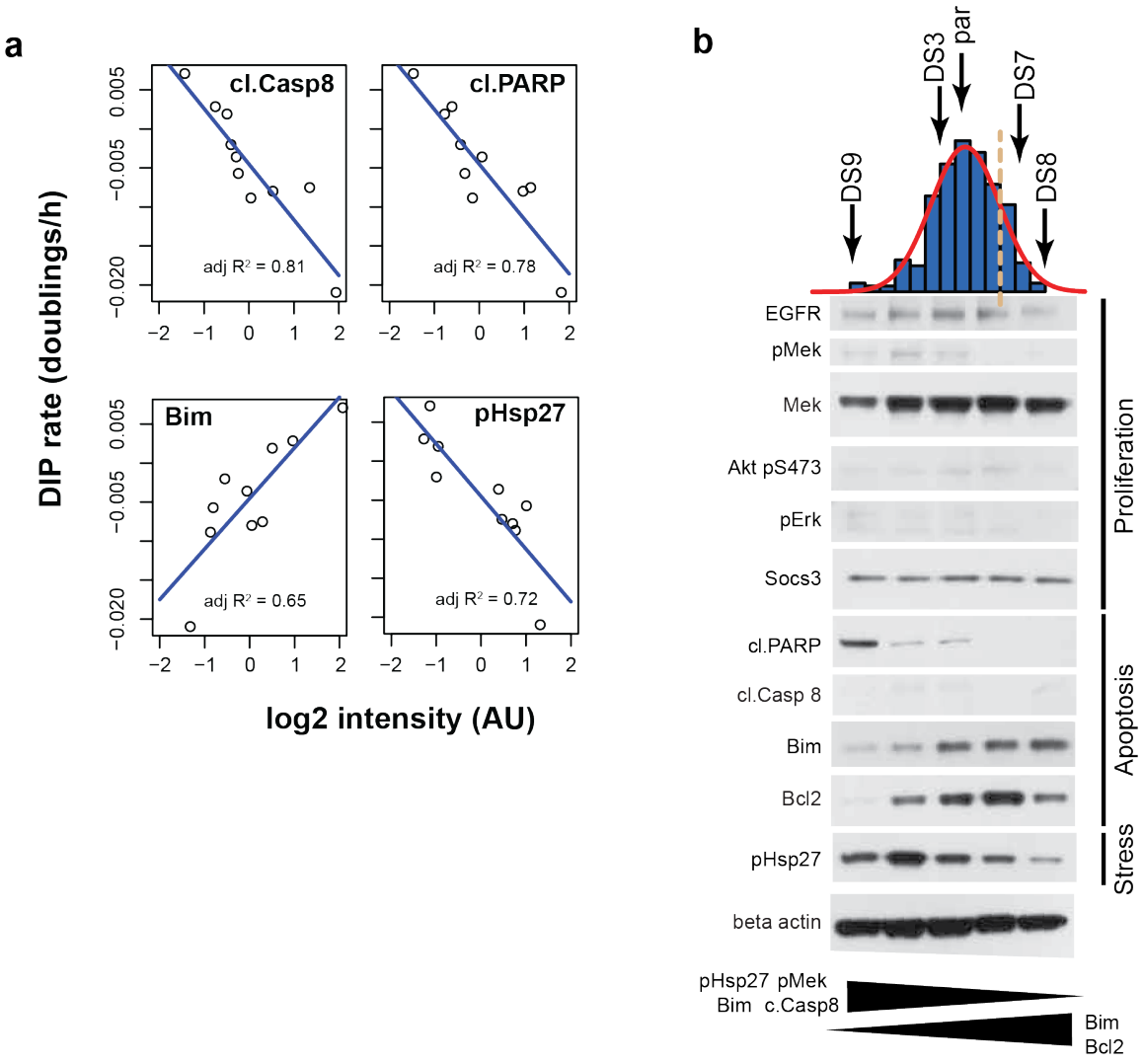

Figure 29 Molecular correlates of clonal DIP rate variation. (a) Cell lysates of DS1–9 and parental PC9 were obtained after 5 days of continuous erlotinib treatment. Samples were probed using 48 unique antibodies by microwestern array (Ciaccio et al., 2010). For each analyte, $\log_2$ intensity values were scaled to the standard deviation across all cell lines and median-centered to 0. Correlation between molecular signals and mean DIP rates was considered significant if a linear model fit the data (R^2>0.6). Each linear fit is shown with a blue line. (b) Western blot validation of analytes for PC9 DS sublines. Mean DIP rates, indicated in the parental (par) PC9 clonal DIP rate distribution, were treated with erlotinib for 5 days and examined for (phospho-) protein levels by western blotting.

Analyte	Phosphorylation site	Functional readout	Correlation with DIP rate (R^2)
p-EGFR	Y1068	EGFR activity	ns
EGFR		EGFR expression	ns
p-Erk	(T202/204)	Mapk activity	ns
p-Akt	(S473)	PI3K-Akt activity	ns
p-S6	(S235/236)	Mapk activity	ns
p-S6	(S240/244)	PI3K-Akt activity	ns
p-mTOR	(S2448)	PI3K-Akt activity	ns
p-4E-BP1	(S65)	Inhibit cell cycling	ns
p-Mek	(S217/221)	Mapk activity	0.71
p-p70 S6K	(S389)	PI3K-Akt activity	ns
p-p90 S6K	(S380)	Mapk activity	ns
p-PDK1	(S241)	PI3K-Akt activity	ns
p-Src	(Y416)	Src activity	ns
Bcl2		pro-survival	ns
p-p38	(T180, Y182)	Stress response protein	ns
Bad		pro-death	ns
p-Bad	(S136)	pro-survival	ns
p-Bad	(S112)	pro-survival	ns
Xiap		pro-survival	ns
Bax		pro-death	ns
Bim		pro-death	0.65
p-Bim		pro-survival	ns
Bcl-xl		pro-survival	ns
p-Sapk/Jnk	(T183/Y185)	DNA damage/stress	ns
Cleaved Caspase-8		Extrinsic apoptosis activity	0.81
p-IKK α/β	(Ser176/180)	NF-kB activity	0.72
p-Hsp27 (Ser82)		Stress response	0.72
p-Mapkapk2	(T222)	Stress response	ns
p-IGF-1R beta	(Y1135/1136)	Survival	ns
p-Met	(Y1234/1235)	Survival	ns
p-GSK3-alpha/beta	(S21/S9)	pro-death	ns
SOCS3		Stat inhibition	0.63
p-Stat5	(Y694)	Stat activity	ns
p-Stat3	(Y705)	Stat activity	ns
p-Stat1	(Y701)	Stat activity	ns
p-AMPK	(T172)	Energy sensor	ns
p-Rb	(S795)	Cell proliferation	ns
cleaved PARP		Apoptosis	0.78
p-Chk2	(T68)	DNA damage	ns
Bid		pro-death	ns
Mcl-1		pro-survival	ns
Survivin		pro-survival	ns
p-FoxO-3a	(S318/321)	pro-death	ns
p-GSK3-alpha/beta	(S21/9)	pro-death	ns
p-SMAD2/SMAD3	(S465/467//423/425)	TGF-beta signaling	ns
Myc		Myc transcriptional regulation	ns
p-MNK	T197/202	Stress response	ns
p-Myc	T58	Myc activity	ns
beta actin		Loading control	ns

Table 3 List of molecular analytes used in microwestern array. Shown are the different antibodies used in microwestern (Ciaccio et al., 2010) analysis of the DS sublines. For phospho-proteins, the specific phosphorylation sites detected are listed. Signaling analytes were designed to test multiple categorized signaling pathways that influence cell fate in response to drugs. Each analyte was tested for individual linear correlation with DIP rate across a panel of ten samples (9 sublines and PC9 parental, see Fig. 19b). Analytes with linear model fits $R^2 > 0.6$ are shown, otherwise not significant (ns).

In the apoptosis pathway, an expected finding was that higher levels of cleaved PARP associated with lower DIP rates (and vice versa). Somewhat unexpectedly, higher levels of pro-apoptotic Bim (Fig. 29a, R^2=0.65) were found in sublines with higher DIP rates. However, in these sublines the anti-apoptotic protein Bcl-2 was concomitantly increased, likely negating the effects of Bim (Fig. 29b). We reasoned that chemically disrupting this interaction would increase apoptosis. In agreement, addition of ABT-737 (a BH3 mimetic that inhibits the pro-survival activity of Bcl-2) to erlotinib-treated PC9 cells, measured by cFP, shifted the mean of the clonal DIP rate distribution to the left, presumably by increasing the rate of cell death (Fig. 30a). The left-shift (-5.8e-3 to -1.2e-2 doublings/hour) of the mean substantially improves the depth of response predicted by the HG model (Fig. 30b). However, by HG mode predictions it produces only a modest benefit on TTR (20 days to 27 days), the more critical element of anticancer drug efficacy.

In the stress response pathway, the levels of phospho heat-shock protein 27 (pHsp27) is relatively diminished in positive DIP rate DS sublines. Hsp27 function is modulated by phosphorylation (Rogalla et al., 1999); however, no direct chemical modulators have been made. Thus, we modulated its phosphorylation by inhibiting a prominent upstream kinase (p38), for which both activators and inhibitors are available (Kuma et al., 2005). We reasoned that if high Hsp27 activity were associated with low DIP rate, then inhibiting p38 would shift the DIP rate distribution mean positively. To inhibit p38, we utilize a specific p38 inhibitor, SB-203580. In contrast to our expectation of shifting the mean, the primary effect of inhibiting p38 was to dramatically increased the DIP rate variance (Fig. 30a; 5.7e-3 to 9.2e-3). By HG model predictions, this would substantially shorten TTR (Fig. 30b; 20 days to 11 days). If p38 inhibition shortens the predicted TTR by increasing the DIP rate variance, we next tested if p38 activation

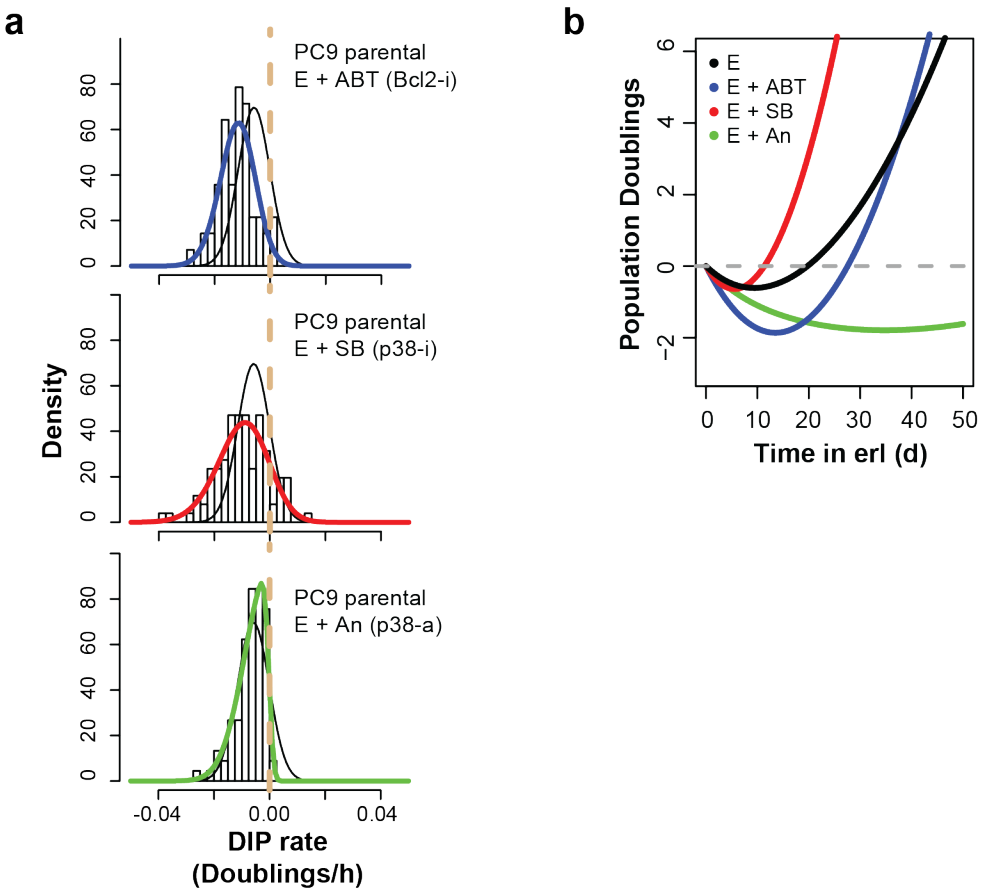

Figure 30 Manipulating clonal DIP rate distribution with combination treatment. (a) clonal DIP rate distributions of erlotinib-treated PC9 cells co-treated with ABT-737 (2.5 μM, top), SB203580 (1 μM, middle), or anisomycin (1μM, bottom). The data are overlaid with estimates of distributions fit to data from erlotinib alone (black curve) or in combination with indicated drug (colorized curve). To better account for the contribution of positive DIP rate colonies, skew-normal distributions were assumed for all model fits. (b) HG model predictions of population-level drug-response dynamics based on the experimental data fits in *a*.

would shrink the variance. Strikingly, through a combination of narrowing the variance (Fig. 30a; 5.7e-3 to 5.2e-3) and increasing the negative skew (4.1e-1 to -9e-1), the p38 activator anisomycin greatly reduces the proportion of positive DIP rate clones. This results in the highly desirable effect of lengthening TTR (Fig. 30b; 20 days to 87 days) in parental PC9 treated in combination with erlotinib. These results show that constraining intra-cell line clonal variance is an attractive means to delay rebound.

Methods

Heterogeneous Growth (HG) model. The Heterogeneous Growth (HG) model is comprised of two parts. In the case of clonal DIP rates induced by a single perturbation, the model uses the assumption of a normal distribution. In the case where multiple perturbations are used simultaneously, the model assumes a skew-normal distribution of clonal DIP rates.

The derivation of the HG model is as follows: The basic exponential growth equation is

$$\mathrm{y} = \mathrm{y}_0 \mathrm{e}^{\alpha \mathrm{t}}$$

where y_0 is the initial size of the population, α is the growth rate of the population and t is time. In the single-agent HG model, we assume that rather than a single growth rate (α), the population consists of a range of growth rates described by the DIP rate Gaussian.

$$f(x|\mu,\sigma) = \frac{1}{2\sigma\sqrt{2\,\pi}} e^{\frac{-(x-\mu)^2}{2\,\sigma^2}}$$

We replaced $\alpha$with the Gaussian distribution function and integrated (Riemann-Stieltjes (Hildebrandt, 1938)) over the basic exponential growth equation to obtain the following equation:

$$y = y_0 e^{\mu t + \sigma^2 t^2/2}$$ (Figure Legends Fig. 24a)

where μ and σ are the mean and standard deviation of the normal distribution. This is equivalent to the moment-generating function of normality. When the ratio *y/y*0 is 1, the population has rebounded to its initial level, and growth will continue above the original population size.

$$\frac{y}{y_0} = e^{\mu t + \sigma^2 t^2/2} = 1 \quad (4)$$

$$\log\left(\frac{y}{y_0}\right) = \mu t + \frac{\sigma^2 t^2}{2} = \log 1 = 0 \quad (5)$$

$$\mu t + \frac{\sigma^2 t^2}{2} = 0 \quad (6)$$

Application of the quadratic formula leads to the trivial root of 0, and

$$t = \frac{-2\mu}{\sigma^2} \quad (7)$$

When the ratio y/y_0 is 1, the population has returned to its initial level, and growth will continue above the original population size. Application of the quadratic formula leads to the trivial root of 0, and $t = \frac{-2\mu}{\sigma^2}$. We define t as the time to rebound.

In the case of multi-agent therapy, the HG model substitutes a skew-normal distribution(Azzalini, 2005), rather than a normal distribution, into the exponential growth equation. While the TTR and depth of response are not evaluated analytically, they can

be computed numerically using R. To compute TTR, the *uniroot* function is used. To find the minimum value, or depth of response, the *optimize* function is applied.

Statistical analysis. All statistical analysis was performed using the R statistical software (R-project.org). Statistical estimations for the skew normal distribution parameters were performed using the *selm* function contained within the *SN* package (http://azzalini.stat.unipd.it/SN, version 1.0). Other statistical model fits were performed as described previously (Tyson et al., 2012).

Fitting rebound data. The green line in Figure 26a, is the predicted curve using a modified HG model curve. The HG model parameters were taken directly from cFP data estimates of the mean growth rate (μ = -0.147 doublings/day) and standard deviation (σ = 0.142 doublings/day). The modification was integrating the HG model using a truncated normal using a cutoff in growth rate of 0.165 doublings/day, which was the maximum observed growth rate of the PC9 cells. The mathematical integration was done using R (http://r-project.org)

Experimental validation of relapse. PC9 parental cells were plated at 4,000 cells per well in full growth media and allowed to adhere overnight. Then 3μM erlotinib was added (with media changes every 3 days) and whole wells of a 96-well plate were imaged daily using the same imaging conditions as in the cFP. The freely available ImageJ (http://imagej.nih.gov/ij/) software was used to quantify cell number for all measurements.

Discussion

Here we present the Heterogeneous Growth (HG) model, a mathematical framework to translate measurements of clonal DIP rates to predict the dynamics of

population rebound following prolonged continuous therapy. This approach enabled the novel conclusion that clonal DIP rate variance is sufficient to drive rebound to targeted therapy. Furthermore, describing distributions in terms of all the relevant parameters enables predictive classification of different cell populations. Applying multiplexed genetic (Su et al., 2011a) and proteomic (Ciaccio et al., 2010) approaches enables discovery of molecular factors determining the clonal DIP rate distribution. Thus combining cFP with the HG model is a powerful approach to measure and manipulate the clonal DIP rate distribution to find therapeutic strategies that provide optimal TTR.

Previous studies of clonal dynamics in drug-treated populations infer the functional behavior of clones retrospectively. Models of acquired resistance mutations can explain patient tumor response time-course data (Bozic et al., 2013; Leder et al., 2014; Michor et al., 2005), and can be used to predict therapy improvement by dual drugging (Bozic et al., 2013; Kanagasabai et al., 2010). While demonstrating the potential of modeling to improve targeted therapy, these models do not quantify underlying cell behavior that results in rebound but rather infer it from the relative proportions of mutant cell types *a posteriori*. These models are based on clinical data outcomes and therefore testing model simulations experimentally is challenging. Likewise, elegant descriptions of functional impact of clonal heterogeneity *in vivo* (Almendro et al., 2014b; Kreso et al., 2013; Tran et al., 2011) are provocative, yet difficult to fully interpret without direct measures of clonal drug response. Here, we join the strength of these approaches using a systems biology approach of combined modeling and experimentation. The HG model predicts population dynamics based on experimental measures that integrate cell fate heterogeneity and clonal structure (Fig. 24). Due to the simplicity of the cFP assay, such measurements are easily accessible

and may eventually be taken on tumor specimens and allow for predictions of TTR *before* treatment.

An unsuspected outcome of the HG model is that the clonal DIP rate variance itself is an attractive drug target, since collapsing it disproportionally affects TTR lengthening (Fig. 25). Simply put, positive DIP rate clones drive rebound and clonal DIP rate variance is sufficient to drive rebound. Finding the molecular factors responsible for DIP rate variance may therefore be advisable. Indeed, by studying panels of PC9 clonal sublines spanning the full distribution and using their mean DIP rate as a continuous variable, we found molecules that correlate with DIP rate. We focused in on p38, which appears to modify the variance. A p38 inhibitor, SB203580, broadens the DIP rate variance while decreasing the mean. Thus, while appearing to be an effective combination on average, our results suggest that increased DIP rate variance would result in a faster rebound. In contrast, a p38 activator, anisomycin, reduces the fraction of positive DIP rate clones by compressing the variance and adding a negative skew to the data. This in turn lengthens the TTR. Stress response signals have been shown both to be pro-survival in response to irradiation (Kanagasabai et al., 2010), or pro-death in response to oncogene deprivation (Tran et al., 2011). In our case, they appear to modify the variance of DIP rate distribution in the presence of a targeted drug. This observation will have to be substantiated by in-depth mechanistic studies, given the large number of p38 substrates. Nevertheless, this is a proof-of-principle that provides a feasible approach to improving erlotinib treatment outcomes. Thus, while the source of clonal DIP rate variation in cancer cell populations remains to be determined at a mechanistic level, using DIP rate variation as a continuous variable, as we do here, should facilitate identification of its molecular underpinnings.

Our findings effectively cast tumor heterogeneity in a new light: While remaining a formidable foe, it can be harnessed by quantitative tools and used to predict depth and duration (dynamics) of drug effects with a model based on simple in vitro measurements. For now, our report is by necessity limited to *in vitro* cell lines, but its conclusions transform intratumor heterogeneity into a tractable problem. A key challenge is to apply the clonal DIP rate metric and the cFP assay to preclinical or clinical tumor material, to enable predictions of clinical time course and provide a rationale for *a priori* choice of treatment. The utility of combining modeling approaches with clinical data has been shown (Bozic et al., 2013; Leder et al., 2014; Michor et al., 2005). Finally, our data unequivocally show that ablating the positive DIP rate clones, by reducing variance or negatively skewing the distribution, is the best strategy to lengthen TTR. Novel drug screening strategies based on clonal DIP rate variance containment can be readily envisioned.

Chapter V

CONCLUSION

Discussion

Overall, our work presents a framework to quantify cancer heterogeneity and relate it to the overall change in size of a cell population in a predictive fashion (Fig. 31). While cancer heterogeneity is widely recognized, it remains a clinically-relevant problem, divided into two categories: intertumor heterogeneity, or differences between patients; and intratumor heterogeneity, which describes variability within a tumor that is responsible for either primary resistance to therapy (Pao and Chmielecki, 2010) or the eventual rebound. Intertumor heterogeneity aims to define population subsets that will respond to certain therapies. Whereas efforts in intratumor heterogeneity are divided across two fronts: the first is to find new molecular susceptibilities to amplify the initial tumor responses; the second is to try to find ways to resensitize resistant cell populations after the onset of resistance. While molecular mechanisms underlying cancer rebound are increasingly discovered (Sequist et al., 2011), data regarding the dynamics of treatment failure are only available retrospectively. This leaves a glaring hole in the research literature: can the dynamics of tumor rebound to treatment be predicted and improved prospectively? The overall goal of this work was use a systems biology approach (Kohl et al., 2010) in order to link quantitative measurements of drug response heterogeneity to population dynamics. To address this question, we developed quantitative imaging tools to measure dynamic cellular response to perturbations (Chapter 2), examined how drug responses cross biological scales (Chapter 3), and then subsequently developed a mathematical model to enable data-driven predictions of cancer rebound (Chapter 4).

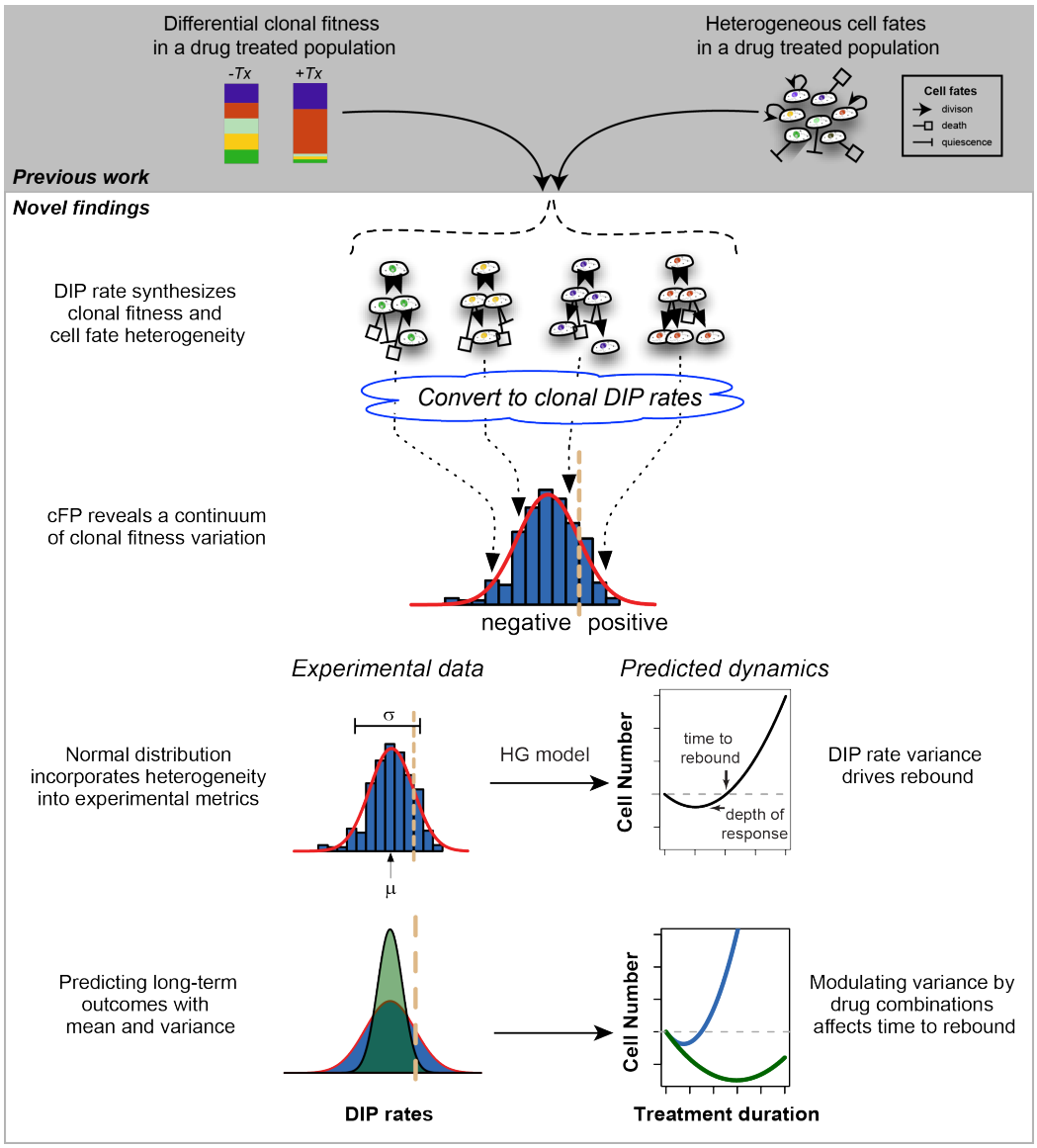

Figure 31 Conceptual Schematic. *Top*, Previous understanding of cell-to-cell heterogeneity. A cell population responds to perturbations with multiple cell fates that, when combined, form a population-level proliferation rate. New insights from this manuscript are shown on the left. Cell-to-cell heterogeneity in response to targeted therapy is encapsulated in DIP rates within single-cell lineages. Positively proliferating lineages within a cell lineage are a result of DIP rates being distributed normally. The HG model predicts relapse by integrating the contribution of lineage-specific heterogeneity across a population. The HG model predicts that reducing the DIP rate variance will prolong rebound during continuous treatment.

To understand the dynamics of an intrinsically heterogeneous response to perturbations, we devised the clonal Fractional Proliferation (cFP) assay. Previously, relating cell heterogeneity to population dynamics has been problematic because cell-to-cell heterogeneity is so widespread (Gascoigne and Taylor, 2008). Our previous work, FPM, (Tyson et al., 2012) described a framework to relate single-cell measurements to changes in population size over time. However, FPM is limited in throughput and best applied retrospectively to accommodate comprehensive single-cell datasets. The DIP rate is the key simplifying device that enables cFP to relate single-cell and clonal fitness to population response. In FPM hundreds of single-cells were tracked to estimate the rates of transition into different cell fates. Now, in cFP, those single-cell fates are encapsulated into steady rates of proliferation. Thus, sparse temporal measurements of population size are sufficient to integrate the functional contribution of multiple cell fates into a single, dynamic metric of fitness (DIP rate) for clonal response.

The vast reduction in information content required enables the analysis to be proportionally scaled up in terms of sample throughput. Therefore, cFP reports the clonal structure of a population over time in multiple conditions (Fig. 9a). Previously, single-cell measurements were either unrelated to population size changes or limited to short-term effects. Alternatively, other assays designed to quantify heterogeneity, such as the colony formation assay (Franken et al., 2006), report as an all-or-none value, the number of clones to adhere, survive perturbation, and grow to a certain size. By contrast, cFP reveals the clonal fitness structure that underlies long-term population dynamics by integrating single-cell fates into a continuous distribution of clonal fitness.

Then, by cFP, we measured the clonal fitness profile of EGFR-addicted cells to erlotinib, an EGFR small molecule inhibitor. EGFR-addicted cell lines are a good model

system because they recapitulate many relevant features of heterogeneity observed clinically. In response to erlotinib, all measured cell lines display a clonal DIP rate profile described by a normal distribution with a negative mean (Fig. 26). In each case, a small number of clones, without drug selection, display a positive DIP rate. Thus, the biological variation, intrinsic to the cell population, is sufficient to ensure the existence of minor clonal populations that grow in the presence of drug, despite a primarily apoptotic response within the cell line.

This finding is intriguing, because clonal variation, in the context of tumor evolution, usually refers to clonal outgrowth of single cells with a completely unique drug response (Nowell, 1976). For example, random mutations that occlude drug binding, thus effectively reversing drug sensitivity in a binary manner. However, in each of the positive DIP rate clones that we isolate, they remain sensitive to the drug, evidenced by nanomolar IC50 values (Fig. 20), and comprise a continuous, unimodal distribution of phenotype (Fig. 24a). Thus the clones we describe are likely not due to the accumulation of random genetic mutations. Furthermore, single-cell isolated clones also yield normal distributions of DIP rates with variance similar to that of the parental population (Fig. 22). The diversification of clonal DIP rates far exceeds the error rate in DNA repair. More likely, we speculate that the clonal variance of a cell population is the extrapolation of cellular heterogeneity (Altschuler and Wu, 2010) to its full range of phenotypic effects. Supporting this view, clones with a stable DIP rate display a wide range of single-cell fates in response to perturbation (Fig. 21b–c). While DIP rate variance appears intrinsic, the mean of clonal populations remains stable over many cell generations (Fig. 19c). This is in contrast to previous reports (Sharma et al., 2010b), where response drug sensitivity cannot be cloned out and is mediated by an all-or-nothing reversible transition from a unique, stem-like state.

In an attempt to translate these dynamic measurements of clonal behavior into predictions of cell population response, we devised the heterogeneous growth (HG) model. The HG model is a mathematical model that combines the DIP rate distribution into the basic exponential growth equation and integrates across the full range of rates. Thus the HG model takes experimentally-derived measurements of clonal fitness and predicts the resultant population dynamics (Fig. 24b). In each EGFR-addicted cell line examined (Fig. 26), the population undergoes an expected initial decrease in population size, and owing to the reduction of the bulk population and the exponential expansion of positive DIP rate clones, steadily increases the population size. This leads to the unexpected finding that DIP rate variance alone is sufficient to drive rebound (Fig. 24b). In addition, the HG model gives predictions of the maximal reduction in population size (depth of response) and the time until the population regrows to its initial size (TTR). Previously, the only other model used to predict rebound based on experimental measurements operates under the assumption that rebound is driven by a rare subset of cells, selected by drug treatment, that proliferate similar to the untreated condition (Chmielecki et al., 2011; Michor et al., 2005). Our experimental validation of rebound is consistent with the HG model (Fig. 27b), but does not match the model assuming rare mutations (Fig. 27c).

By expanding single cells into discrete sublines (DS), we obtained an experimental system suitable for examining the molecular underpinning of DIP rate (Fig. 19a). We conducted, in a panel of ten erlotinib-treated sublines, a molecular screen of analytes involved in several signaling pathways (Fig. 29a–b). As a confirmation of the validity of the screen, we discovered anti-apoptotic proteins, previously implicated in protecting cells from apoptosis (Fan et al., 2011). In addition, our screen uncovered phospho-Hsp27, involved in the stress response pathway (Kanagasabai et al., 2010;

Rogalla et al., 1999). We sought to shift the mean clonal DIP rate by modulating Hsp27 indirectly by an upstream activator, p38, since direct chemical modulators of Hsp27 are not available. In contrast to shifting the mean, p38 inhibition broadened the DIP rate variance (Fig. 30a).

This result is intriguing because DIP rate variance has a disproportionately strong effect on TTR relative to the DIP rate mean (Fig. 25). Indeed, p38 activation had the reverse effect: combining erlotinib with anisomycin greatly reduced the number of positive DIP rate clones and drastically increased the TTR (Fig. 30). Large-scale unbiased screening efforts of gene knock-out (Bivona et al., 2011) or knock-in (Johannessen et al., 2013) have uncovered molecular pathways that modify mean drug effects. As an alternative, our data suggests that discovery efforts characterizing the full clonal DIP rate profile may uncover chemicals that and reduce the number of positive DIP rate clones and be more fruitful at identifying compounds that achieve optimal TTR.

Single-cell responses to perturbation are widespread and interpreting their effects is problematic. As a result, the primary means to estimate their effects is restricted to biology that groups cell fates into distinct biological groups. For example, most models of treatment of drug-sensitive cells assume that all cells that do not die have acquired an additional genetic (Chong and Jänne, 2013) or epigenetic (Sharma et al., 2010b) event that makes them resistant to therapy. As both our data and previous publications demonstrate, biological variation is enormous and this framework is likely not universally applicable. In contrast, the DIP rate metric accommodates cellular heterogeneity by placing it into the context of clonal DIP rates, a dynamic metric of fitness that can in turn be translated into predictions of rebound by the HG model. The molecular mechanistic origins of single-cell fate variability are an active area on interest.

However, the sheer amount of information that can be collected from a heterogeneous population at the single cell level, e.g., proteomic (Bendall et al., 2011), genomic (Macaulay and Voet, 2014), etc., is staggering. We show an alternative way to comprehensively account for widespread cell heterogeneity. Instead of going down in scale to measure molecular differences at the single-cell level (Gascoigne and Taylor, 2008), we exploit dynamic measurements of clonal behavior as a proxy for prohibitively large single-cell datasets. Thus, while accepting single-cell behavior as phenomenological, we utilize clonal DIP rates to bridge the scales from complex single-cell behaviors to the clinically-relevant predictions of population rebound.

The overall goal of this work was to use systems biology to understand the contribution of intratumor heterogeneity. The cFP assay and the HG model together translate single-cell heterogeneity into clinically-relevant predictions of rebound. This finding was enabled by joining modeling and experimentation as a model system. cFP effectively bridged biological scales, quantitatively linking single-cell, clonal, and population level responses. Applying cFP to EGFR-addicted cells treated with erlotinib revealed that the clonal profile can be represented as a normal distribution. The HG model subsequently interpreted the effects of positive DIP rate clones. Furthermore, the model highlighted the parameters that are most beneficial to modulate. These insights in turn guided experimental screens that uncovered drug combinations that optimally lengthened the TTR. Thus, incorporating measurements of clonal heterogeneity into further drug screens could identify therapeutics leading to better outcomes. In summary, these results demonstrate that while intratumor heterogeneity is formidable, addressing it quantitatively can uncover novel strategies to constrain its influence in undermining therapy.

Future directions

The cFP assay was designed using an *in vitro* setting to enable single-cell measurements. Several factors may enable the conclusions to more closely mimic *in vivo* growth conditions. For example, growing cells on a layer of extracellular matrix proteins could provide more realistic microenvironmental conditions. Alternatively, cFP could be performed in a 3D culture system that still retains fluorescence imaging capacity. Other studies have quantified DNA barcoding to measure the relative clonal abundance during tumorigenesis and drug treatment in patient-derived xenografts (Kreso et al., 2013). It would be interesting to expand cFP as a model system by using DNA barcoding to directly relate direct *in vitro* measurements of clonal population size to indirect measurements of relative clonal abundance *in vivo*.

cFP relies on fluorescence imaging of cell nuclei to quantify cell nuclei. In cFP we showed that a secondary fluorescence marker, geminin-mAg (Sakaue-Sawano et al., 2008a), could be added as a binary readout of cell cycle position (Fig. 9c). The extra fluorescence channel could be chosen to be more specific to the biological system of interest. For example, a differentiation marker could be added in the case of stem cells or a readout of signaling activity associated with the phenotype of interest could be added (Spencer et al., 2009).

Examining the clonal profile of EGFR-addicted cells in response to erlotinib revealed that DIP rates are normally distributed (Fig. 26). Culturing these clones as discrete sublines revealed that the mean clonal DIP rate is conserved across experiments (Fig. 19c), thus they display phenotypic stability over the time scales we examined. However, the rapid repopulation of the DIP rate variance within clonal populations (Fig. 22) suggests that DIP rate is not determined solely by genetic

mutations. Still, we acknowledge that the underlying factors governing DIP rate mean remain unknown. Furthermore, while p38 signaling appears to modify the width of the DIP rate distribution (Fig. 30a), the molecular mechanisms involved in this process are largely unknown. Based on the timescale of change, DIP rate mean and variance may be due to epigenetic factors (Rando and Verstrepen, 2007). Others have speculated that the range of phenotypes within a cell population is governed by greater mRNA entropy that could be restricted by the chromatin state of cells (MacArthur and Lemischka, 2013). It would be interesting to see if p38 activity modifies the overall profile of epigenetic events within a cell population (Taiwo et al., 2012), and if so, which modifications are most critical. Alternatively, independent of epigenetic analysis, the role of p38 in driving the DIP rate variance could be modified through RNA interference, overexpression, or conditional expression. Other inhibitors or activators of p38 (Kuma et al., 2005), with different mechanisms of actions, may further uncover the role of p38 signaling in modifying clonal DIP rate variance.

Positive growth in the presence of drug treatment is most commonly associated with the acquisition of genetic resistance mutations. This is well established to the high frequency of conserved mutations occurring upon failure to targeted therapy (Sequist et al., 2011). But how do positive DIP rate clones relate to genetic resistance mutations? It could be that positive DIP rate clones act as a reservoir of cells that eventually become resistant during prolonged therapy. Culturing sublines during continuous erlotinib treatment could test this directly. Regarding genetic resistance mutations, recent work has focused on targeting mutant proteins for the express purpose of overcoming acquired resistance (Zhou et al., 2009). However, it is increasingly clear that cancer cells develop resistance to inhibitors that were designed specifically to overcome resistance. Some have lamented this therapeutic strategy as a “whack-a-mole” approach to treating

cancer (Gillies et al., 2012). Further studies are warranted to examine if clonal DIP rate variance also underlies the failure of these next-generation therapeutics.

The HG model suggests that compounds that can reduce the σ parameter of the DIP rate distribution would be highly effective at prolonging rebound (Fig. 25). High throughput screens currently focus on finding chemical or or genetic combinations that modulate the initial number of cells killed by therapy (Bivona et al., 2011; Johannessen et al., 2013), likely affecting the μ parameter. Creating parallel high-throughput screens to modify the variance may provide an attractive alternative to discover therapeutic strategies that delay rebound. Several strategies may enable such an approach. First, cFP relies on minimal user input to ensure accurate colony segmentation. Currently cFP relies on sparse plating to separate clones. Separating clones by culturing them in 384- or 1536 well plates may remove concerns about colony separation. If user input is removed entirely, this could allow the sample throughput and data analysis to be scaled up to generate DIP rate distributions for many compounds.

Utilizing the panel of DS sublines could also help to identify rebound. The microwestern array analysis was designed to screen for molecular differences between compounds with differing μ agnostic of σ (Fig. 29). However, sublines may have different σ as well. For example, DS6 has a narrower clonal DIP rate variance than DS9 (Fig. 22). Designing the screen using sublines with similar μ but variable σ may help to identify the molecular components underlying DIP rate variance.

Alternatively, a direct protein readout may enable the discovery of combinations to decrease variance. If a single protein could be predictive of DIP rate, then the clonal DIP rate distribution could be measured as a proxy by flow cytometry. Then high

throughput screens, using a single sample, could measure the effect of a treatment on the protein concentration variance (Dar et al., 2014). In addition to the subline screening technique described above, multiplex mass cytometry assays (Bendall et al., 2011) could identify analytes with changing variance matching that of the DIP rate distributions from drug combination experiments (Fig. 30).

REFERENCES

Almendro, V., Cheng, Y.-K., Randles, A., Itzkovitz, S., Marusyk, A., Ametller, E., Gonzalez-Farre, X., Muñoz, M., Russnes, H.G., Helland, Å., et al. (2014a). Inference of Tumor Evolution during Chemotherapy by Computational Modeling and In Situ Analysis of Genetic and Phenotypic Cellular Diversity. Cell Reports.

Almendro, V., Kim, H., Cheng, Y.-K., Gonen, M., Itzkovitz, S., Argani, P., van Oudenaarden, A., Sukumar, S., Michor, F., and Polyak, K. (2014b). Genetic and phenotypic diversity in breast tumor metastases. Cancer Res.

Almendro, V., Marusyk, A., and Polyak, K. (2013). Cellular heterogeneity and molecular evolution in cancer. Annu Rev Pathol *8*, 277–302.

Altschuler, S.J., and Wu, L.F. (2010). Cellular Heterogeneity: Do Differences Make a Difference? Cell *141*, 559–563.

Anaka, M., Hudson, C., Lo, P.-H., Do, H., Caballero, O.L., Davis, I.D., Dobrovic, A., Cebon, J., and Behren, A. (2013). Intratumoral genetic heterogeneity in metastatic melanoma is accompanied by variation in malignant behaviors. BMC Med Genomics *6*, 40.

Anderson, A.R.A., and Quaranta, V. (2008). Integrative mathematical oncology. Nat. Rev. Cancer *8*, 227–234.

Anderson, A.R., Weaver, A.M., Cummings, P.T., and Quaranta, V. (2006). Tumor Morphology and Phenotypic Evolution Driven by Selective Pressure from the Microenvironment. Cell *127*, 905–915.

Avraham, R., and Yarden, Y. (2011). Feedback regulation of EGFR signalling: decision making by early and delayed loops. Nat. Rev. Mol. Cell Biol. *12*, 104–117.

Azzalini, A. (2005). The Skew-normal Distribution and Related Multivariate Families*. Scand J Stat *32*, 159–188.

Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial persistence as a phenotypic switch. Science *305*, 1622–1625.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–607.

Bedard, P.L., Hansen, A.R., Ratain, M.J., and Siu, L.L. (2013). Tumour heterogeneity in the clinic. Nature *501*, 355–364.

Bendall, S.C., Simonds, E.F., Qiu, P., Amir, E.-A.D., Krutzik, P.O., Finck, R., Bruggner, R.V., Melamed, R., Trejo, A., Ornatsky, O.I., et al. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science *332*, 687–696.

Bivona, T.G., Hieronymus, H., Parker, J., Chang, K., Taron, M., Rosell, R., Moonsamy, P., Dahlman, K., Miller, V.A., Costa, C., et al. (2011). FAS and NF-κB signalling modulate dependence of lung cancers on mutant EGFR. Nature *471*, 523–526.

Bozic, I., Reiter, J.G., Allen, B., Antal, T., Chatterjee, K., Shah, P., Moon, Y.S., Yaqubie, A., Kelly, N., Le T, D., et al. (2013). Evolutionary dynamics of cancer in response to targeted combination therapy. eLife *2*, e00747.

Brock, A., Chang, H., and Huang, S. (2009). Non-genetic heterogeneity--a mutation-independent driving force for the somatic evolution of tumours. Nature Reviews Genetics *10*, 336–342.

Burgess, A.W., Cho, H.-S., Eigenbrot, C., Ferguson, K.M., Garrett, T.P.J., Leahy, D.J., Lemmon, M.A., Sliwkowski, M.X., Ward, C.W., and Yokoyama, S. (2003). An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Mol. Cell *12*, 541–552.

Chakrabarty, A., Sanchez, V., Kuba, M.G., Rinehart, C., and Arteaga, C.L. (2012). Feedback upregulation of HER3 (ErbB3) expression and activity attenuates antitumor effect of PI3K inhibitors. Proceedings of the National Academy of Sciences *109*, 2718–2723.

Chandarlapaty, S. (2012). Negative Feedback and Adaptive Resistance to the Targeted Therapy of Cancer. Cancer Discovery *2*, 311–319.

Chang, H.H., Hemberg, M., Barahona, M., Ingber, D.E., and Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature *453*, 544–547.

Chapman, P.B., Hauschild, A., Robert, C., Haanen, J.B., Ascierto, P., Larkin, J., Dummer, R., Garbe, C., Testori, A., Maio, M., et al. (2011). Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N. Engl. J. Med. *364*, 2507–2516.

Chin, L., Tam, A., Pomerantz, J., Wong, M., Holash, J., Bardeesy, N., Shen, Q., O'Hagan, R., Pantginis, J., Zhou, H., et al. (1999). Essential role for oncogenic Ras in tumour maintenance. Nature *400*, 468–472.

Chin, T.M., Quinlan, M.P., Singh, A., Sequist, L.V., Lynch, T.J., Haber, D.A., Sharma, S.V., and Settleman, J. (2008). Reduced Erlotinib sensitivity of epidermal growth factor receptor-mutant non-small cell lung cancer following cisplatin exposure: a cell culture model of second-line erlotinib treatment. Clin. Cancer Res. *14*, 6867–6876.

Chmielecki, J., Foo, J., Oxnard, G.R., Hutchinson, K., Ohashi, K., Somwar, R., Wang, L., Amato, K.R., Arcila, M., Sos, M.L., et al. (2011). Optimization of dosing for EGFR-mutant non-small cell lung cancer with evolutionary cancer modeling. Science Translational Medicine *3*, 90ra59–90ra59.

Chong, C.R., and Jänne, P.A. (2013). The quest to overcome resistance to EGFR-targeted therapies in cancer. Nat. Med. *19*, 1389–1400.

Ciaccio, M.F., Wagner, J.P., Chuu, C.-P., Lauffenburger, D.A., and Jones, R.B. (2010). Systems analysis of EGF receptor signaling dynamics with microwestern arrays. Nature Methods *7*, 148–155.

Citri, A., and Yarden, Y. (2006). EGF–ERBB signalling: towards the systems level. Nature Reviews Molecular Cell Biology *7*, 505–516.

da Cunha Santos, G., Shepherd, F.A., and Tsao, M.S. (2011). EGFR mutations and lung cancer. Annu Rev Pathol *6*, 49–69.

Dar, R.D., Hosmane, N.N., Arkin, M.R., Siliciano, R.F., and Weinberger, L.S. (2014). Screening for noise in gene expression identifies drug synergies. Science.

Dexter, D.L., Kowalski, H.M., Blazar, B.A., Fligiel, Z., Vogel, R., and Heppner, G.H. (1978). Heterogeneity of tumor cells from a single mouse mammary tumor. Cancer Res. *38*, 3174–3181.

Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., McLellan, M.D., et al. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature *481*, 506–510.

Doebele, R.C., Pilling, A.B., Aisner, D.L., Kutateladze, T.G., Le, A.T., Weickhardt, A.J., Kondo, K.L., Linderman, D.J., Heasley, L.E., Franklin, W.A., et al. (2012). Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small cell lung cancer. Clin. Cancer Res. *18*, 1472–1482.

Druker, B.J., Talpaz, M., Resta, D.J., Peng, B., Buchdunger, E., Ford, J.M., Lydon, N.B., Kantarjian, H., Capdeville, R., Ohno-Jones, S., et al. (2001). Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N. Engl. J. Med. *344*, 1031–1037.

Eisenhauer, E.A., Therasse, P., Bogaerts, J., Schwartz, L.H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., et al. (2009). New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). European Journal of Cancer *45*, 228–247.

Ercan, D., Xu, C., Yanagita, M., Monast, C.S., Pratilas, C.A., Montero, J., Butaney, M., Shimamura, T., Sholl, L., Ivanova, E.V., et al. (2012). Reactivation of ERK Signaling Causes Resistance to EGFR Kinase Inhibitors. Cancer Discovery *2*, 934–947.

Faber, A.C., Corcoran, R.B., Ebi, H., Sequist, L.V., Waltman, B.A., Chung, E., Incio, J., Digumarthy, S.R., Pollack, S.F., Song, Y., et al. (2011). BIM expression in treatment-naive cancers predicts responsiveness to kinase inhibitors. Cancer Discov *1*, 352–365.

Fan, W., Tang, Z., Yin, L., Morrison, B., Hafez-Khayyata, S., Fu, P., Huang, H., Bagai, R., Jiang, S., Kresak, A., et al. (2011). MET-Independent Lung Cancer Cells Evading EGFR Kinase Inhibitors Are Therapeutically Susceptible to BH3 Mimetic Agents. Cancer Research *71*, 4494–4505.

Felsher, D.W., and Bishop, J.M. (1999). Reversible tumorigenesis by MYC in hematopoietic lineages. Mol. Cell *4*, 199–207.

Felsher, D.W., Weinstein, I.B., and Joe, A. (2008). Oncogene Addiction versus Oncogene Amnesia: Perhaps More than Just a Bad Habit? Cancer Research *68*, 3081–3086.

Fidler, I.J., and Kripke, M.L. (1977). Metastasis results from preexisting variant cells within a malignant tumor. Science *197*, 893–895.

Franken, N.A.P., Rodermond, H.M., Stap, J., Haveman, J., and van Bree, C. (2006). Clonogenic assay of cells in vitro. Nat Protoc *1*, 2315–2319.

Fraser, D., and Kaern, M. (2009). A chance at survival: gene expression noise and phenotypic diversification strategies. Molecular Microbiology *71*, 1333–1340.

Gascoigne, K.E., and Taylor, S.S. (2008). Cancer cells display profound intra- and interline variation following prolonged exposure to antimitotic drugs. Cancer Cell *14*, 111–122.

Georgescu, W., Wikswo, J.P., and Quaranta, V. (2012). CellAnimation: an open source MATLAB framework for microscopy assays. Bioinformatics *28*, 138–139.

Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England Journal of Medicine *366*, 883–892.

Gillies, R.J., Verduzco, D., and Gatenby, R.A. (2012). Evolutionary dynamics of carcinogenesis and why targeted therapy does not work. Nat. Rev. Cancer *12*, 487–493.

Greaves, M., and Maley, C.C. (2012). Clonal evolution in cancer. Nature *481*, 306–313.

Gridelli, C., Bareschino, M.A., Schettino, C., Rossi, A., Maione, P., and Ciardiello, F. (2007). Erlotinib in non-small cell lung cancer treatment: current status and future development. Oncologist *12*, 840–849.

Haber, D., Gray, N., and Baselga, J. (2011). The Evolving War on Cancer. Cell *145*, 19–24.

Hanahan, D., and Coussens, L.M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. Cancer Cell *21*, 309–322.

Hidalgo, M., Siu, L.L., Nemunaitis, J., Rizzo, J., Hammond, L.A., Takimoto, C., Eckhardt, S.G., Tolcher, A., Britten, C.D., Denis, L., et al. (2001). Phase I and pharmacologic study of OSI-774, an epidermal growth factor receptor tyrosine kinase inhibitor, in patients with advanced solid malignancies. J. Clin. Oncol. *19*, 3267–3279.

Hildebrandt, T. (1938). Definitions of Stieltjes integrals of the Riemann type. The

American Mathematical Monthly *45*, 265–278.

Huang, S. (2009). Non-genetic heterogeneity of cells in development: more than just noise. Development *136*, 3853–3862.

Jackman, D., Pao, W., Riely, G.J., Engelman, J.A., Kris, M.G., Janne, P.A., Lynch, T., Johnson, B.E., and Miller, A.V. (2010). Clinical Definition of Acquired Resistance to Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Non-Small-Cell Lung Cancer. Journal of Clinical Oncology *28*, 357–360.

Janku, F., Stewart, D.J., and Kurzrock, R. (2010). Targeted therapy in non-small-cell lung cancer--is it becoming a reality? Nat Rev Clin Oncol *7*, 401–414.

Jänne, P.A., Gray, N., and Settleman, J. (2009). Factors underlying sensitivity of cancers to small-molecule kinase inhibitors. Nat Rev Drug Discov *8*, 709–723.

Johannessen, C.M., Johnson, L.A., Piccioni, F., Townes, A., Frederick, D.T., Donahue, M.K., Narayan, R., Flaherty, K.T., Wargo, J.A., Root, D.E., et al. (2013). A melanocyte lineage program confers resistance to MAP kinase pathway inhibition. Nature *504*, 138–142.

Johnston, R.J., Jr, and Desplan, C. (2010). Stochastic mechanisms of cell fate specification that yield random or robust outcomes. Annual Review of Cell and Developmental Biology *26*, 689.

Kanagasabai, R., Karthikeyan, K., Vedam, K., Qien, W., Zhu, Q., and Ilangovan, G. (2010). Hsp27 Protects Adenocarcinoma Cells from UV-Induced Apoptosis by Akt and p21-Dependent Pathways of Survival. Molecular Cancer Research *8*, 1399–1412.

Kohl, P., Crampin, E.J., Quinn, T.A., and Noble, D. (2010). Systems biology: an approach. Clin. Pharmacol. Ther. *88*, 25–33.

Kreso, A., Kreso, A., O'Brien, C.A., van Galen, P., Gan, O.I., Notta, F., Brown, A.M., Ng, K., Ma, J., Wienholds, E., et al. (2013). Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. Science *339*, 543–548.

Kuma, Y., Sabio, G., Bain, J., Shpiro, N., Márquez, R., and Cuenda, A. (2005). BIRB796 inhibits all p38 MAPK isoforms in vitro and in vivo. J. Biol. Chem. *280*, 19472–19479.

Leder, K., Pitter, K., Laplant, Q., Hambardzumyan, D., Ross, B.D., Chan, T.A., Holland, E.C., and Michor, F. (2014). Mathematical Modeling of PDGF-Driven Glioblastoma Reveals Optimized Radiation Dosing Schedules. Cell *156*, 603–616.

Lee, M.J., Ye, A.S., Gardino, A.K., Heijink, A.M., Sorger, P.K., MacBeath, G., and Yaffe, M.B. (2012). Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks. Cell *149*, 780–794.

Linardou, H., Dahabreh, I.J., Bafaloukos, D., Kosmidis, P., and Murray, S. (2009).

Somatic EGFR mutations and efficacy of tyrosine kinase inhibitors in NSCLC. Nature Reviews Clinical Oncology *6*, 352–366.

Loeb, L.A. (2011). Human cancers express mutator phenotypes: origin, consequences and targeting. Nature Reviews Cancer *11*, 450–457.

Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., et al. (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N. Engl. J. Med. *350*, 2129–2139.

MacArthur, B.D., and Lemischka, I.R. (2013). Statistical mechanics of pluripotency. Cell *154*, 484–489.

Macaulay, I.C., and Voet, T. (2014). Single cell genomics: advances and future perspectives. PLoS Genet. *10*, e1004126.

Marusyk, A., and Polyak, K. (2013). Cancer cell phenotypes, in fifty shades of grey. Science *339*, 528–529.

Marusyk, A., Almendro, V., and Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? Nat. Rev. Cancer *12*, 323–334.

McGranahan, N., Burrell, R.A., Endesfelder, D., Novelli, M.R., and Swanton, C. (2012). Cancer chromosomal instability: therapeutic and diagnostic challenges. EMBO Reports *13*, 528–538.

Mendoza, M.C., Er, E.E., and Blenis, J. (2011). The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. Trends Biochem. Sci. *36*, 320–328.

Merlos-Suárez, A., Barriga, F.M., Jung, P., Iglesias, M., Céspedes, M.V., Rossell, D., Sevillano, M., Hernando-Momblona, X., da Silva-Diz, V., Muñoz, P., et al. (2011). The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. Cell Stem Cell *8*, 511–524.

Michor, F., Hughes, T.P., Iwasa, Y., Branford, S., Shah, N.P., Sawyers, C.L., and Nowak, M.A. (2005). Dynamics of chronic myeloid leukaemia. Nature *435*, 1267–1270.

Michor, F., Iwasa, Y., and Nowak, M.A. (2004). Dynamics of cancer progression. Nat. Rev. Cancer *4*, 197–205.

Mumenthaler, S.M., Foo, J., Leder, K., Choi, N.C., Agus, D.B., Pao, W., Mallick, P., and Michor, F. (2011). Evolutionary Modeling of Combination Treatment Strategies To Overcome Resistance to Tyrosine Kinase Inhibitors in Non-Small Cell Lung Cancer. Molecular Pharmaceutics *8*, 2069–2079.

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011). Tumour evolution inferred by single-cell sequencing. Nature *472*, 90–94.

Nicholson, S.E., Metcalf, D., Sprigg, N.S., Columbus, R., Walker, F., Silva, A., Cary, D., Willson, T.A., Zhang, J.-G., Hilton, D.J., et al. (2005). Suppressor of cytokine signaling (SOCS)-5 is a potential negative regulator of epidermal growth factor signaling. Proc. Natl. Acad. Sci. U.S.a. *102*, 2328–2333.

Niepel, M., Spencer, S.L., and Sorger, P.K. (2009). Non-genetic cell-to-cell variability and the consequences for pharmacology. Curr Opin Chem Biol *13*, 556–561.

Nowell, P.C. (1976). The clonal evolution of tumor cell populations. Science *194*, 23–28.

Ohashi, K., Sequist, L.V., Arcila, M.E., Moran, T., Chmielecki, J., Lin, Y.-L., Pan, Y., Wang, L., de Stanchina, E., Shien, K., et al. (2012). Lung cancers with acquired resistance to EGFR inhibitors occasionally harbor BRAF gene mutations but lack mutations in KRAS, NRAS, or MEK1. Proceedings of the National Academy of Sciences of the United States of America *109*, E2127–E2133.

Okabe, T., Okamoto, I., Tamura, K., Terashima, M., Yoshida, T., Satoh, T., Takada, M., Fukuoka, M., and Nakagawa, K. (2007). Differential Constitutive Activation of the Epidermal Growth Factor Receptor in Non-Small Cell Lung Cancer Cells Bearing EGFR Gene Mutation and Amplification. Cancer Research *67*, 2046–2053.

Paez, J.G. (2004). EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy. Science *304*, 1497–1500.

Pao, W., and Chmielecki, J. (2010). Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. Nature Reviews Cancer *10*, 760–774.

Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L., et al. (2004). EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc. Natl. Acad. Sci. U.S.a. *101*, 13306–13311.

Park, B.O., Ahrends, R., and Teruel, M.N. (2012). Consecutive positive feedback loops create a bistable switch that controls preadipocyte-to-adipocyte conversion. Cell Reports *2*, 976–990.

Pepperkok, R., and Ellenberg, J. (2006). High-throughput fluorescence microscopy for systems biology. Nature Reviews. Molecular Cell Biology *7*, 690–696.

Piccart-Gebhart, M.J., Procter, M., Leyland-Jones, B., Goldhirsch, A., Untch, M., Smith, I., Gianni, L., Baselga, J., Bell, R., Jackisch, C., et al. (2005). Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N. Engl. J. Med. *353*, 1659–1672.

Pisco, A.O., Brock, A., Zhou, J., Moor, A., Mojtahedi, M., Jackson, D., and Huang, S. (2013). Non-Darwinian dynamics in therapy-induced cancer drug resistance. Nat Commun *4*, 2467.

Quaranta, V., Tyson, Garbett, S., and Weidow, B. (2009). Trait variability of cancer cells quantified by high-content automated microscopy of single cells. Methods in

Enzymology *467*, 23–57.

Rando, O.J., and Verstrepen, K.J. (2007). Timescales of Genetic and Epigenetic Inheritance. Cell *128*, 655–668.

Roesch, A., Fukunaga-Kalabis, M., Schmidt, E.C., Zabierowski, S.E., Brafford, P.A., Vultur, A., Basu, D., Gimotty, P., Vogt, T., and Herlyn, M. (2010). A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. Cell *141*, 583–594.

Rogalla, T., Ehrnsperger, M., Preville, X., Kotlyarov, A., Lutsch, G., Ducasse, C., Paul, C., Wieske, M., Arrigo, A.P., Buchner, J., et al. (1999). Regulation of Hsp27 oligomerization, chaperone function, and protective activity against oxidative stress/tumor necrosis factor alpha by phosphorylation. J. Biol. Chem. *274*, 18947–18956.

Rosell, R., Carcereny, E., Gervais, R., Vergnenegre, A., Massuti, B., Felip, E., Palmero, R., Garcia-Gomez, R., Pallares, C., Sanchez, J.M., et al. (2012). Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. Lancet Oncol. *13*, 239–246.

Sakaue-Sawano, A., Kurokawa, H., Morimura, T., and Hanyu, A. (2008a). Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. Cell.

Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., Kashiwagi, S., Fukami, K., Miyata, T., Miyoshi, H., et al. (2008b). Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. Cell *132*, 487–498.

Sáez-Ayala, M., Montenegro, M.F., Sánchez-Del-Campo, L., Fernández-Pérez, M.P., Chazarra, S., Freter, R., Middleton, M., Piñero-Madrona, A., Cabezas-Herrera, J., Goding, C.R., et al. (2013). Directed phenotype switching as an effective antimelanoma strategy. Cancer Cell *24*, 105–119.

Sebaugh, J.L. (2011). Guidelines for accurate EC50/IC50 estimation. Pharm Stat *10*, 128–134.

Sequist, L.V., Waltman, B.A., Dias-Santagata, D., Digumarthy, S., Turke, A.B., Fidias, P., Bergethon, K., Shaw, A.T., Gettinger, S., Cosper, A.K., et al. (2011). Genotypic and Histological Evolution of Lung Cancers Acquiring Resistance to EGFR Inhibitors. Science Translational Medicine *3*, 75ra26–75ra26.

Sharma, S.V., Bell, D.W., Settleman, J., and Haber, D.A. (2007). Epidermal growth factor receptor mutations in lung cancer. Nature Reviews Cancer *7*, 169–181.

Sharma, S.V., Haber, D.A., and Settleman, J. (2010a). Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents. Nature Reviews Cancer *10*, 241–253.

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. (2010b). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell *141*, 69–80.

Sheiner, L.B., Stanski, D.R., Vozeh, S., Miller, R.D., and Ham, J. (1979). Simultaneous modeling of pharmacokinetics and pharmacodynamics: application to d-tubocurarine. Clin. Pharmacol. Ther. *25*, 358–371.

Singh, A.B., and Harris, R.C. (2005). Autocrine, paracrine and juxtacrine signaling by EGFR ligands. Cell. Signal. *17*, 1183–1193.

Singh, D.K., Ku, C.-J., Wichaidit, C., Steininger, R.J., Wu, L.F., and Altschuler, S.J. (2010). Patterns of basal signaling heterogeneity can distinguish cellular populations with different drug sensitivities. Molecular Systems Biology *6*.

Slack, M.D., Martinez, E.D., Wu, L.F., and Altschuler, S.J. (2008). Characterizing heterogeneous cellular responses to perturbations. Proc. Natl. Acad. Sci. U.S.a. *105*, 19306–19311.

Slamon, D.J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., et al. (2001). Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N. Engl. J. Med. *344*, 783–792.

Snijder, B., Sacher, R., Rämö, P., Damm, E.-M., Liberali, P., and Pelkmans, L. (2009). Population context determines cell-to-cell variability in endocytosis and virus infection. Nature *461*, 520–523.

Spencer, S.L., Cappell, S.D., Tsai, F.-C., Overton, K.W., Wang, C.L., and Meyer, T. (2013). The proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. Cell *155*, 369–383.

Spencer, S.L., Gaudet, S., Albeck, J.G., Burke, J.M., and Sorger, P.K. (2009). Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. Nature *459*, 428–432.

Spencer, S., and Sorger, P. (2011). Measuring and Modeling Apoptosis in Single Cells. Cell *144*, 926–939.

Su, Z., Dias-Santagata, D., Duke, M., Hutchinson, K., Lin, Y.-L., Borger, D.R., Chung, C.H., Massion, P.P., Vnencak-Jones, C.L., Iafrate, A.J., et al. (2011a). A platform for rapid detection of multiple oncogenic mutations with relevance to targeted therapy in non-small-cell lung cancer. J Mol Diagn *13*, 74–84.

Su, Z., Dias-Santagata, D., Duke, M., Hutchinson, K., Lin, Y.-L., Borger, D.R., Chung, C.H., Massion, P.P., Vnencak-Jones, C.L., Iafrate, A.J., et al. (2011b). A platform for rapid detection of multiple oncogenic mutations with relevance to targeted therapy in non-small-cell lung cancer. The Journal of Molecular Diagnostics : JMD *13*, 74–84.

Suda, K., Murakami, I., Katayama, T., Tomizawa, K., Osada, H., Sekido, Y., Maehara,

Y., Yatabe, Y., and Mitsudomi, T. (2010). Reciprocal and Complementary Role of MET Amplification and EGFR T790M Mutation in Acquired Resistance to Kinase Inhibitors in Lung Cancer. Clinical Cancer Research *16*, 5489–5498.

Taiwo, O., Wilson, G.A., Morris, T., Seisenberger, S., Reik, W., Pearce, D., Beck, S., and Butcher, L.M. (2012). Methylome analysis using MeDIP-seq with low DNA concentrations. Nat Protoc *7*, 617–636.

Takezawa, K., Pirazzoli, V., Arcila, M.E., Nebhan, C.A., Song, X., de Stanchina, E., Ohashi, K., Janjigian, Y.Y., Spitzler, P.J., Melnick, M.A., et al. (2012). HER2 amplification: a potential mechanism of acquired resistance to EGFR inhibition in EGFR-mutant lung cancers that lack the second-site EGFRT790M mutation. Cancer Discovery *2*, 922–933.

Tracy, S. (2004). Gefitinib Induces Apoptosis in the EGFRL858R Non-Small-Cell Lung Cancer Cell Line H3255. Cancer Research *64*, 7241–7244.

Tran, P.T., Bendapudi, P.K., Lin, H.J., Choi, P., Koh, S., Chen, J., Horng, G., Hughes, N.P., Schwartz, L.H., Miller, V.A., et al. (2011). Survival and death signals can predict tumor response to therapy after oncogene inactivation. Science Translational Medicine *3*, 103ra99.

Turke, A.B., Zejnullahu, K., Wu, Y.-L., Song, Y., Dias-Santagata, D., Lifshits, E., Toschi, L., Rogers, A., Mok, T., Sequist, L., et al. (2010). Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. Cancer Cell *17*, 77–88.

Tyson, D.R., Garbett, S.P., Frick, P.L., and Quaranta, V. (2012). Fractional proliferation: a method to deconvolve cell population dynamics from single-cell data. Nature Methods *9*, 923–928.

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.

Waddington, C. (1953). Genetic assimilation of an acquired character. Evolution *7*, 118–126.

Wang, L., Brugge, J.S., and Janes, K.A. (2011). Intersection of FOXO- and RUNX1-mediated gene expression programs in single breast epithelial cells during morphogenesis and tumor progression. Proceedings of the National Academy of Sciences *108*, E803–E812.

Webb, G.F., MC D'Agata, E., Magal, P., and Ruan, S. (2005). A model of antibiotic-resistant bacterial epidemics in hospitals. Proc. Natl. Acad. Sci. U.S.a. *102*, 13343–13348.

Weinstein, I.B., Joe, A., and Felsher, D. (2008). Oncogene Addiction. Cancer Research *68*, 3077–3080.

Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J., et al. (2012). The origin and evolution of

mutations in acute myeloid leukemia. Cell *150*, 264–278.

Yao, G., Lee, T.J., Mori, S., Nevins, J.R., and You, L. (2008). A bistable Rb-E2F switch underlies the restriction point. Nat. Cell Biol. *10*, 476–482.

Zhou, W., Ercan, D., Chen, L., Yun, C.-H., Li, D., Capelletti, M., Cortot, A.B., Chirieac, L., Iacob, R.E., Padera, R., et al. (2009). Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. Nature *462*, 1070–1074.